Adequacy of the In Situ Observing System in the Satellite Era for Climate SST

Huai-Min Zhang, Richard W. Reynolds, and Thomas M. Smith

NOAA NESDIS National Climatic Data Center

February 18, 2005

(Revision, submitted to J. Atmos. Oceanic Tech.)

Corresponding Address:

Huai-Min Zhang

NOAA NESDIS National Climatic Data Center

151 Patton Avenue

Asheville, NC 28801

E-mail: huai-min.zhang@noaa.gov

Phone: 828-271-4090

Abstract

A method is presented to evaluate the adequacy of the recent in situ network for climate sea surface temperature (SST) analyses using both in situ and satellite observations. Satellite observations provide superior spatiotemporal coverage but with biases; in situ data are needed to correct the satellite biases. Recent NOAA/U.S. Navy operational AVHRR satellite SST biases were analyzed to extract typical bias patterns and scales. Occasional biases of 2°C were found during large volcano eruptions and near the end of the satellite instruments' lifetime. Since future biases could not be predicted, the in situ network was designed to reduce the large biases that have occurred to a required accuracy. Simulations with different buoy density were used to examine their ability to correct the satellite biases and to define the residual bias as a potential satellite bias error (PSBE).

The PSBE and buoy density (BD) relationship was found to be nearly exponential, resulting in an optimal BD range of 2-3/10°x10^{o} box for efficient PSBE reduction. A BD of 2 buoys/10°x10^{o} reduces a 2°C maximum bias to below 0.5^{o}C and reduces a 1°C maximum bias to about 0.3^{o}C. The present in situ SST observing system was evaluated to define an equivalent buoy density (EBD), allowing ships to be used along with buoys according to their random errors. Seasonally averaged monthly EBD maps were computed to determine where additional buoys are needed for future deployments. Additionally, a PSBE was computed from the present EBD to assess the in situ system's adequacy to remove potential future satellite biases.

1. Introduction

Historically, ocean observations have been made for a variety of purposes by various groups. Although these observations provided important information about the ocean, the purposes and groups did not necessarily provide an optimal observing system for climate monitoring and study. Over the last few decades, international groups have begun designing a Global Ocean Observing System (GOOS) as a component of the Global Climate Observing System (GCOS). An international "Ocean Observing Conference for Climate" was held in Saint Raphael, France in October 1999 to help produce a better integrated system (Koblinsky and Smith, 2001). Sea surface temperature (SST) was one of the important parameters considered at the conference. The purpose of this paper is to examine the present in situ and satellite observing system and to recommend how future in situ observing system should be improved to efficiently correct satellite biases for climate SST.

The present SST observing system consists of in situ and satellite observations as discussed by Reynolds et al. (2002). In situ observations are made from ships and buoys (both moored and drifting). For historical reasons, error information is not always available for each observation. To our knowledge, the most extensive studies on the in situ SST random error are Kent et al. (1993 and 1999), Parker et al. (1995), and Emery et al. (2001). They studied the effects of different instrumentations, instrument layouts (e.g., at different depths, drifting versus moored buoys), and different observation techniques. Based on the above, Reynolds and Smith (1994) and Reynolds et al. (2002) estimated that typical random errors are 0.5 and 1.3°C for buoy and ship observations, respectively.

Satellite observations have provided dramatically improved coverage in time and space. Over the last two decades, the satellite coverage has expanded from one infrared (IR) instrument to the present day array of multiple IR and microwave instruments. Satellite random error for the Advanced Very High Resolution Radiometers (AVHRR) SST was discussed by McClain et al. (1985) and May et al. (1998). Reynolds and Smith (1994) and Reynolds et al. (2002) estimated that typical random errors are 0.5 and 0.3°C for daytime and nighttime operational NOAA (National Oceanic and Atmospheric Administration)/U.S. Navy AVHRR SSTs, respectively.

The high satellite data coverage reduces sampling and random errors in SST analyses using combined in situ and satellite data. For example, application of the optimal averaging (OA) procedure of Smith et al. (1994; also Kagan 1979) on 5° monthly ship, buoy and AVHRR SST observations shows that the OA random and sampling errors are typically less than 0.2^{o}C for the 1995-2002 period (not shown). In regions with sparser satellite data due to cloud cover, the monthly OA analysis errors increase only to 0.3^{o}C. Including more recent satellite data [the Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI, Kummerow et al. 1998; Wentz et al. 2000), the Along Track Scanning Radiometer (ATSR, Mutlow et al., 1994), the Advanced Microwave Scanning Radiometer (AMSR-E, e.g., Wentz et al. 2003), and the Moderate Resolution Imaging Spectroradiometer (MODIS, Esaias et al. 1998)] further reduces the OA errors. However, satellite bias error remains significant as will be shown below.

