

An Improved In Situ and Satellite SST Analysis for Climate

RICHARD W. REYNOLDS

NOAA/NESDIS/National Climatic Data Center, Asheville, North Carolina

NICK A. RAYNER

Hadley Centre for Climate Prediction and Research, Met Office, Bracknell, Berkshire, United Kingdom

THOMAS M. SMITH

NOAA/NESDIS/National Climatic Data Center, Asheville, North Carolina

DIANE C. STOKES

NOAA/NWS/National Centers for Environmental Prediction, Camp Springs, Maryland

WANQIU WANG

Science Applications International Corporation, Camp Springs, Maryland

(Manuscript received 27 July 2001, in final form 21 November 2001)

ABSTRACT

A weekly 1° spatial resolution optimum interpolation (OI) sea surface temperature (SST) analysis has been produced at the National Oceanic and Atmospheric Administration (NOAA) using both in situ and satellite data from November 1981 to the present. The weekly product has been available since 1993 and is widely used for weather and climate monitoring and forecasting. Errors in the satellite bias correction and the sea ice to SST conversion algorithm are discussed, and then an improved version of the OI analysis is developed. The changes result in a modest reduction in the satellite bias that leaves small global residual biases of roughly -0.03°C . The major improvement in the analysis occurs at high latitudes due to the new sea ice algorithm where local differences between the old and new analysis can exceed 1°C . Comparisons with other SST products are needed to determine the consistency of the OI. These comparisons show that the differences among products occur on large time- and space scales with monthly rms differences exceeding 0.5°C in some regions. These regions are primarily the mid- and high-latitude Southern Oceans and the Arctic where data are sparse, as well as high-gradient areas such as the Gulf Stream and Kuroshio where the gradients cannot be properly resolved on a 1° grid. In addition, globally averaged differences of roughly 0.05°C occur among the products on decadal scales. These differences primarily arise from the same regions where the rms differences are large. However, smaller unexplained differences also occur in other regions of the midlatitude Northern Hemisphere where in situ data should be adequate.

1. Introduction

Sea surface temperature (SST) is an important indicator of the state of the earth's climate system. Thus, accurate knowledge of SST is essential for climate monitoring, research, and prediction. SSTs are also used as a surface boundary condition for numerical weather prediction and for other atmospheric model simulations. There are many different SST analyses produced

throughout the world that resolve different temporal and spatial scales and use different sets of satellite and/or in situ data. In this paper we focus on climate-scale analyses, which we define as spatial scales of 1° and larger and temporal scales of 1 week and longer. Our period of interest is 1982 to the present when both in situ and satellite data are available.

We will cover two interrelated topics here. The first is a discussion of the weekly optimum interpolation (OI) climate-scale in situ and satellite SST analyses produced at the National Oceanic and Atmospheric Administration (NOAA; Reynolds and Smith 1994, hereafter RS). We show that there are errors in this OI analysis (hereafter OI.v1, for OI version 1) due to an undercorrection

Corresponding author address: Dr. Richard W. Reynolds, NOAA/NESDIS/National Climatic Data Center, 151 Patton Ave., Asheville, NC 28801.
E-mail: richard.w.reynolds@noaa.gov

of satellite bias and the choice of the sea ice to SST conversion algorithm. We then present a new version of the OI analysis (hereafter OI.v2) that reduces the errors.

The second topic is an assessment of SST analysis errors for our period of interest. This is a necessary step in discussing errors in the OI.v1 and OI.v2. It is also important so that users have a better idea of the size of SST errors and the locations where they differ from the average. Here, we focus on global differences in time and mean differences in space.

In the sections that follow we first briefly discuss in situ and satellite data. In section 3, we summarize the OI.v1 procedure and introduce the other analyses and data summaries used for the intercomparisons. In the next two sections the problems in the OI.v1 are discussed followed by a description of the improvements found in OI.v2. In section 6, we present intercomparisons of the two OI versions and the other products. The results of the intercomparisons suggest the size of the error by the differences among analyses. However, the results are not completely satisfactory because all analyses and data summaries share some of the data in common. Concluding comments are presented in the last section.

2. SST data

The in situ SST data are determined from observations from ships and buoys (both moored and drifting). Most ship observations in our period of interest were made from insulated buckets, hull contact sensors, and engine intakes at depths of one to several meters. (A relatively small fraction of SST observations, roughly 5%, is obtained from profiles of temperature variations with depth. These observations must be used with caution because unrepresentative surface temperatures may occur.) Although selected SST observations can be very accurate (see Kent et al. 1993, 1999), typical rms errors of individual observations from ships are larger than 1°C and may have daytime biases of a few tenths of a degree Celsius.

SST observations from drifting and moored buoys were first used in the late 1970s. Buoy observations became more plentiful following the start of the Tropical Ocean Global Atmosphere (TOGA) program (McPhaden et al. 1998) in 1985. These observations are typically made by thermistor or hull contact sensor and usually relayed in real time by satellites. Although the accuracy of the buoy SST observations varies, the random error is usually smaller than 0.5°C and, thus, is better than ship error. In addition, typical depths of the measurements are roughly 0.5 m rather than the 1 m and deeper measurements from ships. A typical recent distribution of ship and buoy SST observations (see Fig. 1) shows that the deployment of the buoys has been designed to fill in some regions with few ship obser-

vations. This process had the most impact in the tropical Pacific Ocean and the Southern Hemisphere.

In late 1981, Advanced Very High Resolution Radiometer (AVHRR) satellite retrievals improved the data coverage over that of in situ observations alone. The satellite retrievals allowed better resolution of small-scale features such as Gulf Stream eddies. Because the AVHRR cannot see the surface in cloud-covered regions, the biggest challenge in retrieving SST is to eliminate cloud contamination. The cloud-clearing algorithms are different during the day and the night because the AVHRR visible channels are useful in detecting clouds but can only be used during the day. Once clouds have been eliminated, the SST retrieval algorithm is designed to minimize the effects of atmospheric water vapor. The algorithms are “tuned” by regression against quality-controlled buoy data using the multichannel SST technique of McClain et al. (1985). This procedure converts the retrieval of the temperature of the “skin” (roughly a micron in depth) to a “bulk” (roughly 0.5 m in depth) SST. The tuning is redone when a new satellite becomes operational or when comparison with the buoy data shows increasing errors. The AVHRR instrument has three infrared (IR) channels. Due to noise from reflected sunlight (sun glint), only two channels can be used during the day. However, at night the three IR channels are used because the residual noise is lower. Thus, the algorithm is usually tuned and applied separately for daytime and nighttime using two channels during the day and three at night (Walton et al. 1998). The algorithms are computed globally and are not a function of position or time.

If the satellite SST retrievals are partially contaminated by clouds, they have a negative bias because cloud temperatures are colder than the SSTs. Negative biases can also be caused by atmospheric aerosols, especially stratospheric aerosols from large volcanic eruptions (see Reynolds et al. 1989; Reynolds 1993). Although these negative biases are the most frequent, biases of either sign can also occur due to instrument problems (e.g., due to errors in the onboard blackbody calibration). Because some biases remain in satellite SST data, in situ data are critical not only for satellite calibration and validation but also for any final bias corrections needed at the time of the analysis.

To illustrate the changes in SST data distributions over time, we show time series in Fig. 2 of the number of 1° grid squares from 60°S to 60°N, which have data for each week from the beginning of 1982 to the end of 2000. As expected from the discussion on buoy SSTs, the buoy coverage is low at the beginning of the time series but gradually increases to roughly 1000 squares per week. The ship coverage is initially much larger (roughly 10 000 per week) and gradually decreases to between 7000 and 8000 by the end of 1997. For the 1998–2000 period, the ship coverage is between 5000 and 6000 squares per week. The dropoff at the end of 1997 is due to a change in the source of the in situ data.

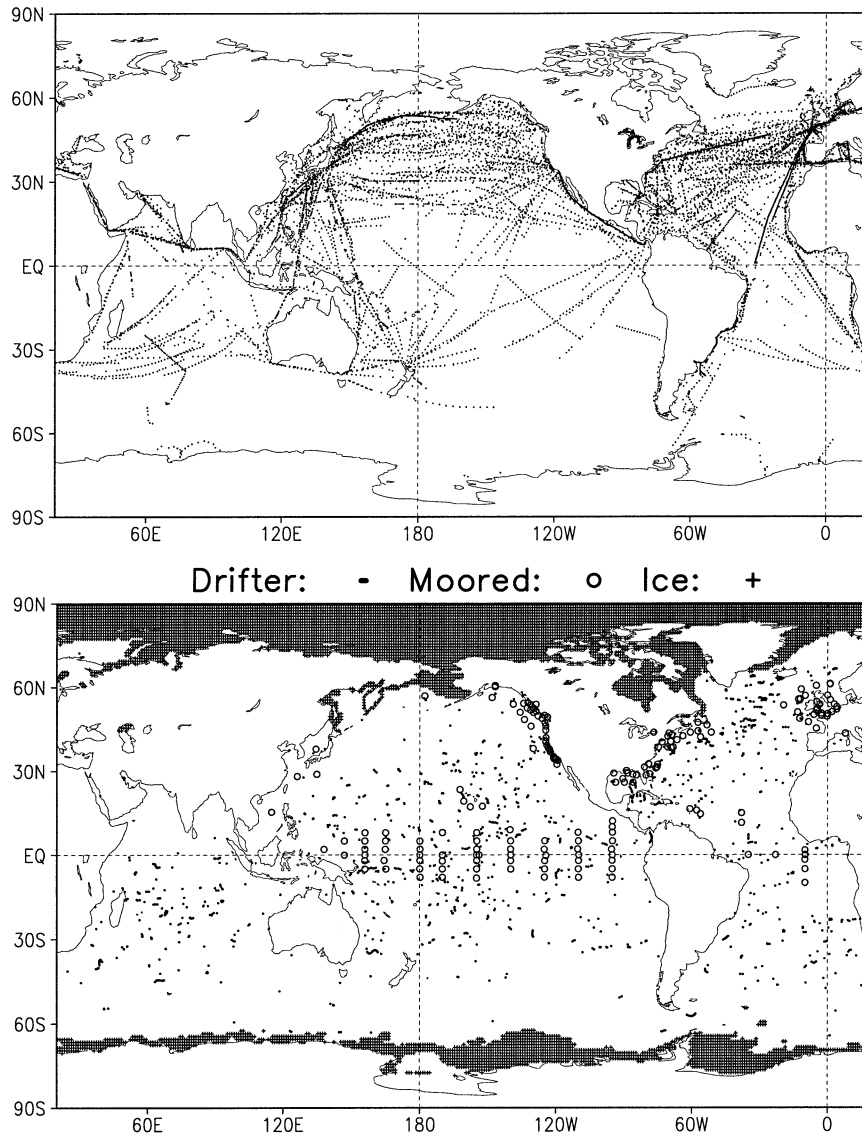


FIG. 1. (top) Distribution of SST in situ observations from ships and (bottom) buoys for the week of 2–8 Jan 2000. Sea ice data (at concentrations ≥ 0.5 , 50%) are also shown (by hatching) in the bottom panel.

