Improvements of the Daily Optimum Interpolation Sea Surface Temperature (DOISST) Version 2.1

BOYIN HUANG,[a] CHUNYING LIU,[b] VIVA BANZON,[a] ERIC FREEMAN,[b] GARRETT GRAHAM,[c] BILL HANKINS,[b] TOM SMITH,[d] AND HUAI-MIN ZHANG[a]

[a] *NOAA/National Centers for Environmental Information, Asheville, North Carolina*
[b] *Riverside Technology, Asheville, North Carolina*
[c] *North Carolina Institute for Climate Studies, North Carolina State University, Asheville, North Carolina*
[d] *NOAA/NESDIS/Center for Satellite Applications and Research, College Park, Maryland*

(Manuscript received 11 March 2020, in final form 30 September 2020)

ABSTRACT: The NOAA/NESDIS/NCEI Daily Optimum Interpolation Sea Surface Temperature (SST), version 2.0, dataset (DOISST v2.0) is a blend of in situ ship and buoy SSTs with satellite SSTs derived from the Advanced Very High Resolution Radiometer (AVHRR). DOISST v2.0 exhibited a cold bias in the Indian, South Pacific, and South Atlantic Oceans that is due to a lack of ingested drifting-buoy SSTs in the system, which resulted from a gradual data format change from the traditional alphanumeric codes (TAC) to the binary universal form for the representation of meteorological data (BUFR). The cold bias against Argo was about −0.14°C on global average and −0.28°C in the Indian Ocean from January 2016 to August 2019. We explored the reasons for these cold biases through six progressive experiments. These experiments showed that the cold biases can be effectively reduced by adjusting ship SSTs with available buoy SSTs, using the latest available ICOADS R3.0.2 derived from merging BUFR and TAC, as well as by including Argo observations above 5-m depth. The impact of using the satellite *MetOp-B* instead of *NOAA-19* was notable for high-latitude oceans but small on global average, since their biases are adjusted using in situ SSTs. In addition, the warm SSTs in the Arctic were improved by applying a freezing point instead of regressed ice-SST proxy. This paper describes an upgraded version, DOISST v2.1, which addresses biases in v2.0. Overall, by updating v2.0 to v2.1, the biases are reduced to −0.07° and −0.14°C in the global ocean and Indian Ocean, respectively, when compared with independent Argo observations and are reduced to −0.04° and −0.08°C in the global ocean and Indian Ocean, respectively, when compared with dependent Argo observations. The difference against the Group for High Resolution SST (GHRSST) Multiproduct Ensemble (GMPE) product is reduced from −0.09° to −0.01°C in the global oceans and from −0.20° to −0.04°C in the Indian Ocean.

KEYWORDS: Sea surface temperature; Data processing; In situ oceanic observations; Satellite observations

1. Introduction

Sea surface temperature (SST) is an essential climate indicator and has been widely used in climate monitoring, assessment, and simulation, as well as in applications related to environmental protection, agriculture, and industry (IPCC 2013, 2018; EPA 2014). There are several types of SST products. The first is based purely on in situ observations from ships, drifting buoys, and moored buoys (and Argo floats in some products). Examples of this type of product include the Extended Reconstructed SST (ERSST; Smith et al. 1996; Smith and Reynolds 2003, 2004; Huang et al. 2015a, 2017, 2020), Met Office Hadley Centre SST (HadSST; Kennedy et al. 2011a,b, 2019), and Japan Meteorological Agency Centennial Observation-Based Estimates of SSTs (COBE-SST; Ishii et al. 2005; Hirahara et al. 2014). The second type of SST product is based solely on satellite observations; one example is the European Space Agency (ESA) Climate Change Initiative SST (CCI SST; Merchant et al. 2014, 2019). The third type of product is based on the blending of both in situ and satellite observations. These products include the Met Office Hadley Centre Ice and SST (HadISST; Rayner et al. 2003), Weekly Optimum Interpolation SST (WOISST; Reynolds et al. 2002), Daily Optimum Interpolation, version 2.0 (DOISST v2.0; Reynolds et al. 2007), and the Group for High Resolution SST (GHRSST) Multiproduct Ensemble (GMPE) product (Martin et al. 2012; Fiedler et al. 2019). Here we discuss improvements to the blended DOISST. The DOISST has been widely used in climate assessments and monitoring, such as the NOAA/NCEP Climate Forecast System Reanalysis (CFSR; Saha et al. 2010).

A common impediment in producing these SST datasets is the difficulty in homogenizing different types of observations, since biases of SST observations are dependent on instruments and observing platforms that vary with space and time (Kent et al. 2019). For example, historic ship observations are typically cold biased because of the cooling from buckets before the 1940s and are warm biased as a result of heating of engine room intake (ERI) after the 1940s (Kennedy et al. 2011a,b; Huang et al. 2015a, 2017; Kent et al. 2019). Satellite

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

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JCLI-D-20-0166.s1.

Corresponding author: Boyin Huang, boyin.huang@noaa.gov

DOI: 10.1175/JCLI-D-20-0166.1

observations may be biased by contaminations from clouds, volcano eruptions, and continental dust and aerosols (Zhang et al. 2004). The large-scale biases of ship SSTs may be adjusted toward more reliable observations such as nighttime marine air temperature, or SSTs from buoy and Argo measurements (Huang et al. 2015a 2017), or they may be adjusted by simulating the heat loss or gain (Folland and Parker 1995; Kennedy et al. 2011a,b; Carella et al. 2018). Satellite SSTs are calibrated against collocated buoy observations or by using a reduced-state-vector optimal estimation algorithm (Merchant et al. 2008, 2014, 2019).

Satellite and in situ systems measure the ''sea surface temperature'' at different depths. A perfectly accurate satellite SST retrieval of the sea surface skin temperature would be different from a perfectly accurate collocated buoy measurement of SST at a depth of approximately 0.2 m or a ship measuring an intake temperature (at a depth of several meters). Studies indicate that most satellite SSTs deviate slightly from in situ SSTs, and therefore a further adjustment of satellite SSTs is needed before they can be blended with in situ SSTs (Reynolds et al. 2002, 2007; Huang et al. 2015b, 2016). A common practice in these studies is to adjust satellite SSTs toward in situ SSTs. This suggests that the adjusted satellite SSTs are dependent on the accuracy and area coverage of in situ SSTs. Ship SSTs are noisier, while buoy and Argo SSTs are relatively accurate and homogeneous.

Since 2016, the in situ SSTs from the traditional alphanumeric codes (TAC) data stream have been gradually transitioned to the binary universal form for representation (BUFR) of meteorological data. Because DOISST v2.0 depended on TAC data, the buoy coverage ingested into the DOISST system decreased from near 30% of the global oceans in 2016 to near 5% in 2019 (Fig. 1a, solid green). The lower buoy coverage is mostly attributed to the lower drifting-buoy coverage, while the moored-buoy coverage does not change much. The lower drifting-buoy coverage decreased the total coverage of in situ SSTs by about 15%, from 55% in 2016 to 40% in 2019, because Argo SSTs were not included in v2.0. The transition of transmission of drifting-buoy measurements from TAC to BUFR lowered our ability to feed in situ measurements into the DOISST v2.0 system and therefore lowered the quality of DOISST v2.0.

DOISST v2.0 is also likely affected by an overcorrection of an assumed warm bias of ship SSTs since 2005, particularly after 2016. In DOISST v2.0, the bias of ship SSTs is assumed to be 0.14°C, based on its average between 1982 and 2000 (Reynolds et al. 2002). However, since 2004, the biases of ship SSTs against buoy SSTs have been clearly decreasing (Huang et al. 2017; Kennedy et al. 2019) and they approach zero after 2016 (Fig. 2). The lower coverage of buoy SSTs and the overcorrection of ship SSTs may have resulted in the recently identified cold bias in DOISST v2.0. The cold bias may also result from the use of degraded satellite SSTs from *NOAA-19* and the residual bias of satellite SSTs from *MetOp-A*, which cannot be resolved by the algorithms of bias correction in DOISST v2.0. In contrast, DOISST v2.0 exhibited a warm bias in the Arctic and Southern Oceans (Castro et al. 2016; Banzon et al. 2020), which may result from an inaccurate method for estimating SST by proxy from sea ice concentration (Banzon et al. 2020).

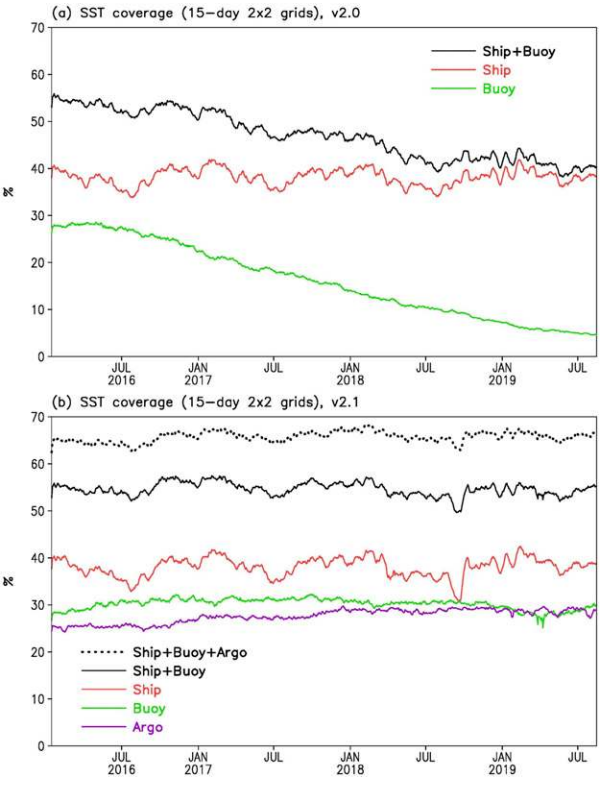

FIG. 1. Global SST coverages (%) on 15-day 2° × 2° grids in (a) OISST v2.0 and (b) OISST v2.1. The coverages of ship, buoy, Argo, ship + buoy, and ship + buoy + Argo are indicated in solid red, green, purple, black, and dotted black lines, respectively. The reason for using 15-day 2° × 2° grids is that the biases of satellite SSTs are calculated at these grids as described in section 2.

In this study, we document improvements in an upgraded version, DOISST v2.1, which addresses identified problems in v2.0 and assesses the relative contribution of each progressive improvement from v2.0 to v2.1. The rest of the paper is organized as follows: DOISST v2.0 and v2.1 are briefly described in section 2. Data and experiment design for each improvement are described in section 3. Comparisons with Argo observations in each experiment are presented in section 4, and comparisons with GMPE are presented in section 5. Our results are summarized and discussed in section 6.

