Sea surface temperature intercomparison in the framework of the Copernicus Climate Change Service (C3S)

Article

Yang, C., Leonelli, F. E., Marullo, S., Artale, V., Beggs, H., Bunogiorno Nardelli, B., Chin, T. M., De Toma, V., Good, S., Huang, B., Merchant, C. J. ORCID: https://orcid.org/0000-0003-4687-9850, Sakurai, T., Santoleri, R., Vazquez-Cuervo, J., Zhang, H.-M. and Pisano, A. (2021) Sea surface temperature intercomparison in the framework of the Copernicus Climate Change Service (C3S). Journal of Climate, 34 (13). pp. 5257-5283. ISSN 1520-0442 doi: 10.1175/JCLI-D-20-0793.1 Available at https://centaur.reading.ac.uk/97502/

To link to this article DOI: http://dx.doi.org/10.1175/JCLI-D-20-0793.1

Publisher: American Meteorological Society

Sea Surface Temperature Intercomparison in the Framework of the Copernicus Climate Change Service (C3S)

CHUNXUE YANG,[a] FRANCESCA ELISA LEONELLI,[a,b] SALVATORE MARULLO,[a,c] VINCENZO ARTALE,[a,c] HELEN BEGGS,[d] BRUNO BUONGIORNO NARDELLI,[e] TOSHIO M. CHIN,[f] VINCENZO DE TOMA,[c,g] SIMON GOOD,[h] BOYIN HUANG,[i] CHRISTOPHER J. MERCHANT,[j] TOSHIYUKI SAKURAI,[k] ROSALIA SANTOLERI,[a] JORGE VAZQUEZ-CUERVO,[e] HUAI-MIN ZHANG,[g] AND ANDREA PISANO[a]

[a] *Institute of Marine Sciences, National Research Council of Italy, Rome, Italy*
[b] *Department of Mathematics Guido Castelnuovo, University of Rome La Sapienza, Rome, Italy*
[c] *Italian National Agency for New Technologies, Energy, and Sustainable Economic Development (ENEA), Frascati, Italy*
[d] *Bureau of Meteorology, Melbourne, Victoria, Australia*
[e] *Institute of Marine Sciences, National Research Council of Italy, Naples, Italy*
[f] *Jet Propulsion Laboratory–California Institute of Technology, Pasadena, California*
[g] *Department of Physics and Istituto Nazionale di Fisica Nucleare, Tor Vergata University of Rome, Rome, Italy*
[h] *Met Office, Exeter, United Kingdom*
[i] *NOAA National Centers for Environmental Information, Asheville, North Carolina*
[j] *University of Reading, and National Centre for Earth Observation, Reading, United Kingdom*
[k] *Japan Meteorological Agency, Tokyo, Japan*

(Manuscript received 14 October 2020, in final form 17 March 2021)

ABSTRACT: A joint effort between the Copernicus Climate Change Service (C3S) and the Group for High Resolution Sea Surface Temperature (GHRSST) has been dedicated to an intercomparison study of eight global gap-free sea surface temperature (SST) products to assess their accurate representation of the SST relevant to climate analysis. In general, all SST products show consistent spatial patterns and temporal variability during the overlapping time period (2003–18). The main differences between each product are located in the western boundary current and Antarctic Circumpolar Current regions. Linear trends display consistent SST spatial patterns among all products and exhibit a strong warming trend from 2012 to 2018 with the Pacific Ocean basin as the main contributor. The SST discrepancy between all SST products is very small compared to the significant warming trend. Spatial power spectral density shows that the interpolation into 1° spatial resolution has negligible impacts on our results. The global mean SST time series reveals larger differences among all SST products during the early period of the satellite era (1982–2002) when there were fewer observations, indicating that the observation frequency is the main constraint of the SST climatology. The maturity matrix scores, which present the maturity of each product in terms of documentation, storage, and dissemination but not the scientific quality, demonstrate that ESA-CCI and OSTIA SST are well documented for users' convenience. Improvements could be made for MGDSST and BoM SST. Finally, we have recommended that these SST products can be used for fundamental climate applications and climate studies (e.g., El Niño).

KEYWORDS: Sea surface temperature; Climate records; Remote sensing; Satellite observations; Climate variability; Climate services

1. Introduction

Sea surface temperature (SST), as one of the essential ocean variables and essential climate variables, plays a crucial role in heat, freshwater, and momentum flux exchange at the ocean–atmosphere interface. The variation of SST at different temporal and spatial scales modulates the atmospheric lower boundary layer (e.g., Renault et al. 2019) eventually driving small- and large-scale changes at interannual to decadal time scales in the atmosphere (Perlin et al. 2014; McPhaden 2012). Additionally, the SST changes can influence the biogeochemical marine environment, contributing to modulating the primary production and related carbon absorption in the ocean (Behrenfeld et al. 2006). Besides its importance for assessing and monitoring the state of the global climate system, SSTs are widely used as boundary conditions in weather and climate operational forecast systems (Robinson et al. 2012) and as initial conditions in ocean operational forecast systems (Le Traon et al. 2019). Therefore, assessing the quality of SST data is critical from several perspectives, from operational to climate studies, marine environment, and related services.

SST observations are mainly obtained from low-Earth orbit infrared and microwave satellite imagery and geostationary infrared imagery, and from various in situ platforms including moored and drifting buoys, Argo floats, ships of opportunity, autonomous sailing drones, and radiometers (O'Carroll et al. 2019). All these instruments provide observations characterized by different representativeness, resolution, and accuracy. Different retrieval methods and reanalysis techniques are thus applied to obtain temporally and spatially consistent long-term SST products with global coverage (Minnett et al. 2019).

Denotes content that is immediately available upon publication as open access.

Corresponding author: Chunxue Yang, chunxue.yang@cnr.it

DOI: 10.1175/JCLI-D-20-0793.1

The Group for High-Resolution Sea Surface Temperature (GHRSST; www.ghrsst.org; Donlon et al. 2009) is an international initiative aimed at coordinating the provision of SST products developed and distributed by different agencies and research institutes. Among GHRSST products, level 4 data (L4) provide gap-free SST maps at regional and global scales, obtained with different algorithms that combine and interpolate satellite-based SST data, acquired by a variety of different sensors, sometimes also including in situ observations. Different interpolation techniques and related configurations (e.g., observation/background error correlation scales), interpolation grid size, input data bias correction, and the sampling adopted by GHRSST data providers induce a significant diversity among L4 SST products (Dash et al. 2012). Understanding the consistency and discrepancy of the different SST L4 products will not only help data providers to improve their algorithms, but also represents an important step to inform users about the characteristics of the different products, helping them to select the one that may better suit their applications.

Several previous global SST analysis intercomparison studies have already been performed, among which, most noticeably, the Global Climate Observing System (GCOS) SST–Sea Ice Intercomparison Project (https://www.nodc.noaa.gov/SatelliteData/ghrsst/intercomp.html), and the GMPE (Group for High-Resolution SST, GHRSST, Multi-Product Ensemble) system, performed as a contribution to GHRSST activities. The initial work by Martin et al. (2012) and Dash et al. (2012), which was focused on a relatively short time series over the satellite period (for the year 2010), has recently been extended to intercompare longer-term analyses over the overlapping period of 1991–2010 (Fiedler et al. 2019a). A much shorter period (one year) is considered in the intercomparison of satellite-based analyses performed by Okuro et al. (2014), while a comparison study on the historical SST datasets based on in situ data alone is described in Yasunaka and Hanawa (2011). With the recent reprocessing of several global high-resolution daily L4 products from the start of the operational satellite SST era (1981) to recent years, it is now timely to perform an intercomparison of additional SST analyses over a significantly longer period.

In the framework of the European Copernicus Climate Change Service (C3S), an independent assessment of essential climate variables (ECVs) present in the C3S Climate Data Store (CDS) is foreseen. The C3S CDS distributes and provides access to quality-assured climate datasets and tools in the clouds for users. The independent assessment aims to evaluate the quality, usability, and consistency of available ECVs for different applications, ranging from scientific studies (e.g., on climate change) to commercial and private sector uses. SST is one of the ECVs considered in the assessment framework of C3S and the intercomparison of SST products available in the CDS will help the users to understand the quality of different SST products and choose the right one for their specific applications.

The study presented hereafter represents the joint effort between the GHRSST SST Analysis Intercomparison Task Team (https://www.ghrsst.org/about-ghrsst/task-teams/) and the C3S SST assessment activities. The objective of this study is to evaluate the basic characteristics and the maturity of eight state-of-the-art global SST analysis products; to describe how SST climatology and variability is represented in each SST product; to understand the consistency and discrepancy between all these long-term eight SST analyses available in or outside of CDS (some of the SST products are provided in GHRSST L4 format); and eventually to provide guidance on which product might be better suited for users' applications.

The paper is organized as follows: section 2 introduces the characteristics of SST analysis products included in this study, the basic diagnostics are presented in section 3, the data maturity of all SST products is described in section 4, and finally, the summary of the evaluation and the recommendations to users are presented in sections 5 and 6.

2. Datasets

Currently, two global SST analysis datasets are distributed through the CDS, namely European Space Agency (ESA) Climate Change Initiative (ESA CCI) version 2.1 and the fifth-generation atmospheric reanalysis produced by the European Centre for Medium-Range Weather Forecasts (ERA5). They are compared here with a selection of six state-of-the-art SST analyses distributed outside the CDS, obtained from different input data and analysis system configurations:

- U.K. Met Office Hadley Centre Sea Ice and Sea Surface Temperature (HadISST1) (Rayner et al. 2003);
- U.K. Met Office Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system (Good et al. 2020)
- NOAA Daily OISST v2.1 daily reanalysis, also referred to as Reynolds SST (Reynolds et al. 2007; Banzon et al. 2016; Huang et al. 2020);
- Multiscale Ultrahigh Resolution 0.25° (MUR25) SST analysis v.4.2 (Chin et al. 2017);
- Merged satellite and in situ data Global Daily Sea Surface Temperature (MGDSST) (Sakurai et al. 2005; Kurihara et al. 2006);
- Australian Bureau of Meteorology Global Monthly SST Analysis (BoM Monthly SST) (Smith et al. 1999).

These eight datasets combine satellites and in many cases in situ temperature measurements to generate gap-free (optimally interpolated) SST fields at the global scale. All these datasets are specifically designed to provide accurate high spatial and temporal resolution SST estimates that can be used in operational applications such as assimilation and/or boundary conditions in numerical weather prediction models (e.g., MGDSST and OSTIA SST) and/or analyzed for climate applications (e.g., HadISST1, NOAA Daily OISST analysis, MUR25, BoM Monthly SST). Some of the selected datasets, namely ESA CCI v2.1, OSTIA, NOAA Daily OISST v2.1, MUR25, and BoM Monthly, are provided in GHRSST L4 format (GHRSST Science Team 2012).

Below, we detail the characteristics of all the SST products included in this intercomparison study (see Table 1).

a. ESA-CCI SST

The ESA CCI SST dataset (version 2.1) provides global daily SST estimates based on observations acquired from different satellite sensors covering the period from September 1981 to December 2018 (at the time of the study, including an extension produced by C3S). The CCI SSTs are designed to provide a stable, low-bias climate data record derived from different infrared sensors, namely the Advanced Very-High-Resolution Radiometer (AVHRR), Advanced Along Track Scanning Radiometer (AATSR), and Sea and Land Surface Temperature Radiometer (SLSTR) series of sensors (Merchant et al. 2019, 2014). These data are provided at different processing levels: single-sensor data on the native swath grid (Level-2); uncollated single-sensor (Level-3U) and collated multisensor (Level-3C) gridded data; and blended multisensor and optimally interpolated (Level-4) data.

The ESA CCI Level-4 product considered here consists of gap-free (optimally interpolated) maps of daily average SST at 20-cm depth on a 0.05° × 0.05° latitude–longitude grid (approximately 5 km × 5 km at the equator). The Level-4 data have been produced by running the OSTIA system (Donlon et al. 2012) using CCI Level-3U SSTs as inputs; no in situ data are included. Estimates of standard uncertainty (considered as the standard deviation of the estimated error distribution) are provided for every SST at all product levels. The evaluated global median uncertainty is 0.18 K (Merchant et al. 2019). The multiannual stability of the whole time series, evaluated relative to drifting buoy measurements, is within $0.003\,\mathrm{K\,yr^{-1}}$ (Merchant et al. 2019). Given the high temporal and spatial resolution and the performance statistics, this dataset gives an accurate representation of SST spatiotemporal variability of relevance to climate applications. Target applications of the ESA CCI SST dataset include climate and ocean model assessment; accurate definitions of climatic indices; and quantification of climate variability and its impacts on weather extremes (including marine heatwaves), marine ecosystems, and related services.

b. ERA5

The ERA5 SST dataset is produced by ECMWF to be used for ERA5 atmospheric reanalysis (Hirahara et al. 2016). It consists of hourly global gap-free SST data at a 0.25° × 0.25° latitude–longitude grid covering the period from 1979 to the present. ERA5 SST data are based on the HadISST2 (Titchner and Rayner 2014) product from 1979 to August 2007, and the daily operational OSTIA (Donlon et al. 2012) product from September 2007 to present. The HadISST1 version 2 was developed by the U.K. Met Office Hadley Centre, and its "pentad" dataset consists of spatially complete, 5-day mean fields on a 0.25° spatial resolution grid. OSTIA is a high-resolution (0.05° × 0.05°) operational daily product developed by the U.K. Met Office and distributed through the Copernicus Marine Environment Monitoring Service (CMEMS). These two SST datasets are aggregated into one continuous data record and interpolated onto the ERA5 model grid (Dee et al. 2011) to be used as boundary conditions for ERA5 atmospheric reanalysis. There are two types of sea surface temperature in ERA5: sea surface skin temperature and sea surface temperature. In this study we have used monthly ERA5 sea surface temperature. ERA5 SST is calculated as the SST from an ocean model with increment as the difference between OSTIA SST and the ocean analysis. Since the input of SST comes from both OSTIA and HadISST2, the ERA5 SST is a mixture of SST in the absence of diurnal variation, "foundation SST" (OSTIA), and SST at indeterminate depth, "SSTdepth" (HadISST2), following the SST definitions in Minnett and Kaiser-Weiss (2012). Here we give the SST type as SSTdepth for ERA5 SST.

c. HadISST1

Hadley Centre Sea Ice and Sea Surface Temperature dataset (HadISST1) is available at https://www.metoffice.gov.uk/hadobs/hadisst/data/download.html. This dataset includes a combination of monthly globally complete fields of SST and sea ice concentration on a 1° × 1° latitude–longitude grid from 1870 to present. HadISST1 data have been produced using SST measurements from the Met Office Marine Data Bank (MDB), mainly ship tracks, and a blend of in situ and adjusted satellite-derived SSTs for 1982 onward. A bias adjustment of the satellite SST data is performed by subtracting the in situ fields from the AVHRR fields. Specifically, the difference fields are smoothed using a moving window average with a radius of 2224 km (20° of latitude). The smoothed bias fields are then subtracted from the monthly AVHRR SST [see appendix C in Rayner et al. (2003) for further details].

To enhance data coverage, monthly median SSTs for 1871 onward from the Comprehensive Ocean–Atmosphere Dataset (COADS) (now ICOADS) were also used where MDB data were not available. Information on sea ice concentrations is also included in the HadISST product. This information is derived from several sources that include digitized sea ice charts and satellite data. Temperatures are reconstructed using a two-stage reduced-space optimal interpolation procedure (Kaplan et al. 1997), followed by superposition of quality-improved gridded observations onto the reconstructions to restore local detail (Rayner et al. 2003).

d. NOAA (Daily OISST)

The NOAA Daily OISST v2.1 dataset (Reynolds et al. 2007; Banzon et al. 2016; Huang et al. 2020), also known as the "Reynolds" Daily Optimum Interpolation SST analysis, consists of global daily spatially complete SST data on a 0.25° × 0.25° latitude–longitude grid from 1981 to present (https://www.ncdc.noaa.gov/oisst). This dataset is routinely produced by NOAA/NESDIS/NCEI and publicly provided at https://www.ncei.noaa.gov/data/sea-surface-temperature-optimum-interpolation/v2.1/.

GHRSST GDS2 L4 format (GHRSST Science Team 2012) files are also available from 1981 to 2015 from https://podaac.jpl.nasa.gov/dataset/AVHRR_OI-NCEI-L4-GLOB-v2.0 and 2016 to present from https://podaac.jpl.nasa.gov/dataset/AVHRR_OI-NCEI-L4-GLOB-v2.1.

The NOAA optimal interpolation analysis uses both in situ and satellite-derived SST data. Satellite SSTs are estimated from NOAA/AVHRR and MetOp/AVHRR observations. This dataset also utilizes the in situ ICOADS dataset to correct the residual satellite SST biases. OISST has been updated from v2.0 to v2.1 from January 2016 onward. The updates include the following five aspects: 1) *MetOp-B* replaces *NOAA-19* while *MetOp-A* remains unchanged, 2) freezing-point temperature replaces ice-SST regression in SST proxy in ice-covered oceans, 3) the estimated ship SST bias is reduced from 0.14° to 0.01°C, 4) ship and buoy observations from ICOADS-D

TABLE 1. Descriptive product comparison summary for the described products from section 2. Input observations are derived from satellite infrared (IR) and/or microwave (MW) sensors and/or in situ measurements.