The in situ data prior to 1998 are obtained from the Comprehensive Ocean–Atmosphere Data Set (COADS; see Slutz et al. 1985; Woodruff et al. 1998). After this period the data are obtained in real time from the Global Telecommunication System (GTS). Because almost all of the buoy data are transmitted in real time, there is little difference in the number of buoy observations between COADS and GTS data. However, some ship reports are only available via hard copy. Thus, COADS contains significantly more ship data than does the GTS. At the time of the development of the OI.v1 (the version of the OI described by RS) COADS data were being digitized and processed for the early 1990s. Thus, a preliminary version of the COADS data was used prior to 1990 in OI.v1, and GTS data were used from January

1990 onward. (The new version of the analysis will use COADS through 1997.)

Figure 2 also shows the daytime and nighttime satellite data coverage for the period. (For almost all of this period, two satellites were operational, one with a morning and one with an afternoon equatorial crossing time. However, operational SST retrieval algorithms were usually only available for the afternoon satellite. Thus, only afternoon satellite data were used in the OI and shown in the figure.) The results show that satellite coverage was generally much greater than the ship coverage, as expected. The satellite coverage approached the maximum possible value of 32 347 squares for our land–sea mask. However, there were some problems in the beginning of the record when satellite data were not

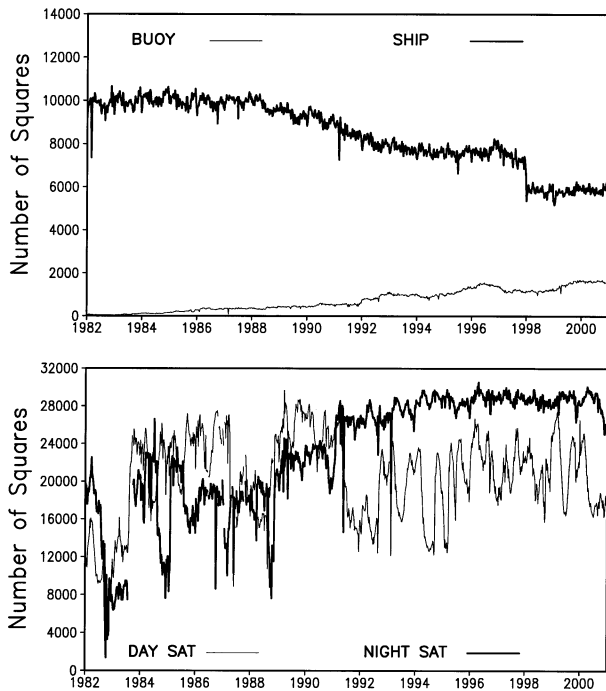


FIG. 2. Weekly number of 1° data squares (60°S – 60°N) with SST data from: (top) ships and buoys, and (bottom), daytime and nighttime satellites. The ship and buoy data were taken from COADS through 1997 and from GTS after 1997. This change results in a drop in the ship coverage beginning in Jan 1998 (see text). The satellite data were obtained using similar algorithms but the sources (see text) changed in 1990 and 1996. (In the 1980s there were 10 and 19 weeks of missing day and night satellite retrievals, respectively.) There are 32 347 ocean 1° squares for this region using our land–sea mask.

available. The figure shows that the night coverage increased over the period and became smoother relative to the day coverage. This difference in behavior is primarily due to the different cloud clearing algorithms developed over the period. The changes in the daytime retrievals were also influenced by the tendency of the equator crossing to drift from early afternoon toward evening as the satellite aged. This change strongly reduced the coverage of the daytime retrievals in the winter hemisphere where daylight was more limited.

All the satellite SST retrievals were processed using the algorithms mentioned above. However, the sources differed. For the period prior to 1990, the satellite SST retrieval data shown in the figure and used in the OI.v1 were obtained from weekly data summaries of National Environmental Satellite, Data, and Information Service (NESDIS) data produced at the University of Miami's Rosenstiel School of Marine and Atmospheric Sciences (RSMAS; R. Evans 1993, personal communication). Data from 1990 to present were processed from the OI data archives. For 1990–95 we used satellite SST retrieval data obtained directly from NOAA/NESDIS. After 1995 we used data processed by the U.S. Navy (May et al. 1998). The change after 1995 was due to an agreement between NOAA and the navy to share resources.

Although, we did not find any important differences between the two datasets, we found that direct comparison of the NESDIS and navy retrievals occasionally showed problems with one or both algorithms. This intercomparison was not only useful for identifying problems but also for correcting them.

3. SST analyses

We briefly discuss the current OI in situ and satellite SST analysis and then introduce other SST products that we will use for intercomparisons.

a. OI analysis (OI.v1)

Optimum interpolation or OI was developed by Gandin (1963) as an objective analysis method for irregularly spaced data. The analysis begins with a background or first-guess field, which in our case is the previous week's analysis. Differences between data and the first-guess field, data increments, are computed. At each analysis grid point, the analysis method objectively determines a series of weights for each of the data increments. The weights are based on distances between the grid point and the data as well as the variance and covariance errors of the first-guess field and the data increments. Once the weights are defined, they are multiplied by the data increments and summed to produce an analysis increment at each grid point. The completed analysis is determined by adding the analysis increment to the first guess. The analysis is objective if all these errors are known. Of course, they are not perfectly known and are only estimated. Thus, the analysis is never as perfect as the name, optimum interpolation, suggests.

The OI.v1 analysis is computed weekly on a 1° latitude by 1° longitude grid using satellite and in situ data. In RS the weeks were defined to be centered on Wednesday. Unfortunately, our satellite archive did not cover all of the 1980s. The easiest access to the satellite data for this period was from the RSMAS fields, which were centered on Sunday. Thus, we defined the center of the week as Sunday in the 1980s and Wednesday from January 1990 to the present. (There is one week of partial overlap at the end of 1989 and no SST discontinuity between the end of 1989 and the beginning of 1990.) Data are quality controlled as reported in appendixes A and C in Reynolds (1988) with improved ship and buoy track checks to eliminate observations with bad locations.

The OI method assumes that the data do not contain long-term biases (e.g., see Lorenc 1981). Because satellite biases occur in our period of interest, as discussed in more detail in RS, a preliminary step using Poisson's equation is carried out to remove satellite biases relative to in situ data before the OI analysis is begun. This method adjusts any large-scale satellite biases and gradients relative to the boundary conditions defined by the in situ field. The correction is derived separately for

the daytime and nighttime satellite data because, as mentioned above, different satellite algorithms (both for SST retrieval and cloud removal) are used for day and night measurements. The satellite data are adjusted by these smoothed bias corrections before they are used in the OI step.

In the OI procedure, various random error statistics are assigned that are a function of latitude and longitude. These errors are discussed in RS and only the global averages are mentioned here. The globally averaged guess error was 0.3°C ; the globally averaged data error was 1.3°C for ship data, 0.5°C for buoy and daytime satellite data, and 0.3°C for nighttime satellite data and SST data generated from sea ice concentrations. Clearly the ship data have the largest error and hence the lowest relative weight. It was assumed by RS that the error correlations in space were Gaussian in shape. The globally averaged zonal and meridional e -folding values for the correlations were 850 and 615 km, respectively. Because the satellite distribution is so much better than the in situ distribution, the satellite data usually overwhelm the in situ data in the OI. However, as discussed above, the in situ data are critically important in correcting any satellite biases before these data are used in the OI.

The daytime satellite data have been assigned higher error values than the nighttime data. The diurnal cycle is not resolved in the OI analysis, and any diurnal signal will appear as noise to the analysis. Additional errors in the satellite data can occur when the assumed correlation between skin and bulk temperatures begins to break down. In this case, the satellite retrieval algorithm also breaks down. This occurs most frequently during the daytime in regions with light winds and high insolation when the surface diurnal warming impacts the skin temperature but does not fully impact the bulk temperature 0.5 m or more beneath the skin (Weller and Anderson 1996). This decoupling can also occur at night although it is much less likely.

For intercomparison studies, there is also a special analysis of the OI.v1, which is computed without the satellite bias correction step, hereafter labeled OLNO.v1 for "no" bias correction. Monthly fields of the OI.v1 and OLNO.v1 are computed by linearly interpolating the weekly fields to produce daily fields and then averaging the appropriate days within a month to produce monthly averages.

b. Additional SST products

To better understand the problems in the OI.v1, we will use five additional SST data summaries and two additional SST analyses. All data summaries are monthly averages of the observations within latitude by longitude boxes. These summaries are not referred to as analyses here because there are no assigned SST values in ocean boxes without data.

The first data summary is the Met Office Historical SST dataset (MOHSST), version 6, of Parker et al.

(1994), which is computed on a 5° grid using both ship and buoy data. The quality control (QC) for MOHSST is complex, as described in appendix 1 of Parker et al. (1995b).

The second and third summaries are derived from the OI.v2 in situ input sources. Thus, COADS data are used through 1997, and GTS data are used after 1997. The second product, hereafter labeled OI-SHIP, is a 2° averaged summary of the ship data that passed the OI QC; the third product, hereafter labeled OI-SITU, is similar to OI-SHIP but merges both ships and buoys. As mentioned above, the QC is described in Reynolds (1988) with an additional ship and buoy tracking procedure to eliminate observations with bad locations.

The last two data summaries are COADS 2° gridded SST summaries. To QC COADS the actual distribution of the data is used instead of assuming a Gaussian distribution. As discussed in Slutz et al. (1985), this method develops robust estimates of the mean and standard deviation that are statistically more stable when outliers are present. The main difference between the two summaries is that one product uses only surface marine observations from ships, while the other product adds data from buoys and other in situ sources to the ship data. Following the COADS definitions, we will refer to the two COADS summaries as standard (hereafter COADS-S) for the ship-only product and enhanced (hereafter COADS-E) for the product using all in situ data. There are additional QC steps in the procedures that cross link SST and other marine variables. For complete details see Slutz et al. (1985). The COADS summaries end in 1997. Input data for both OI and COADS data summaries are identical through 1997. However, the summary products differ because the QC procedures differ.