2. DOISST v2.0 and v2.1

DOISST v2.0 and v2.1 have a resolution of daily and 0.25° × 0.25° (Table 1). DOISST v2.0 did not include BUFR ship, buoy and Argo float observations, but v2.1 does, starting from January 2016. Studies indicate that the ship SSTs exhibit biases that are due to changes in instruments and observing platforms (Kennedy et al. 2011a,b; Huang et al. 2015a, 2017). The biases of ship SSTs are assumed to be 0.14°C (the average of 1982–2000) in v2.0 from September 1981 to December 2015

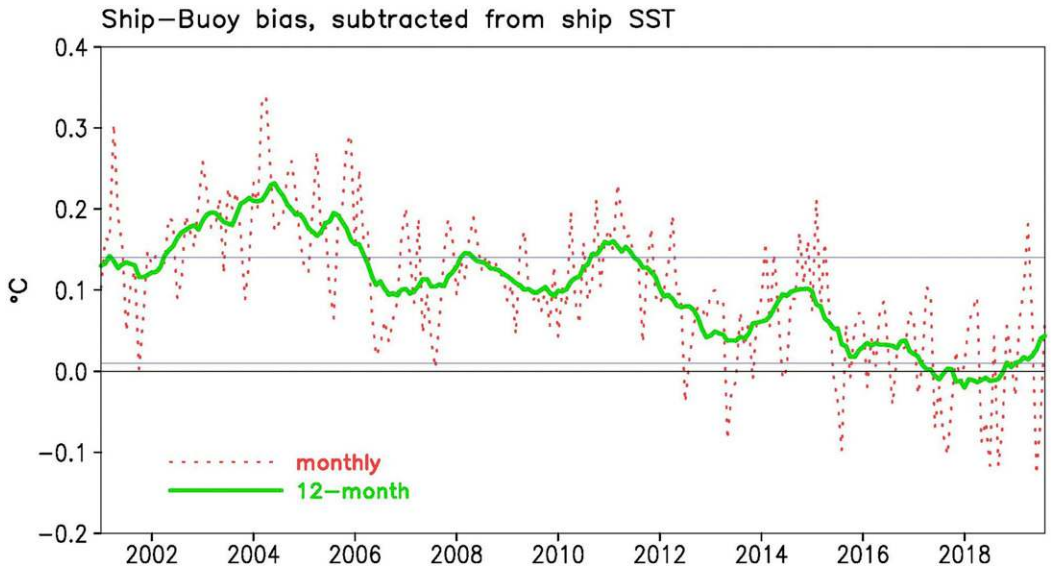

FIG. 2. Monthly (dotted red) and 12-month running averaged (solid green) ship SST biases (°C) defined as the difference between ship and buoy SSTs. SSTs are from ICOADS R3.0.0 before 2015 and from ICOADS-D R3.0.2 for 2016 and onward. Biases of 0.14° and 0.01°C are indicated by two gray horizontal lines.

(Reynolds et al. 2007) but are set to 0.01°C (the average from January 2016 to August 2019) in this study (v2.1) from January 2016 to the present (Fig. 2 and Table 1). Note that the ship SST bias decreases from 0.14°C to about 0.10°C between 2006 and 2015, which may result in a slight cold bias in DOISST v2.0. However, comparisons with other similar SST products indicate (not shown in figure) that the cold bias does not appear until mid-2016. The small impact of the slightly reduced ship bias may have been mitigated by a relatively small signal-to-noise ratio of ships (1.0) as compared with buoys and Argo floats (7.0; Huang et al. 2015a, 2017) when these in situ observations are merged.

The SSTs from the Advanced Very High Resolution Radiometer (AVHRR) measurements have a near-global

TABLE 1. A brief description of DOISST v2.0 and DOISST v2.1.

	DOISST v2.0	DOISST v2.1
Time period	Sep 1981–25 Apr 2020	Jan 2016–present; Sep 1981–Dec 2015: converted from v2.0 NetCDF 3.6 to NetCDF 4
Resolution	Daily 0.25° × 0.25°	Daily 0.25° × 0.25°
AVHRR sources	*NOAA-7* (1982–85); *NOAA-9* (1985–88); *NOAA-11* (1988–94); *NOAA-9* (1994–95); *NOAA-14* (1995–2000); *NOAA-16* (2000–02); *NOAA-17* (2003–05); *NOAA-18* (2005–06); *NOAA-17* (2006–08); *NOAA-18* (2007–11); *NOAA-19* (2011–15); *MetOp-A* (2009–present) from U.S. Navy	*MetOp-A* (2016–present) from the U.S. Navy; *MetOp-B* (2016–present) from the U.S. Navy
In situ data used to correct AVHRR biases	Ship and buoy observations	Ship, buoy, and Argo observations
In situ ship and buoy observations	ICOADS R2.5 (1981–2006); NCEP GTS receipts since 2007	ICOADS-D R3.0.2
In situ Argo observations	Not applicable	Argo GDAC; https://www.seanoe.org/data/00311/42182
Assumed bias of ship SSTs	0.14°C	0.01°C
Sea ice concentration	NASA (1981–2004); NCEP (2005–present)	NASA (1981–2004); NCEP (2005–present)
Proxy SST over ice-covered regions	Regression between SST and ice concentration	Freezing point based on climatological sea surface salinity

coverage but may contain systematic biases relative to in situ observations. AVHRR SSTs are retrieved from the radiance measurements, and therefore SST accuracy is critically dependent on retrieval algorithms. The infrared AVHRR can penetrate only the top millimeters of the sea surface and represent the skin temperature of the sea surface water, which is different from the in situ measurement at a nominal depth of 0.2 m. Therefore, AVHRR SSTs are more variable and sensitive to diurnal heating. AVHRR SSTs may also be contaminated by clouds and aerosols (Zhang et al. 2004; Huang et al. 2015b). The AVHRR SSTs are calibrated by in situ buoy observations, but the calibration in satellite SST retrieval algorithms may not work perfectly, for various reasons. For example, the calibration was done for averaged atmospheric conditions and might not be able to catch the spatially and temporally varying components. These are the reasons why the AVHRR SSTs are adjusted when they are ingested into the DOISST system (Reynolds and Smith 1994; Reynolds et al. 2002, 2007).

The large-scale biases of AVHRR SSTs are corrected against the available in situ observations from ships and buoys in DOISST v2.0 (Reynolds et al. 2007) and from ships, buoys, and Argo floats in this study (v2.1; Table 1). Daily biases of AVHRR SSTs (e.g., 15 January 2020) are calculated in the following procedures: 1) daily AVHRR and in situ SSTs are bin-averaged separately to 2° × 2° grids; 2) daily AVHRR and in situ SSTs are averaged separately within a 15-day running window (e.g., 8–22 January 2020); 3) the averaged AVHRR and in situ SSTs are projected onto a common set of empirical orthogonal teleconnection (EOT) functions; 4) the difference between EOT-filtered AVHRR and in situ SSTs is defined as AVHRR biases; and 5) the daily biases on 2° × 2° grids are interpolated linearly to 0.25° × 0.25° grids and applied to AVHRR SST [see more details in Reynolds et al. (2007); Huang et al. 2015b]. By adjusting the biases of AVHRR SSTs, the DOISST v2.0 and v2.1 represent the SSTs measured by in situ instruments at 0.2-m nominal depth. We note that residual biases may remain in bias-corrected AVHRR SSTs since biases at grid scales may not be resolved by EOTs and because in situ data are not available everywhere to inform the correction.

In the ice-covered regions of the Arctic and Southern Ocean, the proxy SST from ice concentration is blended with in situ and satellite SSTs. The sea ice concentration is from NASA (1981–2004; Cavalieri et al. 1996, 1999) and NCEP (2005–present; Grumbine 2014). The proxy SST is calculated using linear regressions over 28 regions in DOISST v2.0 (Reynolds et al. 2007) in two separate time periods, 1981–2014 and 2005–present. The proxy SST is replaced by freezing points of seawater in v2.1 when ice concentration is higher than 35%, based on the recent study of Banzon et al. (2020). The freezing points are calculated on the basis of climatological sea surface salinity.

3. Data and experiments

DOISST v2.0 (Table 1) used ship and buoy SSTs from the International Comprehensive Ocean–Atmosphere Datasets Release 2.4 (ICOADS R2.4) between September 1981 and December 2006 (Woodruff et al. 2011) and from the National Centers for Environmental Prediction (NCEP) Global Telecommunication System (GTS) receipts thereafter. DOISST v2.1 uses SSTs from ship and buoy observations from a merged TAC–BUFR ICOADS daily stream (ICOADS-D R3.0.2; C. Liu et al. 2020, unpublished manuscript) as well as Argo observations (Argo 2000; Roemmich et al. 2001) above 5-m depth from January 2016 to present. ICOADS-D R3.0.2 is a merging of TAC and BUFR data provided by the National Centers for Environmental Information (NCEI).

TABLE 2. Descriptions for DOISST experiments.

Expt	Description
MA+N19	*MetOp-A* and *NOAA19*, which is the same as DOISST v2.0
MA+MB	Same as MA+N19 but using *MetOp-A* and *MetOp-B*
FrzPnt	Same as MA+MB but using freezing-point (FrzPnt) over the regions of sea ice
Ship01	Same as FrzPnt but replacing ship bias of 0.14°C with 0.01°C
R3.0.2	Same as Ship01 but replacing ship and buoy SSTs from NCEP with SSTs from ICOADS-D R3.0.2
ALL	Same as R3.0.2 but including Argo temperatures above 5-m depth, which is the same as DOISST v2.1
Argo90%	Same as ALL but including 90% of Argo

DOISST v2.0 used satellite-based AVHRR observations from Pathfinder v5.0 (*NOAA-7*, 1982–85) and v5.1 (*NOAA-9*, 1985–88; *NOAA-11*, 1988–94; *NOAA-9*, 1994–95; *NOAA-14*, 1995–2000; *NOAA-16*, 2000–02; *NOAA-17*, 2003–05; *NOAA-18*, 2005–06), and from the U.S. Navy (*NOAA-17*, 2006–08; *NOAA-18*, 2007 to August 2011; *NOAA-19*, August 2011–15; *MetOp-A*, 2009–present) (Banzon et al. 2016). In DOISST v2.1 (Table 1), *NOAA-19* is replaced by *MetOp-B* from January 2016 to the present. *MetOp-A* and *MetOp-B* SSTs are retrieved by the U.S. Navy Naval Oceanographic Office.

The impacts of input data changes were progressively tested with the following six experiments from January 2016 to August 2019 (Table 2): MA+N19, using AVHRR SSTs from *MetOp-A* (MA) and *NOAA-19* (N19), which is the same as DOISST v2.0; MA+MB, which is the same as MA+N19 but replacing *NOAA-19* with *MetOp-B* (MB); FrzPnt, which is the same as MA+MB but using freezing-point (FrzPnt) proxy SST instead of ice-concentration proxy SST in ice-covered regions; Ship01, which is the same as FrzPnt but using ship SST bias of 0.01°C (Ship01) instead of 0.14°C; R3.0.2, which is the same as Ship01 using ship and buoy SSTs from ICOADS-D R3.0.2 instead of NCEP GTS receipts; and ALL, which is the same as R3.0.2 but including all (ALL) in situ observations from ships, buoys, and Argo floats above 5-m depth and has been implemented in operational DOISST v2.1.