Bias error is the systematic difference between one instrument or a set of instruments (e.g., those onboard satellites) and another (e.g., those used for in situ observations). As shown in Zhang et al. (2004), the NOAA/U.S. Navy operational

AVHRR SST biases have changed over time, and they can be as large as 2°C. Similar biases can occur with other satellites [e.g., Reynolds et al. (2004) discussed biases in the Tropical Rainfall Measuring Mission (TRMM) Microwave Imager retrievals]. Satellite biases change with time due to orbit changes, aging of satellite instruments, and changes in atmospheric conditions which may differ from those used in the development of the satellite SST retrieval algorithms (e.g., unexpected volcanic aerosols). These biases needed to be corrected to minimize systematic errors in climate SST analyses.

There are a large number of SST analyses on different time and space scales produced from combinations of different parts of the in situ and satellite observing system. This paper focuses on climate scales. Following Reynolds et al. (2002), these scales are defined to have temporal resolutions of one week or longer and spatial resolutions of 1° or larger. A version of higher resolution analysis (daily and 0.25^{o}) has been planned. However, that will not change the results of this study on the needed buoy density since the conclusions are drawn from satellite biases of larger scales (section 3).

The Reynolds and Smith (1994) and Reynolds et al. (2002) SST product, often referred as Reynolds SST, has been widely used by the climate community, although it has been produced operationally on weekly basis. In the present study, we use their optimum interpolation (OI) and bias correction techniques to determine the optimal in situ data density for efficient bias reduction. Consistent with the operational Reynolds SST product, the NOAA/U.S. Navy operational AVHRR SST is used in this study. Note that other reprocessed AVHRR SSTs exist, such as the Pathfinder AVHRR SST (Kilpatrick et al. (2001). Early versions of the Pathfinder AVHRR SST had biases as well and the biases were not necessarily smaller than the operational AVHRR SST (Reynolds

et al. 2002). Since then, improved Pathfinder AVHRR SST has been produced with reduced biases (information and data available at http://www.nodc.noaa.gov/sog/pathfinder4km). It is planned that the next version of the Reynolds SST will be based on the new Pathfinder AVHRR SST together with other satellite SST. Once this is done, the work presented here will be re-evaluated.

Details of the OI and bias correction techniques can be found in Reynolds and Smith (1994) and Reynolds et al. (2004). Briefly, the OI objectively determines a series of weights for SST data increments at each grid point. The data increment is the difference between each observation and the analysis first guess value. Reynolds and Smith (1994) used the previous week's analysis as the first guess in their weekly analysis on a 1°spatial grid. The OI method assumes that the data do not contain long-term biases (e.g., see Lorenc, 1981). Because satellite biases occur, an optional step using a correction based on Poisson's equation can be carried out to remove satellite biases relative to in situ data prior to the OI. This step produces an adjustment of the satellite data, anchoring it to the in situ data and matching the gradients of the two fields. In the OI procedure, various error statistics are assigned that are functions of latitude and longitude. In this work, two changes were made in the OI procedures. First, the analysis was computed on monthly time scales instead of weekly. Second, the first guess was taken as the monthly climatology in the buoy need simulations (see section 4 for details).

The remainder of the paper is organized as the following. Section 2 presents the strategy for this study. Section 3 presents the typical satellite SST bias patterns extracted from an empirical orthogonal function (EOF) analysis. Section 4 presents the relationship between in situ data density and satellite bias reduction rate. Section 5 evaluates the

adequacy of the present in situ system, and recommends where to deploy additionally needed buoys. Section 6 presents a discussion and summary.

2. Study Strategy

This section outlines the overall strategy of this study. To correct the satellite biases with respect to in situ data, we first need to understand the spatial scales and magnitudes of the biases. The in situ data density requirement for bias correction depends on the spatial patterns of the satellite biases. For example, if the biases were globally constant, only one accurate in situ observation would be needed to correct the constant bias. Generally, the more complicated the bias spatial pattern, the more in situ observations are needed. The spatial patterns of the major biases are characterized by an EOF analysis in section 3 for the historical NOAA/U.S. Navy operational AVHRR SST, the longest satellite SST record. Because future satellite biases cannot be predicted, we simulate them using the historical bias patterns, reconstructed from the EOF spatial modes. We then design an in situ network to sufficiently correct all the bias to a desired accuracy.