Dataset	Institution	Type of product	Time range	Observation input	Type of SST	Horizontal grid spacing	Vertical resolution	Temporal resolution	Main reference
ESA CCI SST (v.2.0)	Met Office	SST analysis	1981–2018	IR	SST at 0.2 m	Global 0.05° × 0.05°	Surface	Daily	Merchant et al. (2019)
ERA5	ECMWF	SST analysis	1979–2018	IR + MW + in situ	SST depth	Global 0.25° × 0.25°	Surface	Hourly	Hirahara et al. (2016)
HadISST1	Met Office	SST analysis	1870–2018	IR + in situ	SST depth	Global 1° × 1°	Surface	Monthly	Rayner et al. (2003)
NOAA Daily OISST v2.1	NOAA	SST analysis	1981–2018	IR + in situ	SST at 0.2 m	Global 0.25° × 0.25°	Surface	Daily	Huang et al. (2020)
MUR25 (v.4.2)	PO.DACC	SST analysis	2003–18	IR + MW + in situ	Foundation SST	Global 0.25° × 0.25°	Surface	Daily	Chin et al. (2017)
MGDSST	Japanese Met. Agency (JMA)	SST analysis	1982–2018	IR + MW + in situ	SST depth	Global 0.25° × 0.25°	Surface	Daily	Sakurai et al. (2005)
BoM Monthly SST	Australian Bureau of Met. (BoM)	SST analysis	2002–18	IR + in situ	SST depth	Global 1° × 1° (weekly/monthly)	Surface	Weekly/monthly	Smith et al. (1999)
OSTIASST	U.K. Met Office	SST analysis	1981–2018	IR + MW + in situ	Foundation SST	0.05° × 0.05°	Surface	Daily/weekly/ monthly	Good et al. 2020

R3.0.2 are used instead of NCEP GTS receipts, and 5) Argo observations above 5-m depth are included. The Argo observations were first used as independent data to validate the improvements in the updates from 1 to 4, and the Argo observations were finally included in OISST in update 5.

e. MUR25

The Multiscale Ultrahigh Resolution 0.25° (MUR25) SST analysis (v.4.2) is a global daily spatially complete SST dataset on a 0.25° × 0.25° grid covering the period from mid-2002 to the present. The analyzed SST is representative of the foundation temperature (i.e., the temperature free, or nearly free, of any diurnal cycle; Minnett and Kaiser-Weiss 2012). This dataset is a reprocessed version of the MUR dataset v.4.1 (Chin et al. 2017), which provides global daily spatially complete SST analyses at 0.01° spatial resolution. MUR25 is provided by NASA's Jet Propulsion Laboratory (JPL) Physical Oceanography Distributed Active Archive Center (PO.DAAC) and is available at https://podaac.jpl.nasa.gov/dataset/MUR25-JPL-L4-GLOB-v04.2.

The MUR L4 analysis is built by using only nighttime SST observations derived from different types of satellite sensors, which include microwave and infrared measurements from, e.g., Advanced Microwave Scanning Radiometer (AMSR) for Earth Observing System (AMSR-E) and NOAA/AVHRR observations. In addition, MUR25 ingests in situ SST measurements from the NOAA iQuam dataset (Xu and Ignatov 2014) to improve the estimate of the foundation temperature, and ice concentration data from the EUMETSAT Ocean and Sea Ice Satellite Application Facility (OSI SAF), which are used for an improved SST parameterization in the polar regions. Satellite and in situ data are combined using MRVA, a meshless multiscale interpolation method that uses wavelets as basis functions in order to build the daily MUR SST analysis (Chin et al. 2017).

f. MGDSST

The merged satellite and in situ data global daily SST (MGDSST) analysis dataset provides global daily spatially complete SST fields on a 0.25° × 0.25° latitude–longitude grid covering the period from 1982 to the present. This dataset is derived from infrared satellite sensors (NOAA/AVHRR and MetOp/AVHRR), microwave satellite sensors (*Coriolis*/WINDSAT, *GCOM-W1*/AMSR-2), and in situ temperature measurements (from buoys and ships). This dataset is provided by the Japanese Meteorological Agency (JMA) and is available at https://ds.data.jma.go.jp/gmd/goos/data/rrtdb/jma-pro/mgd_sst_glb_D.html.

SSTs from the microwave sensor *Aqua*/AMSR-E are used in the analysis from May 2002 through 5 October 2011. In the reanalysis data, SSTs under sea ice are determined according to the statistical relation between sea ice concentration and SST. The lowest SST is −1.8°C where the sea ice concentration is 100%. Additional information is provided by Kurihara et al. (2006) and Sakurai et al. (2005).

g. BoM monthly

The Monthly Optimal Interpolation (OI) SST analysis is the global monthly spatially complete SST dataset on a 1° × 1° grid produced by the Australian Bureau of Meteorology (BoM), covering the period of 1994 to present (Smith et al. 1999), formed by averaging the BoM Weekly OI SST analyses over each month. In this study, we use the GHRSST version 1 L4 format files of this dataset covering the period from 2002 to present (Beggs and Pugh 2009). The SST observations are obtained from in situ SST observations from drifting and moored buoys, ships, Argo floats, conductivity–temperature–depth (CTD) and expendable bathythermographs (XBTs), and satellite-derived SST from infrared AVHRR sensors aboard NOAA Polar-Orbiting Environmental Satellites (POES) and ESA/EUMETSAT MetOp satellites. Weekly OI analyses of the in situ data are used to correct for biases in the satellite data (Smith et al. 1999), similar to the method used in the NOAA Weekly 1° × 1° OISST v2 (Reynolds et al. 2002). The resulting outputs of the weekly and monthly OI analyses of in situ and satellite data are therefore SST values of indeterminate depth, referred to as SSTdepth.

At high latitudes, the BoM weekly analysis system uses the daily sea ice concentration analysis from NOAA/NCEP (https://polar.ncep.noaa.gov/seaice/Analyses.shtml) to constrain the SST, by setting SST at a given grid point to −1.8°C if the concentration of NCEP ice data in that grid cell is greater than 50%. Until 12 March 2008, the 0.5° resolution sea ice analysis was used and after that date, the 1/12° resolution sea ice analysis (Grumbine 1996).

Maps of these weekly and monthly SST analyses are available at http://www.bom.gov.au/marine/sst.shtml, and they are used operationally by BoM to generate El Niño indices, monitor the Indian Ocean dipole, and produce SST anomaly maps for climate applications (http://www.bom.gov.au/climate/enso/#tabs=Sea-surface). The BoM Weekly and Monthly OI SST analysis GHRSST L4 format files are available on request (http://www.bom.gov.au/climate/data-services/data-requests.shtml). It should be noted that higher-resolution (0.25° × 0.25°) global daily OI SST analyses have been produced operationally at the Bureau of Meteorology since 2008 (Zhong and Beggs 2008; http://www.bom.gov.au/marine/sst.shtml) but these only cover the period from 2008 to the present and so were not included in this study.

h. Met Office OSTIA SST

The U.K. Met Office OSTIA (Good et al. 2020) system is a daily global SST product with a grid resolution of 1/20° (approximately 5–6 km). The version of OSTIA SST we use in this study is the CMEMS reprocessed SST analysis based on the OSTIA configuration reported in Good et al. (2020), covering the period 1 October 1981–31 December 2018. This OSTIA reanalysis is formed by the combination of satellite SST data provided by the GHRSST project with additional AATSR, SLSTR, and AVHRR data from ESA CCI SST v2.0, C3S projects, and in situ observations from the HadIOD by using NEMOVAR, a variational assimilation (Fiedler et al. 2019b), instead of the optimal interpolation algorithm originally used to generate OSTIA (Martin et al. 2007; Donlon et al. 2012). Note that ESA CCI SST v2.0 and V2.1 only differ in the file specification, but no scientific differences. Bias correction is performed for all the input satellite data (except the satellite data in the reference dataset) by carrying out match-ups between satellite and reference measurements. The depth of the SST analysis represents the subskin

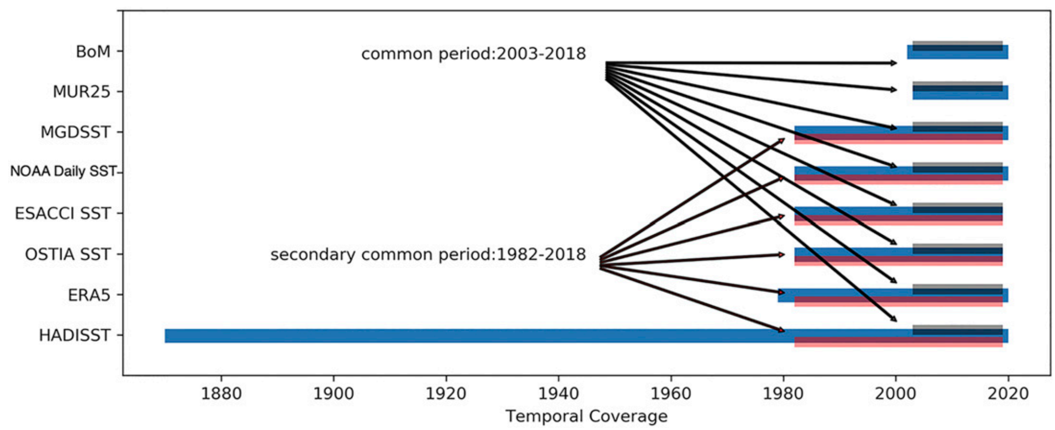

FIG. 1. Temporal range (years) covered by each SST dataset. The common period for all datasets is highlighted (2003–18), and the secondary common period is 1982–2018 with fewer SST products included.

temperature immediately before sunrise also referred to as foundational SST that is free of diurnal variability (Donlon et al. 2012). The OSTIA reanalysis is publicly available from https://resources.marine.copernicus.eu/?option=com_csw&task=results?option=com_csw&view=details&product_id=SST_GLO_SST_L4_REP_OBSERVATIONS_010_011.

To verify the accuracy of the reprocessed SST analysis, near-surface Argo data, which are not included in SST analysis are used as independent data for quality assessment as shown in CMEMS quality information documentation of OSTIA SST (https://resources.marine.copernicus.eu/documents/QUID/CMEMS-SST-QUID-010-011.pdf). Note that the drifting buoy SSTs used for validation are ingested into the analyses; however, the validation process uses OSTIA background fields to compare with drifting buoys from analysis day plus 1 day to avoid the validation data independence issue.

OSTIA SST has been used as boundary conditions for operational forecast models at the U.K. Met Office and European Centre for Medium-Range Weather Forecasting (ECMWF) and is also part of the CMEMS project. The validation, assessment activities update regularly through the CMEMS project, the data, and relevant documentations are available at https://resources.marine.copernicus.eu/?option=com_csw&view=details&product_id=SST_GLO_SST_L4_REP_OBSERVATIONS_010_011.

3. Basic diagnostics

To compare the selected datasets (see section 2) especially against global SST climatology, all the SST products need to be mapped on a common temporal and spatial resolution (regular $1° \times 1°$ latitude–longitude grid.). Apart from HadISST1, the majority of the SST products have higher resolution than $1° \times 1°$ and the advantage of high resolution is to resolve small-scale ocean processes. The interpolation from higher resolution to low resolution may exclude the impacts of important small-scale signals in the SST products. Before we present the basic diagnostics such as mean climatology and variability, we have performed spatial spectral analysis [see section 3a(1) for methods and section 3b(1) for results] to quantify the impact of interpolation to the common $1° \times 1°$ resolution we have performed in our basic diagnostics.

The grid of HadISST1 has been chosen as the reference grid (at $1° \times 1°$ nominal resolution). The HadISST1 land–sea mask has then been applied to all products. In addition, a sea ice mask was built from HadISST1 and used as a common sea ice mask for all datasets.

To homogenize the datasets' temporal and spatial resolution we have used CDO (Climate Data Operator) command line operators (see the user guide at https://code.mpimet.mpg.de/projects/cdo/embedded/cdo.pdf). In particular, we have chosen a bilinear interpolation for gridding all datasets on the HadISST1 spatial grid.

For all the selected SST products, the overlapping period is 2003–18 (Fig. 1) and the intercomparison of all SST products is performed for the period 2003–18, when observations are abundant compared to the beginning of the satellite era. Recent period increased quantities of observations ingested in the SST analysis may reduce the spread of ensemble SST products produced with different algorithms. To understand more deeply the discrepancy and consistency between all the SST analyses produced with different algorithms, similar intercomparison diagnostics of SST products (ESA-CCI, ERA5, OSTIA, NOAA OISST, MGDSST, and HadISST1) that have the common period of 1982–2018 (Fig. 1) are also carried out for the earlier period of the satellite era (1982–2002) when the observations are scarce compared to the later period of the satellite era.

In this section, we first introduce the methodologies we applied to produce the basic diagnostics, and the spatial spectral analysis method used to investigate the impact of spatial resolution is also presented. Then we present the results generated by these diagnostics in terms of intercomparison for the period 2003–18, and the intercomparison of SST products that cover the period 1982–2002 is presented at the end of this section.

a. *Statistical methods*

A set of basic diagnostics have been defined to evaluate the similarity and disagreements between selected SST datasets, as detailed in the following subsections. Some of these metrics, such as the mean climatology, quantify the long-term mean spatial distribution (climatology) of the SST for each single dataset and can be used to qualitatively evaluate the capability of SST in representing the climatological spatial patterns and the temporal variability of globally averaged SSTs. Other metrics, such as difference, root-mean-square difference (RMSD), and correlation, measure the distance between a single product and a "reference." The latter can be either a previously validated dataset (if available) or any other dataset that is arbitrarily chosen as reference. In this report, we have taken the median of all datasets (hereafter the "ensemble median") as a reference and used it to measure the difference among different SST products. Finally, we choose a specific case study of the El Niño–Southern Oscillation (ENSO) Niño-3.4 index to evaluate the capability of representing ENSO events in all SST products. Niño-3.4 is the average SST anomaly in the region 5°N–5°S, 170°–120°W.

1) Spatial spectral analysis

The spectral analysis method we adopted in this study is the multitaper power spectral density estimate (MTM) (Thomson 1982), which is a very useful tool for the analysis of relatively short and noisy series that may contain both broadband and line components. Different from several other techniques, MTM multiplies the data by a small set of orthogonal tapers rather than a single taper to minimize the spectral leakage due to the finite length of the series.

MTM power spectral estimates were performed using the pmtm matlab function (https://www.mathworks.com/help/signal/ref/pmtm.html). For more details please refer to Ghil et al. (2002, section 3d).

We have chosen four datasets, ESA-CCI and OSTIA with the original spatial resolution of 0.05° and MGDSST and NOAA Daily OISST (Reynolds 0.25° × 0.25° SST) with the original resolution of 0.25° all covering the same period 1982–2018 with daily frequency. Meanwhile, we chose the Pacific equator pixel line, spanning from Indonesia to South America as the study region (0°N, 120°E–80°W). For each dataset the spatial power spectral density has been estimated on a daily basis over the common period (1982–2018) and then time averaged. The detailed results and discussion are given in section 3b(1).

2) Trend analysis

SST trends have been estimated by using the X-11 seasonal adjustment procedure (see, e.g., Pezzulli et al. 2005). Given X_t as the input time series (i.e., an SST time series), the X-11 procedure generates the following decomposition:

$$X_t = T_t + S_t + I_t,$$

where T_t is the trend component, S_t the seasonal component, and I_t the irregular component, which accounts for the residual irregular variations such as subannual fluctuations. The decomposition is obtained through iterative application of different running means, which have the effect of a low-pass filter for T_t estimation and a seasonal filter for S_t estimation.

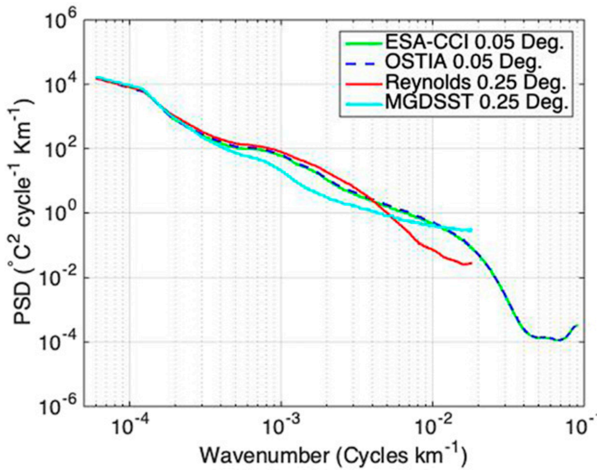

FIG. 2. Power spectral density at the equator in the Pacific Ocean (0°N, 120°E–80°W) for ESA-CCI (green), OSTIA (dashed dark blue), NOAA Daily OISST (Reynolds 0.25°; red), and MGDSST (cyan) based on the daily temporal and original spatial resolution for the period 1982–2018.