These progressive experiments were designed to assess each revision in Table 2 on the improvement of DOISST

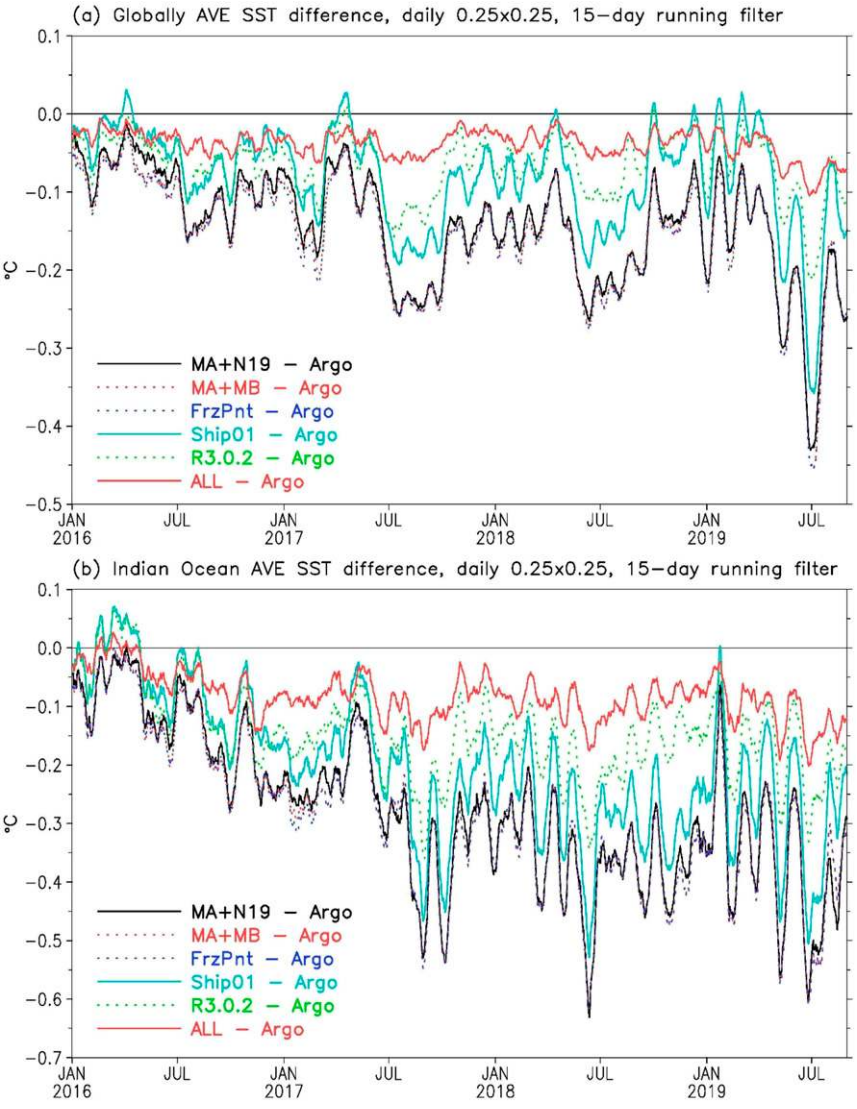

FIG. 3. (a) Globally averaged and (b) Indian Ocean averaged SST differences (°C) in daily 0.25° × 0.25° grids against Argo for experiments MA+N19 (solid black), MA+MB (dotted red), FrzPnt (dotted black), Ship01 (solid light blue), R3.0.2 (dotted green), and ALL (solid red). A 15-day running average is applied when plotted.

v2.1 over v2.0. The impacts of the MA+N19, MA+MB, FrzPnt, Ship01, and R3.0.2 experiments were assessed by comparison with independent Argo observations in section 4 and with GMPE SST in section 5. These impacts may change when the order of experiments is switched. The inclusion of Argo observations in final experiment ALL and operational DOISST v2.1 was intended to maximize the use of available observations. We note that the comparison between experiment ALL and Argo is not independent. To verify the role of Argo in DOISST v2.1, we ran an additional experiment Argo90% (Table 2), which is the same as experiment ALL except that only 90% of Argo profiles are randomly selected and included, leaving the other 10% of Argo profiles for independent verification.

4. Improvements toward DOISST v2.1

a. Biases in DOISST v2.0

The bias in DOISST v2.0 (equivalent to experiment MA+N19) was assessed by comparing globally averaged SST differences with Argo (Argo float data are not assimilated into this experiment, and thus serve as an independent validation dataset; Fig. 3a, solid black). Experiment MA+N19 showed a cold bias in the global oceans, which increases from about −0.1°C in 2016 to between −0.1° and −0.2°C during 2017–18, and finally to between −0.1° and −0.4°C during January–August 2019. The averaged bias is about −0.14°C (Table 3). Averaged between January 2016 and August 2019, the cold bias is approximately from −0.4° to −0.8°C in the

TABLE 3. Comparisons of DOISST experiments with Argo in averaged (from January 2016 to August 2019) biases and root-mean-square difference in the global oceans and Indian Ocean.

Expt	Mean global bias	Global RMSD	Mean Indian Ocean bias	Indian Ocean RMSD
MA+N19	−0.14°C	0.43°C	−0.28°C	0.50°C
MA+MB	−0.15°C	0.43°C	−0.28°C	0.50°C
FrzPnt	−0.15°C	0.43°C	−0.29°C	0.50°C
Ship01	−0.08°C	0.42°C	−0.20°C	0.50°C
R3.0.2	−0.07°C	0.38°C	−0.14°C	0.46°C
ALL	−0.04°C	0.24°C	−0.08°C	0.26°C

Indian Ocean and from −0.1° to −0.4°C in the western tropical Pacific, the subtropical North and South Pacific, and the tropical–subtropical Atlantic (Fig. 4a). The large bias in the Indian Ocean happens over the entire period of 2016–19, particularly after June 2017 (Fig. 3b, solid black). The averaged bias in the Indian Ocean from January to August 2019 is about −0.28°C, which is about 2 times the global averaged bias (Table 3), although the root-mean-square difference (RMSD) values are very close to those in the global average.

In addition to the overall increased cold bias with time in DOISST v2.0, the bias was relatively strong during the Northern Hemisphere summer (June–August; Fig. 3a, solid black). These stronger biases may result from the seasonal nature of higher cloudiness and dust aerosols in the tropical oceans. The higher cloudiness and dust aerosols can result in a cold bias of satellite AVHRR measurements as indicated in Zhang et al. (2004) and many other studies.

In contrast to the overall cold bias in MA+N19, there are warm biases (0.2°–0.6°C) in the western tropical Indian, eastern tropical Pacific, northwestern North Pacific, and northern North Atlantic Oceans (Fig. 4a). The RMSDs between MA+N19 and Argo are small (about 0.1°C) in most of the global oceans except for the western North Pacific and western North Atlantic (about 0.6°C) (Fig. 5a). Our analysis indicates that those mean biases and RMSDs are associated with sparse in situ observations and higher SST variabilities in these regions, particularly when most buoy observations were not ingested to the analysis during the transition from TAC to BUFR (Fig. 1a) in OISST v2.0.

b. *Role of* MetOp-B

By replacing *NOAA-19* with *MetOp-B* in experiment MA+MB (Fig. 3, dotted red), the area-averaged biases against Argo over the global and Indian Ocean are almost identical to that in MA+N19 (again, Argo float data are not assimilated into this experiment and thus serve as an independent validation dataset; Fig. 3, solid black). The spatial distributions of biases are very close in MA+MB and MA+N19 (Figs. 4a,b), as well as the RMSDs in MA+MB and MA+N19 (Figs. 5a,b). The similar biases in MA+MB and MA+N19 suggest that the use of *MetOp-B* has not improved the DOISST performance. This may seem counterintuitive, since *NOAA-19* has developed increasing bias in recent years. However, in blending different SST observations, the satellite SSTs are bias corrected by adjusting them relative to in situ SSTs, as described in section 2. Therefore, the magnitude and spatial distribution of adjusted large-scale satellite SSTs are in theory close to those of large-scale in situ SSTs (Huang et al. 2013, 2015b, 2016). However, since the coverage of buoy SSTs used in these experiments and OISST v2.0 was increasingly reduced after 2016 (Fig. 1a), the satellite biases have not been corrected well in either MA+N19 or MA+MB experiments. Therefore the large biases remain (Figs. 4a,b and 5a,b).

Another way to assess the role of *MetOp-B* is to calculate the SST difference between experiments MA+MB and MA+N19 (Fig. 6a). By using *MetOp-B*, the SST becomes slightly warmer (<0.05°C) in the Southern Ocean south of 50°S and coastal regions of the Arctic, and becomes slightly colder in the tropical and subtropical oceans. Overall, the globally averaged impact of using *MetOp-B* is small (Fig. 7, solid blue), which is consistent with the comparisons with Argo observations over the entire experiment period between January 2016 and August 2019 (Fig. 3, dotted red and solid black).

c. *Role of freezing point as an SST proxy*

Studies indicate that DOISST v2.0 over the ice-covered Arctic regions had a warm bias, sometimes as large as 2°C (Castro et al. 2016) when and where ice-SST regressions did not work well. A recent study (Banzon et al. 2020) indicated that the freezing point–based ice-SST proxy can reduce the warm bias over the ice-covered regions, which is validated by independent buoy observations. Therefore, the ice-SST regressions are replaced by freezing points when ice concentration is higher than 35% in v2.1 (Table 1).

By using freezing-point temperature as an SST proxy in DOISST v2.1 (Fig. 6b), SSTs decrease by from −0.1° to −0.3°C in most of the Arctic region and by −0.1°C along the coasts of the Southern Ocean. The lower SSTs in these regions are consistent with the conclusion that DOISST v2.0 exhibited a warm bias by applying ice-SST regressions (Castro et al. 2016; Banzon et al. 2020). In contrast, SSTs increase by 0.1°–0.3°C along the coasts of the European and Asian continents and the Queen Elizabeth Islands of Canada. The higher SSTs may be associated primarily with a higher freezing point due to a lower salinity forced by river runoff (e.g., the Queen Elizabeth Islands of Canada; https://www.nodc.noaa.gov/OC5/regional_climate/arctic).