The two analyses are produced at the Met Office Hadley Centre for Climate Prediction and Research using variations on MOHSST as input data. These analyses also use AVHRR satellite data beginning in 1982. The first is the Global Sea Ice and SST dataset (GISST), version 2.3b, of Rayner et al. (1996). The second is the more recent Hadley Centre Sea Ice and SST dataset (HadISST), version 1, which is described in Parker et al. (1999). Both GISST and HadISST are computed monthly on a 1° grid. An overview of these SST products is given in Table 1.

4. OI analysis problems (OI.v1)

Intercomparisons of different SST products have shown important differences among analyses (e.g., Hurrell and Trenberth 1999). These differences give an estimate of the overall analysis error. However, the most important consequence of the intercomparisons is to indicate where analyses need to be improved.

TABLE 1. SST summary and analysis information. Summaries using only ship data are indicated by “ship only.” Analyses using a more limited in situ dataset (see text) are indicated by “reduced.” Analyses using satellite data are indicated by “yes” if used or “corrected” if used with additional bias corrections. Analyses using sea ice data converted to SSTs are indicated by yes.

Name	Type	Resolution	In situ data	Satellite data	Ice data
COADS-E	Summary	2°	All	No	No
COADS-S	Summary	2°	Ship only	No	No
OI-SITU	Summary	1°	All	No	No
OI-SHIP	Summary	1°	Ship only	No	No
MOHSST	Summary	5°	All	No	No
OI.v1	Analysis	1°	Reduced	Corrected	Yes
OLNO.v1	Analysis	1°	Reduced	Yes	Yes
OI.v2	Analysis	1°	All	Corrected	Yes
GISST	Analysis	1°	All	Corrected	Yes
HadISST	Analysis	1°	All	Corrected	Yes

a. SST biases

To examine overall SST biases, Fig. 3 shows time series for the monthly OI.v1, OLNO.v1, and MOHSST differences with respect to COADS-E for 1982–97. We have selected COADS-E as our standard of comparison for this paper. We wanted to select one of the COADS summaries because they are not produced at a center doing SST analyses and would be more independent as a standard. COADS-E was selected over COADS-S because COADS-E includes the growing contribution of buoy data (see Fig. 2). The averaging region from 60°S to 60°N was selected to minimize the impact of sea ice. Because in situ data are not available for all parts of the globe, the differences were computed only over regions where MOHSST and COADS-E both had values. The figure shows that the MOHSST tends to be slightly more positive, roughly 0.05°C, than COADS-E from 1990 onward. In contrast, the OI.v1 tends to be slightly more negative, roughly -0.05°C, than COADS-E with larger values in the 1990s. The differences between the OLNO.v1 and COADS-E are several times larger than the differences with the other products. This shows the importance of the real-time bias correction of the satellite data. In particular, impacts of the large negative satellite biases resulting from the volcanic aerosols from El Chichón (1982–83) and Mount Pinatubo (1991–92) are clearly evident (for more details see Reynolds 1993). These results suggest that the bias correction step in the OI.v1 has left residual satellite biases, especially in the 1990s.

To focus on the uncertainty of the in situ observations considered as ground truth for satellite bias correction, we show in Fig. 4 average differences (60°S–60°N) of 4 of the SST summaries relative to COADS-E. In general, the differences in Fig. 4 are within $\pm 0.02^\circ\text{C}$ from 1982 to about the beginning of 1993. During this period, MOHSST is usually warmer than COADS-E by roughly 0.01°C, while the OI-SHIP and OI-SITU summaries are roughly -0.02°C too cold. The COADS-S is overall the

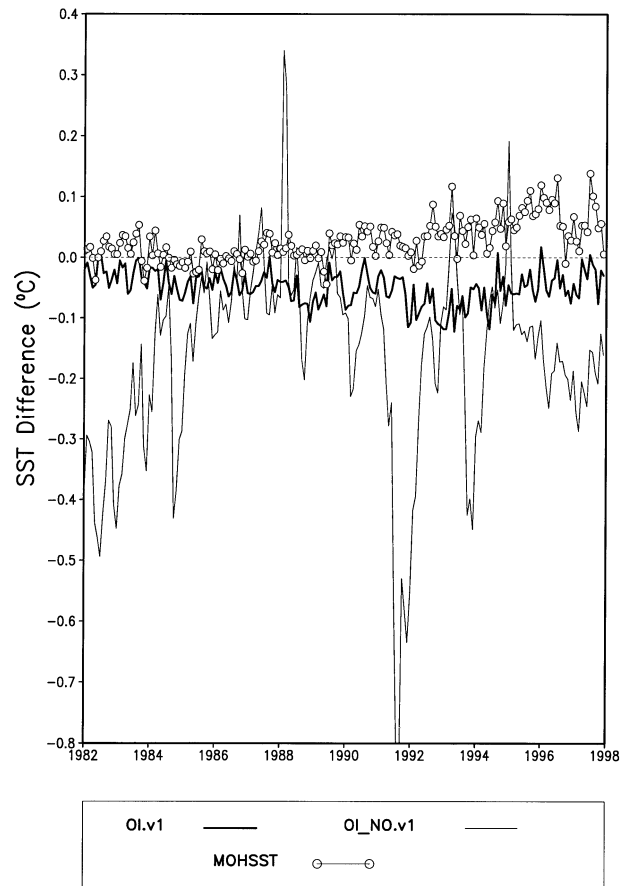


FIG. 3. Monthly averaged (60°S–60°N) SST differences for the OI.v1, OLNO.v1, and MOHSST relative to COADS-E. The averages are computed over common areas where MOHSST and COADS-E are defined.

closest to COADS-E for the 1982–92 period. After 1992 OI-SHIP and COADS-S are warm relative to COADS-E. This is most likely due to the growing influence of buoy SSTs in COADS-E. In the 1990s MOHSST shows an overall warming trend with the biggest differences in 1996. Please note that COADS-E and the OI-SITU as well as COADS-S and the OI-SHIP use exactly the same input data. Thus, differences between these two summary pairs must be due to the in situ QC procedures used to eliminate bad data.

Folland et al. (1993) sorted ship observations into bucket and nonbucket categories. They reported that ship nonbucket temperatures were 0.11°C warmer than bucket temperatures. This is most likely due to inclusion of ship engine room intake temperatures with a warm bias of roughly 0.3°C (Saur 1963) that is included in the nonbucket category. For 1997, COADS reports that there were 1.2 million ship SST observations of which 20% were classified as bucket, 35% intake, and 4% hull contact. The remaining 41% were unknown. Unfortunately, the percentage of unknown observations is much larger in years prior to 1993. However, a division

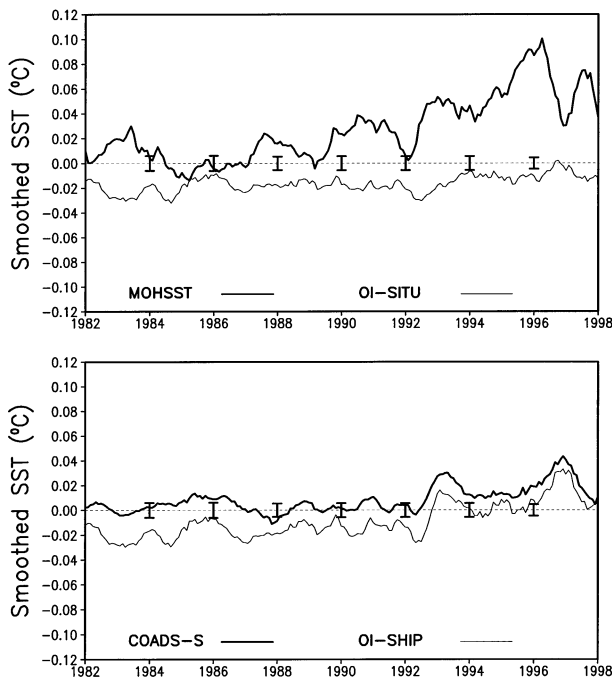


FIG. 4. Monthly averaged (60°S – 60°N) in situ SST differences for the (top) MOHSST and OI-SITU, and (bottom) COADS-S and the OI-SHIP and OI-SITU data summaries relative to COADS-E (see text). A 7-point running mean filter was used to smooth the time series. Estimated 95% confidence limits for in situ summaries are shown for selected years for Jan. The confidence limits for COADS-E show 2 times the std dev assuming uncorrelated errors with values of 0.5°C for buoy SSTs and 1.3°C for ship SSTs. For each panel, differences are computed over common areas where all products are defined.

into ship and buoy categories is possible over our period of interest. We used the weekly averaged 1° summaries prepared for the OI and computed a difference value if there were collocated 1° buoy and ship summaries for each week. The average of the differences for all weeks from 1982 to 2000 is summarized in Table 2. There is an overall ship-minus-buoy difference of 0.13°C with an indication of a latitudinal dependence. We also show in the table that there is a seasonal variation of the Northern Hemisphere midlatitude difference.

Table 2 indicates that different combinations of buoy and ship data can clearly lead to biases. There are undoubtedly other data differences. A careful intercomparison of the data procedures using a common input

dataset is needed to help determine the causes of the differences. Figure 4 suggests that present in situ bias error limits can exceed 0.05°C for globally averaged monthly SSTs. We have not corrected the OI.v2 in situ data by the factors in Table 2 because of the uncertainties of the biases in the table. However, any correction of satellite data is further complicated by in situ biases and their uncertainties.

b. Sea ice biases and sea ice to SST conversion algorithms

A large potential error occurs near the sea ice edge where in situ observations tend to be sparse because of navigation hazards and satellite observations tend to be sparse due to cloud cover. Thus, using sea ice data to generate simulated SSTs in the marginal ice zone (MIZ) helps fill in a region with limited data. We define the fraction of sea ice contribution as I with a range of 0–1 (0%–100%). In the OI.v1, an SST value of -1.8°C was generated for $I \geq 0.5$. (The value of -1.8°C is the freezing point of seawater with a salinity of 34 psu.) For GISST and HadISST a climatological relation between SST and I is defined by a quadratic equation

$$\text{SST} = aI^2 + bI + c, \quad (1)$$

where a , b , and c are constant coefficients. The coefficients are determined by collocated matchups between SST and sea ice concentration with the constraint that $\text{SST} = -1.8^{\circ}$ or 0°C when $I \geq 0.9$ over the ocean or freshwater lakes, respectively. The choice of the quadratic relationship in (1) was determined (see Rayner et al. 1996) by inspection of seasonal scatterplots of collocated SST and sea ice concentrations. These results showed that sea ice concentrations within the 0.5–0.8 range were usually considerably warmer than -1.8°C . Thus, (1) agreed with the data better than the binary assumptions made for the OI.v1.