In addition, the Mann–Kendall test is used to assess whether a monotonic upward or downward trend in T_t exists (against the null hypothesis of no trend), Sen's method is applied to estimate the slope of T_t (i.e., the trend) as the median of the slopes of all pairs of sample points), and a bootstrap procedure is used to estimate the 95% confidence interval of the trend (Mann 1945; Sen 1968; Kendall 1975; Efron and Tibshirani 1993).

b. Results

1) Spatial spectrum analysis

To verify the suitableness of our choice of interpolation, we have performed spatial power spectral analysis [section 3a(1)] based on the chosen SST products (Fig. 2). With rapid growth of computing power and storage capacity, along with advancement of scientific knowledge and users' needs, spatial resolution of SST gap-free analyses has increased dramatically to resolve smaller-scale features in the ocean. The spatial resolution of SST products used in this study spans from 1° to 0.05°, meaning that the highest resolution is 20 times the lowest resolution. In the high-resolution SST products, the mesoscales might be resolved, by contrast in the low-resolution SST products only large-scale features are represented.

All of the SST products we chose for the spectral analysis cover the same period from 1982 to 2018 with daily frequency. OSTIA and ESA-CCI SST have the original spatial resolution of 0.05° and MGDSST and NOAA Daily OISST have the spatial resolution of 0.25°. If the power spectra gradient becomes flat at a certain wavelength it means that the analysis carried out at a wavelength shorter than this certain wavelength contains only noise. The power spectrum density of these four datasets shows that even though all of these SST products have higher grid resolution than the chosen common grid, 1°, the power density of all SST products starts to decline at spatial wavelengths greater than their grid resolution. The

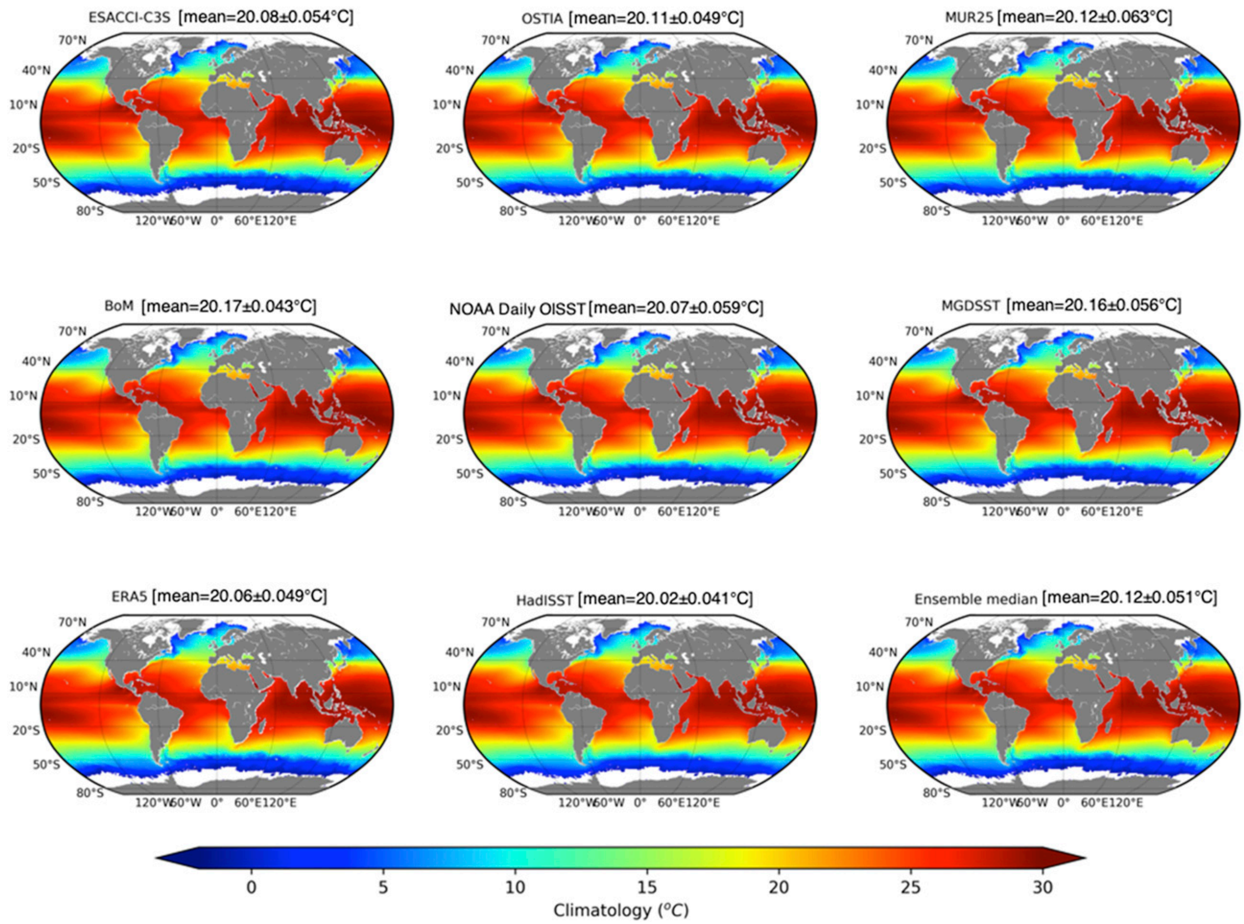

FIG. 3. Global SST climatologies for the period 2003–18. Global SST average value and its 95% confidence interval is also shown.

prominent differences between NOAA OISST and MGDSST are most likely due to different background correlation length scales being used in the optimal interpolation and different methodology used to correct satellite-based observations. For high-resolution datasets, the 0.05° products, the power density significantly declined after ~100 km (wavenumber 10^{-2}), which is close to 1° spatial resolution near the equator and the gradient becomes flat at wavelengths ~70 km. It means that the signals within a wavelength of 100 km are noise, with no physical meaning in 0.05° SST products, and that also applies to 0.25° resolution SST products. Similar results were shown in Fiedler et al. (2019a): in the Gulf Stream regions for the 2017 northern winter the spectral density of SST starts to depart from the $k^{-11/3}$ cascade of SST field [equivalent to kinetic energy power spectrum cascade of $k^{-5/3}$ based on Le Traon et al. (1990, 2008)] at wavelengths around 90 km. This confirms that the interpolation to 1° does not undermine the interpretation of results presented in our study.

Additionally, the diagnostics performed in the following sections mainly focus on the general features (mean climatology and long-term temporal variability) of the representation of all the SST products. We believe the interpolation of all SST products to 1° brings minor issues to the interpretation of the results. Certainly, the intercomparison between all the SST products in terms of specific details; for example, the representation of the Gulf Stream and mesoscale features are not in the scope of this study. Related activities are underway and will be presented by the GHRSST SST Analysis Intercomparison Task Team in the near future.

2) MEAN AND VARIABILITY (2003–18)

In terms of the basic diagnostics, we have first calculated the mean climatology of the global SST distribution of the eight selected SST datasets during 16 years from 2003 to 2018 plus the median of all the eight SST products (i.e., the climatology of the ensemble median) (Fig. 3). In all eight cases, the average correctly reveals the dominant latitudinal spatial SST pattern: higher at the tropics, milder at middle latitudes, and lower in the polar regions. Regions impacted by occasional or persistent presence of sea ice are flagged; that is, only complete years have been considered for the average estimate in each grid point.

A first qualitative inspection of the eight mean SST fields suggests that all products reproduce a very similar spatial distribution of SST with minor differences not appreciable from Fig. 3. Considering a confidence level of 95%, the eight global mean SST estimates for the period 2003–18 range in an interval

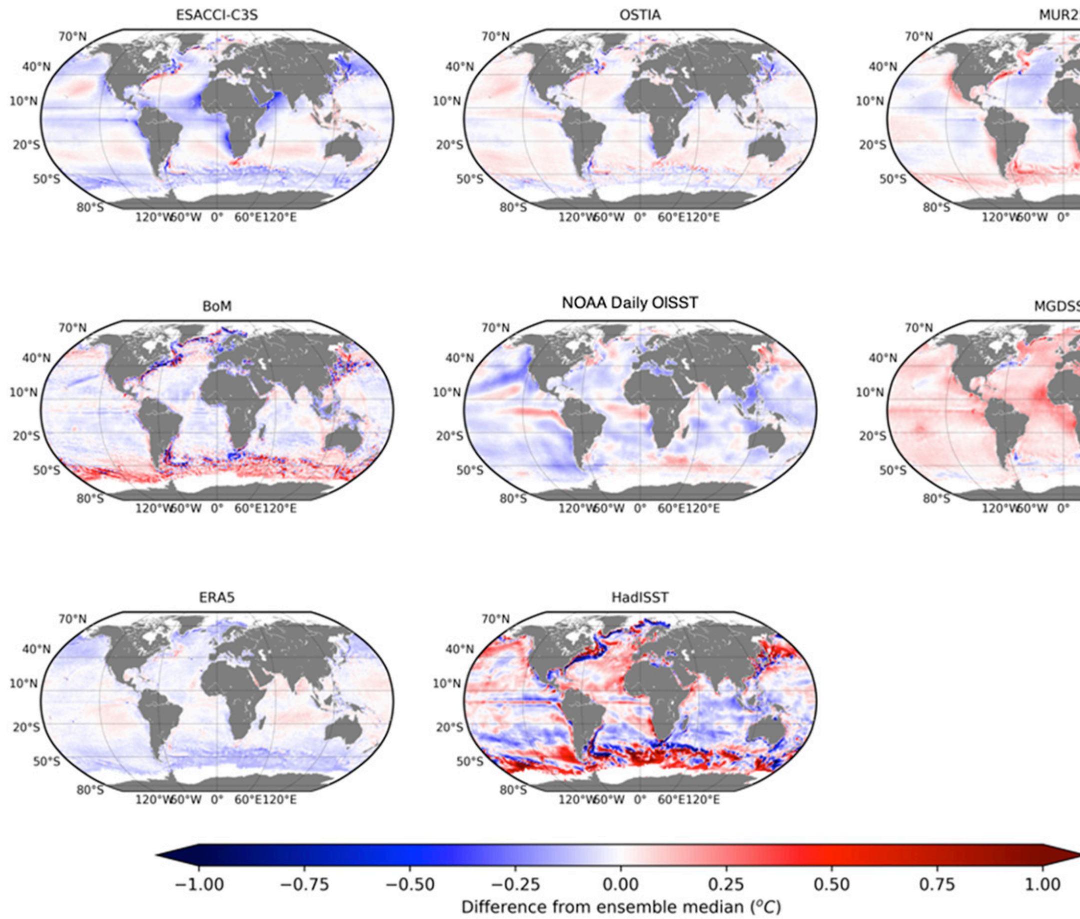

FIG. 4. The difference between each SST product and the ensemble median for the period of 2003–18.

between 20.02° and 20.17°C. The ensemble median obviously falls close to the middle of this range (i.e., 20.12°C).

To have a further investigation of the consistency and discrepancy between all SST products, we calculated the difference between each SST product and the ensemble median displayed in Fig. 4. Considering a 95% confidence interval, the global mean difference between each single product and the ensemble median ranges between −0.05° and 0.1°C with relevant spatial variability (Fig. 4). In fact, differences are more pronounced in the Southern Ocean where distances between single product values and the ensemble median reach values higher than 1°C. This is particularly evident in the case of HadISST1 data. In general, higher difference areas correspond to the western boundary current systems such as the Gulf Stream Current, the Kuroshio in the Northern Hemisphere, the Brazil currents in the southern Atlantic Ocean, and the Antarctic Circumpolar Current (ACC), where eddies are extremely active. In some datasets, especially ESA-CCI SST, MGDSST, and OSTIA, the greatest differences from the ensemble median are also located within eastern boundary currents which represent the main upwelling systems (e.g., Peru–Chile Current, the Benguela Current, along the northern West African coast, and along the southern Saudi Arabia coast). These discrepancies could be due to mismatch in the position of the main streams, especially the eddy representation in different SST products. Along the coast, the disagreement may come from the interpolation methodology implemented in different SST datasets by data providers. Especially regions where upwelling is active add difficulties to retrieving satellite observations for representing SST patterns and variability. For the case of ESA CCI SSTs, it has been shown that cool biases off the northern West African coast and in the Arabian Sea arise from influences of mineral dust aerosol on IR retrievals of SST, and a large-scale adjustment (not used here) for the dust-related biases has been devised (Merchant and Embury 2020).

The RMSD is defined as the square root of the average squared difference between the SST value of each dataset and the ensemble median, which is an absolute measure of the distance between each single product value and the ensemble median. Considering the 95% confidence interval, the global average RMSD ranges from 0.02° to 0.18°C. Extreme RMSD values (Fig. 5) are concentrated in the Southern Ocean and correspond to the ACC, as also evidenced by the mean difference (Fig. 4), particularly evident in HadISST1 data. These higher RMSD values are also observed in correspondence to large differences between each SST product and the ensemble median that are mainly located in the western boundary

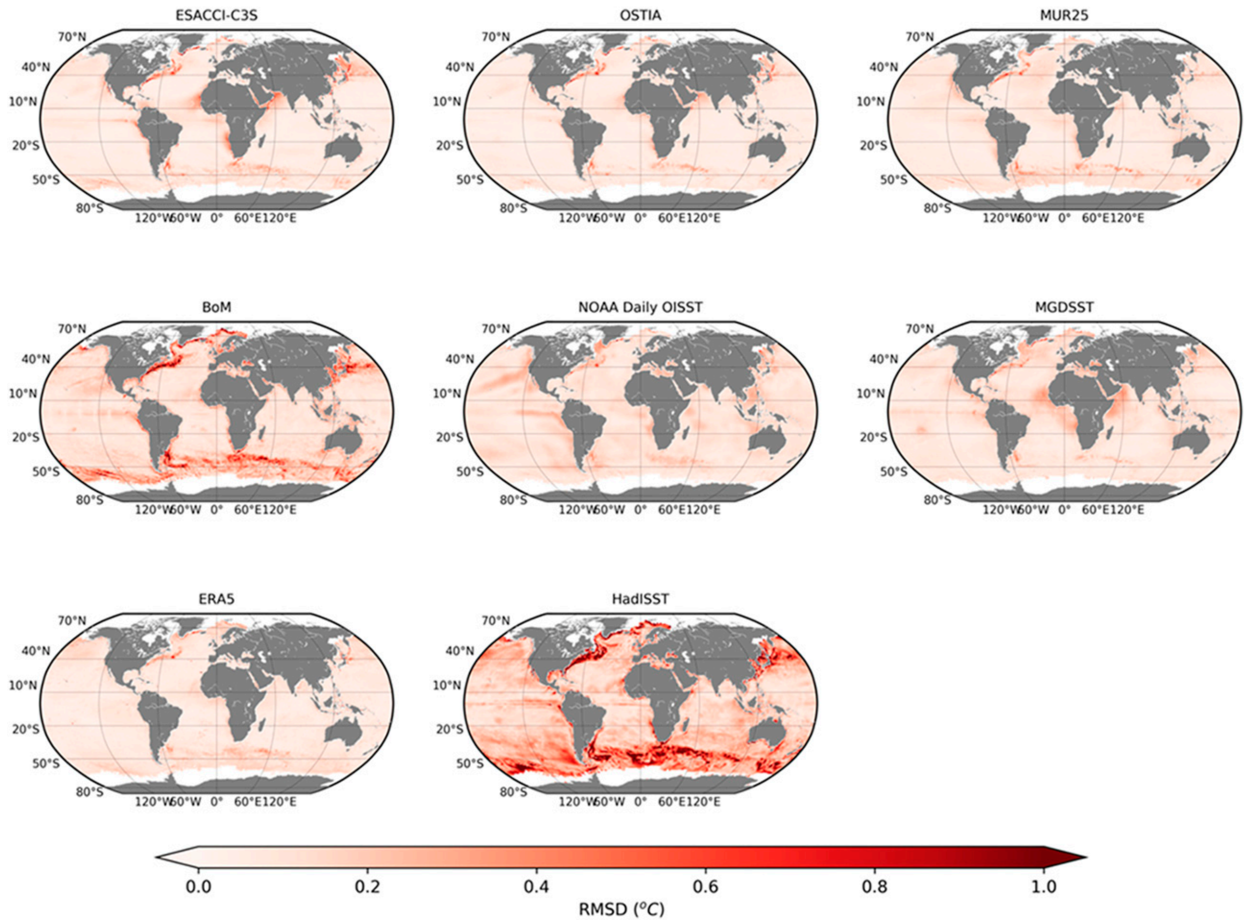

FIG. 5. The RMSD between each SST product and the ensemble median for the period of 2003–18.

currents, namely the Gulf Stream in the North Atlantic Ocean and the Kuroshio in the North Pacific Ocean, and the ACC regions.