In section 3 we show that the major bias EOF modes are linked to physical phenomena in the atmosphere, such as aerosols from volcano eruptions and seasonal variations of clouds and desert dust aerosols. Because each of the physical phenomena can occur independently, the designed in situ network has to be able to correct bias caused by each EOF as well as the composite. Thus, in the determination of the needed in situ data density, we simulate the bias regime corresponding to each EOF spatial mode separately. To assure that the in situ system can deal with biases associated with

simultaneous multiple phenomena, each simulated bias case (each EOF spatial pattern) is scaled to a maximum total (composite) bias magnitude. This is certainly a conservative strategy since under normal conditions not all types of biases occur at the same time. However, this is necessary since future occurrences of biases cannot be predicted and the designed in situ system has to be able to deal with the worst cases.

Once the representative biases (EOF spatial modes) and magnitudes are chosen (section 3), a Monte Carlo simulation is used to simulate the bias variations in time (section 4). Simulations are run with various in situ data density to determine an optimal in situ density for efficient error reduction.

3. Satellite SST bias scales and patterns

To design a better in situ network to correct potential satellite biases, it is necessary to examine the typical scales and patterns of historical biases. The objective is to extract the dominant components of the biases, and to use them to simulate typical bias regimes. This is done using EOF analysis (e.g., Davis 1976) of the satellite biases. Briefly, the EOF analysis decomposes the multivariate bias into orthogonal modes, where a small number of the modes often contain the major part of the data variance.

Ideally, the biases could be computed at co-located positions of satellite and in situ data. However, the sparseness of the in situ data hinders extracting the spatial scales and patterns. Thus the OI technique of Reynolds and Smith (1994) was used to compute the OI SSTs on a regular 1° spatial grid with monthly in situ and AVHRR satellite data (Zhang et al. 2004). To define the satellite biases, two OI SSTs were computed, one with and one without the satellite bias correction step. The satellite bias was then defined as

the difference between the two OI SSTs. Because the bias was defined with reference to the in situ data, which was sparser at high southern latitudes, the bias might be underestimated there. This underestimation was not critical because our goal was to extract the bias spatial scales. As will be shown in section 4, it is the bias spatial pattern, not the bias magnitude, that determines the optimal in situ data density range for efficient bias reduction.

The bias analysis was detailed in Zhang et al. (2004) for 1982-2002. Here the bias EOFs for 1990-2002 were recomputed for the design of the in situ network. The 1982-1989 data were not used because the buoy data, which are of critical importance in the tropics and Southern Hemisphere, were sparser in the 1980s, as discussed in Reynolds et al. (2002). However, the major EOF features for the two time periods are very similar.

The total bias was first separated into the time mean and the deviation from the mean. Fig. 1a shows the bias mean for the 1990-2002 period. Fig. 1b shows the root-mean-square (RMS) of the bias with the mean removed, and is referred as the bias standard deviation. The mean bias is primarily negative in the open ocean, most likely due to cloud and aerosol contamination on the satellite retrievals. The magnitudes of the mean bias are larger in the tropics, especially in the eastern Pacific and Atlantic and western Indian Ocean, with largest magnitudes found in the tropical Atlantic where they exceed 0.5^{o}C, largely due to dust aerosol (e.g., Haywood et al. 2001). Positive means can be seen along the coasts of North America, with weaker values off the coasts of Asia roughly north of 40°N as well as off parts of the coasts of Africa and South America. The time variation of the biases, represented by the bias standard deviation (Fig. 1b), is generally smoother in space, and the variations are small along 20°N and 20°S and large

along the equator and 40°N and 40°S. The largest variations exceed 0.4^{o}C. The drop in both bias mean and standard deviation south of about 50°S is most likely due to uncorrected satellite biases due to the lack of in situ data. When the bias mean and standard deviation are combined into the total RMS bias (not shown), the maximum values can exceed 0.6^{o}C.

The bias EOFs were computed with the bias mean removed. Fig. 2 shows the EOF spatial patterns and time series for modes 1 to 6. Mode 1 reveals biases primarily due to stratospheric aerosols from volcanic eruptions, with large scale zonal tropical biases. These biases can last for several months. Modes 2 and 3 are seasonal biases, which are strongly related to local weather phenomena, such as seasonal dust aerosols and cloud covers. Mode 4 represents a bias trend in time. Modes 5 and 6 indicate interannual variations. Mode 6 shows the strongest loadings in the middle latitude Southern Hemisphere. This is an important mode to include in the simulations below, because in situ data are sparse south of roughly 40°S. Immediately higher modes show similar spatial scales. Much higher modes show smaller scales and more scattered spatial features that may reflect the increased role of data noise. Taken together, the first six EOFs represent 52.7% of the total variance (Fig. 3). As shown in the next section, the buoy need density for bias correction generally increases as the mode number increases from 1 to 4, and the buoy need density is similar for modes 4 through 6.