In addition to uncertainties in these methods, the analyzed value of the ice concentration as defined in different sea ice analyses differs especially in the Arctic in Northern Hemisphere summer. Climatological sea ice concentrations for July are shown in Fig. 5 for two analyses. The first, combined from Nomura (1995) and Grumbine (1996), the Nomura–Grumbine analysis, is an objective analysis of retrievals of sea ice concentration from observations made by passive microwave in-

TABLE 2. Ship minus buoy biases for 1° gridded weekly collocated averages computed by region for all seasons from Jan 1982 to Dec 2000 and by season for 20° – 60°N .

Region (all seasons)	60° – 20°S	20°S – 20°N	20° – 60°N	60°S – 60°N
No. of pairs	17 753	45 605	160 180	223 538
Bias	0.08°C	0.04°C	0.17°C	0.14°C
Season (20° – 60°N)	Winter	Spring	Summer	Fall
No. of pairs	37 967	38 446	40 877	42 890
Bias	0.20°C	0.27°C	0.11°C	0.11°C

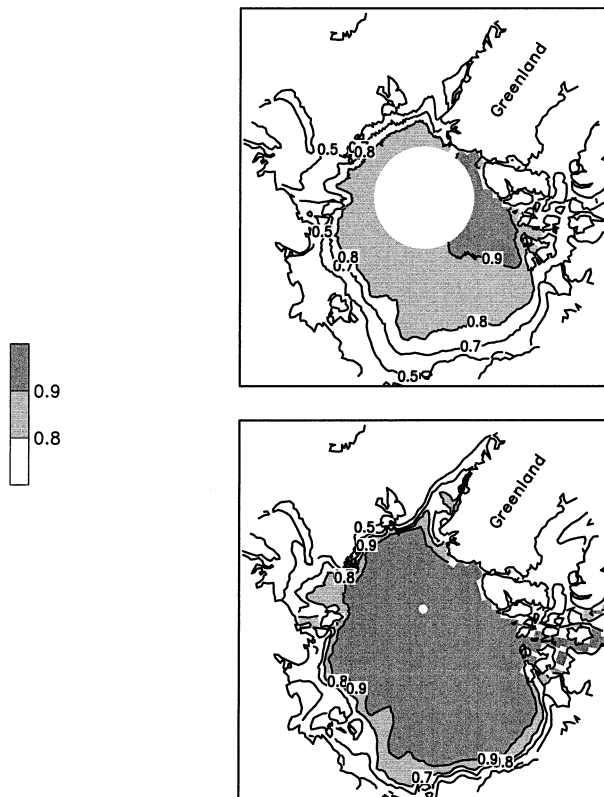


FIG. 5. Climatological sea ice concentrations for the Arctic for Jul for the period 1979–92. (top) The analysis from Nomura and Grumbine. The missing data near the Pole occurs because of lack of satellite observations. (bottom) The analysis from the NIC (see text). The range of ice concentration is 0 (0%)–1 (100%). Contours are 0.9, 0.8, 0.7, and 0.5.

struments [the Scanning Multichannel Microwave Radiometer (SMMR) and the Special Sensor Microwave Imager (SSM/I)]. The second, the National Ice Center analysis (NIC; Knight 1984), is a subjective analysis of in situ, aircraft, and satellite-based passive microwave and infrared observations. The concentrations of the purely passive microwave derived analysis are much lower in summer primarily because the instruments interpret meltwater on top of the sea ice as open water. In addition, they are also lower because the NIC analysts are less interested in the inner ice pack and often assume the concentrations to be 1 (100%).

Because we set SST to -1.8°C in grid boxes where the sea ice concentration is at least 0.9 (90%), a difference of a few percent in ice concentrations in the interior of the pack has no effect on the SST field. However, these concentration differences are very important when the fields are used as a boundary condition for atmospheric models. In these models, the air–sea fluxes vary significantly between ice-covered and partially ice-free grid boxes.

In an attempt to objectively determine which sea ice to SST conversion algorithm is more accurate, we computed the average monthly difference of the OI.v1 and

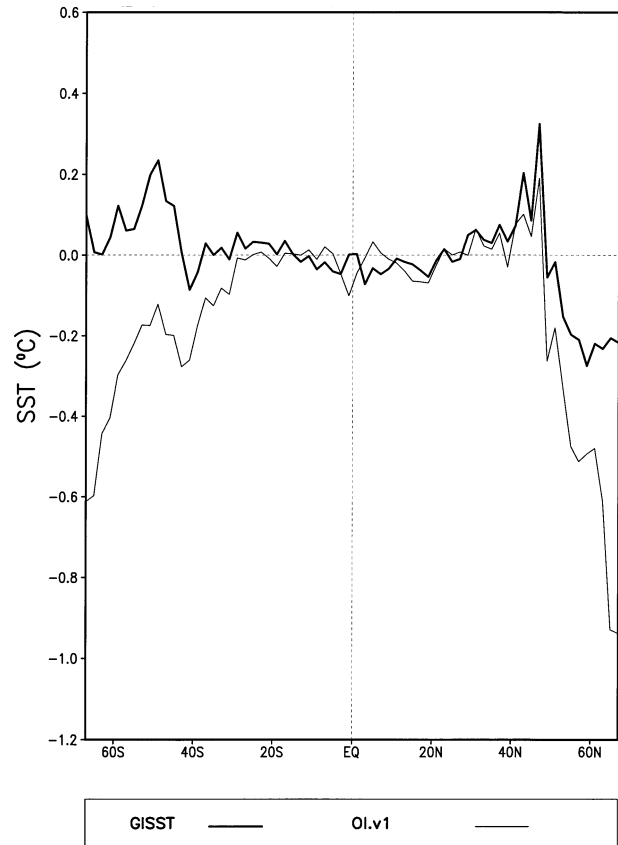


FIG. 6. Zonally averaged monthly SST differences for the OI and GISST relative to COADS-E for 1982–97. The differences at high latitudes are due to differences in the sea ice-to-SST conversion algorithms. Differences are computed over common areas where COADS-E is defined.

GISST relative to COADS-E for 1982–97. The differences were computed only where COADS-E observations were available. Thus, the differences can be considered to be averages that include only the ice-free portions of the grid boxes. In Fig. 6 we show the average zonal differences for January 1982–December 1997. The results show that GISST is generally closer to COADS-E at high latitudes than the OI.v1. In addition, the OI.v1 is much too cold relative to COADS-E as expected. Thus, the U.K. method is superior although some biases remain.

5. OI analysis improvements (OI.v2)

In the sections that follow, we will discuss a possible change in the satellite SST bias correction algorithm and the algorithm to convert sea ice concentration to SST. The sea ice procedure was also used in the HadISST analysis.

a. Satellite bias correction

As mentioned earlier, there is a preliminary step in the OI.v1 procedure (described in more detail in RS)

where in situ and satellite analyses are produced to generate satellite bias correction fields. The corrections are obtained formally by solving Poisson's equation so that any large-scale satellite biases and gradients can be adjusted relative to the in situ data. The Poisson method has been successfully used for SST satellite bias correction since it was first introduced by Reynolds (1988). However, the major problem with this method is that the correction is performed independently for each week's analysis. Thus, there is no continuity of the correction between weeks. In most cases, the cause of the bias, for example, the presence of volcanic aerosols, does persist in time. This is clearly evident in the monthly OLNO.v1 difference with respect to COADS-E shown in Fig. 3. Thus, knowledge of the previous bias corrections can be useful. Of course if the in situ data were of high quality with adequate coverage, this limitation would not be serious. However, in many regions in the Southern Hemisphere in situ data are extremely sparse. This may lead to corrections that are locally noisy due to changes in the sparse in situ data network with time.

In our SST analysis procedure, the previous analysis is the first guess. The analysis is performed using data increments, which are defined as the difference between the data and the first guess. Thus, each analysis is built upon previous analyses. The specification of the data and the guess error determines the relative weight of the guess and the data. The ability to use the information from the previous analysis suggests using an OI type of analysis for the bias correction as an alternative to the Poisson method. However, as will be discussed in the appendix both methods gave very similar results. Switching from the preliminary version of COADS to the current COADS release for 1981–89 and from GTS to COADS for 1990–97 increased the ship SST coverage. This change in coverage caused a stronger impact on the bias correction than the change in procedure.

b. Sea ice bias correction and sea ice to SST conversion algorithm

A method has been developed (see also Parker et al. 1999) to first correct the passive microwave derived sea ice concentration fields in the Arctic and then convert the corrected concentration to SSTs. This method is used here and in the HadISST analysis. GISST used a similar sea ice to SST conversion algorithm without the sea ice correction (Rayner et al. 1996). Thus, we can expect the OI.v2, HadISST, and (to a lesser degree) GISST to be in good agreement in the MIZ.

Sea ice concentration data from different sources are particularly heterogeneous. Thus, the available hemispheric-scale sea ice analyses were cross calibrated to provide a self-consistent history of observed sea ice concentration. Our method is to correct the passive microwave derived fields, which would be used operationally, so that they are consistent with the historical

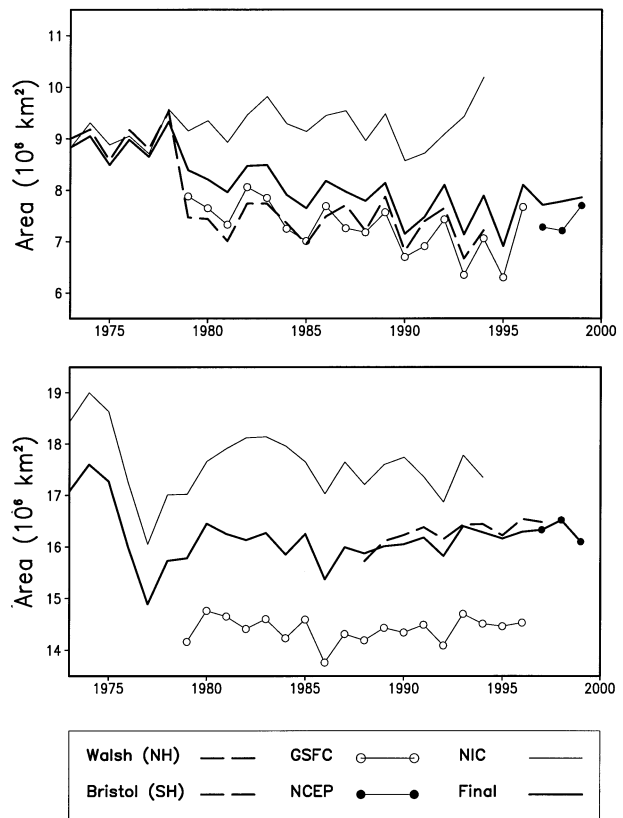


FIG. 7. Monthly averaged sea ice coverage (10^6 km^2) from different sea ice products (see text) including our adjusted final product. (top) The Northern Hemisphere (NH) coverage for Jul; (bottom) the Southern Hemisphere (SH) coverage for Sep. Only ice concentrations greater than 0.15 (15%) are included in the computation.

fields. An example is shown in Fig. 7 for July in the Northern Hemisphere and September in the Southern Hemisphere. These months were selected because of large analysis differences. We have shown analyses before 1982 to illustrate the difficulty of eliminating bias errors caused by combining different analyses while retaining climate signals. (The HadISST and GISST analyses extend backward in time to 1870.) The figure shows sea ice areas for six analyses. Here, the sea ice area is defined as the sum of the products of the concentration and the area of the box for sea ice concentrations of at least 0.15 (15%). These analyses are the Walsh (Walsh and Chapman 2001), NASA Goddard Space Flight Center (GSFC; Cavalieri et al. 1999), NIC, Bristol algorithm (Hanna and Bamber 2001), and the National Centers for Environmental Prediction (NCEP; Grumbine 1996) analyses and our final analysis. GSFC, NCEP, and the Bristol fields were derived from passive microwave retrievals, which give sea ice concentrations in each grid box. Prior to late 1978, the Walsh analysis was derived (see Walsh and Chapman 2001) from charts of in situ measurements of sea ice extent that extend back to 1901.