The spatial distribution of the Pearson correlation coefficient (Fig. 6) highlights the different behavior of HadISST1 with respect to the other seven products. In particular, in the Southern Ocean region, the correlation falls down to values as low as 0.5 or even less. Similar but less extended discrepancies are also observed for BoM, NOAA Daily OISSTs, ESA-CCI, MUR25, ERA5, OSTIA, and MGDSST. In particular, ESA-CCI seems well representative of the ensemble median. MUR25, ERA5, MGDSST, and OSTIA are well representative of the ensemble median as well but with slightly higher discrepancies than other SST products. However, the low correlation especially along the coastal regions could be due to the interpolation method adopted during the SST production by data providers because it is still a challenge to correctly retrieve satellite observations at the coastal upwelling regions where SST is highly variable.

The temporal variability of globally averaged monthly mean SSTs (Fig. 7) clearly exhibits the annual oscillation around the mean value of 20.12°C (Fig. 3). This oscillation has an amplitude of about 0.6°C as a result of the opposite seasonal cycle in the Southern and Northern Hemispheres. SST anomalies from 2003 to 2018 (Fig. 8) are obtained by subtracting from all SST products the annual cycle of the ensemble median (i.e., the mean of each month over the whole period, 2003–18). Two main periods are observed with distinct mean values: the first period before 2012 where the temperature oscillates around a constant mean value of about 20.1°C and a second period where a positive (warming) trend is observed. All the eight datasets show temperatures that vary coherently over all time scales but with relative absolute biases in the range from zero to 0.4°C.

3) GLOBAL LINEAR TRENDS (2003–18)

Global SST trend maps have been computed for each product over the common 16-yr period from 2003 to 2018 (Fig. 9). All the datasets exhibit a global mean warming SST trend ranging from 0.012°C yr^{-1} (HadISST1) to 0.022°C yr^{-1} (MGDSST), with an average value of 0.019°C yr^{-1} (ensemble median). Within the 95% confidence interval, these results are close to the global ocean warming trend of 0.011°C yr^{-1} from 1980 to 2005 reported in the last IPCC report (Pachauri et al. 2014) and the differences are due to the different calculating period. The prominent warming trends shown in all SST

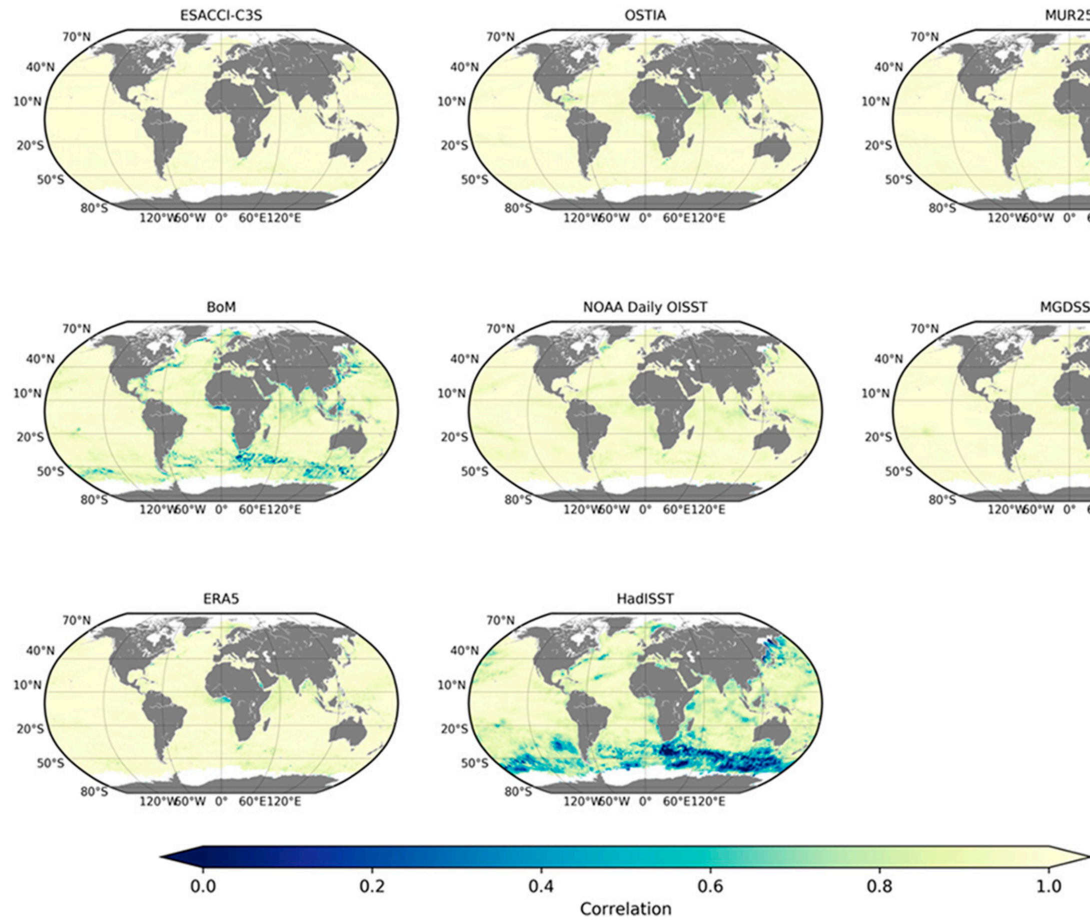

FIG. 6. The correlation between each SST product and the ensemble median for the period of 2003–18.

products are located in the subtropical North Atlantic Ocean, south Indian Ocean, and eastern tropical Pacific Ocean close to the American continent. Especially at the Gulf Stream area all SST products (apart from HadISST1, which has slightly weaker signals compared to other datasets) exhibit distinct warming trends for the period of 2003 to 2018.

In the North Atlantic Ocean, between 40° and 70°N, negative trends are observed in the subpolar gyre region extending up to the coastal areas of Ireland. A second common negative trend area is present in the Southern Ocean at longitudes centered around the Drake Passage. In the tropical Atlantic Ocean, a large area of negative trends is observed only in ERA5 and a smaller area in BoM, OSTIA, and HadISST1. For all the other products this area is characterized by no significant trends (i.e., areas where, given the $p = 0.05$ limit, the null hypothesis cannot be refuted) with few sparse negative trend points.

The Mediterranean Sea shows an evident positive trend in all products in contrast with a close to zero trend region in the adjacent northeast Atlantic Ocean. This is in agreement with what was recently published by Pisano et al. (2020), who observe that, after 1990, SST in the Mediterranean Sea continues to increase in contrast with the adjacent areas of the Atlantic Ocean where a pause of the general warming trend occurred. The larger area of positive SST trends is present in the Indian Ocean. Intense (positive) trends cover more uniformly and densely the reddish areas in ESA CCI, MUR, NOAA OISST, and MGDSST data, while a more patchy and less intense positive trend coverage is observed in ERA5, BoM, OSTIA, and HadISST1 data.

Besides a bias that separates the curves by a maximum of 0.2°C, the trend component of the eight spatially averaged global SST time series (Fig. 10a), obtained using the X-11 procedure with a 2-yr low-pass filter [section 3a(2)], shows a very similar behavior for all the products. The time evolution of the trend component reveals an apparently neutral period until 2011 included with a single maximum centered on the year 2009. After this period, a continuous warming phase is observed with an increase of the temperature of nearly 0.3°C, that is, about $0.06°\mathrm{C\,yr^{-1}}$, which is consistent with the signal observed in the time series anomalies (Figs. 7 and 8).

To understand better the contribution to the significant warming trends for the period of 2012–18 observed in all SST products, we have calculated the SST trend component in different ocean basins, the Pacific Ocean (Fig. 10b), Atlantic Ocean (Fig. 10c), and Indian Ocean (Fig. 10d). Quantitatively, the warming trends for the period 2012–18 ranges from $0.036°\mathrm{C\,yr^{-1}}$ (BoM) to $0.062°\mathrm{C\,yr^{-1}}$ (MUR25) with

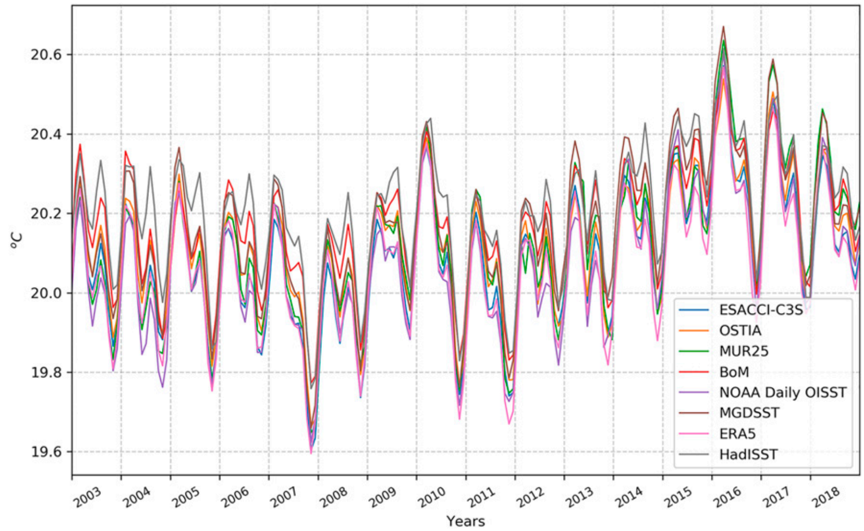

FIG. 7. Global monthly mean SST time series from 2003 to 2018.

0.049°C yr^{-1} in the ensemble median. The major contributor to this warming trend comes from the Pacific Ocean where warming trends span from 0.045°C yr^{-1} (BoM) to 0.084°C yr^{-1} (MUR25) with 0.064°C yr^{-1} in the ensemble median. The contribution from the Atlantic (0.02°C yr^{-1} from BoM to 0.52°C yr^{-1} from MUR25) is smaller compared to the Pacific Ocean, and the warming trends in the Indian Ocean from 2012 to 2018 are relatively very small (from 0.002°C yr^{-1} for MGDSST to 0.030°C yr^{-1} for BoM), which are evidently exhibited in Fig. 10d.

4) INTERCOMPARISON DURING THE EARLY PERIOD (1982–2002)

In this section, we present the intercomparison of all SST products covering the period 1982–2002. First we have shown the global mean SST time series (Fig. 11) that covers the time period originally obtained in each SST product allows us to detect the consistency and disagreement between all SST products for a longer period to fully take advantage of SST products which covers the period beyond 2003 and 2018. As we have discussed, all the SST products are very similar to the period of 2003–18 when there are abundant observations. On the contrary, during the period of early satellite era (1982–2002), the disagreement between all the SST products is larger compared to the later period (2003–18), which may be due to fewer observations ingested in the SST analysis.

To quantify the consistency and discrepancy of SST products for the early satellite era (1982–2002) we have calculated the mean climatology (Fig. 12) for all SST products which cover the period back to 1982 (Fig. 1), including ESA-CCI, OSTIA, ERA5, NOAA OISST, MGDSST, and HadISST1

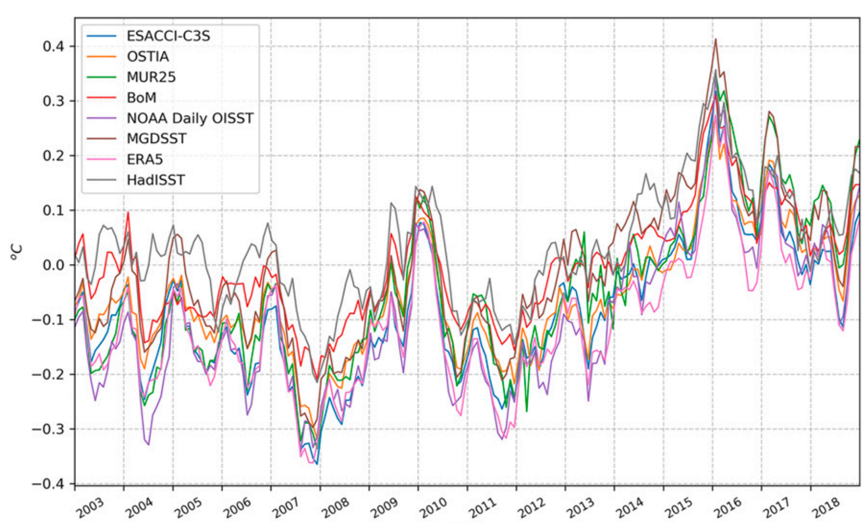

FIG. 8. Global SST monthly anomalies time series, obtained by subtracting the climatology of the ensemble median from all the SST ensemble members from 2003 to 2018.

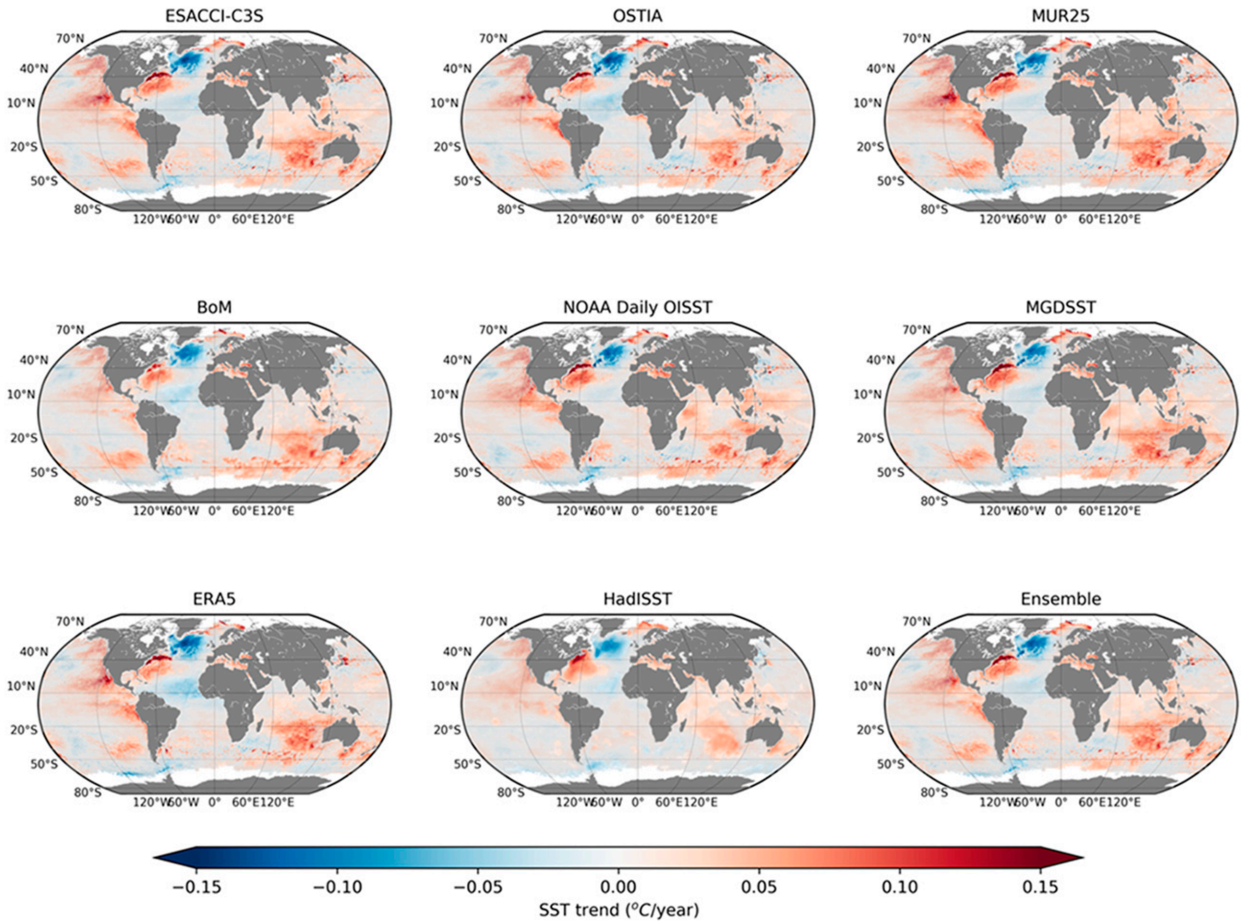

FIG. 9. Global linear trend maps (2003–18) (°C yr^{-1}) of each ensemble member and ensemble median. Areas with no significant (95% significance level) trends are covered by gray points.

and the differences between each member with the ensemble median (Fig. 13). The mean climatology of SST during the period of 1982–2002 spans the range from 19.76°C (NOAA OISST) to 20.05°C (HadISST1) with the ensemble median as 19.79°C. The differences of each member relative to the ensemble median for the period of 1982–2002 range from 0.03° to 0.26°C that is much higher than those during the period of 2003–18, which range from 0.01° to 0.1°C. The discrepancies of all SST products (Fig. 13) are located in the areas that are similar to the period of 2003–18 (Fig. 4), but with amplified signals. However, in some SST products, the differences relative to the ensemble median change signs. For example, during the period of 2003–18 the MGDSST mean climatology is higher than the ensemble median in the eastern Indian Ocean. On the contrary, the mean climatology differences between MGDSST and the ensemble median became negative during the period of 1982–2002. ERA5 SST is based on OSTIA SST; however, there are differences between them because ERA5 is forced by SST from an ocean model with increments based on the difference between ocean analysis and OSTIA, which contains information from the OSTIA SST but is not exactly identical.