4. Optimal buoy density for efficient satellite bias reduction

In this section, we study the relationship between the in situ data density and the bias error reduction rate. The relationship was quantified for various bias patterns,

represented by the bias EOF modes computed in the previous section. As outlined in the strategy in section 2, each of the EOF modes was treated as a typical bias regime that could happen independently, and each bias representation was scaled to a composite bias magnitude to assure that the in situ system can reduce each bias mode as well as composite bias from multiple modes to a required accuracy. The simulations used a maximum bias of 2°C to simulate a worse case as discussed in section 3 and detailed in Zhang et al. (2004).

To study the response of bias correction to changes in buoy density, the satellite SST values were simulated as the monthly climatology (as the assumed ground truth) plus the representative biases (scaled EOFs) at the locations of actual satellite observations. This was done monthly between 1990 and 2002. The monthly climatology was formed using the 1990-2002 monthly OI SSTs, which was computed using the version with the bias corrected data. The buoy need density for satellite bias correction is not sensitive to the choice of the climatology. To simulate the satellite data over the 156-month (1990-2002) period, a time varying amplitude function is needed for the EOF spatial patterns for more realistic simulations. The time amplitude is defined as a Gaussian random time series, with a mean of zero and a standard deviation of one. Consequently, the RMS (in time) of the simulated bias has a maximum value of 2°C over the global ocean. The simulated satellite SST can be expressed as

$$T_{si}(x,t) = T_g(x,t) + EOF_i(x)\, a(t), \qquad for\ i=1\ to\ 6. \tag{1}$$

Here $T_{si}(x,t)$ is the simulated satellite SST, and $T_g(x,t)$ is the climatology, which is used as the ground truth. $EOF_i(x)$ is the EOF spatial mode i with the simulations run for each mode separately as outline in section 2. The Gaussian random time amplitude is

represented by $a(t)$. The variable t is the time in months from $t = 1$ to 156 (January 1990 to December 2002), and x is the vector location of each satellite observation. Random noise was not added to satellite data because of the high data density, which reduces the random errors to insignificant levels compared to the bias errors.

Buoy SST values were simulated as the ground truth $T_g(x,t)$ plus typical random buoy SST error. To study the response of bias correction to changes in buoy density, the buoy data were placed on regular grids with various grid resolutions for the multiple simulations. For each grid resolution, one buoy was placed at each grid point. For sparser buoy data, data noise becomes more important. Thus a Gaussian random noise was added to the simulated buoy SST value. As discussed in section 1, the buoy SST error in the OI was defined to be 0.5°C (Reynolds and Smith 1994). Thus, the simulated buoy SSTs were expressed as

$$T_b(x,t) = T_g(x,t) + 0.5\ e(t). \tag{2}$$

Here $T_b(x,t)$ is the simulated buoy SST, and $e(t)$ is the Gaussian random time series with a zero mean and standard deviation of 1.

The OI analysis with bias correction was then applied on the simulated satellite and buoy data. By design, the monthly climatology would be the expected result of the OI SST if the satellite biases were completely removed. The difference between the OI SST with bias correction and the monthly climatology is the uncorrected residual bias, which is evaluated as a function of buoy density. Intuitively, if there were no in situ data available, any satellite biases would be uncorrected. In contrast, if the buoy data were densely distributed over the global ocean, all the satellite biases would be nearly completely removed with reference to the in situ data. The objective here is to determine

the response of bias reduction rate to the buoy density and to define a threshold buoy density for a required SST accuracy.

Residual uncorrected SST biases were computed over the global ocean for each month over the simulation period. The residual biases were any biases that could not be eliminated by the simulated buoy density. The RMS of the residual bias was defined as the potential satellite bias error (PSBE). We refer to this as potential bias because the simulated satellite bias was set to a maximum of 2°C over the global open ocean, which is representative of the larger satellite biases observed over the satellite period.

The PSBE was computed over the simulation period. Fig. 4 shows examples of the PSBE for simulated bias regime of EOF mode 1 (scaled to 2°C in spatial maximum). If there were no buoy data, the PSBE would be the absolute value of the simulated bias (scaled EOF mode 1, Fig. 4a). For a buoy grid resolution of 20° (Fig. 4b), the biases were greatly reduced, but the PSBE exceeded 0.5°C at a few locations. At buoy grid resolution of 7° (Fig. 4c), all PSBEs were reduced to less than 0.4°C.