The lower SSTs in Fig. 6b occur primarily in the Northern Hemisphere summer (June–August; Fig. 7, dotted red) when ice concentration is lower and ice-SST regressions do not work well. In contrast, the seasonal variations of ice concentration and freezing point are small in the Southern Ocean. Since there are few Argo observations in the regions of the ice-covered

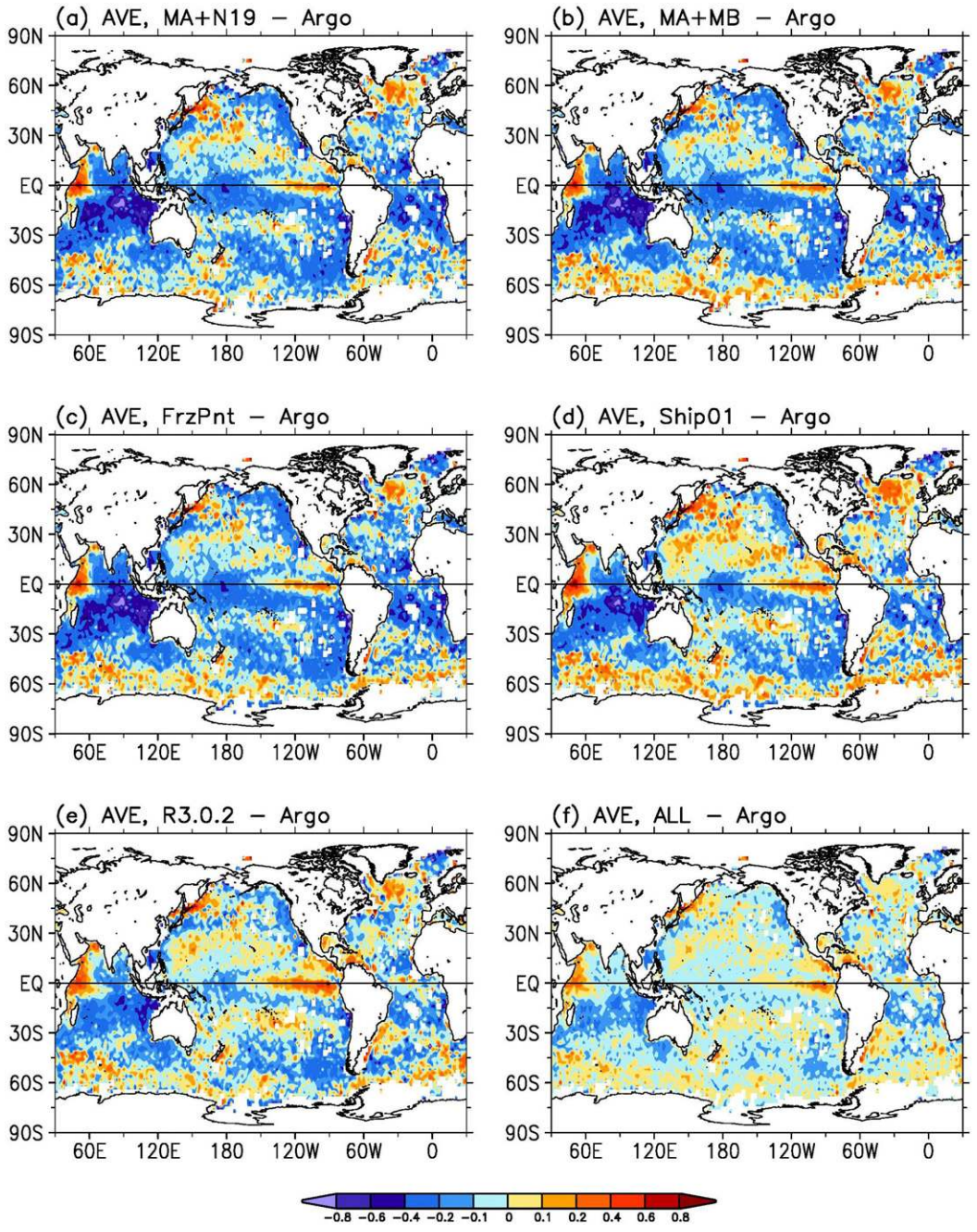

FIG. 4. Averaged (from January 2016 to August 2019) SST differences (°C) against Argo in (a) MA+N19, (b) MA+MB, (c) FrzPnt, (d) Ship01, (e) R3.0.2, and (f) ALL. The differences are averaged to 2° × 2° for visualization purposes.

Arctic and Southern Ocean, the comparisons of experiment FrzPnt with Argo (Fig. 3, dotted blue and red; Figs. 4c and 5c) remain almost the same as those in experiment MA+MB.

d. Role of ship SST bias correction

SST observations were obtained purely from ships before buoys and Argo floats became available, with buoys beginning to gain importance for global sampling in the mid-1980s (Freeman et al. 2017). Earlier studies indicated that ship SSTs were cold biased (from −0.2° to −0.3°C) because of the heat loss of buckets before the 1940s and warm biased (about 0.1°C) because of the heating from ERI after the 1940s (Kennedy et al. 2011a,b; Huang et al. 2015a, 2017). Recent studies showed that the biases of ship SST could be more

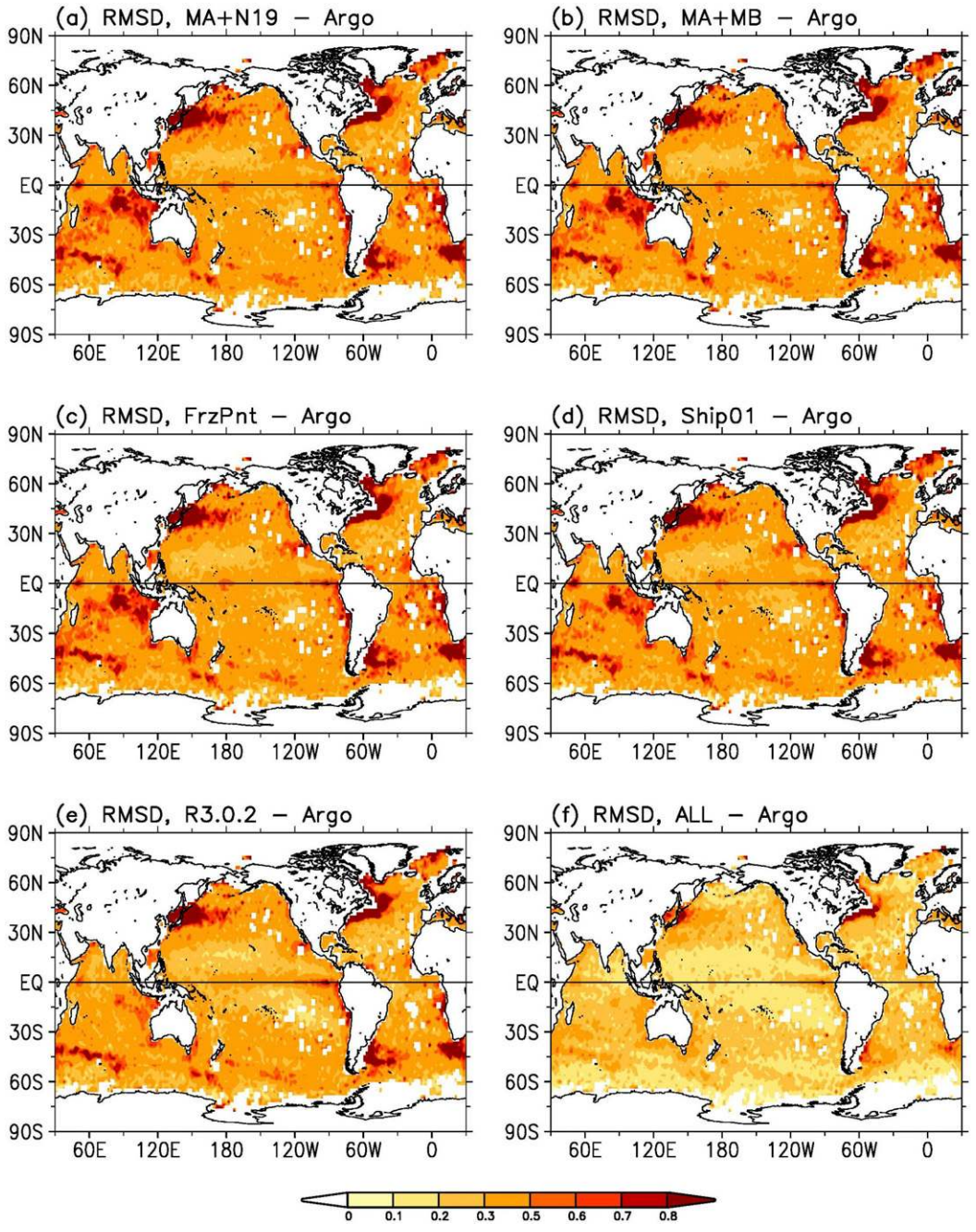

FIG. 5. SST RMSDs (°C; from January 2016 to August 2019) against Argo in daily 0.25° × 0.25° grids in (a) MA+N19, (b) MA+MB, (c) FrzPnt, (d) Ship01, (e) R3.0.2, and (f) ALL. The RMSDs are averaged to 2° × 2° for visualization purposes.

variable at the level of individual decks and at a global scale (Chan et al. 2019; Chan and Huybers 2019; Kennedy et al. 2019). In DOISST v2.0, a bias of 0.14°C was estimated from early years of data and removed from ship SSTs; however, this early estimate became clearly an overestimate after 2012 (Fig. 2). In particular, the biases are near zero after 2016. Therefore, the averaged bias of 0.01°C was applied in DOISST v2.1 after 2016 based on recent years of data (Table 1). The bias correction to ship SSTs may need to be refined further in the future development of DOISST to correctly quantify its variabilities at different geographic locations at seasonal to interannual time scales.

By setting the biases of ship SSTs to near zero (0.01°C) in experiment Ship01, the globally and time-averaged SST bias against Argo observations (again, Argo float data are not