In the Northern Hemisphere, we assumed that the winter concentrations in the passive microwave datasets

were correct, but that the summer concentrations were biased low because of surface melt ponds (Comiso and Kwok 1996). We used comparisons with the NIC analysis (while allowing for their biases) to adjust the summer concentrations in the passive microwave analysis to our final product. The extent of the GSFC data was not changed, just the concentrations away from the MIZ. This explains why the resultant area is still much more like that of the unadjusted GSFC data than the NIC data, which have greater extents (not shown). Sea ice coverage for the Caspian Sea was obtained from a climatology (used originally in the GISST1 dataset; Parker et al. 1995a) because none of our datasets contained information there. [GISST and HadISST, see Parker et al. (1999), used the Walsh analysis prior to the beginning of the GSFC analysis in late 1978.]

In the Southern Hemisphere, the NCEP ice had higher concentrations and greater areal extents than the GSFC even though both used the National Aeronautics and Space Administration (NASA) Team algorithm (Cavaliere et al. 1999). The Bristol analysis was used as an independent check and indicated that GSFC concentrations in the Antarctic are biased low. We recalibrated our final product by adjusting the GSFC mean bias to agree with the Bristol analysis, which is more consistent with in situ observations (E. Hanna 2000, personal communication). [GISST and HadISST, see Parker et al. (1999), used monthly sea ice extent climatologies, e.g., the Russian atlas of Tolstikov (1966) for 1947–62, to define the sea ice prior to 1973.]

In the MIZ, there are few in situ observations of SST. Thus, we utilized our homogenized sea ice concentration fields to specify SST in grid boxes with both sea ice and open water. First, we used recent in situ and bias-adjusted AVHRR SST observations to statistically determine by regression the unknown coefficients (a , b , and c) in (1), which specify the functional relationship between sea ice concentration and SST. We then used (1), along with the sea ice concentration fields, to generate simulated SSTs in the MIZ, wherever sea ice concentration was less than 90% and at least 15%. Where concentrations were at least 90%, simulated SSTs were set to -1.8°C (0°C in the freshwater of the Great Lakes). Using (1), the coefficients were calculated separately for each hemisphere and calendar month in each of 360 overlapping 31° longitude sectors. For each calendar month, data from that month and the two neighboring months were used to determine the coefficients and ensure a smooth transition from month to month. Where there were too few data pairs to form stable coefficients, we interpolated regression coefficients from neighboring regions. In the Northern Hemisphere, independent coefficients were determined for the Great Lakes, the Baltic Sea, the Seas of Okhotsk and Japan, and the Gulf of Alaska. (Because of data sparsity, average coefficients for the Northern Hemisphere were used for the Caspian Sea with the open ocean constraint of $\text{SST} = -1.8^{\circ}\text{C}$ for $I \geq 0.9$.)

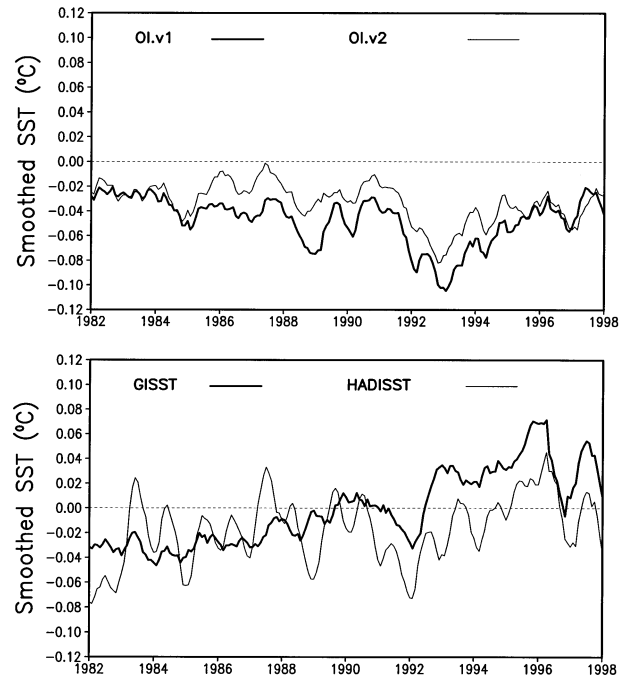


FIG. 8. Monthly averaged (60°S – 60°N) SST differences for (top) the OI.v1 and OI.v2, and (bottom) GISST, and HadISST analyses relative to COADS-E. Differences are computed only over grid boxes where COADS-E is defined. A 7-point running mean filter was used to smooth the time series.

6. Intercomparison of SST analyses

We recomputed the new version of the OI (OI.v2) using COADS data through 1997 and GTS data beginning in 1998 with the improved sea ice to SST conversion algorithm for the entire period (November 1981 to present). We now compare the OI.v2 with the other analyses. These results are based on monthly fields where the OI.v1 and OI.v2 were interpolated from the weekly analyses. As mentioned earlier, the OI.v1 uses a preliminary version of COADS prior to 1990 and GTS data beginning in 1990 with the old sea ice to SST conversion algorithm. Our first comparison is shown in Fig. 8 (for 60°S – 60°N), which shows time series of the difference of the OI.v1, OI.v2, GISST, and HadISST analyses with respect to COADS-E. The figure shows that the OI.v2 analysis is closer to COADS-E than the OI.v1 in almost every month. Thus, the residual negative bias in the OI.v1 analysis relative to COADS-E has been reduced in the OI.v2 analysis but not completely eliminated. This difference is due to the increase in the number of in situ observations (primarily from ships) that are available to the OI.v2. The magnitude of the residual bias is similar to the MOHSST to COADS-E differences after 1994 (see Fig. 4), but the sign is reversed. The residual bias of GISST with respect to COADS-E is initially similar to both versions of the OI. However, the GISST difference gradually changes sign. It is influenced by differences in MOHSST, version 6,

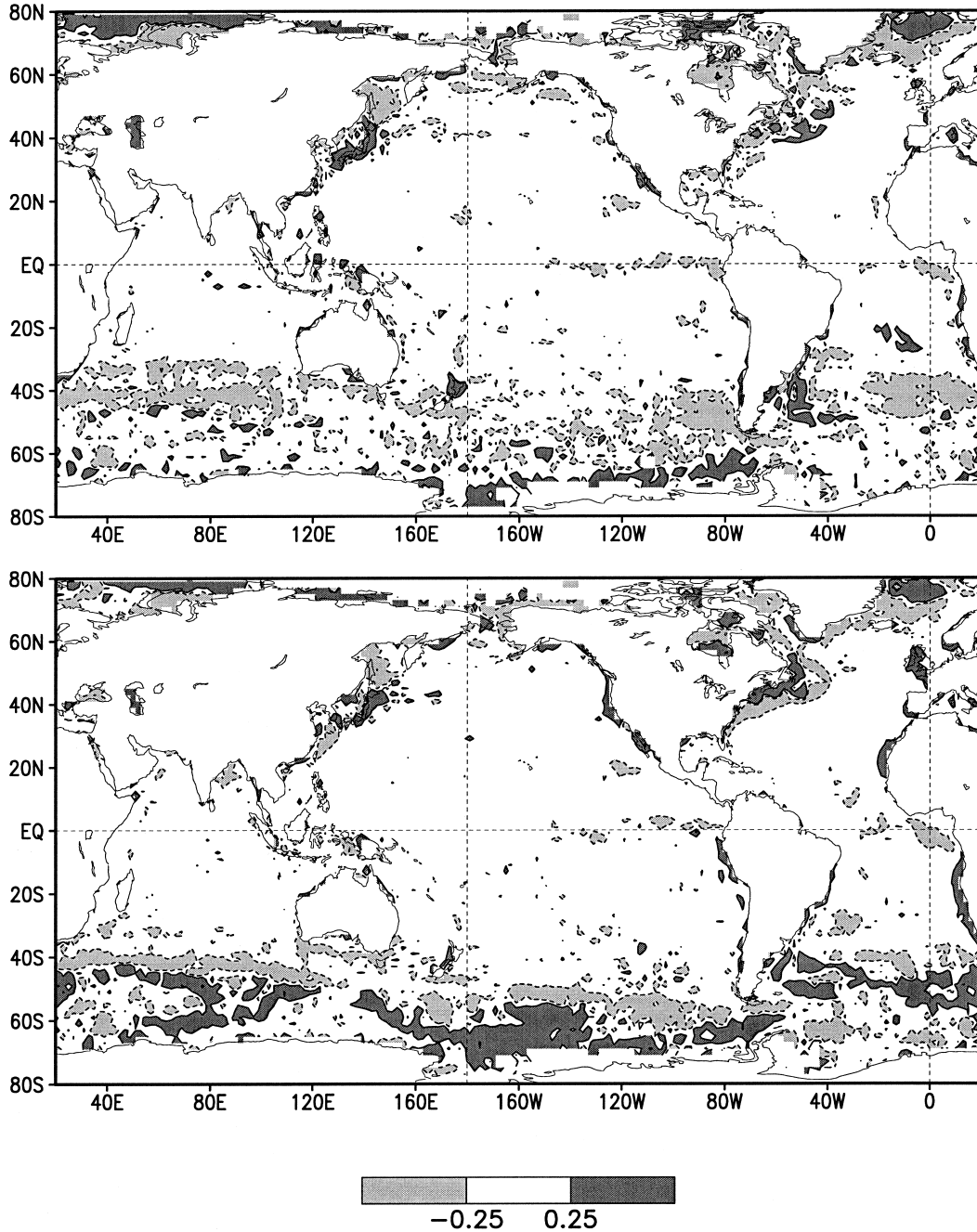


FIG. 9. Bias differences ($^{\circ}\text{C}$) between the (top) OI.v2 and (bottom) HadISST analyses relative to COADS-E for all months for 1982–97. Differences are shown on the COADS 2° spatial grid.

which are used as in situ input to GISST. HadISST incorporates a newer version of MOHSST (called HadSST; see Jones et al. 2001) and shows smaller differences relative to COADS-E than GISST.