These results are consistent with what is shown in Fig. 11, namely that during the early period of the satellite era (1982–2002; fewer SST observations) all the SST products have larger differences compared to the later period (2003–18, more SST observations), indicating that observation number is the main factor to constrain the climatology of all the SST products developed with different algorithms. The total number of valid in situ SST observations from drifting buoys, ships, Argo floats, and moorings, used for bias-correcting satellite SST ingested into ERA5, HadISST1, OSTIA, Daily OISST, and BoM Monthly, indeed increases over time (Xu and Ignatov 2014; https://www.star.nesdis.noaa.gov/socd/sst/iquam).

In 2002, the microwave radiometer AMSR-E, which measures ocean brightness temperatures through clouds, commenced operation on the *Aqua* satellite. This improvement in spatial coverage is another important factor affecting SST data ingested into OSTIA, ERA5, MGDSST, and MUR25, and it is notable that all SST products studied converge more after 2003 compared to before 2003.

5) NIÑO-3.4 INDEX

To have a deeper evaluation of the quality of the SST for climate studies, we investigated the capability of representing the climate modes in all SST products for the period of 1982–2018 in order to include more ENSO events, here the Niño-3.4 index

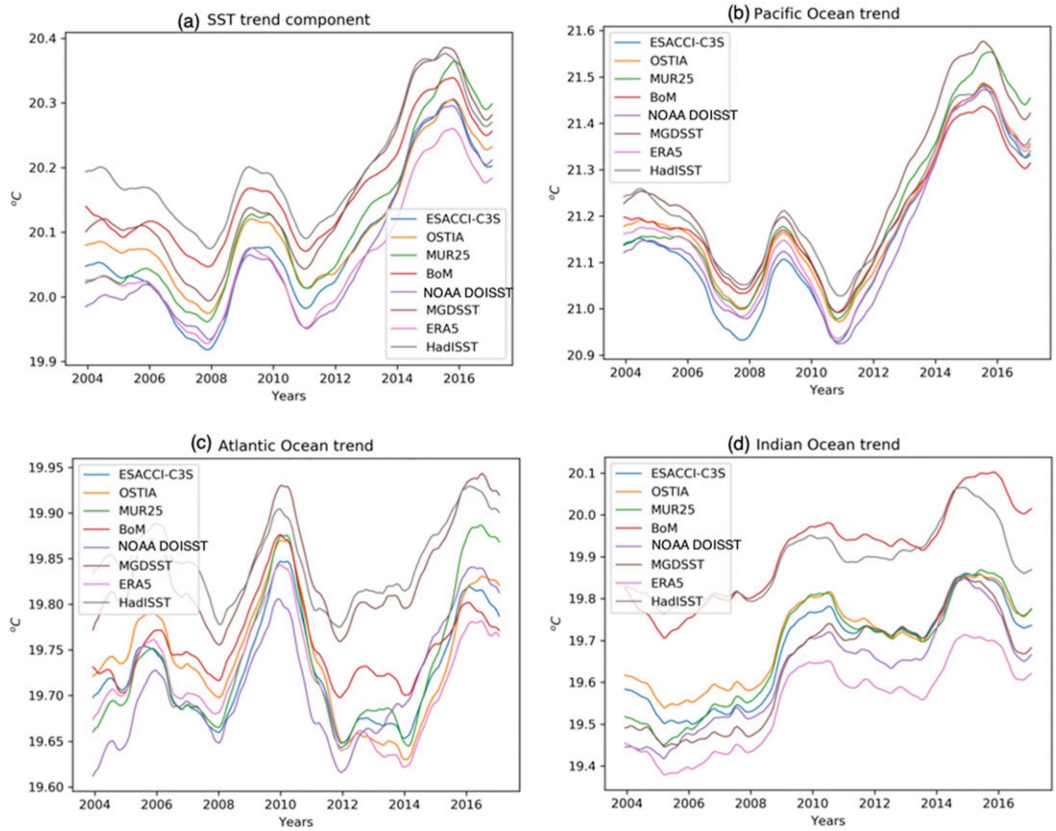

FIG. 10. (a) Global average SST trend component deduced from the global average monthly mean time series (Fig. 7) using the X-11 procedure [section 3a(2)], and the same calculation, but for the (b) Pacific Ocean basin (c) Atlantic Ocean basin, and (d) Indian Ocean Basin for the period of 2003–18.

(Trenberth and NCAR Research Staff 2020). Niño-3.4 is one of the most used indexes to monitor the occurrence and variability of El Niño and La Niña events, defined as the average equatorial SST anomalies across the Pacific in the region 5°S–5°N, 170°–120°W. Figure 14 show the time evolution of the Niño-3.4 index during the 1982–2018 ''common period'' for each product time series after applying a 5-month running mean filter.

All products give evidence of the very strong El Niño events in the period selected. The procedure used here to independently compute the Niño-3.4 index for all the datasets is the same applied by Trenberth and NCAR Research Staff (2020). The time evolution of the Niño-3.4 SST anomaly is nearly identical for all the products with minor differences (Fig. 14). The three strong El Niño events that occurred during this investigation period, namely 1982–83, 1997–98, and 2015–16, are reproduced, with a similar intensity, by all products. Moreover, the larger intensity of the El Niño positive anomalies with respect to the negative La Niña events confirms the asymmetry hypothesis of Monahan and Dai (2004).

4. Data maturity matrix

The concept of the data maturity matrix is to evaluate the basic characteristics of a dataset initiated by the World Meteorological Organization (WMO) to develop technical guidance and standards for collecting, processing, and managing datasets. The assessment of the maturity of the individual dataset is essential to guarantee and further improve the documentation, storage, and dissemination of datasets that are applicable for users (Peng et al. 2019).

The system maturity matrix (SMM) for climate data records (CDRs) is first developed in the Coordinating Earth Observation Data Validation for Reanalysis for Climate Services project (CORE-CLIMAX) (Su et al. 2018). The objective is to develop a tool to evaluate different aspects of the CDRs combining scientific and engineering views. In the SMM framework assessments are made in six major category areas and a score of 1–6 is assigned that reflects the maturity of the CDR with respect to a specific category, namely

1) Software readiness
2) Metadata
3) User documentation
4) Uncertainty characterization
5) Public access, feedback, and update
6) Usage

However, the assessment of maturity can only reflect aspects of process maturity. It does not interpret the scientific

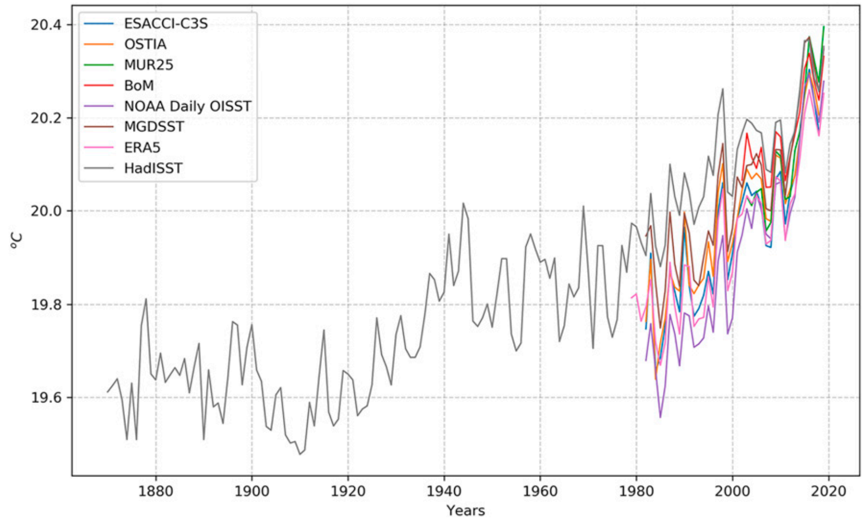

FIG. 11. Global monthly mean SST time series for all the ensemble members for the whole covered period originally obtained in each SST product.

quality of a dataset. For example, a mature product may not be scientifically reliable thus the maturity matrix only provides the assessment of fitness-of-purpose of a given product for climate service practitioners in terms of the categories mentioned above.

Additionally, the SMM scores recognize that at the early evaluation stage in the life cycle of the product the low scores in some of the categories do not demonstrate the possible future maturity of the dataset. Instead, low SMM scores indicate a recently released and evolving product at a less mature stage being made available to users.

In the context of the C3S_511 project, the aim of our assessment is to evaluate the maturity of the dataset instead of the whole CDR. We have adopted the SMM methodology of the CORE-CLIMAX for our use to evaluate individual datasets. We defined our matrix as the maturity matrix (MM) since we evaluate the dataset instead of the system of the dataset. Not all the categories from CORE-CLIMAX are included because some of them are not suitable for our usage. A guidance document is developed in the framework of C3S_511 project, and the assessment scores given in this study are based on our guidance document (https://confluence.ecmwf.int/display/CKB/Guidance+document+on+applying+the+Maturity+Matrix+as+part+of+the+Evaluation+and+Quality+Control). The MM, as important as the scientific quality, provides data providers important information in which aspects they need to improve their dataset for potential easy access and usage for users.

The MM of ESA-CCI and ERA5 SST (Table 2) shows that ESA-CCI SST is much more mature compared to ERA5 SST in terms of documentation, uncertainty characterization, and usage. As we mentioned above, low MM scores do not suggest that the scientific quality of ERA5 SST is lower than ESA-CCI SST. However, in terms of the documentation of the dataset, ESA-CCI SST is much more advanced than ERA5 SST.

In this study we have extended the evaluation of the MM to the dataset outside of CDS (Table 2). Due to the length limit, detailed defensible traces to score MM for SST products are given in the appendix. In terms of metadata, MGDSST has a lower score because it is provided in text format not following any standards with limited global attributes. The rest of the SST analysis products follow the NetCDF format and CF compliance with detailed information on metadata. Compared with other datasets, BoM, MGDSST, and MUR25 lack user documentation including the formal description of scientific methodology, validation report, and product user guide. A formal user guide is not found for HadISST1 either. Very few SST products (OSTIA and ESA-CCI SST) have automated quality monitoring in terms of the uncertainty characterization category. Thanks to GHRSST activities, all GHRSST L4 products follow internationally agreed GHRSST specifications, which provide uncertainty calculations. Several SST analysis products (HadISST1, MGDSST, BoM, and ERA5) have very limited validation, standards, or uncertainty quantification documentation. All SST products are publicly available via the online portal, except that BoM SST is available on request from the data provider via their website. However, the versioning, user feedback, and updates to records in the category of public access to SST products are not fully developed for BoM and MGDSST. All SST products except ERA5 are widely used in multiple research fields, and most of them either support decision support systems or usage and benefits of the SST products are emerging.

Overall, most of the SST products are well documented and user friendly. As we mentioned before, this scoring does not judge the scientific quality of the SST product. However, the low scoring of some products might give data providers important information to improve the documentation of their products in order to make the product more user friendly.

5. Summary of evaluations

SST is an essential climate variable (ECV) to assess the state of the global climate system and monitor its variations on

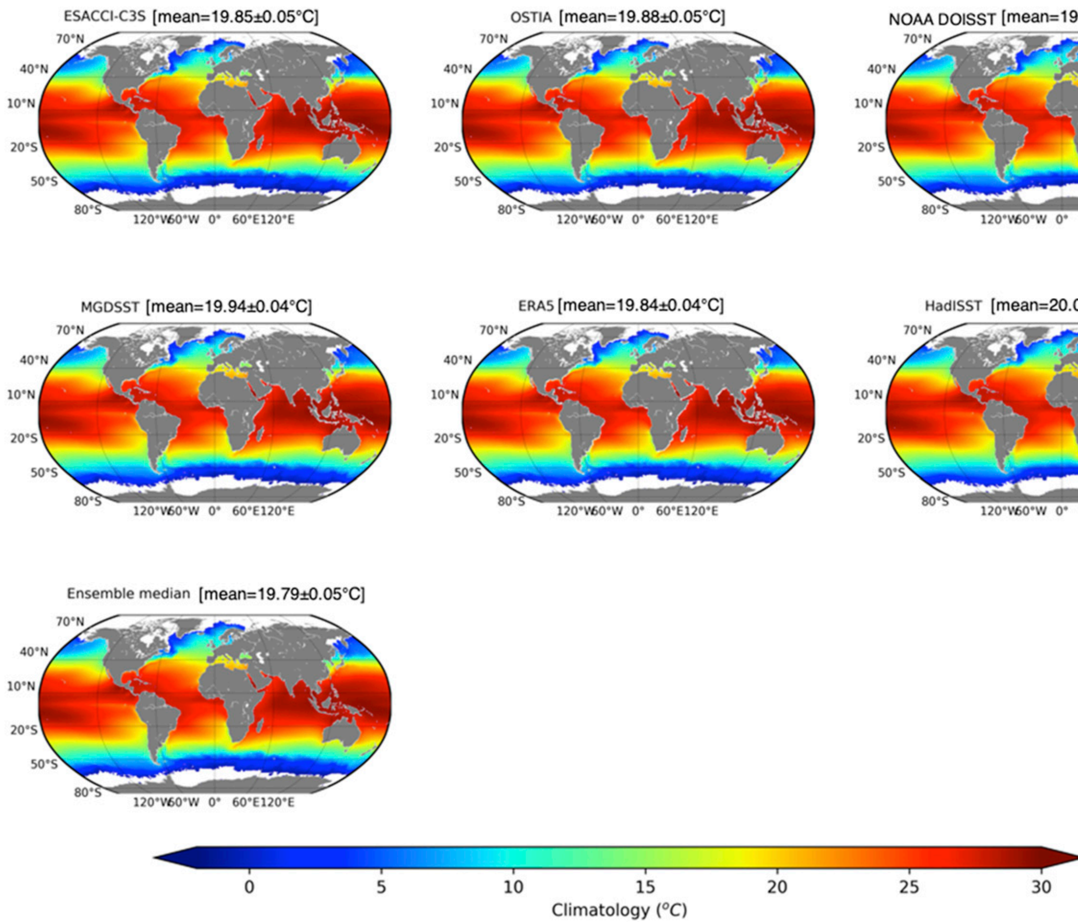

FIG. 12. Global SST climatologies for the period 1982–2002. Global SST average and its 95% confidence interval is also shown in brackets above each map.

interannual and (multi)decadal time scales. Accurate SST observations at high spatial and temporal resolution over a long-term period are needed to evaluate the present state of the oceans and the impact of global surface warming.

In this report, eight different SST datasets have been analyzed and intercompared for the overlapping period 2003–18. The ESA CCI SST v.2.1 and ERA5 reanalysis are available through the C3S Climate Data Store while the remaining six datasets (OSTIA, HadISST1, NOAA Daily OISST, MUR, MGDSST, and BoM) are provided outside the CDS. All these datasets provide global gap-free (optimally interpolated) SST maps but at different spatial and temporal resolutions. Then, to be comparable, all the datasets have been gridded to a common grid (i.e., 1° ×1°) and averaged to a common temporal frequency (i.e., monthly) over the overlapping period from 2003 to 2018. Finally, the average of the median of all the datasets (i.e., the ensemble median) has been defined in order to analyze differences among these datasets.