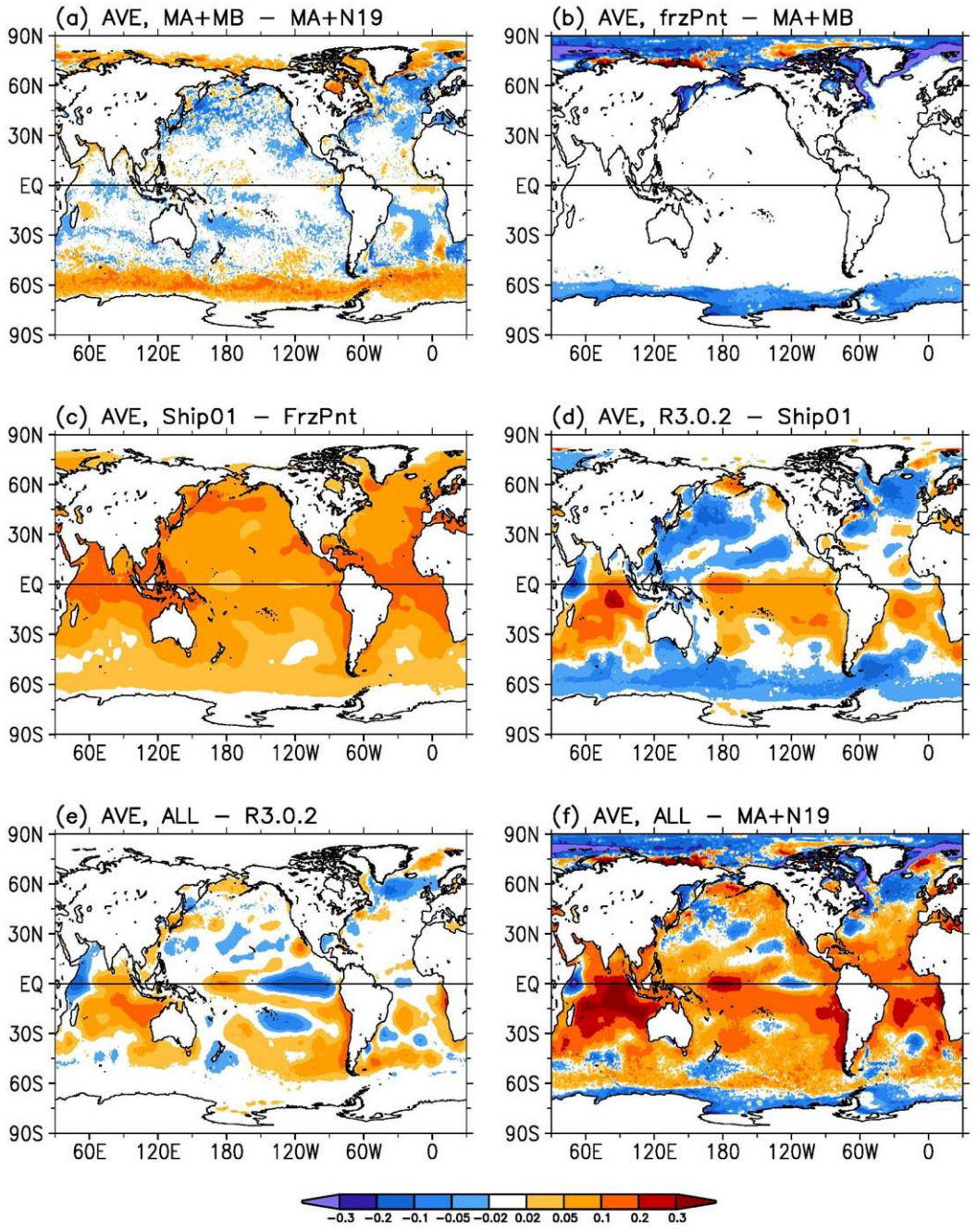

FIG. 6. Averaged (from January 2016 to August 2019) SST differences (°C) in daily 0.25° × 0.25° grids between progressive experiments of (a) MA+MB minus MA+N19, (b) FrzPnt minus MA+MB, (c) Ship01 minus FrzPnt, (d) R3.0.2 minus Ship01, (e) ALL minus R3.0.2, and (f) ALL minus MA+N19.

assimilated into this experiment, and thus serve as an independent validation dataset in this step) decreases from −0.15° to −0.08°C (Table 3), particularly after June 2017 and in the Indian Ocean (Fig. 3, solid light blue and dotted blue; Table 3). The reduction of the cold biases is evident not only in the Indian Ocean but also in the tropical Atlantic (Figs. 4c,d), although weak warm biases increase slightly in the western equatorial Indian Ocean, eastern equatorial Pacific, and North Atlantic south of Greenland. However, the RMSDs remain similar in FrzPnt and Ship01 (Figs. 5c,d; Table 3).

The impact of reducing ship SST bias correction is evident (~0.10°C) over the global oceans (Fig. 6c) where ship observations are available, particularly in the Indian Ocean, the tropical–subtropical Atlantic, and the North Pacific near 50°N.

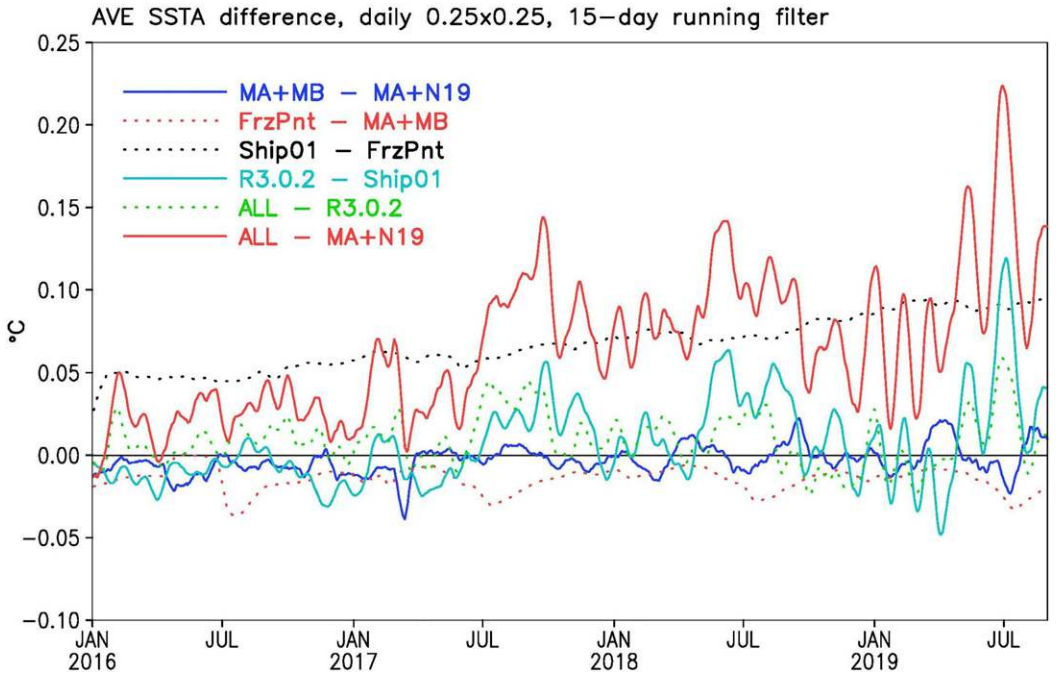

FIG. 7. Globally averaged SST differences (°C) in daily 0.25° × 0.25° grids between progressive experiments of MA+MB minus MA+N19 (solid blue), FrzPnt minus MA+MB (dotted red), Ship01 minus FrzPnt (dotted black), R3.0.2 minus Ship01 (solid light blue), ALL minus R3.0.2 (dotted green), and ALL minus MA+N19 (solid red). A 15-day running average is applied when plotted.

For the global average, SSTs increased by 0.05°C in 2016 and by approximately 0.10°C in 2019 (Fig. 7, dotted black). The enhancement of SST increase is associated with the reduction of buoy SST coverage (Fig. 1a, solid green), since the coverage of ship SST is near a constant (Fig. 1a, solid red). This implies that ship SSTs become more important in the merged in situ SST and the bias correction of satellite SSTs, as the observations from buoys ingested into DOISST v2.0 become fewer and fewer since the transition from TAC to BUFR lowered our ability to feed in situ measurements into the DOISST v2.0 system. That problem has been corrected in v2.1.

e. Role of ICOADS-D R3.0.2

One of the most serious problems in DOISST v2.0 is that the coverage of the ingested buoy SST decreased from nearly 30% in 2016 to only 5% by 2019 (Fig. 1a, solid green), which decreased the total coverage of Ship+Buoy from about 55%–40% by 2019 (Fig. 1a, solid black). As discussed earlier, the decrease of buoy SSTs was caused by the transition of GTS data transmissions from TAC to BUFR (C. Liu et al. 2020, unpublished manuscript). The reduction of buoy SSTs first affects the accuracy and spatial coverage of merged in situ data from ship and buoy SSTs, and then affects the accuracy of bias correction to satellite SSTs, as indicated in our earlier studies (Huang et al. 2013, 2015b, 2016).

By using ICOADS-D R3.0.2 derived from merging TAC and BUFR (C. Liu et al. 2020, unpublished manuscript) the coverage of merged in situ SSTs increases from a range of 40%–55% in DOISST v2.0 (Fig. 2a, solid black) to about 65% in v2.1 (Fig. 2, dotted black). SST biases in experiment R3.0.2 decrease significantly over the global oceans during June–November 2017, May–August 2018, and May–August 2019 (Fig. 3a, dotted green and solid light blue) and over the Indian Ocean during January–April 2017 and from July 2018 to August 2019 (Fig. 3b). For the time average from January 2016 to August 2019, SST biases decrease slightly from −0.08°C in experiment Ship01 to −0.07°C in experiment R3.0.2 over the global oceans and from −0.20° to −0.14°C in the Indian Ocean (Table 3). The reduction of the cold biases is evident not only in the Indian Ocean but also in the central equatorial Pacific near the date line and in the South Atlantic (Fig. 4e). The use of ICOADS-D R3.0.2 also reduces the warm biases in the west equatorial Indian Ocean and North Atlantic south of Greenland, but the warm bias in the eastern equatorial Pacific increases slightly. Overall, the RMSD against Argo (again, Argo float data are not assimilated into this experiment, and thus serves as an independent validation dataset in this step; Figs. 5d,e) decreases between experiments Ship01 and R3.0.2. The RMSD decreased from 0.42° to 0.38°C (Table 3) in global average and from 0.50° to 0.46°C in the Indian Ocean, indicating a clear improvement in DOISST biases by using ICOADS-D R3.0.2. Similar features are found when these experiments are compared with GMPE in section 5.

It is interesting that the impact of using ICOADS-D R3.0.2 is not uniformly distributed over the global oceans or in time (Fig. 6d). SSTs increase by about 0.02°C in the tropical and South Pacific, by about 0.02°C in the tropical and South Atlantic, and by about 0.10°–0.30°C in the Indian Ocean except

TABLE 4. Comparisons of DOISST experiments with Argo10% in averaged (from January 2016 to August 2019) biases and root-mean-square difference in the global oceans and Indian Ocean.

Expt	Mean global bias	Global RMSD	Mean Indian Ocean bias	Indian Ocean RMSD
MA+N19	−0.14°C	0.41°C	−0.24°C	0.48°C
MA+MB	−0.15°C	0.41°C	−0.26°C	0.47°C
FrzPnt	−0.15°C	0.41°C	−0.26°C	0.47°C
Ship01	−0.08°C	0.40°C	−0.18°C	0.44°C
R3.0.2	−0.07°C	0.36°C	−0.12°C	0.38°C
Argo90%	−0.06°C	0.31°C	−0.09°C	0.33°C

for the region along the Somalia coast. The increase of SSTs in these regions demonstrates the capability of using ICOADS-D R3.0.2 in correcting the cold biases of AVHRR satellite measurements due to the cloudiness and dust aerosols over the tropical oceans, especially during the Northern Hemisphere summer (Zhang et al. 2004). In contrast, SSTs decrease by 0.05°C in most of the Northern Hemisphere oceans and in the Southern Ocean south of 45°S. These opposite changes result in an overall small increase in globally averaged SST during periods of July–December 2017, May–August 2018, and May–August 2019 (Fig. 7, solid light blue), which is consistent with a small reduction in globally averaged bias against Argo as shown in Fig. 3a (dotted green and solid light blue).

f. Role of Argo

Our previous studies indicated that Argo observations can play an important role in SST reconstruction, particularly in data-sparse regions such as the equatorial oceans and the Southern Ocean where ships and surface drifters are hard to sustain (Huang et al. 2017, 2019). The importance of Argo observations can be seen clearly from the increase of in situ data coverage: the total coverage of in situ SSTs is about 65% (Fig. 1b, dotted black) when Argo SSTs are included, and about 55% (Fig. 1b, solid black) when Argo SSTs are not included.

By using all available in situ observations including Argo, SSTs in experiment ALL increase by 0.05°–0.10°C in the Indian Ocean, the western tropical Pacific, the South Pacific between 30° and 50°S and along the coasts of South America, and the South Atlantic (Fig. 6e). It should be pointed out that SSTs in experiment R3.0.2, which does not include Argo, suffer from cold biases exactly in these regions as shown in Fig. 4e. Therefore, it is easy to understand that the inclusion of Argo should have improved the SST quality in experiment ALL as suggested by the comparison with Argo, particularly in the Indian Ocean (Figs. 4f and 5f).