We now compare the average spatial differences relative to COADS-E for the OI.v2 and HadISST as shown in Fig. 9. The most important difference occurs in the midlatitude Southern Hemisphere (roughly 60° – 30°S). There HadISST has both warm and cool regions relative

to COADS-E that tend to balance each other. However, the OI.v2 tends to be consistently cooler than COADS-E. This is the region that strongly contributes to the residual negative bias in the OI.v2 analysis as shown in Fig. 8. The differences between the two analyses and COADS-E are similar at high latitudes as expected because they use the same sea ice and sea ice to SST conversion algorithms. The GISST differences with respect to COADS-E (not shown) are very similar to those

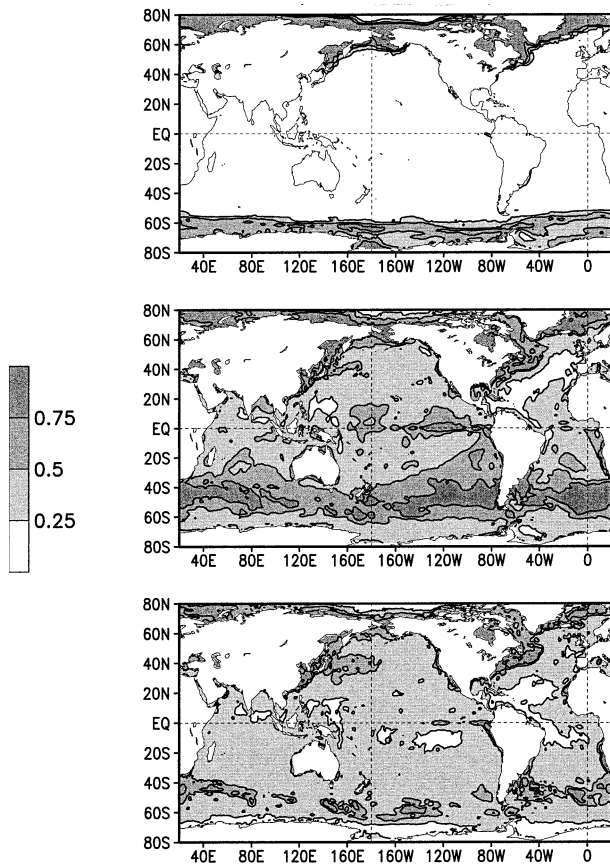


FIG. 10. The rms differences ($^{\circ}\text{C}$) between (top) the OI.v1, (middle) GISST, and (bottom) HadISST analyses relative to the OI.v2 analysis for all months for 1982–99. Differences are shown on a 1° spatial grid.

shown for HadISST including the high latitudes. The OI.v1 differences (also not shown) have a slightly larger negative bias. Otherwise they are very similar to the OI.v2 differences except at high latitudes where the change in the sea ice to SST conversion algorithm has a large impact as expected from Fig. 6.

The monthly rms differences of the GISST, HadISST, and OI.v1 analyses relative to the OI.v2 are shown in Fig. 10. The intercomparisons show that the OI.v1 and the OI.v2 analyses are similar except in the MIZ. Thus, the change in the in situ data has a relatively small effect. The differences between the OI.v2 and GISST analyses are greater. In addition, the differences tend to be lower along shipping routes where more in situ data are available. Hurrell and Trenberth (1999) showed that GISST had relatively low 1-month lagged autocorrelations for the 1982–97 period compared to other SST analyses and to GISST fields prior to 1982. This problem has been corrected in HadISST (Parker et al. 1999). This correction is also suggested by our figure, which shows that the OI.v2 to HadISST rms differences are much smaller than the OI.v2 to GISST differences. In addition, the OI.v2 and HadISST rms differences do not show

any ship track signatures. It is important to note that the *minimum* OI.v2 and HadISST rms difference is approximately 0.25°C . If the differences are assumed Gaussian, then at least one-third of the months should have larger differences. In the Gulf Stream, the rms differences exceed 0.6°C and can reach values of 1.4°C (not shown). This may be due to the different spatial and temporal analysis resolutions. HadISST fields for this period were analyzed as anomalies on a 2° latitude by longitude grid and then added back to the 1° resolution climatology to give the final 1° resolution fields. The OI.v2 is consistently analyzed on the finer 1° spatial resolution (and on weekly rather than monthly temporal resolution). This inevitably leads to large differences in regions of tight spatial SST gradients such as the Gulf Stream.

We computed the OI.v2 and HadISST differences by month to show the seasonal differences as suggested by Fig. 8. To illustrate the maximum seasonal range, we show the January and July differences in Fig. 11. The most systematic difference occurs in the Northern Hemisphere midlatitudes where the OI.v2 to HadISST difference is often positive ($>0.2^{\circ}\text{C}$) in January and often negative ($<-0.2^{\circ}\text{C}$) in July. Although these results are based on only 19 years, the seasonal variations are surprising in this region of dense in situ data (see Fig. 1). This is the region that contributes much of the seasonal variation shown in the HadISST in Fig. 8. It should be noted that there are also large differences in the Arctic in July even though the same sea ice concentration and sea ice to SST conversion algorithm have been used in both HadISST and the OI.v2. This is most likely related to the sparse SST data found in the MIZ, which limits analysis accuracy.

Our results show that improvements in the OI.v2 analysis relative to the OI.v1 are modest except at high latitudes. Because the SST climatology derived by Smith and Reynolds (1998) is strongly dependent on the OI.v1 at high latitudes, a new climatology was needed. Smith and Reynolds derived a monthly SST climatology on a 1° grid from two sources: the average 1° OI.v1 monthly fields for the period 1982–96 and the monthly average of the 2° gridded fields of Smith et al. (1996) for the period 1961–90. The Smith et al. (1996) fields were computed using a fit of empirical orthogonal functions to COADS data between 45°S and 69°N . In Smith and Reynolds (1998), a smoothed difference between the two climatologies was produced. The difference became a correction that was used to adjust the 1° OI-based climatology. This adjusted climatology maintained the 1° resolution while using a 1961–90 base period wherever the Smith et al. (1996) fields were defined.

Because of the large OI differences at high latitudes, we recomputed the climatology following the Smith and Reynolds (1998) method using the OI.v2 analysis (1982–2000) instead of the OI.v1. The adjusted climatology is again based on the period 1961–90. There

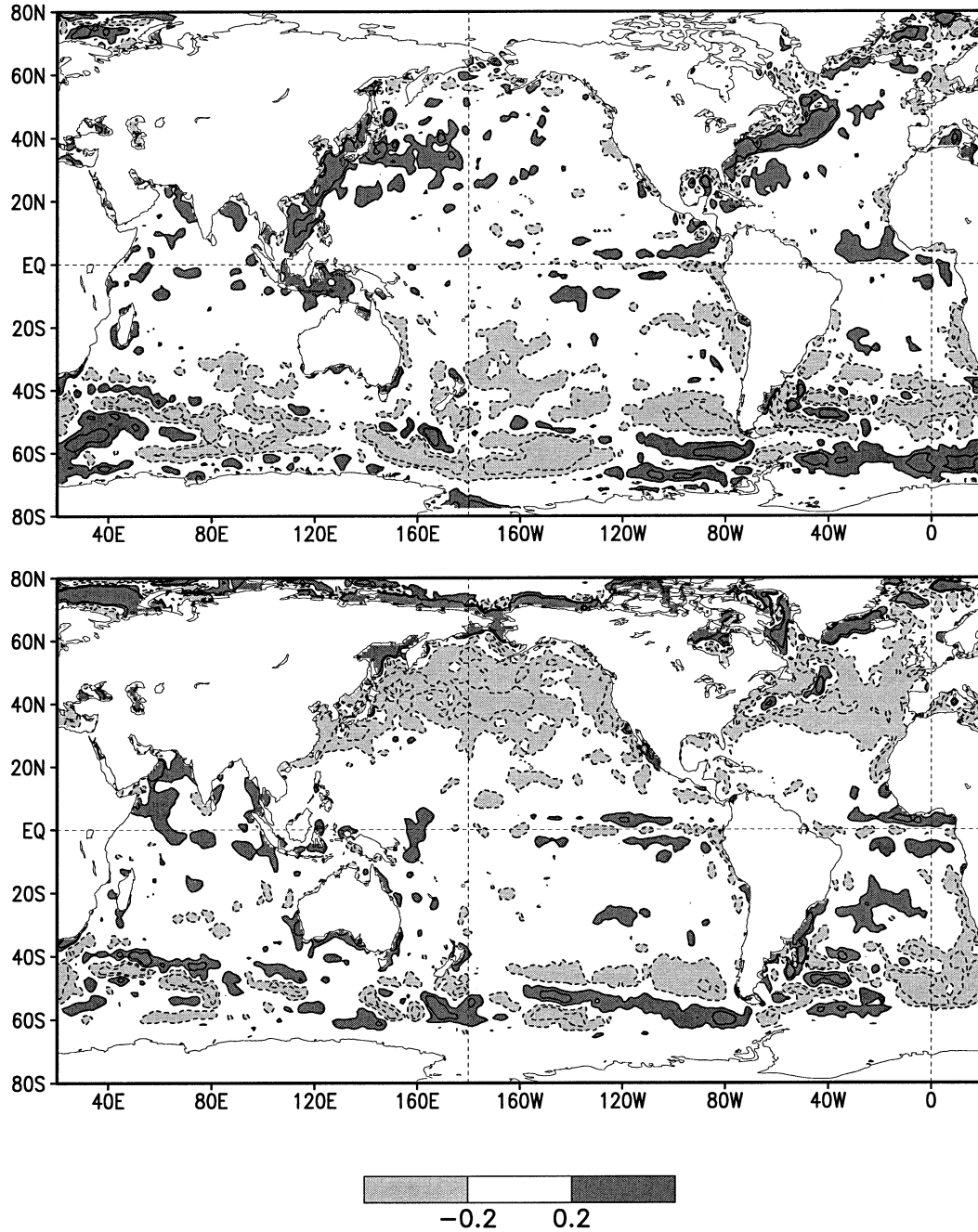


FIG. 11. Bias differences ($^{\circ}\text{C}$) between the OI.v2 and HadISST analyses for 1982–99 for (top) Jan and (bottom) Jul. Differences are shown on a 1° spatial grid. Contours are -0.5 , -0.2 , 0.2 , and 0.5 .

is little change between the two climatologies (not shown) except at high latitudes where the new climatology is warmer especially in summer months. There, the differences can exceed 1°C .