In general, all the SST datasets show consistent climatological spatial patterns (section 3b). The global monthly mean and anomaly SST time series of these datasets show very good agreement. When compared to the ensemble median, higher differences (in terms of mean difference, root-mean-square difference, and correlation) are found in correspondence to the main current systems, such as the Gulf Stream, the Kuroshio, and the Antarctic Circumpolar Current. These discrepancies are due to the different retrieval methods used to derive the spatially complete SST analyses. Differences can originate from several factors: interpolation technique and related configuration (e.g., observation/background error correlation scales), interpolation grid size, input data bias-correction, and, if present, the correction applied to obtain the foundation temperature or the temperature at 0.2 m. As an example, OSTIA, MUR25, MGDSST, and ERA5 (via OSTIA from 2007 onward) are the only L4 analyses included in the study that ingested microwave SST data. Since these datasets (OSTIA, MUR25, MGDSST, and ERA5) would ingest possibly cooler daytime SST observations over cloudy regions, they may therefore exhibit slightly cooler biases after 2002 compared with the other analyses that ingest only infrared SST observations and in situ data. This effect may be offset in some analyses, such as BoM Monthly and NOAA Daily OISST v2.1, where in situ data from 0.2 m to several meters depth are used to bias correct the infrared AVHRR SST data. However, on

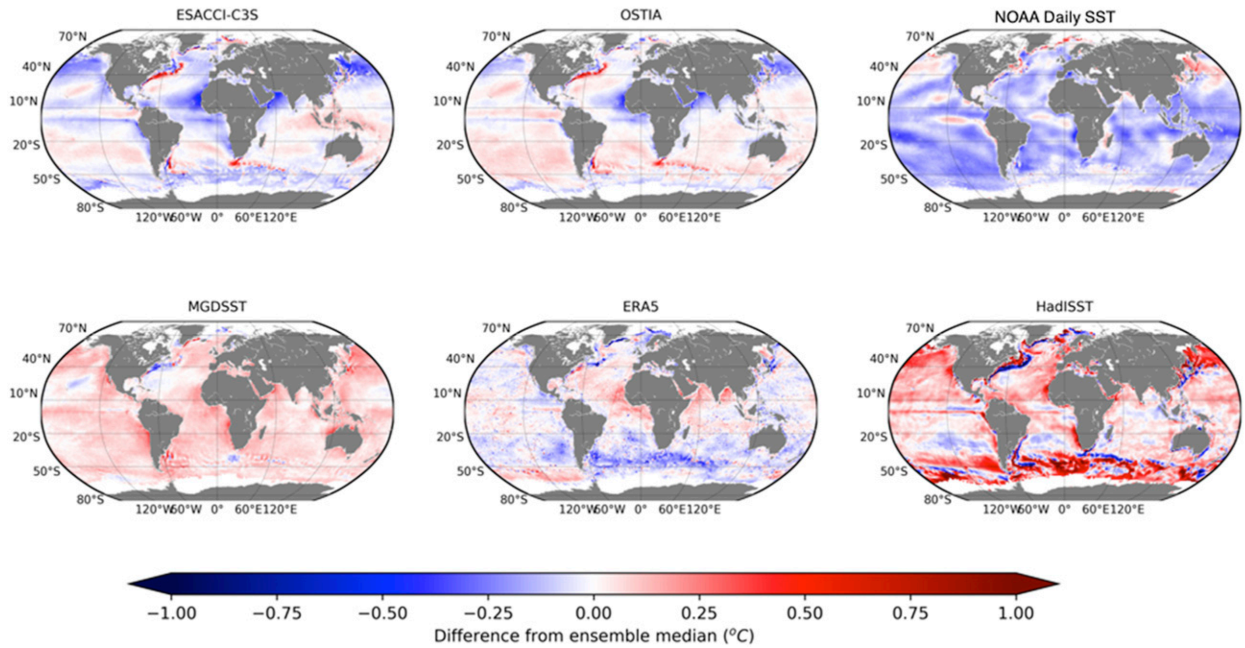

FIG. 13. The difference between each SST product and the ensemble median for the period of 1982–2002.

average, the Taylor diagram (Taylor 2001) confirms the very close similarity between the different datasets.

All the datasets reproduce very similar spatial patterns of global SST trends (section 3c). In addition, global mean warming trends as estimated from all the datasets are consistent (within the 95% confidence interval) with the global ocean warming trend as reported in the last IPCC report, estimated at 0.011°C yr^{-1} from 1980 to 2005. The linear trend in different basins shows that the main contributor from 2012 to 2018 is the Pacific Ocean.

The global mean SST time series for the whole period originally covered by all the SST products reveals that the disagreement between all SST products is larger in the early period (1982–2002) of the satellite era during which fewer observations are available compared to the later period (2003–18) of the satellite era. Specifically, the difference between each ensemble member and the ensemble median ranges from 0.03° to 0.26°C during the early period (1982–2002) and from 0.01° to 0.1°C during the later period (2003–18), respectively. It indicates that the observations ingested into each SST analysis play a significant role in constraining the SST climatology. Satellite sensor improvements (e.g., the launch in 2002 of AMSR-E, which could measure ocean brightness temperatures through clouds) is another important factor affecting SST quality after 2003. Note that the impact of natural variability on SST climatology is embedded

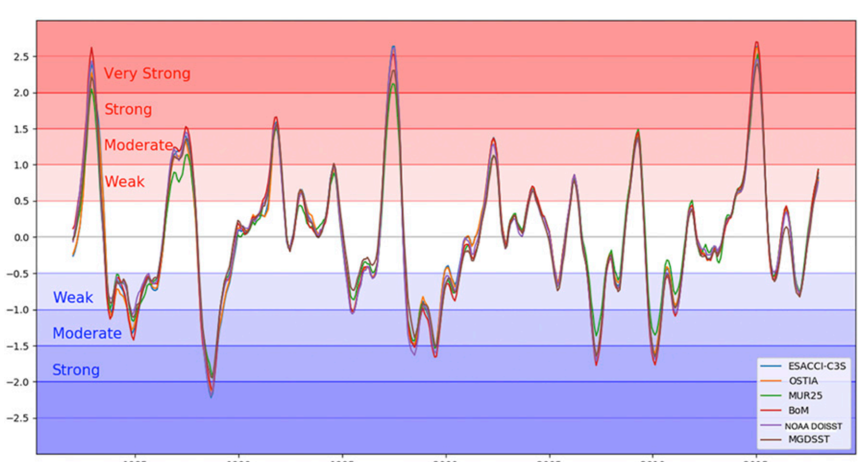

FIG. 14. Intercomparison between Niño-3.4 time series of the five SST products: HadISST1, ERA5, ESA CCI SST, MGDSST, and NOAA OISST.

TABLE 2. Maturity matrix for all SST products.

C3S_511 MM category	Name							
	ESA CCI SST	ERA5 SST	OSTIA	BoM	MGDSST	MUR25	NOAA Daily OISST	HadISST1
Metadata								
Standards	6	6	6	6	3	6	6	6
Collection level	6	5	6	5	2	6	6	6
User documentation								
Formal description of scientific methodology	6	6	6	4	3	6	6	6
Formal validation report	6	3	6	2	4	4	6	6
Formal product user guide	6	6	6	4	3	2	6	3
Uncertainty characterization								
Standards	6	3	6	6	1	6	6	1
Validation	6	3	6	5	6	6	6	6
Uncertainty quantification	6	3	6	6	1	6	6	1
Automated quality monitoring	6	2	6	1	2	4	4	1
Public access, feedback, and update								
Public access/archive	5	5	5	4	4	5	5	5
Version	6	6	6	2	2	6	6	6
User feedback mechanism	6	6	6	3	3	6	3	6
Updates to record	6	6	6	5	4	5	6	6
Usage								
Research	6	3	6	4	6	6	6	3
Decision support system	6	1	6	6	6	3	3	6

in the analysis (i.e., it is difficult to differentiate from the constraint of SST observations on the SST climatology). Additionally, the discrepancy between each product due to algorithms, observations ingested, etc. is very small compared to the significant warming trends shown in the linear trends and time series.

Finally, the tropical Pacific region has been selected, as a test case, to assess the capability of the different SST products, with a longer common temporal period, to capture the main modes of variability of a well-known climate oscillatory mode (ENSO). This analysis confirmed the close similarity of all the five datasets selected and their capability to reproduce, in the same way, the main components of the tropical Pacific region space and time variability at time scales compatible with the length of the selected time series.

The maturity matrix score of all SST products (Table 2), which aims to demonstrate the maturity of data documentation during the life cycle of one product, shows that most SST products are user friendly and provide sufficient information. Low scores of some SST products (Table 2) indicate a direction where data providers could improve their products in terms of data documentation, storage, and dissemination for users. Thanks to the GHRSST effort, all GHRSST products are well documented for their uncertainty characteristics (GHRSST Science Team 2012).

6. Recommendations to users

All the datasets presented here provide state-of-the-art spatially complete SST products at the global scale. These datasets are characterized by different spatial and temporal resolutions and temporal coverage that can fulfil the requirements of a large variety of users.

Intercomparison results and a test case analysis suggest these datasets provide an accurate representation of the SST spatiotemporal variability. These datasets can then be used for fundamental climate applications compatible with the length of each time series, such as long-term monitoring of SST changes (e.g., trends) and comparison to or initialization of numerical models. Other target applications include the use of these datasets in the definition of climatic indices, assessment and monitoring of weather extreme events (including marine heatwaves) and their impact on marine ecosystem, and related services.

In this study we have interpolated all SST products into 1° and monthly frequency in order to facilitate intercomparison studies. However, to understand which dataset is suitable for specific case studies where spatial and/or temporal resolution are critical, such as the separation of the Gulf Stream and the diurnal cycle of the SST products, specific intercomparison studies are required. Indeed, in the framework of the GHRSST intercomparison team, several such intercomparison tasks are ongoing and scientific findings will be available in the near future.

Finally, users are strongly encouraged to also consider the type of SST offered by each producer and to distinguish between, say, skin SST, subskin or SSTdepth, and foundation SST according to the specific application for which the data are intended to be used. For example, in conditions of high insolation and low surface ocean mixing skin SST is strongly impacted by diurnal warming, SST at 0.2-m depth somewhat impacted, SSTdepth below 1 m minimally impacted, and foundation SST has no diurnal signature (Gentemann et al.

2009; Minnett and Kaiser-Weiss 2012). In our study, we have used SSTdepth, foundation SST, and SST at 0.2-m depth, which appears to have had minor impacts on the results.

Acknowledgments. This work is funded by European Copernicus Climate Change Service (C3S) implemented by the European Centre for Medium-Range Weather Forecasts (ECMWF) under the service contract Independent Assessment on ECVs (C3S_511), especially Chunxue Yang, Bruno Buongiorno Nardelli, Francesca Leonelli, Salvatore Marullo, Andrea Pisano, Vincenzo Artale, Rosalia Santoleri, and Vincenzo De Toma.

The BoM Monthly OI SST analysis GHRSST L4 format files were provided by the Australian Bureau of Meteorology and are available on request from http://www.bom.gov.au/climate/data-services/data-requests.shtml.

A portion of this research was carried out by Toshio M. Chin and J. Vazquez-Cuervo at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004). The authors declare no conflict of interest.

Data availability statement. The download website of all datasets used in this study has been included in the manuscript in section 2.

APPENDIX

Detailed Defensible Traces to Score MM for SST Products

This section provides defensible traces for maturity matrix score given to all SST products shown in Table 2 based on the guidance document (https://confluence.ecmwf.int/display/CKB/Guidance+document+on+applying+the+Maturity+Matrix+as+part+of+the+Evaluation+and+Quality+Control) developed within the C3S independent assessment project (C3S_511).

a. *ESA-CCI SST*

1) METADATA

(i) Standard (Score: 6/6)

The ESA CCI SST data files follow the GHRSST Data Specification v2.0 (GDS) and are provided in NetCDF-4 format CompactFlash (CF)-1.5 compliant. File specifications are fully detailed in the ESA CCI Product User Guide (PUG). The NetCDF files contain detailed metadata describing the data by means of global attributes, which are applicable to the whole file, and variable attributes, which apply to a specific data field.

(ii) Collection level (Score: 6/6)

The ESA CCI SST data files follow the GHRSST Data Specification v2.0 (GDS). Global attributes provide all information available on the data and relative references. In addition the product specification document (PSD) with detailed information of metadata is available.

2) USER DOCUMENTATION

(i) Formal description of scientific methodology (Score: 6/6)

The formal description of the ESA CCI SST product is detailed in the Algorithm Theoretical Background Document (ATBD), published by the data provider, which describes and justifies the algorithms used for obtaining SST estimates. A synthesis of the formal ATBD is also available in the CDS. In addition, the ESA CCI SST dataset has been published in *Nature Scientific Data* (Merchant et al. 2019).

(ii) Formal validation report (Score: 6/6)

For the formal validation report of the ESA CCI SST L4 product users can refer to Merchant et al. (2019), PUG, and Climate Assessment Report (CAR).

(iii) Formal product user guide (Score: 6/6)

The formal product user guide ESA CCI SST product is published by the data provider (PUG). A synthesis of the formal user product guide is also available in the CDS.

3) UNCERTAINTY CHARACTERIZATION

(i) Standards (Score: 6/6)

Uncertainty characterization follows the internationally agreed GHRSST standard specifications, which are detailed in the GHRSST Data Specification v2.0 (GDS) document.

(ii) Validation (Score: 6/6)

A detailed and comprehensive validation of the ESA CCI SST L4 product is provided in the PUG, CAR, and in Merchant et al. (2019). The validation of the ESA CCI SST L4 product is based on different procedures, from automated and visual inspection to comparison of SST data with collocated in situ measurements.

(iii) Uncertainty quantification (Score: 6/6)

Uncertainty in the ESA CCI SST L4 data at each location (i.e., the analysed_sst field in the NETCDF file) is quantified and provided (i.e., in the analysis_error field) through an analysis quality methodology. The methodology used to derive the uncertainty is based on the optimal interpolation theory and described in the ATBD and PUG, giving comprehensive information of validation of the quantitative uncertainty estimates and error covariance.

(iv) Automated quality monitoring (Score: 6/6)

The identification of valid observations for SST estimation and algorithms used in the preparatory preprocessing are described in the ATBD and PUG. Moreover, a confidence level on a scale of 0–5 is provided for each SST as a quality indicator, following the international GHRSST conventions. Five indicates the highest confidence. Quality levels 4 and 5 should be used for climate applications. An automated check is implemented to valid the data quality (Merchant et al. 2019).

4) Public access, feedback, and update

(i) Public access/archive (Score: 5/6)

The ESA CCI SST dataset v2.0 is available on the data provider's website. Detailed information available in the PUG. However, the source code is not publicly available.

(ii) Version (Score: 6/6)

The version is fully established by the data provider.

(iii) User feedback (Score: 6/6)

The ESA CCI SST dataset v2.0 is also provided through the CMEMS and is part of GHRSST. Within CMEMS, a Multi-Year Product Quality Working Group is established with the aim of periodically assessing the status of the CMEMS climate data records, including ESA CCI SST, integrating users' needs and feedback. Feedback from users is also included in the CAR. In addition, the ESA CCI data provider provides an e-mail contact to collect users' feedback.

(iv) Updates to record (Score: 6/6)

Currently the ESA CCI SST dataset v2.0 covers the period from late 1981 to 2018. Updates through to the near-present are expected this year (2020). Extensions are expected to be produced by the Copernicus Climate Change Service (C3S) with only ~5 days delay to real time.

5) Usage

(i) Research (Score: 6/6)

The ESA CCI SST dataset v.2.0 is very recent. However, it has already been used in some research publications.

(ii) Decision support system (Score: 6/6)

ESA-CCI SST is part of the ESA Climate Change Initiative, and one of the essential climate variables. The objective of ESA-CCI SST is to establish a long term data record to monitor the global climate system required by UNFCCC (http://cci.esa.int/) for decision making.

b. ERA5 SST

1) Metadata

(i) Standard (Score: 6/6)

ERA5 SST data can be downloaded from the CDS in both GRIB and NetCDF formats. The native data format is GRIB, but they can be converted to NetCDF format through the CDS. In NetCDF global attributes reference to CF-1.6 conventions is made. This represents a mature state-of-the-art metadata standard according to guidance.

(ii) Collection level (Score 5/6)

The standardized attributes on the collection level of the dataset are sufficient to understand the data's origins without further documents, including standardized information on how to obtain raw data and its preprocessing procedures.

Note: The collection level in this case includes the ECMWF confluence wiki (https://confluence.ecmwf.int/display/CKB/ERA5%3A+data+documentation).

2) User documentation

(i) Formal description of scientific methodology (Score 6/6)

The scientific description is comprehensive and publicly available in the form of a scientific report/ATBD and e-library of ECMWF. The description is kept up to date with the updated dataset. There is also a peer-reviewed methodological journal paper published.

(ii) Formal validation report (Score: 3/6)

There is no formal validation report for ERA5 SST. The ERA5 documentation available at confluence wiki can be regarded as a user guide but does not have any clear version number with a publication date and is a document that is changing. Due to the nature of ERA5 being in development it makes sense to have an evolving documentation, but the creation of a formal product validation report in the future is recommended. An assessment report evaluating HadISST2 and OSTIA SST datasets (from which ERA5 SST is built) is available (Hirahara et al. 2016).

(iii) Formal product user guide (Score 6/6)

There is a regularly updated comprehensive formal PUG for the dataset publicly available.

Note: In this case the confluence wiki is regarded as the PUG.

3) Uncertainty characterization

(i) Standards (Score 3/6)

Uncertainty information follows standard nomenclature.

Note: In this case the ensemble members are regarded as uncertainty measures.

(ii) Validation (Score: 3/6)

A formal validation report of ERA5 SST is not available. However, an assessment report evaluating HadISST2 and OSTIA SST datasets (from which ERA5 SST is built) is available (Hirahara et al. 2016), and users can refer to HadISST2 and OSTIA documentation.

(iii) Uncertainty quantification (Score 3/6)

A comprehensive uncertainty quantification of systematic and random effects is available.

Note: In this case the ensemble members are regarded as uncertainty measures.

(iv) Automated quality monitoring (Score 2/6)

There is no automated quality monitoring documented for the dataset.

Note: Although there is no automated quality monitoring documented, data assimilation itself could be regarded as a quality check.