In contrast, by including Argo observations, SSTs decrease by between −0.05° and −0.10°C in the western equatorial Indian Ocean along the Somalia coast, eastern equatorial Pacific, central South Pacific near 25°S, South Pacific near New Zealand, and North Atlantic south of Greenland (Fig. 6e). The SST decrease in these regions has clearly reduced the warm biases in v2.1 (Figs. 4f and 5f) in comparison with experiment R3.0.2 (without including Argo; Fig. 4e).

On the global average, the reduction of biases appears over July–November 2017, May–August 2018, and April–August 2019 (Fig. 3, solid red and dotted green). For the global and time average (January 2016 to August 2019), the inclusion of Argo observations has reduced the mean bias from −0.07° to −0.04°C and reduced the RMSD from 0.38° to 0.24°C (Table 3). In contrast, in the Indian Ocean (Fig. 3b), the biases are improved from −0.14° to −0.08°C on average and from 0.46° to 0.26°C in RMSD (Table 3), indicating an important impact of Argo on improving DOISST biases.

Note that those reductions in mean bias and RMSD are partly associated with comparisons with the same Argo data included in experiment ALL; that is, in this ALL experiment, the Argo data are not an independent dataset, but are merged with ship and buoy data. To verify the role of Argo in DOISST v2.1 shown in Table 3 and Figs. 3, 4f, and 5f by an independent dataset, an additional experiment Argo90% (Table 2) was undertaken. Experiment Argo90% was the same as experiment ALL except that only 90% of Argo profiles were randomly selected and included, leaving the other 10% of Argo profiles (Argo10%) as an independent dataset for verification. Our analyses indicate that the comparisons with independent Argo10% (Table 4) are qualitatively consistent with those with Argo (Table 3; figures are shown in the online supplemental information). Quantitatively, however, the improvements of using Argo on DOISST become smaller in experiment Argo90% than in experiment ALL, since verification is independent in Argo90% but not completely independent in ALL.

5. Comparisons with GMPE

To further illustrate the improvements of DOISST v2.1 over v2.0, DOISST experiments (Table 2) are compared with GMPE SST (Martin et al. 2012; Fiedler et al. 2019). GMPE SST is derived from the median of an ensemble of 16 SST products of GHRSST in daily 0.25° × 0.25° grids and is commonly used as a reference in the SST data intercomparisons within the GHRSST community. Note that DOISST v2.0 is included in GMPE, and thus DOISST v2.1 may not be entirely independent from GMPE. Our comparisons indicate that the globally averaged differences (Fig. 8a) are consistent with those against Argo observations (Fig. 3a): The differences remain large in experiments MA+N19, MA+MB, and FrzPnt, and are clearly reduced in experiments Ship01, R3.0.2, and ALL. The mean differences are −0.09°, −0.09°, −0.09°, −0.02°, −0.02°, and −0.01°C, respectively, in experiments MA+N19, MA+MB, FrzPnt, Ship01, R3.0.2, and ALL (Table 5). Note that these differences against GMPE are smaller than those against Argo in Table 3, which may result from 1) a large sample in the global average since every

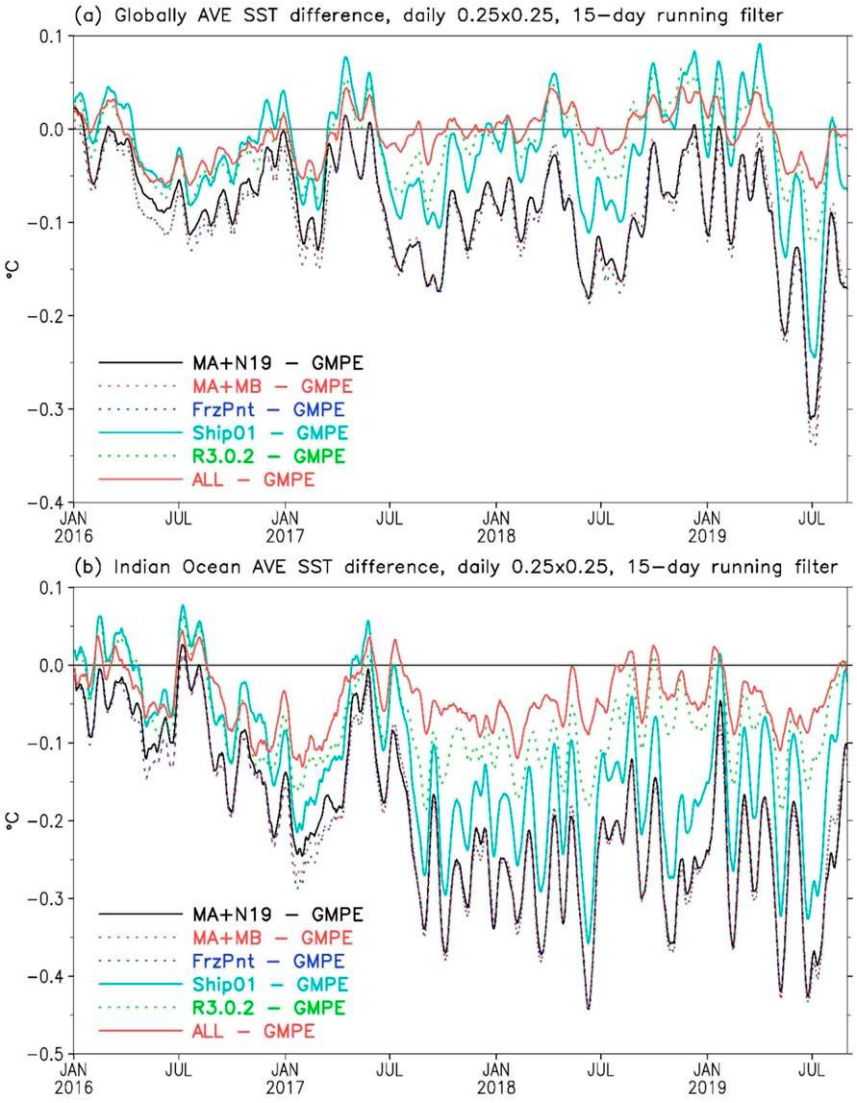

FIG. 8. (a) Globally and (b) Indian Ocean averaged SST differences (°C) in daily 0.25° × 0.25° grids against GMPE for experiments MA+N19 (solid black), MA+MB (dotted red), FrzPnt (dotted black), Ship01 (solid light blue), R3.0.2 (dotted green), and ALL (solid red). A 15-day running average is applied when plotted.

oceanic grid has a valid SST value when compared with GMPE and/or 2) a potential cold bias in GMPE SST since the cold-biased DOISST v2.0 is included in GMPE. The RMSDs against GMPE are 0.40°, 0.39°, 0.40°, 0.40°, 0.36°, and 0.32°C, respectively, in experiments MA+N19, MA+MB, FrzPnt, Ship01, R3.0.2, and ALL (Table 5). Note that the RMSD reduction (0.04°C) is nearly the same between experiments R3.0.2 and Ship01 and between ALL and R3.0.2, indicating

TABLE 5. Comparisons of DOISST experiments with GMPE in averaged (from January 2016 to August 2019) biases and RMSD in the global oceans and Indian Ocean.

Expt	Mean global bias	Global RMSD	Mean Indian Ocean bias	Indian Ocean RMSD
MA+N19	−0.09°C	0.40°C	−0.20°C	0.40°C
MA+MB	−0.09°C	0.39°C	−0.20°C	0.39°C
FrzPnt	−0.09°C	0.40°C	−0.20°C	0.39°C
Ship01	−0.02°C	0.40°C	−0.12°C	0.39°C
R3.0.2	−0.02°C	0.36°C	−0.07°C	0.33°C
ALL	−0.01°C	0.32°C	−0.04°C	0.29°C

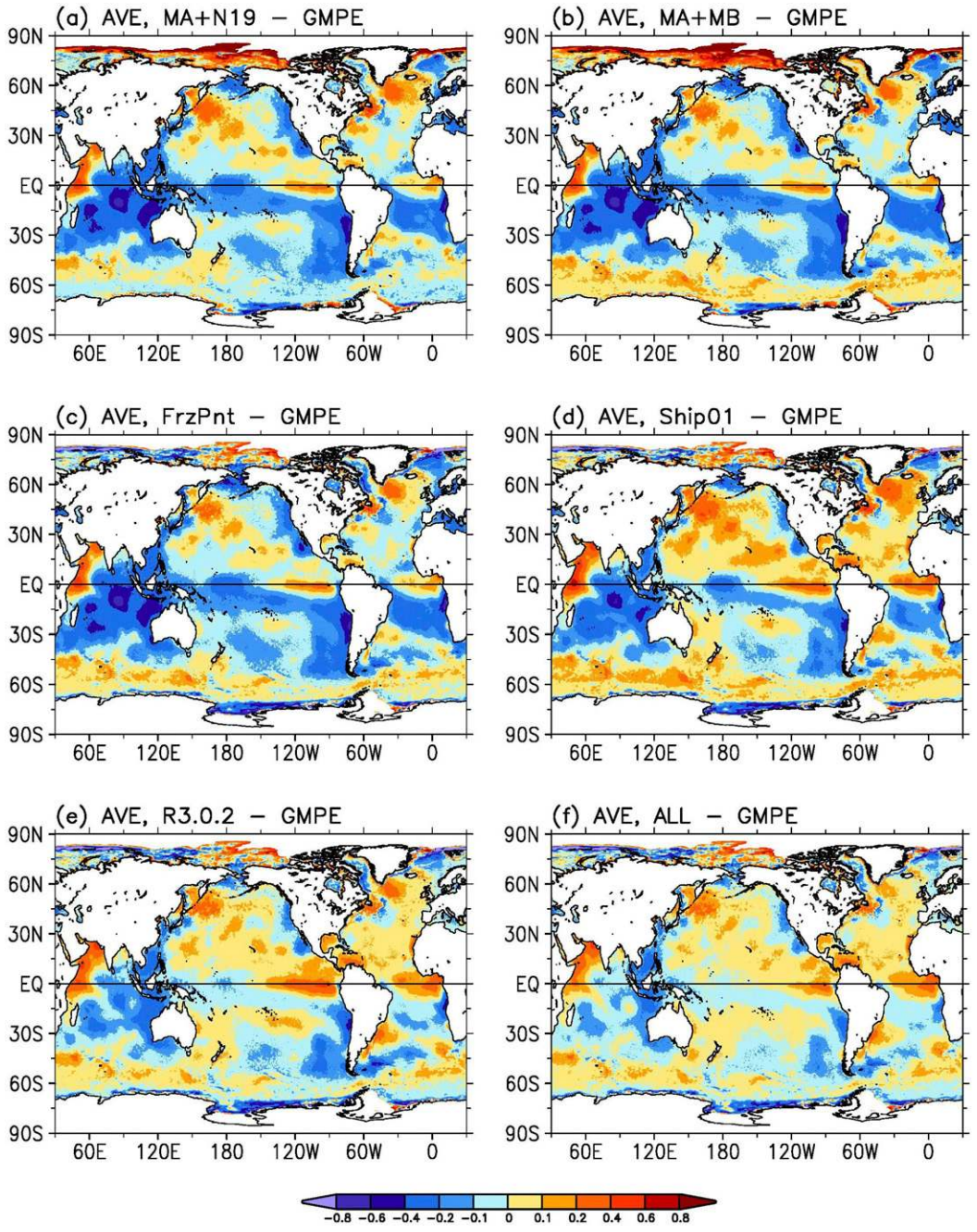

FIG. 9. Averaged (from January 2016 to August 2019) SST differences (°C) in daily 0.25° × 0.25° grids against GMPE in (a) MA+N19, (b) MA+MB, (c) FrzPnt, (d) Ship01, (e) R3.0.2, and (f) ALL.

an equally important role of using ICOADS-D R3.02 and Argo SST in reducing DOISST biases.