7. Final comments

As discussed in the preceding sections, the OI.v2 analysis is a replacement of RS OI.v1. The OI.v2 anal-

ysis has a modest improvement in the bias correction because of the addition of more in situ data. However, a small uncorrected residual bias of roughly -0.03°C remains. In addition, the OI.v2 uses an improved climatological sea ice to SST conversion algorithm that better fits the in situ data from COADS-E (enhanced COADS) as discussed in section 5.

We found it fairly easy to remove the major part of any satellite biases but difficult to remove a persistent

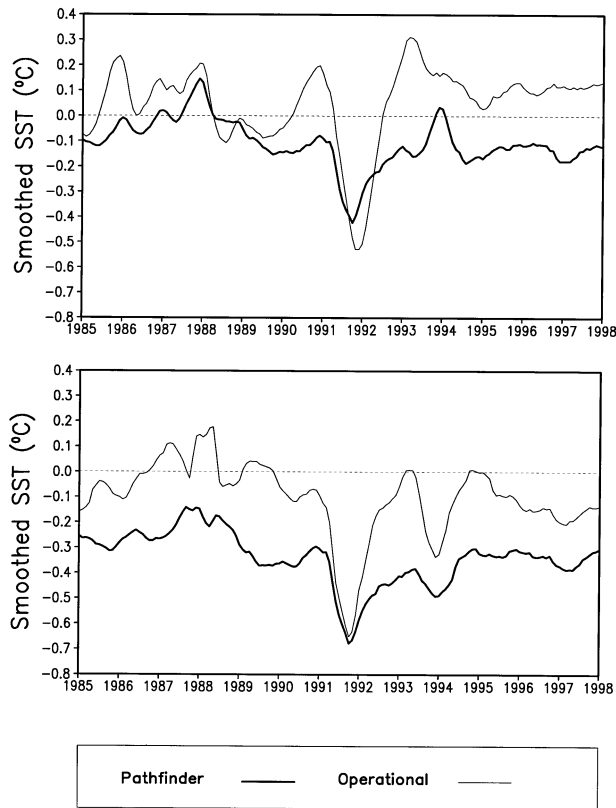


FIG. 12. Monthly averaged (60°S – 60°N) (top) day Pathfinder and operational satellite SSTs, and (bottom) night Pathfinder and operational satellite SSTs, relative to COADS-E. A 7-point running mean filter was used to smooth the time series. For each panel, differences are computed over common areas where all products are defined.

residual. It would, of course be easier if there were no satellite bias. A delayed-mode processing of satellite data, as done for the Pathfinder project (Kilpatrick et al. 2001), could correct these differences and could be a better product for climate studies. In Fig. 12, we compare the original operational day and night summaries and the Pathfinder day and night summaries relative to COADS-E. The figure shows a strong bias in all satellite summaries in late 1991 that relate to the Mount Pinatubo volcanic eruptions. Ignoring this part of the record, the daytime operational product is roughly 0.1°C warmer than COADS-E while the daytime Pathfinder and nighttime operational products are roughly 0.1°C colder. In addition, the nighttime Pathfinder data is roughly 0.3°C colder than COADS-E. This is not a “skin” minus “bulk” difference because the Pathfinder algorithm is determined by regression against buoy data as was done for the operational algorithm. However, the number of Pathfinder observations is much larger than for the operational product. Thus, it is likely that some of the Pathfinder cold bias is due to cloud contamination introduced by extending the number of retrievals retained. [Casey and Cornillon (1999) also determined that Pathfinder SST retrievals were negatively biased by cloud

contamination and developed a filter to reduce the bias.] As discussed in section 3, the estimated daytime satellite errors are larger than the nighttime errors, which emphasize nighttime retrievals in the OI. Thus, the use of Pathfinder satellite data instead of the operational satellite data would most likely increase the OI residual biases.

Our results also show that significant differences remain among analyses during the last two decades. In particular residual globally averaged differences of roughly 0.05°C occur on decadal scales. For monitoring and detecting climate change, these differences are, unfortunately, not negligible. It has been suggested that the OI.v2 be corrected by a constant, possibly ad hoc, offset. This could be done, but does not seem very satisfying. The correction patterns (see, e.g., Figs. 8 and 9) would have to vary with space and time. In addition the corrections are known only roughly because our standard of comparison, COADS-E, is noisy. Furthermore, it is not certain how the real-time analyses would be corrected. It seems to us that it would be much better to quantify the differences among the analyses by more careful comparisons. This could be done by intercomparisons using independent data as discussed in the appendix. This would help quantify differences in both data processing and analyses.

Even if the analysis were perfect, we would still need to improve both in situ and satellite observations. There are now new efforts to carefully monitor and improve the observations from selected ships (see http://www.soc.soton.ac.uk/JRD/MET/VOSCLIM/vos_links.html and http://oceanic.cms.udel.edu/gos/gcos/VOS_CLIM_PROJ.htm). In addition, there are new efforts to produce high-resolution SST analyses that include careful examination of the satellite algorithms, the bulk and skin SST difference, and utilize multiple sensors (see <http://www.bom.gov.au/bmrc/ocean/GODAE/HiResSST/index.html>). (The OI.v2 SST fields are available online at ftp://ftp.ncep.noaa.gov/pub/cmb/sst/oisst_v2 for the weekly version and at ftp://ftp.ncep.noaa.gov/pub/cmb/sst/oimonth_v2 for the monthly version).

Acknowledgments. The Pathfinder SST data was provided by Jorge Vasquez and Ed Armstrong, Jet Propulsion Laboratory. The AVHRR data used for the OI for the 1980s were obtained from the NASA Physical Oceanography Distributed Active Archive Center at the Jet Propulsion Laboratory. The graphics were computed using the Grid Analysis and Display System (GrADS) (<http://grads.iges.org/grads>), Center for Ocean–Land–Atmosphere Studies. The SST climatology for the 1961–90 base period and an additional version for the 1971–2000 base period was computed by Yan Xue, NOAA/NCEP/Climate Prediction Center, (http://www.cpc.noaa.gov/products/predictions/30day/SSTs/sst_clim.html). The sea ice adjustments were developed with invaluable help from Bill Chapman, Bob Grumbine, Jim Maslanik, David Parker, Mark Serreze, and John Walsh. We are grateful

to NCDC and the NOAA Office of Global Programs, which provided partial support for this work. The development of HadISST1 along with the sea ice adjustments and sea ice to SST conversion algorithms was supported by the U.K. Public Meteorological Service Research and Development contract. Chris Folland, Alexey Kaplan, David Parker, Tom Peterson, Yan Xue, and an anonymous reviewer provided valuable scientific and editorial help.

APPENDIX

Analysis Verification

We tested an analysis system that was designed to improve the satellite bias correction as suggested in section 5. This analysis used the same data as the OI.v2 including the improved sea ice algorithms, but was slightly less accurate. We would like to briefly discuss the analysis here because comparisons with the OI.v2 suggest a procedure that would be helpful in quantifying differences among analyses. The tested analysis used a variational method computed in two spatial dimensions and is henceforth called the 2DVAR. Both the OI and 2DVAR methods reduce to solving a linear equation, and the results should be very similar. The main advantage of the 2DVAR is that the solution is an iterative method of steepest descent (see Parrish and Derber 1992) that reduces computer storage requirements. In addition to replacing the OI procedure with the 2DVAR method, we replaced the Poisson bias correction with a preliminary 2DVAR analysis step. This preliminary step was a separate 2DVAR analysis of the weekly collocated satellite and in situ differences on a 1° grid. We expected that the preliminary analysis step would give a stronger bias correction because it would use the previous 2DVAR correction field as a first guess. The bias correction change would be the major contributor to any OI.v2 and 2DVAR difference.

To get quantitative estimates of analysis differences, we produced several analyses while holding back some of the data. We then used the withheld data for independent verification. For this study our standard of comparison will be buoy data. To attempt to avoid any selection bias, we simply exclude any buoy data (both moored and drifting) with a call sign ending in either 4 or 9 from both the analysis and the bias correction. This randomly excludes approximately 20% of the buoy data. The fraction of withheld buoys was selected to exclude enough data for verification while minimizing the impact on the analysis.

We excluded these data from three analyses. They are 1) the OI with the Poisson bias correction (labeled OI.v2), 2) the OI without any bias correction (labeled OLNO.v2), and 3) the 2DVAR analysis with the 2DVAR-based bias correction (labeled 2DVAR). All analyses were produced weekly for the years 1990–97. This period was selected to include periods of El Niño

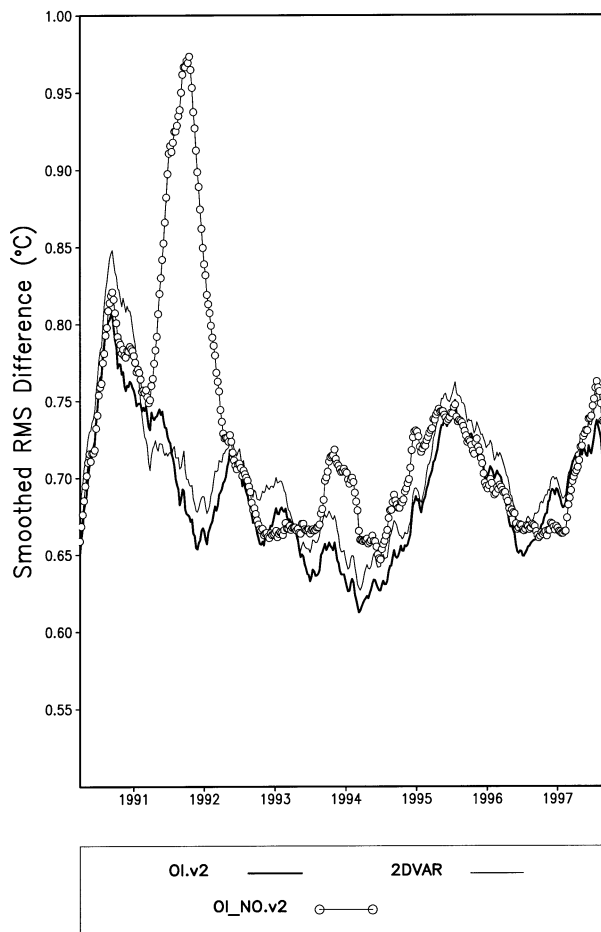


FIG. A1. Weekly rms differences between independent buoy data (65°S – 80°N) and three analyses: OI.v2, OLNO.v2, and 2DVAR. A 23-point running mean filter was used to smooth the time series.

and La Niña as well as a period of high-stratospheric volcanic aerosols. It also is a period with a potential under correction of biases as indicated in Fig. 3. Years prior to the 1990s would be less useful for this comparison because buoy data were much less plentiful. All analyses use exactly the same data.