4) PUBLIC ACCESS, FEEDBACK, AND UPDATES

(i) Access and archive (Score 5/6)

The dataset is publicly available. The different versions of data including documentation and source code is archived by the data provider. Source code is not publicly available.

(ii) Version control (Score 6/6)

There is full information on version control of documentation, data, and/or metadata available for the dataset. The documented version control information is fully traceable from the files.

Note: In this case the version control is referring to the confluence wiki.

(iii) User feedback (Score 6/6)

There is a public reach-out/feedback form/contact point for collecting feedback for the dataset. There are regular events, groups, two-way feedback mechanisms, etc. organized by the data provider. The established feedback fed back into data production is documented, including third-party international data quality assessment results.

(iv) Updates to record (Score 6/6)

There are regular operational updates available for the dataset, depending on the availability of input data and including improved methodology.

5) USAGE

(i) Research (Score: 3/6)

Although ERA5 reanalysis has been largely used in many research publications, it seems that there are few relevant publications based on ERA5 SST data (e.g., Wang et al. 2019). This could arise from the prevalent use of ERA5 in atmospheric research.

(ii) Decision support system (Score: 1/6)

To the evaluators' knowledge the product is not used yet for the decision support system.

c. OSTIA SST

1) METADATA

(i) Standard (Score: 6/6)

The OSTIA SST data files are provided in NetCDF-4 format CF-1.5 compliant through CMEMS and the Recommended GHRSST Data Specification (GDS). File specifications are fully detailed in the OSTIA Product User Manual (PUM) available in CMEMS. The NetCDF files contain detailed metadata describing the data by means of global attributes, which are applicable to the whole file, and variable attributes, which apply to a specific data field.

(ii) Collection level (Score: 6/6)

Global attributes provide all information available on the data and relative references. In addition the PUM with detailed information on metadata is available.

2) USER DOCUMENTATION

(i) Formal description of scientific methodology (Score: 6/6)

The formal description of the OSTIA product is detailed in the peer-reviewed paper (Good et al. 2020), published by the data provider, which describes and justifies the algorithms used for obtaining SST estimates. A synthesis of the PUM is also available in the CMEMS.

(ii) Formal validation report (Score: 6/6)

For the formal validation report of the OSTIA product users can refer to the Quality Information Document (QUID) available in the CMEMS service.

(iii) Formal product user guide (Score: 6/6)

The formal product user guide OSTIA product is published by the data provider (PUM) as a peer-reviewed journal article (Good et al. 2020). A synthesis of the formal user product guide (PUM) is also available in the CMEMS.

3) UNCERTAINTY CHARACTERIZATION

(i) Standards (Score: 6/6)

Uncertainty characterization follows the internationally agreed GHRSST standard specifications, which are detailed in the GHRSST Data Specification v2.0 (GDS) document (GHRSST Science Team 2012).

(ii) Validation (Score: 6/6)

A validation of the OSTIA product is provided in the Quality Information Document through CMEMS. The validation of the OSTIA SST product is based on comparison of SST data with collocated in situ measurements.

(iii) Uncertainty quantification (Score: 6/6)

Uncertainty in the OSTIA data at each location (i.e., the analysed_sst field in the NETCDF file) is quantified and provided (i.e., in the analysis_error field) through an analysis quality methodology. The methodology used to derive the uncertainty is produced using a special "observation influence" analysis (Good et al. 2020).

(iv) Automated quality monitoring (Score: 6/6)

Automatic quality is monitored during the production of the SST product. The real-time OSTIA SST analysis is routinely validated by the U.K. Met Office against the GHRSST Multiproduct ensemble (http://ghrsst-pp.metoffice.gov.uk/ostia-website/gmpe-monitoring.html) and Argo SST (http://ghrsst-pp.metoffice.gov.uk/ostia-website/gmpe-argo-stats.html).

4) PUBLIC ACCESS, FEEDBACK, AND UPDATE

(i) Public access/archive (Score: 5/6)

The OSTIA SST is available on the CMEMS website. Detailed information available in the PUM. However, the source code is not publicly available.

(ii) Version (Score: 6/6)

The version is fully established by the data provider.

(iii) User feedback (Score: 6/6)

The OSTIA is provided through the CMEMS and is part of GHRSST. Within CMEMS, a Multi-Year Product Quality Working Group is established with the aim of periodically assessing the status of the CMEMS data records, including OSTIA, integrating users' needs and feedback.

(iv) Updates to record (Score: 6/6)

Currently the OSTIA SST dataset covers the period from late 1981 to 2018. Updates through to the near-present are expected this year (2020). Extensions are expected to be produced by the CMEMS with only ~5 days delay to real time.

5) USAGE

(i) Research (Score: 6/6)

The current version of OSTIA SST is very recent. However, it has already been used in some research publications.

(ii) Decision support system (Score: 6/6)

OSTIA SST is part of the CMEMS project, and the information derived from SST products is used in the CMEMS ocean state report for decision making.

d. BoM

1) METADATA

(i) Standard (Score: 6/6)

The BoM SST files are provided in the GHRSST Data Specification version 1.7 NetCDF classic format CF-1 (Beggs and Pugh 2009) on request from the data providers. The NetCDF files contain detailed metadata describing the data by means of global attributes, which are applicable to the whole file, and variable attributes, which apply to a specific data field.

(ii) Collection level (Score: 5/6)

Global attributes provide all information available on the data and relative references. However, the reference shown in the metadata (Beggs and Pugh 2009) is not accessible at the moment of writing this report although it is available by request from library@bom.gov.au.

2) USER DOCUMENTATION

(i) Formal description of scientific methodology (Score: 4/6)

The formal description of the BoM Monthly OI SST is published in a conference paper (Smith et al. 1999) and a peer-reviewed paper (Beggs et al. 2011); however, the peer-reviewed paper focuses on the BoM higher-resolution daily 1/12° regional analyses available from 2006, which uses a modified version of the FORTRAN "SIANAL" code used to produce the original BoM Weekly and Monthly OI SST analyses.

(ii) Formal validation report (Score: 2/6)

BoM Monthly OI 1° L4 SST is part of the GHRSST suite of L4 products, and intercomparison of the BoM higher-resolution daily SST analyses with other SST products have been published in peer-reviewed journals (Beggs et al. 2011; Dash et al. 2012; Martin et al. 2012). However, the only previously published comparison of the lower-resolution BoM Weekly 1° OI SST analysis with other SST analysis products is in a BoM Operations Bulletin (Zhong and Beggs 2008).

(iii) Formal product user guide (Score: 4/6)

The description of the BoM Monthly OI SST analysis methodology is published in Smith et al. (1999) and Beggs et al. (2011), and a user guide is provided (Beggs and Pugh 2009). However, Beggs and Pugh (2009) is not accessible at the moment of writing this report although it is available by request from library@bom.gov.au.

3) UNCERTAINTY CHARACTERIZATION

(i) Standards (Score: 6/6)

Uncertainty characterization follows the internationally agreed GHRSST standard specifications (analysis_error), which are detailed in the GHRSST Data Specification v2.0 (GDS) document (GHRSST Science Team 2012).

(ii) Validation (Score: 5/6)

No validation report is found for BoM SST. However, BoM is part of the GHRSST community and intercomparison activities of the BoM Daily Global SST analyses have been performed in the framework of GHRSST (Dash et al. 2012; Martin et al. 2012). Although routine verification of the BoM Global Daily 0.25° OI SST analysis (GAMSSA) are performed by the U.K. Met Office (http://ghrsst-pp.metoffice.gov.uk/ostia-website/gmpe-argo-stats.html) and NOAA/NESDIS/STAR (https://www.star.nesdis.noaa.gov/socd/sst/squam/analysis/l4), there are no routine verifications of the BoM Monthly or Weekly OI SST analyses.

(iii) Uncertainty quantification (Score: 6/6)

Uncertainty in the BoM data at each location (i.e., the analysed_sst field in the NETCDF file) is quantified and provided (i.e., in the analysis_error field) through an analysis quality methodology (Beggs et al. 2011).

(iv) Automated quality monitoring (Score: 1/6)

No automatic quality is provided.

4) PUBLIC ACCESS, FEEDBACK, AND UPDATE

(i) Public access/archive (Score: 4/6)

BoM monthly SST product is available on request from the data provider website for both real-time and archived GHRSST L4 files.

(ii) Version (Score: 2/6)

No information is found for the version control for BoM SST.

(iii) User feedback (Score: 3/6)

Data providers collect and evaluate feedback from the scientific community through the data provider's website, but no feedback mechanisms are set up from data providers.

(iv) Updates to record (Score: 5/6)

BoM daily, weekly, and monthly SST analyses are published in real time for climate monitoring on the BoM website.

5) Usage

(i) Research (Score: 4/6)

The BoM weekly and monthly SST analyses have been used by the BoM for research, especially climate studies.

(ii) Decision support system (Score: 6/6)

BoM monthly SST is an operational SST analysis that serves for climate monitoring, which is an essential service of the Australian Government Bureau of Meteorology.

e. MGDSST

1) Metadata

(i) Standard (Score: 3/6)

The MGDSST is provided in the .txt format and variable attributes are limited.

(ii) Collection level (Score: 2/6)

There is limited information about standard attributes, but extra information published in the data provider's website is needed to use and understand the data.

2) User documentation

(i) Formal description of scientific methodology (Score: 3/6)

Limited information is provided on the data provider's website, but the method is documented in two non-peer-reviewed reports.

(ii) Formal validation report (Score: 4/6)

No JMA validation report is found for MGDSST at the time of writing this report.

However, MGDSST was compared with other SST analyses and independent observations in Martin et al. (2012) and Fiedler et al. (2019a) for the periods 2010 and 1992 to 2011. The U.K. Met Office routinely compares MGDSST with the GHRSST Multiproduct ensemble (http://ghrsst-pp.metoffice.gov.uk/ostia-website/gmpe-monitoring.html) and Argo SST (http://ghrsst-pp.metoffice.gov.uk/ostia-website/gmpe-argo-stats.html).

(iii) Formal product user guide (Score: 3/6)

Limited product user guide from the data provider.

3) Uncertainty characterization

(i) Standards (Score: 1/6)

No information is available at this stage.

(ii) Validation (Score: 6/6)

MGDSST is part of the GHRSST and intercomparison with other SST products has been performed and published in peer-reviewed journals (Fiedler et al. 2019a; Martin et al. 2012).

(iii) Uncertainty quantification (Score: 1/6)

No uncertainty quantification is found.

(iv) Automated quality monitoring (Score: 2/6)

No automatic quality is monitored during the production of the SST product.

4) Public access, feedback, and update

(i) Public access/archive (Score: 4/6)

The MGDSST is publicly accessible from the data provider's website and brief information of the data is provided in the data provider's website.

(ii) Version (Score: 2/6)

No information is found for the version control.

(iii) User feedback (Score: 3/6)

Data providers collect and evaluate feedback from the scientific community through the data provider's website.

(iv) Updates to record (Score: 4/6)

MGDSST is published in real time for climate monitoring and numerical weather prediction on the data provider's website.

5) Usage

(i) Research (Score: 6/6)

The data have already been used in some research publications.

(ii) Decision support system (Score: 6/6)

MGDSST is an operational SST analysis that serves for climate monitoring and numerical weather prediction, which is an essential service of the Japanese Meteorological Agency (JMA).

f. MUR25

1) Metadata

(i) Standard (Score: 6/6)

The MUR25 SST is provided in NetCDF format. The NetCDF files contain detailed metadata describing the data by means of global attributes, which are applicable to the whole file, and variable attributes, which apply to a specific data field.

(ii) Collection level (Score: 6/6)

Global attributes provide all information available on the data and relative references.

2) User documentation

(i) Formal description of scientific methodology (Score: 6/6)

The formal description of the MUR25 product is detailed in the peer-reviewed journal (Chin et al. 2017) that is published by the data provider.

(ii) Formal validation report (Score: 4/6)

No formal validation report is available; however, the validation is performed in a peer-reviewed paper (Chin et al. 2017). Additional validation of the 1-km product occurred with direct comparisons with the Saildrone autonomous vehicle with the published article. The validation focused on an exemplary coastal area, the California/Baja Coast.

(iii) Formal product user guide (Score: 2/6)

No formal product user guide is available for MUR25 SST.

3) UNCERTAINTY CHARACTERIZATION

(i) Standards (Score: 6/6)

Uncertainty characterization follows the internationally agreed GHRSST standard specifications, which are detailed in the GHRSST Data Specification v2.0 (GDS) document.

(ii) Validation (Score: 6/6)

Intercomparison of MUR25 has been performed in the framework of GHRSST.

(iii) Uncertainty quantification (Score: 6/6)

Uncertainty in the MUR25 data at each location (i.e., the analysed_sst field in the NETCDF file) is quantified and provided (i.e., in the analysis_error field) through an analysis quality methodology.

(iv) Automated quality monitoring (Score: 4/6)

No automatic quality monitoring is found for MUR25 SST product, but the 1-km resolution version of the MUR SST analysis is routinely validated with the GHRSST Multiproduct ensemble (http://ghrsst-pp.metoffice.gov.uk/ostia-website/gmpe-monitoring.html; https://www.star.nesdis.noaa.gov/socd/sst/squam/analysis/l4). Since Argo SST are ingested into MUR25 they are not useful for verification.

4) PUBLIC ACCESS, FEEDBACK, AND UPDATE

(i) Public access/archive (Score: 5/6)

The MUR25 SST is published in the data provider's archive center. However, source code is not publicly available.

(ii) Version (Score: 6/6)

The version is fully established by the data provider.

(iii) User feedback (Score: 6/6)

Public contact information is given in the data provider's website for users to give feedback. Users can give all feedback through the Physical Oceanography Distributed Active Archive Center (PO.DAAC) user services and forum. All feedback is publicly available.

(iv) Updates to record (Score: 5/6)

Regular updates are available from the data provider. There is no immediate production of interim data products.

5) USAGE

(i) Research (Score: 6/6)

The MUR25 is used in research in multiple fields.

(ii) Decision support system (Score: 3/6)

No decision support system is found for MUR25 SST; however, use is occurring and benefits are emerging.

g. NOAA Daily OISSTv2.1 SST

1) METADATA

(i) Standard (Score: 6/6)

The NOAA Daily OISST data files are provided in NetCDF-4 format CF-1.0 compliant data provider's website. The NetCDF files contain detailed metadata describing the data by means of global attributes, which are applicable to the whole file, and variable attributes, which apply to a specific data field.

(ii) Collection level (Score: 6/6)

Global attributes provide all information available on the data and relative references.

2) USER DOCUMENTATION

(i) Formal description of scientific methodology (Score: 6/6)

The formal description of the NOAA Daily OISST v2.1 is provided in the data provider's website (https://www.ncdc.noaa.gov/oisst) and third party data resource website (https://podaac.jpl.nasa.gov/dataset/AVHRR_OI-NCEI-L4-GLOB-v2.1) and is also detailed in several peer-reviewed papers (Reynolds et al. 2007; Banzon et al. 2016; Huang et al. 2020), published by the data provider, which describe and justify the algorithms used for obtaining SST estimates.

(ii) Formal validation report (Score: 6/6)

Formal validation report of NOAA Daily OISST is along with data access.

(iii) Formal product user guide (Score: 6/6)

The formal product user guide is provided in a peer-reviewed journal (Banzon et al. 2016).

3) UNCERTAINTY CHARACTERIZATION

(i) Standards (Score: 6/6)

Uncertainty characterization follows the internationally agreed GHRSST standard specifications, which are detailed in the GHRSST Data Specification v2.0 (GDS) document.

(ii) Validation (Score: 6/6)

A validation of NOAA Daily OISST is provided through peer-reviewed journals (Dash et al. 2012; Martin et al. 2012; Banzon et al. 2016; Fiedler et al. 2019a; Huang et al. 2020).

(iii) Uncertainty quantification (Score: 6/6)

Uncertainty in the NOAA Daily OISST data at each location (i.e., the analysed_sst field in the NETCDF file available from https://podaac.jpl.nasa.gov/dataset/AVHRR_OI-NCEI-L4-GLOB-v2.1) is quantified and provided (i.e., in the analysis_error field) through an analysis quality methodology.

(iv) Automated quality monitoring (Score: 4/6)

The Daily OISST v2.1 SST analyses are validated in near real time against the GHRSST Multi-Product Ensemble by NOAA/STAR at https://www.star.nesdis.noaa.gov/socd/sst/squam/analysis/l4. Since Argo SST are ingested into Daily OISST v2.1 they are not useful for verification.

4) PUBLIC ACCESS, FEEDBACK, AND UPDATE

(i) Public access/archive (Score: 5/6)

The data are publicly accessible through the data provider's website and also other data portals with documentation. No source code is available publicly.

(ii) Version (Score: 6/6)

The version is fully established by the data provider.

(iii) User feedback (Score: 3/6)

Contact information of the data provider is publicly available for user feedback.

(iv) Updates to record (Score: 6/6)

Data providers regularly update the data record.

5) USAGE

(i) Research (Score: 6/6)

The NOAA Daily OISST is widely used in multiple research fields.