Similar features in improving the DOISST bias are found in the Indian Ocean (Fig. 8b). The mean differences against GMPE are −0.20°, −0.20°, −0.20°, −0.12°, −0.07°, and −0.04°C, respectively, in experiments MA+N19, MA+MB, FrzPnt, Ship01, R3.0.2, and ALL (Table 5). The RMSDs against GMPE are 0.40°, 0.39°, 0.39°, 0.33°, and 0.29°C. Note that the reductions of mean differences are slightly higher between Ship01 and R3.0.2 (0.05°C) than between R3.0.2 and ALL (0.03°C). Similarly, the reduction of RMSDs is higher between Ship01 and R30.02 (0.06°C) than between R3.0.2 and ALL (0.04°C). The magnitude of these bias reductions suggests that the impact of using ICOADS-D R3.0.2 is larger than that when additionally using Argo SSTs. Similar to the comparisons with Argo, SSTs in DOISST v2.0 (experiment MA+NA;

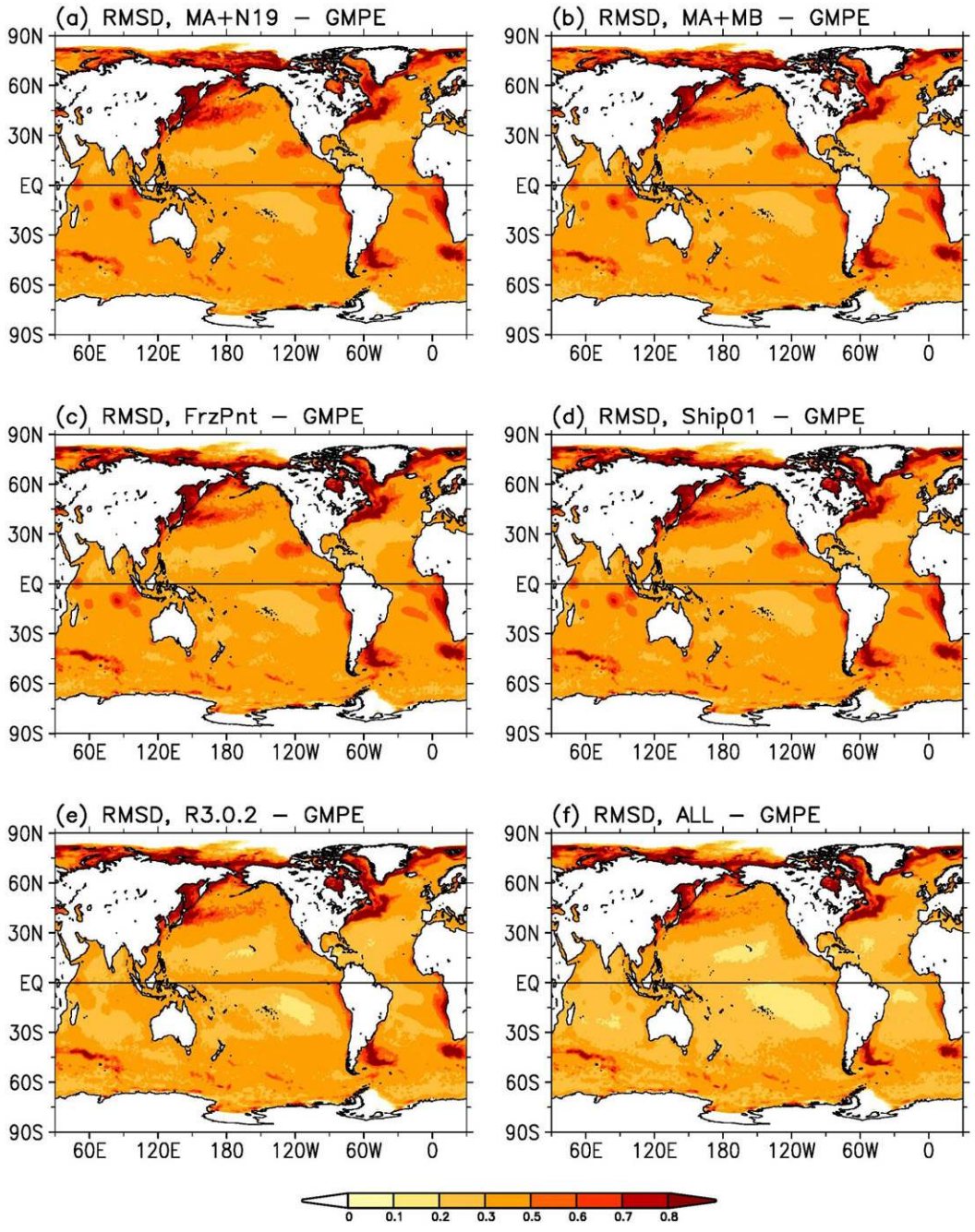

FIG. 10. SST RMSDs (°C; from January 2016 to August 2019) against GMPE in daily 0.25° × 0.25° grids in (a) MA+N19, (b) MA+MB, (c) FrzPnt, (d) Ship01, (e) R3.0.2, and (f) ALL.

Fig. 9a) are clearly colder than those in GMPE in the Indian Ocean (from −0.2° to −0.6°C), western tropical Pacific (−0.2°C), eastern South Pacific (from −0.2° to −0.6°C), and South Atlantic between 10° and 30°S (−0.2°C). In contrast, SSTs in v2.0 are warmer than GMPE along the coasts of the Arctic (as high as 0.8°C), western North Pacific, western North Atlantic, western equatorial Indian Ocean, and eastern equatorial Pacific (0.2°C). These differences do not change much when *MetOp-B* replaces *NOAA-19* (Fig. 9b). However, the positive SST difference along the Arctic coasts is clearly reduced when the use of freezing-point SST proxy replaces the use of ice-SST proxy (Fig. 9c), although its contribution to the reduction of globally averaged difference is small. When biases of ship SSTs are set to 0.01°C (Fig. 9d), the negative SST differences are reduced in the Indian Ocean, western tropical Pacific and South Pacific, and South Atlantic, but positive

SST differences increase slightly in the North Pacific, North Atlantic, eastern equatorial Pacific, eastern equatorial Atlantic, and western equatorial Indian Ocean. These cold and warm SST differences are reduced effectively when ICOADS-D R3.0.2 data are used, except for the warm SST differences in the equatorial oceans (Fig. 9e). The cold SST differences are reduced further by using Argo SSTs in the Indian Ocean, South Pacific, and South Atlantic, but remain in the maritime continental regions (Fig. 9f). The warm SST difference is also reduced in the eastern equatorial Pacific but remains in the western equatorial Indian Ocean and eastern equatorial Atlantic. These results clearly suggest an important role of Argo in DOISST, as described in section 4.

In contrast to the clear reductions in averaged SST difference, the reductions in RMSD are less (Fig. 10; Table 5), although the reduction of RMSD is notable in the tropics of the global ocean basins between 30°S and 30°N, when ICOADS-D R3.0.2 and Argo observations are used (Figs. 10e,f). These results are consistent with those comparisons with Argo in section 4.

6. Summary, discussion, and conclusions

DOISST v2.0 has been updated to v2.1 from January 2016 onward. The updates include the following five aspects: 1) *MetOp-B* replaces *NOAA-19*, 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 R3.0.2 are used instead of NCEP GTS receipts, and 5) Argo observations above 5-m depth are included. It should be kept in mind that the Argo observations were used as independent data to validate the improvements in updates 1–4, but the Argo observations become dependent in the verification of the final update. For an independent validation, an additional experiment was carried out with 10% of Argo float data reserved and not used in DOISST. Among these updates, the ship SST bias revision and the use of ICOADS-D R3.0.2 and Argo have played the most important role in reducing the global biases in v2.0. The SST proxy using freezing-point temperature has reduced the warm bias in the Arctic and Southern Ocean, but does not greatly reduce global average bias. The impact of using *MetOp-B* in reducing the bias in v2.0 is trivial, since satellite SSTs are adjusted against the available in situ SSTs.

By updating DOISST from v2.0 to v2.1, the globally and time-averaged (January 2016 to August 2019) bias is reduced from −0.14° to −0.07°C, and RMSD reduces from 0.43° to 0.38°C when compared with independent Argo observations during the progressive development processes. The bias and RMSD in the final v2.1 are reduced to −0.04° and 0.24°C, respectively, when compared with dependent Argo observations. The reduction of biases is most evident in the Indian Ocean, where v2.0 suffers a serious cold bias since 2016 due to a reduction of drifting buoy observations ingested into v2.0, which was a result of the inability of the system switching from TAC to BUFR-formatted data.

The biases in the Indian Ocean are reduced from −0.28° to −0.14°C on average and from 0.50° to 0.46°C in RMSD when compared with independent Argo observations during the progressive development processes. The average bias and RMSD in the final v2.1 are reduced to −0.08° and 0.26°C, respectively, when compared with dependent Argo observations. Biases are also reduced in the tropical and South Pacific and in the tropical and South Atlantic. These results remain robust when DOISST v2.0 and v2.1 are compared with the GMPE ensemble median of 16 SST products.

In comparison between DOISST v2.1 and v2.0, the globally averaged SST increases from about 0.02° to 0.10°C between January 2016 and August 2019 (Fig. 7, solid red). This SST increase overcomes the increasing cold bias of v2.0 due to the reduction of ingested buoy observations, as indicated in Fig. 3a (solid black). The average increase of SSTs is nearly global except for the Arctic and along the coasts of the Antarctic (Fig. 6f), with the highest magnitude (about 0.3°C) in the Indian Ocean, central equatorial Pacific near the date line, South Pacific along the coasts of South America, South Atlantic, and along the coasts of the Arctic.

However, it should be noted that DOISST v2.1 may still have a residual cold bias of about −0.04°C over the global oceans (Fig. 3a, solid red), particularly near the end of the experiment period (June–August 2019) and about −0.08°C in the Indian Ocean (Fig. 3b, solid red). Comparisons with GMPE SST indicate that the global averaged difference is very small (Fig. 8a, solid red), but a notable difference remains in the Indian Ocean (Fig. 8b, solid red). These differences may result from residual biases of satellite measurements that cannot be resolved by the bias correction algorithm in DOISST. Those differences may also result from nonhomogeneous in situ measurements from ships, buoys, and Argo floats. Our future work will focus on detecting the source of this residual bias and correcting it.