Once the analyses were completed, we used the withheld buoy data as a comparison standard and computed weekly analysis minus buoy statistics including biases, standard deviations, and rms differences. We present the global rms difference between the buoys and each analysis in Fig. A1. The weekly time series have been smoothed by a 23-point running mean filter to reduce the noise. The results show that the OI.v2 has the lowest rms difference with respect to the withheld buoy data for almost all weeks. As discussed in Reynolds (1993), the Mt. Pinatubo eruptions in June 1991 lead to the onset of the biases from the aerosols beginning in July 1991. This caused the large rms differences between the OLNO.v2 and the buoys in late 1991. However, the Pinatubo eruptions cannot explain the period in early

1991 when the 2DVAR had slightly lower rms differences than the OI.v2. It is also important to point out the OLNO.v2 did show the lowest rms difference for part of 1993 and 1996. In these periods, the bias correction actually slightly reduced the analysis accuracy.

The results shown in the figure and other statistical comparisons showed that the OI.v2 usually had a slightly lower rms difference than the 2DVAR. This result convinced us that we should retain the OI at present even though the 2DVAR has important computational advantages. We believe that the accuracy of the 2DVAR was adversely affected because our version required that the correction scales be isotropic and homogenous. In the bias correction step large correlation scales of 1650 km were used. These scales were appropriate for the midlatitude Southern Hemisphere where in situ data are sparse but not for other regions between 20°S and 60°N. It is hoped that the 2DVAR analysis will be improved in the future.

SST analyses have improved over the last 20 years and differences among analyses have tended to become smaller. However, as we have shown, differences remain. In our opinion the best way to continue to improve the analyses is for all SST analysis groups to agree to compute analyses for a limited period using exactly the same input dataset. This dataset should include both dependent data for the analysis and independent data for the verification.

REFERENCES

- Casey, K., and P. Cornillon, 1999: A comparison of satellite and in situ-based sea surface temperature climatologies. *J. Climate*, **12**, 1848–1863.
- Cavalieri, D. J., C. L. Parkinson, P. Gloersen, J. C. Comiso, and H. J. Zwally, 1999: Deriving long-term time series of sea ice cover from satellite passive-microwave multisensor data sets. *J. Geophys. Res.*, **104**, 15 803–15 814.
- Comiso, J. C., and R. Kwok, 1996: Surface and radiative characteristics of the summer Arctic sea ice cover from multisensor satellite observations. *J. Geophys. Res.*, **101**, 28 397–28 416.
- Folland, C. K., R. W. Reynolds, M. Gordon, and D. E. Parker, 1993: A study of six operational sea surface temperature analyses. *J. Climate*, **6**, 96–113.
- Gandin, L. S., 1963: *Objective Analysis of Meteorological Fields*. Gidrometeorizdat, 238 pp. (Translated from Russian by Israeli Program for Scientific Translations in 1965.)
- Grumbine, R. W., 1996: Automated passive microwave sea ice concentration analysis at NCEP/DOC/NOAA/NWS/NCEP/EMC/OMB Tech. Note 120, 13 pp. [Available online at <http://polar.wvb.noaa.gov/seaice/docs/ssmi.auto/ssmi120.html>.]
- Hanna, E., and J. Bamber, 2001: Derivation and optimization of a new Antarctic sea-ice record. *Int. J. Remote Sens.*, **22**, 113–139.
- Hurrell, J. W., and K. E. Trenberth, 1999: Global sea surface temperature analyses: Multiple problems and their implications for climate analysis, modeling, and reanalysis. *Bull. Amer. Meteor. Soc.*, **80**, 2661–1678.
- Jones, P. D., T. J. Osborn, K. R. Briffa, C. K. Folland, E. B. Horton, L. V. Alexander, D. E. Parker, and N. A. Rayner, 2001: Adjusting for sampling density in grid box land and ocean surface temperature time series. *J. Geophys. Res.*, **106**, 3371–3380.
- Kent, E. C., P. K. Taylor, B. S. Truscott, and J. A. Hopkins, 1993: The accuracy of voluntary observing ship's meteorological observations. *J. Atmos. Oceanic Technol.*, **10**, 591–608.
- , P. G. Challenor, and P. K. Taylor, 1999: A statistical determination of the random observational errors present in voluntary observing ships' meteorological reports. *J. Atmos. Oceanic Technol.*, **16**, 905–914.
- Kilpatrick, K. A., G. P. Podesta, and R. Evans, 2001: Overview of the NOAA/NASA advanced very high resolution radiometer Pathfinder algorithm for sea surface temperature and associated matchup database. *J. Geophys. Res.*, **106**, 9179–9198.
- Knight, R. W., 1984: Introduction to a new sea-ice database. *Ann. Glaciol.*, **5**, 81–84.
- Lorenc, A. C., 1981: A global three-dimensional multivariate statistical interpolation scheme. *Mon. Wea. Rev.*, **109**, 701–721.
- May, D. A., M. M. Parmeter, D. S. Olszewski, and B. D. McKenzie, 1998: Operational processing of satellite sea surface temperature retrievals at the Naval Oceanographic Office. *Bull. Amer. Meteor. Soc.*, **79**, 397–407.
- McClain, E. P., W. G. Pichel, and C. C. Walton, 1985: Comparative performance of AVHRR-based multichannel sea surface temperatures. *J. Geophys. Res.*, **90**, 11 587–11 601.
- McPhaden, M. J., and Coauthors, 1998: The tropical ocean global atmosphere (TOGA) observing system: A decade of progress. *J. Geophys. Res.*, **103**, 14 169–14 240.
- Nomura, A., 1995: Global sea ice concentration data set for use in the ECMWF Re-analysis system. ECMWF Tech. Rep. 76, 25 pp.
- Parker, D. E., P. D. Jones, C. K. Folland, and A. Bevan, 1994: Interdecadal changes of surface temperature since the late 19th century. *J. Geophys. Res.*, **99**, 14 377–14 399.
- , C. K. Folland, A. Bevan, M. N. Ward, M. Jackson, and K. Maskell, 1995a: Marine surface data for analysis of climatic fluctuations on interannual to century timescales. *Natural Climate Variability on Decade-to-Century Time Scales*, D. G. Martinson et al., Eds., National Academy Press, 241–250.
- , —, and M. Jackson, 1995b: Marine surface temperature: Observed variations and data requirements. *Climatic Change*, **31**, 559–600.
- , N. A. Rayner, E. B. Horton, and C. K. Folland, 1999: Development of the Hadley Centre sea ice and sea surface temperature data sets (HadISST). *WMO Workshop on Advances in Marine Climatology-CLIMAR99*, Vancouver, BC, Canada, Environment Canada, 194–203.
- Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical–interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763.
- Rayner, N. A., E. B. Horton, D. E. Parker, C. K. Folland, and R. B. Hackett, 1996: Version 2.2 of the global sea-ice and sea surface temperature data set, 1903–1994. Climate Research Tech. Note 74, 43 pp. [Available from Met Office, London Road, Bracknell, Berkshire RG12 2SY, United Kingdom.]
- Reynolds, R. W., 1988: A real-time global sea surface temperature analysis. *J. Climate*, **1**, 75–86.
- , 1993: Impact of Mount Pinatubo aerosols on satellite-derived sea surface temperatures. *J. Climate*, **6**, 768–774.
- , and T. M. Smith, 1994: Improved global sea surface temperature analyses using optimum interpolation. *J. Climate*, **7**, 929–948.
- , C. K. Folland, and D. E. Parker, 1989: Biases in satellite derived sea-surface temperature data. *Nature*, **341**, 728–731.
- Saur, J. F. T., 1963: A study of the quality of sea water temperatures reported in logs of ships' weather observations. *J. Appl. Meteor.*, **2**, 417–425.
- Slutz, R. J., S. J. Lubker, J. D. Hiscox, S. D. Woodruff, R. L. Jenne, D. H. Joseph, P. M. Steurer, and J. D. Elms, 1985: Comprehensive Ocean–Atmosphere Data Set. Release 1, NOAA/ERL, 268 pp. [Available from NOAA/Climate Diagnostics Center, Boulder, CO 80307.]
- Smith, T. M., and R. W. Reynolds, 1998: A high resolution global sea surface temperature climatology for the 1961–90 base period. *J. Climate*, **11**, 3320–3323.
- , —, R. E. Livezey, and D. C. Stokes, 1996: Reconstruction

- of historical sea surface temperatures using empirical orthogonal functions. *J. Climate*, **9**, 1403–1420.
- Tolstikov, E. I., Ed., 1966: *Atlas of the Antarctic*. Vol. 1. Glavnoe Upravlenie po Geodezii i Kartografii, Mivisterstvo Geologii, SSSR, 225 pp.
- Walsh, J. E., and W. L. Chapman, 2001: Twentieth-century sea ice variations from observational data. *Ann. Glaciol.*, **33**, 444–448.
- Walton, C. C., W. G. Pichel, J. F. Sapper, and D. A. May, 1998: The development and operational application of nonlinear algorithms for the measurement of sea surface temperatures with the NOAA Polar-Orbiting Environmental Satellites. *J. Geophys. Res.*, **103**, 27 999–28 012.
- Weller, R. A., and S. P. Anderson, 1996: Surface meteorology and air–sea fluxes in the western equatorial Pacific warm pool during the TOGA Coupled Ocean–Atmosphere Response Experiment. *J. Climate*, **9**, 1959–1990.
- Woodruff, S. D., H. F. Diaz, J. D. Elms, and S. J. Worley, 1998: COADS Release 2 data and metadata enhancements for improvements of marine surface flux fields. *Phys. Chem. Earth*, **23**, 517–527.