(ii) Decision support system (Score: 3/6)

No decision support system is found for NOAA Daily OISST; however, use is occurring and benefits are emerging.

h. HadISST1

1) METADATA

(i) Standard (Score: 6/6)

The HadISST1data files are provided in NetCDF classic format CF compliant through the data provider's website. The NetCDF files contain detailed metadata describing the data by means of global attributes, which are applicable to the whole file, and variable attributes, which apply to a specific data field.

(ii) Collection level (Score: 6/6)

Global attributes provide all information available on the data and relative references.

2) USER DOCUMENTATION

(i) Formal description of scientific methodology (Score: 6/6)

The formal description of the HadISST1 is detailed in a peer-reviewed journal (Rayner et al. 2003), published by the data provider, which describes and justifies the algorithms used for obtaining SST estimates.

(ii) Formal validation report (Score: 6/6)

A formal validation report is published in a peer-reviewed journal.

(iii) Formal product user guide (Score: 3/6)

No formal product user guide is provided. Product information is provided on the data provider's website.

3) UNCERTAINTY CHARACTERIZATION

(i) Standards (Score: 1/6)

No information is available at this stage.

(ii) Validation (Score: 6/6)

The validation is available through a peer-reviewed journal paper.

(iii) Uncertainty quantification (Score: 1/6)

No uncertainty quantification is found.

(iv) Automated quality monitoring (Score: 1/6)

No automatic quality is monitored during the production of the SST product.

4) PUBLIC ACCESS, FEEDBACK, AND UPDATE

(i) Public access/archive (Score: 5/6)

The data are published through the data provider's website, but no source code is publicly available.

(ii) Version (Score: 6/6)

The version is fully established by the data provider.

(iii) User feedback (Score: 3/6)

Contact information of the data provider is given for collecting user feedback.

(iv) Updates to record (Score: 6/6)

The data are regularly updated by the data provider.

5) USAGE

(i) Research (Score: 6/6)

HadISST1 has been widely used in multiple research fields.

(ii) Decision support system (Score: 6/6)

Up to now no decision support system is found for HadISST1; however, influence on decision making is demonstrated.

REFERENCES

Banzon, V., T. M. Smith, T. M. Chin, C. Liu, and W. Hankins, 2016: A long-term record of blended satellite and in situ sea-surface temperature for climate monitoring, modeling and environmental studies. *Earth Syst. Sci. Data*, **8**, 165–176, https://doi.org/10.5194/essd-8-165-2016.

Beggs, H., and T. Pugh, 2009: Format specification for the Australian Bureau of Meteorology's SST analysis L4 files, version 7. Bureau of Meteorology Technical Document, 12 pp., https://imos.org.au/facilities/srs/sstproducts/sstdata0/sstdata-references.

——, A. Zhong, G. Warren, O. Alves, G. Brassington, and T. Pugh, 2011: RAMSSA—An operational, high-resolution, multi-sensor sea surface temperature analysis over the Australian region. *Aust. Meteor. Oceanogr. J.*, **61** (1), 1–22, https://doi.org/10.22499/2.6101.001.

Behrenfeld, M., and Coauthors, 2006: Climate-driven trends in contemporary ocean productivity. *Nature*, **444**, 752–755, https://doi.org/10.1038/nature05317.

Chin, T. M., J. Vazquez-Cuervo, and E. M. Armstrong, 2017: A multi-scale high-resolution analysis of global sea surface temperature. *Remote Sens. Environ.*, **200**, 154–169, https://doi.org/10.1016/j.rse.2017.07.029.

Dash, P., and Coauthors, 2012: Group for High Resolution SST (GHRSST) analysis fields inter-comparisons–Part 2: Near real-time web-based level 4 SST Quality Monitor (L4-SQUAM). *Deep-Sea Res. II*, **77–80**, 31–43, https://doi.org/10.1016/j.dsr2.2012.04.002.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, https://doi.org/10.1002/qj.828.

Donlon, C. J., and Coauthors, 2009: The GODAE high-resolution sea surface temperature pilot project. *Oceanography*, **22**, 34–45, https://doi.org/10.5670/oceanog.2009.64.

——, M. Martin, J. Stark, J. Roberts-Jones, E. Fiedler, and W. Wimmer, 2012: The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system. *Remote Sens. Environ.*, **116**, 140–158, https://doi.org/10.1016/j.rse.2010.10.017.

Efron, B., and R. J. Tibshirani, 1993: *An Introduction to the Bootstrap*. Chapman & Hall/CRC, 456 pp.

Fiedler, E. K., and Coauthors, 2019a: Intercomparison of long-term sea surface temperature analyses using the GHRSST Multi-Product Ensemble (GMPE) system. *Remote Sens. Environ.*, **222**, 18–33, https://doi.org/10.1016/j.rse.2018.12.015.

——, C. Mao, S. A. Good, J. Waters, and M. J. Martin, 2019b: Improvements to feature resolution in the OSTIA sea surface temperature analysis using the NEMOVAR assimilation scheme. *Quart. J. Roy. Meteor. Soc.*, **145**, 3609–3625, https://doi.org/10.1002/qj.3644.

Gentemann, C. L., P. J. Minnett, and B. Ward, 2009: Profiles of ocean surface heating (POSH): A new model of upper ocean diurnal warming. *J. Geophys. Res.*, **114**, C07017, https://doi.org/10.1029/2008JC004825.

Ghil, M., and Coauthors, 2002: Advanced spectral methods for climatic time series. *Rev. Geophys.*, **40**, 1003, https://doi.org/10.1029/2000RG000092.

GHRSST Science Team, 2012: The Recommended GHRSST Data Specification (GDS) 2.0, document revision 5, GHRSST International Project Office, 123 pp., https://www.ghrsst.org/wp-content/uploads/2016/10/GDS20r5.pdf.

Good, S., and Coauthors, 2020: The current configuration of the OSTIA system for operational production of foundation sea surface temperature and ice concentration analyses. *Remote Sens.*, **12**, 720, https://doi.org/10.3390/rs12040720.

Grumbine, R. W., 1996: Automated Passive Microwave Sea Ice Concentration Analysis at NCEP. Tech. Note NOAA/NCEP, 13 pp., https://polar.ncep.noaa.gov/mmab/papers/tn120/ssmi120.pdf.

Hirahara, S., M. A. Balmaseda, E. de Boisseson, and H. Hersbach, 2016: Sea surface temperature and sea ice concentration for ERA5. ERA Rep. Series 26, 27 pp., https://www.ecmwf.int/en/elibrary/16555-sea-surface-temperature-and-sea-ice-concentration-era5.

Huang, B., C. Liu, V. Banzon, E. Freeman, G. Graham, B. Hankins, T. Smith, and H.-M. Zhang, 2020: Improvements of the Daily Optimum Interpolation Sea Surface Temperature (DOISST) version 2.1. *J. Climate*, **34**, 2923–2939, https://doi.org/10.1175/JCLI-D-20-0166.1.

Jolliffe, I. T., 2002: *Principal Component Analysis*, Springer-Verlag, 502 pp.

Kaplan, A., Y. Kushnir, M. Cane, and M. Blumenthal, 1997: Reduced space optimal analysis for historical data sets: 136 years of Atlantic sea surface temperatures. *J. Geophys. Res.*, **102**, 27 835–272009860, https://doi.org/10.1029/97JC01734.

Kendall, M. G., 1975: *Multivariate Analysis*. Charles Griffin & Co, 43 pp.

Kurihara, Y., T. Sakurai, and T. Kuragano, 2006: Global daily sea surface temperature analysis using data from satellite microwave radiometer, satellite infrared radiometer and in-situ observations (in Japanese). *Wea. Service Bull.*, **73**, S1–S18.

Le Traon, P. Y., M. C. Rouquet, and C. Boissier, 1990: Spatial scales of mesoscale variability in the North Atlantic as deduced from Geosat data. *J. Geophys. Res.*, **95**, 20 267–20 285, https://doi.org/10.1029/JC095iC11p20267.

——, K. Klein, and B. L. Hua, 2008: Do altimeter wavenumber spectra agree with the interior or surface quasigeostrophic theory? *J. Phys. Oceanogr.*, **38**, 1137–1142, https://doi.org/10.1175/2007JPO3806.1.

——, and Coauthors, 2019: From observation to information and users: The Copernicus marine service perspective. *Front. Mar. Sci.*, **6**, 234, https://doi.org/10.3389/fmars.2019.00234.

Mann, H. B., 1945: Nonparametric tests against trend. *Econometrica*, **13**, 245–259, https://doi.org/10.2307/1907187.

Martin, M., A. Hines, and M. J. Bell, 2007: Data assimilation in the FOAM operational short-range ocean forecasting system: A description of the scheme and its impact. *Quart. J. Roy. Meteor. Soc.*, **133**, 981–995, https://doi.org/10.1002/qj.74.

——, and Coauthors, 2012: Group for High Resolution Sea Surface temperature (GHRSST) analysis fields inter-comparisons. Part I: A GHRSST multi-product ensemble (GMPE). *Deep-Sea Res. II*, **77–80**, 21–30, https://doi.org/10.1016/j.dsr2.2012.04.013.

McPhaden, M. J., 2012: A 21st century shift in the relationship between ENSO SST and warm water volume anomalies. *Geophys. Res. Lett.*, **39**, L09706, https://doi.org/10.1029/2012GL051826.

Merchant, C. J., and O. Embury, 2020: Adjusting for desert-dust-related biases in a climate data record of sea surface temperature. *Remote Sens.*, **12**, 2554, https://doi.org/10.3390/rs12162554.

——, and Coauthors, 2014: Sea surface temperature datasets for climate applications from phase 1 of the European Space Agency Climate Change Initiative (SST CCI). *Geosci. Data J.*, **1**, 179–191, https://doi.org/10.1002/gdj3.20.

——, and Coauthors, 2019: Satellite-based time-series of sea-surface temperature since 1981 for climate applications. *Sci. Data*, **6**, 223, https://doi.org/10.1038/s41597-019-0236-x.

Minnett, P. J., and A. K. Kaiser-Weiss, 2012: Group for High Resolution Sea-Surface Temperature discussion document: Near-surface oceanic temperature gradients. GHRSST, 7 pp., https://www.ghrsst.org/wp-content/uploads/2016/10/SSTDefinitionsDiscussion.pdf.

——, and Coauthors, 2019: Half a century of satellite remote sensing of sea-surface temperature. *Remote Sens. Environ.*, **233**, 111366, https://doi.org/10.1016/j.rse.2019.111366.

Monahan, A. H., and A. Dai, 2004: The spatial and temporal structure of ENSO nonlinearity. *J. Climate*, **17**, 3026–3036, https://doi.org/10.1175/1520-0442(2004)017<3026:TSATSO>2.0.CO;2.

O'Carroll A. G., and Coauthors, 2019: Observational needs of sea surface temperature. *Front. Mar. Sci.*, **6**, 420, https://doi.org/10.3389/fmars.2019.00420.

Okuro, A., M. Kubota, H. Tomita, and T. Hihara, 2014: Intercomparison of various global sea surface temperature products. *Int. J. Remote Sens.*, **35**, 5394–5410, https://doi.org/10.1080/01431161.2014.926415.

Pachauri, R. K., and Coauthors, 2014: *Climate Change 2014: Synthesis Report*. IPCC, 169 pp.

Peng, G., and Coauthors, 2019. Practical application of a data stewardship maturity matrix for the NOAA OneStop Project. *Data Sci. J.*, **18**, 41, https://doi.org/10.5334/dsj-2019-041.

Perlin, N., S. P. de Szoeke, D. B. Chelton, R. M. Samelson, E. D. Skyllingstad, and L. W. O'Neill, 2014: Modeling the atmospheric boundary layer wind response to mesoscale sea surface temperature perturbations. *Mon. Wea. Rev.*, **142**, 4284–4307, https://doi.org/10.1175/MWR-D-13-00332.1.

Pezzulli, S., D. B. Stephenson, and A. Hannachi, 2005: The variability of seasonality. *J. Climate*, **18**, 71–88, https://doi.org/10.1175/JCLI-3256.1.

Pisano, A., S. Marullo, V. Artale, F. Falcini, C. Yang, F. E. Leonelli, R. Santoleri, and B. Buongiorno Nardelli, 2020: New evidence of Mediterranean climate change and variability from sea surface temperature observations. *Remote Sens.*, **12**, 132, https://doi.org/10.3390/rs12010132.

Rayner, N., D. E. Parker, E. Horton, C. K. Folland, L. V. Alexander, D. Rowell, E. Kent, and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.*, **108**, 4407, https://doi.org/10.1029/2002JD002670.

——, and Coauthors, 2019: SST-CCI-Phase-II SST CCI climate assessment report issue 1. European Space Agency, 153 pp., http://www.esasstcci.org/PUG/pdf/SST_CCI-CAR-UKMO-201_Issue_1-signed.pdf.

Renault, L., S. Masson, V. Oerder, S. Jullien, and F. Colas, 2019: Disentangling the mesoscale ocean–atmosphere interactions. *J. Geophys. Res. Oceans*, **124**, 2164–2178, https://doi.org/10.1029/2018JC014628.

Reynolds, R. W., N. A. Rayner, T. M. Smith, D. C. Stokes, and W. Wang, 2002: An improved in situ and satellite SST analysis for climate. *J. Climate*, **15**, 1609–1625, https://doi.org/10.1175/1520-0442(2002)015<1609:AIISAS>2.0.CO;2.

——, T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax, 2007: Daily high-resolution-blended analyses for sea surface temperature. *J. Climate*, **20**, 5473–5496, https://doi.org/10.1175/2007JCLI1824.1.

Robinson, I., J.-F. Piolle, P. LeBorgne, D. Poulter, C. Donlon, and O. Arino, 2012: Widening the application of AATSR SST data to operational tasks through the Medspiration service. *Remote Sens. Environ.*, **116**, 126–139, https://doi.org/10.1016/j.rse.2010.12.019.

Sakurai, T., Y. Kurihara, and T. Kuragano, 2005: Merged satellite and in-situ data global daily SST. *Proc. 2005 IEEE International, IGARSS'05, Geoscience Remote Sensing. Symp.*, Seoul, South Korea, IEEE, 2606–2608.

Sen, P. K., 1968: Estimates of the regression coefficient based on Kendall's tau. *J. Amer. Stat. Assoc.*, **63**, 1379–1389, https://doi.org/10.1080/01621459.1968.10480934.

Smith, N. S., B. Ebert, and G. Warren, 1999: The Bureau of Meteorology SST analysis system. Report of the OOPC/AOPC Workshop on Global Sea Surface Temperature data sets, GCOS-57, GOOS-79, WMO/TD No. 978, Annex III, 22–31, https://library.wmo.int/doc_num.php?explnum_id=3911.

Su, Z., and Coauthors, 2018: An overview of European efforts in generating climate data records. *Bull. Amer. Meteor. Soc.*, **99**, 349–359, https://doi.org/10.1175/BAMS-D-16-0074.1.

Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.*, **106**, 7183–7192, https://doi.org/10.1029/2000JD900719.

Thomson, D. J., 1982: Spectrum estimation and harmonic analysis. *Proc. IEEE*, **70**, 1055–1096, https://doi.org/10.1109/PROC.1982.12433.

Titchner, H. A., and N. A. Rayner, 2014: The Met Office Hadley Centre sea ice and sea surface temperature data set, version 2: 1. Sea ice concentrations. *J. Geophys. Res. Atmos.*, **119**, 2864–2889, https://doi.org/10.1002/2013JD020316 (data available at https://www.metoffice.gov.uk/hadobs/hadisst2/).

Trenberth, K., and NCAR Research Staff, Eds., 2020: The Climate Data Guide: Niño SST Indices (Nino 1+2, 3, 3.4, 4; ONI and TNI). Accessed June 2020, https://climatedataguide.ucar.edu/climate-data/nino-sst-indices-nino-12-3-34-4-oni-and-tni.

Wang, C., R. M. Graham, K. Wang, S. Gerland, and M. A. Granskog, 2019: Comparison of ERA5 and ERA-Interim near-surface air temperature, snowfall and precipitation over Arctic sea ice: Effects on sea ice thermodynamics and evolution. *Cryosphere*, **13**, 1661–1679, https://doi.org/10.5194/tc-13-1661-2019.

Xu, F., and A. Ignatov, 2014: In situ SST Quality Monitor (iQuam). *J. Atmos. Oceanic Technol.*, **31**, 164–180, https://doi.org/10.1175/JTECH-D-13-00121.1.

Yasunaka, S., and K. Hanawa, 2011: Intercomparison of historical sea surface temperature datasets. *Int. J. Climatol.*, **31**, 1056–1073, https://doi.org/10.1002/joc.2104.

Zhong, A., and H. Beggs, 2008: Analysis and prediction operations bulletin No. 77—Operational implementation of global Australian multi-sensor sea surface temperature analysis, web document, 2 October 2008, http://www.bom.gov.au/australia/charts/bulletins/apob77.pdf.