In conclusion, DOISST v2.1 has overcome most of the cold biases exhibited in the previous v2.0, particularly in the Indian Ocean, South Pacific, and South Atlantic, by recovering previously unused buoy SSTs using the latest ICOADS-D R3.0.2, by including Argo SSTs that were not used in v2.0, by revising ship SST bias correction according to available ship and buoy SSTs, and by using a new SST proxy scheme in the Arctic region.

Acknowledgments. The authors thank the three anonymous reviewers, whose comments have greatly improved the paper in scientific presentation and readability. The contents of this paper are solely the opinions of the authors and do not constitute a statement of policy, decision, or position on behalf of NOAA or the U.S. government.

Data availability statement. ICOADS-D R3.0.2 data are provided at NOAA/NCEI (https://doi.org/10.7289/V5CZ3562; access date: 1 October 2019). OISST v2.1 data are available at NOAA/NCEI (https://www.ncdc.noaa.gov/oisst/optimum-interpolation-sea-surface-temperature-oisst-v21). *NOAA-19*, *MetOp-A*, and *MetOp-B* SST were provided by the NOAA/NCEI Common Ingest system (ftp://data-internal.ncei.noaa.gov with access permission; access date: October 2019). Argo data were provided by the Global Data Assembly Centre (GDAC; https://doi.org/10.17882/42182 and http://www.seanoe.org/data/00311/42182; access date: October 2019). Sea ice concentrations were retrieved from NCEP (http://ftpprd.ncep.noaa.gov/data/nccf/com/omb/prod; access date October 2019) and NASA https://nsidc.org/data/NSIDC-0051/versions/1; access date January 2006). The

GMPE data were provided by GHRSST, Met Office, and CMEMS (ftp://nrt.cmems-du.eu; access date 5 February 2020).

REFERENCES

Argo, 2000: Argo float data and metadata from Global Data Assembly Centre (Argo GDAC). SEANOE, accessed 1 January 2020, https://www.seanoe.org/data/00311/42182/.

Banzon, V., T. M. Smith, 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.

——, ——, M. Steele, B. Huang, and H.-M. Zhang, 2020: Improved estimation of proxy sea surface temperature in the Arctic. *J. Atmos. Oceanic Technol.*, **37**, 341–349, https://doi.org/10.1175/JTECH-D-19-0177.1.

Carella, G., J. J. Kennedy, D. I. Berry, S. Hirahara, C. J. Merchant, S. Morak-Bozzo, and E. C. Kent, 2018: Estimating sea surface temperature measurement methods using characteristic differences in the diurnal cycle. *Geophys. Res. Lett.*, **45**, 363–371, https://doi.org/10.1002/2017GL076475.

Castro, S. L., G. A. Wick, and M. Steele, 2016: Validation of satellite sea surface temperature analyses in the Beaufort Sea using UpTempO buoys. *Remote Sens. Environ.*, **187**, 458–475, https://doi.org/10.1016/j.rse.2016.10.035.

Cavalieri, D. J., C. L. Parkinson, P. Gloersen, and H. J. Zwally, 1996: Sea ice concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS passive microwave data, version 1. NASA National Snow and Ice Data Center Distributed Archive Center, accessed January 2006, https://doi.org/10.5067/8GQ8LZQVL0VL.

——, ——, ——, J. C. Comiso, and H. J. Zwally, 1999: Deriving long-term time series of sea ice cover from satellite passive-microwave multisensory data sets. *J. Geophys. Res.*, **104**, 15 803–15 814, https://doi.org/10.1029/1999JC900081.

Chan, D., and P. Huybers, 2019: Systematic differences in bucket sea surface temperature measurements among nations identified using a linear-mixed-effect method. *J. Climate*, **32**, 2569–2589, https://doi.org/10.1175/JCLI-D-18-0562.1.

——, E. C. Kent, D. I. Berry, and P. Huybers, 2019: Correcting datasets leads to more homogeneous early-twentieth-century sea surface warming. *Nature*, **571**, 393–397, https://doi.org/10.1038/s41586-019-1349-2.

EPA, 2014: Climate change indicators in the United States, 2014. Environmental Protection Agency Rep. EPA 430-R-14-004, 112 pp., https://www.epa.gov/sites/production/files/2016-07/documents/climateindicators-full-2014.pdf.

Fiedler, E. K., and Coauthors, 2019: 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.

Folland, C. K., and D. E. Parker, 1995: Correction of instrumental biases in historical sea surface temperature data. *Quart. J. Roy. Meteor. Soc.*, **121**, 319–367, https://doi.org/10.1002/qj.49712152206.

Freeman, E., and Coauthors, 2017: ICOADS release 3.0: A major update to the historical marine climate record. *Int. J. Climatol.*, **37**, 2211–2232, https://doi.org/10.1002/joc.4775.

Grumbine, R. W., 2014: Automated sea ice concentration analysis history at NCEP: 1996–2012. NOAA Modeling Branch Tech. Note 321, 39 pp., https://polar.ncep.noaa.gov/mmab/papers/tn321/MMAB_321.pdf.

Hirahara, S., M. Ishii, and Y. Fukuda, 2014: Centennial-scale sea surface temperature analysis and its uncertainty. *J. Climate*, **27**, 57–75, https://doi.org/10.1175/JCLI-D-12-00837.1.

Huang, B., M. L'Heureux, J. Lawrimore, C. Liu, V. Banzon, Z.-Z. Hu, and A. Kumar, 2013: Why did large differences arise in the sea-surface temperature datasets across the tropical Pacific during 2012? *J. Atmos. Oceanic Technol.*, **30**, 2944–2953, https://doi.org/10.1175/JTECH-D-13-00034.1.

——, and Coauthors, 2015a: Extended reconstructed sea surface temperature version 4 (ERSST.v4). Part I. Upgrades and intercomparisons. *J. Climate*, **28**, 911–930, https://doi.org/10.1175/JCLI-D-14-00006.1.

——, W. Wang, C. Liu, V. F. Banzon, J. Lawrimore, and H.-M. Zhang, 2015b: Bias adjustment of AVHRR SST and its impacts on two SST analyses. *J. Atmos. Oceanic Technol.*, **32**, 372–387, https://doi.org/10.1175/JTECH-D-14-00121.1.

——, C. Liu, V. F. Banzon, H.-M. Zhang, T. R. Karl, J. H. Lawrimore, and R. S. Vose, 2016: Assessing the impact of satellite-based observations in sea surface temperature trends. *Geophys. Res. Lett.*, **43**, 3431–3437, https://doi.org/10.1002/2016GL068757.

——, and Coauthors, 2017: Extended reconstructed sea surface temperature, version 5 (ERSSTv5): Upgrades, validations, and intercomparisons. *J. Climate*, **30**, 8179–8205, https://doi.org/10.1175/JCLI-D-16-0836.1.

——, C. Liu, G. Ren, H.-M. Zhang, and L. Zhang, 2019: The role of buoy and Argo observations in two SST analyses in the global and tropical Pacific oceans. *J. Climate*, **32**, 2517–2535, https://doi.org/10.1175/JCLI-D-18-0368.1.

——, and Coauthors, 2020: Uncertainty estimates for sea surface temperature and land surface air temperature in NOAAGlobalTemp version 5. *J. Climate*, **33**, 1351–1379, https://doi.org/10.1175/JCLI-D-19-0395.1.

IPCC, 2013: *Climate Change 2013: The Physical Science Basis*. Cambridge University Press, 1535 pp., https://doi.org/10.1017/CBO9781107415324.

——, 2018: *Special Report: Global Warming of 1.5°C*. IPCC, 630 pp., https://www.ipcc.ch/site/assets/uploads/sites/2/2019/06/SR15_Full_Report_Low_Res.pdf.

Ishii, M., A. Shouji, S. Sugimoto, and T. Matsumoto, 2005: Objective analyses of sea-surface temperature and marine meteorological variables for the 20th century using ICOADS and the Kobe Collection. *Int. J. Climatol.*, **25**, 865–879, https://doi.org/10.1002/joc.1169.

Kennedy, J. J., N. A. Rayner, R. O. Smith, D. E. Parker, and M. Saunby, 2011a: Reassessing biases and other uncertainties in sea surface temperature observations measured in situ since 1850: 1. Measurement and sampling uncertainties. *J. Geophys. Res.*, **116**, D14103, https://doi.org/10.1029/2010JD015218.

——, ——, ——, ——, and ——, 2011b: Reassessing biases and other uncertainties in sea surface temperature observations measured in situ since 1850: 2: Biases and homogenization. *J. Geophys. Res.*, **116**, D14104, https://doi.org/10.1029/2010JD015220.

——, ——, C. P. Atkinson, and R. E. Killick, 2019: An ensemble data set of sea surface temperature change from 1850: The Met Office Hadley Centre HadSST.4.0.0.0 data set. *J. Geophys. Res. Atmos.*, **124**, 7719–7763, https://doi.org/10.1029/2018JD029867.

Kent, E. C., and Coauthors, 2019: Observing requirements for long-term climate records at the ocean surface. *Front. Mar. Sci.*, **6**, 441, https://doi.org/10.3389/fmars.2019.00441.

Martin, M., 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.

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

——, 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.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. 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.

Reynolds, R. W., and T. M. Smith, 1994: Improved global sea surface temperature analyses using optimum interpolation. *J. Climate*, **7**, 929–948, https://doi.org/10.1175/1520-0442(1994)007<0929:IGSSTA>2.0.CO;2.

——, 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.

Roemmich, D., and Coauthors, 2001: The global array of profiling floats. *Observing the Ocean in the 21st Century*, C. J. Koblinsky and N. R. Smith, Eds., Australian Bureau of Meteorology, 248–258.

Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015–1057, https://doi.org/10.1175/2010BAMS3001.1.

Smith, T., and R. W. Reynolds, 2003: Extended reconstruction of global sea surface temperature based on COADS data (1854–1997). *J. Climate*, **16**, 1495–1510, https://doi.org/10.1175/1520-0442-16.10.1495.

——, and ——, 2004: Improved extended reconstruction of SST (1854–1997). *J. Climate*, **17**, 2466–2477, https://doi.org/10.1175/1520-0442(2004)017<2466:IEROS>2.0.CO;2.

——, ——, R. E. Livezey, and D. C. Stokes, 1996: Reconstruction of historical sea surface temperatures using empirical orthogonal functions. *J. Climate*, **9**, 1403–1420, https://doi.org/10.1175/1520-0442(1996)009<1403:ROHSST>2.0.CO;2.

Woodruff, S. D., and Coauthors, 2011: ICOADS Release 2.5: Extensions and enhancements to the surface marine meteorological archive. *Int. J. Climatol.*, **31**, 951–967, https://doi.org/10.1002/joc.2103.

Zhang, H.-M., R. W. Reynolds, and T. M. Smith, 2004: Bias characteristics in the AVHRR sea surface temperature. *Geophys. Res. Lett.*, **31**, L01307, https://doi.org/10.1029/2003GL018804.