The spatial maximum of the PSBE over the global ocean was computed as a function of the buoy grid resolution. This was done for each of the six EOF bias representations. It was first found that the global one-point maximum of PSBE was somewhat unstable at larger (>12°) buoy grid resolutions. To obtain a more stable curve, an average PSBE was computed over a defined area surrounding the single point maximum. The averaging area was defined as the locations where $|EOF| > 1^{o}$C. (Note that as stated earlier, the maximum for each EOF had been scaled to 2°C.) To convert the area averaged PSBE to the corresponding maximum residual bias error, the averaged

PSBE were scaled up so that the averaged PSBE value would still be 2°C when there were no buoys.

Fig. 5 shows the adjusted PSBE as a function of the buoy grid resolution for the six EOF bias representations. In this figure, biases of modes 1 to 3 are relatively easy to correct because of their larger spatial structures. Mode 1 has near global zonal scales (Fig. 2). Modes 2 and 3 are of basin scales in the northwest North Pacific and North Atlantic. The curves for modes 4 to 6 are similar to each other due to their similar spatial scales even though their individual spatial patterns are different. Modes 4 to 6 have more stringent requirements because of their smaller spatial scales.

Because of their similarity, the curves for modes 4, 5 and 6 in Fig. 5 are averaged, with the average curve shown in Fig. 6a as the solid line with circles. Overall, the PSBE and buoy data grid resolution has a nearly linear relationship, although the mid section has a slightly steeper slope (quicker error reduction). The dashed line is a linear least-square fit to the data:

$$PSBE = 0.060 d_0 + 0.075 \qquad (3)$$

where d_0 is the grid resolution (decreasing from 20° to 2°). This fit has a RMS residual error of 0.04°C compared to the actual PSBE values (circles).

Surface drifting SST buoys are normally deployed in clusters for practical reasons. Hence it is desirable to plot the PSBE curve as a function of buoy density (*BD*) in a defined grid box. Fig. 6b shows the PSBE with respect to the *BD* with a fixed grid resolution of 10°x10° (denoted as BD_{10}). Mathematically, the conversion from the horizontal axis of Fig. 6a [i.e. one buoy at each grid point (BD_0=1) with grid resolution d_0] to the horizontal axis of Fig. 6b (i.e. number of buoys per 10°x10° box, BD_{10}) is

$$BD_{10} = BD_0\,(d_{10}^2/d_0^2), \qquad \text{for } d_0 = 20 \text{ to } 2. \qquad (4)$$

Here $d_{10} = 10°$.

In Fig. 6b, the solid curve with circles and the thick dashed curve were directly converted from those in Fig. 6a through (4). With BD_0=1, a combination of (3) and (4) results in the dashed curve in Fig. 6b as:

$$PSBE = \frac{0.06 * d_{10}}{\sqrt{BD_{10}}} + 0.075\,. \qquad (5)$$

Fig. 6b shows a very rapid (near exponential) bias reduction for $BD_{10} < 3$, and levels off thereafter. Thus $BD_{10} < 3$ can be defined as the optimal bias reduction range. The optimal buoy density for bias reduction can be defined between 2 and 3 in a 10° grid because it is the buoy density at the end of the rapid bias reduction.

5. Equivalent buoy density (EBD) and new buoy requirement

In this section, the data density of the past and present in situ (ship and buoy) network is evaluated to determine where more buoys are needed to meet a desired SST accuracy. On climate scales, Needler et al. (1999) suggested a SST accuracy of 0.2-0.5^{o}C for satellite bias correction on a 500 km grid and on a weekly time scale. This suggestion was endorsed by the World Meteorological Organization (WMO, http://www.wmo.ch/web/gcos/gcoshome.html). Because satellite biases do not change greatly from weekly to monthly periods and because a 5° latitude-longitude box is close to 500 km box (only 10% larger at the equator), the minimal bias accuracy used here is 0.5^{o}C on a monthly 5° grid. This modification is for the convenience of computation and to simplify buoy deployment plans.

According to Figs. 6a and 6b, to reduce a 2°C bias to 0.5°C (the minimum desired accuracy mentioned above), the needed buoy grid resolution is about 7° and the needed buoy density is about 2 per 10° grid box. This coincides with the optimal buoy density for bias reduction defined in section 4. As a reminder, satellite biases of 2°C have occurred during extreme conditions (large volcano eruptions). However, under normal conditions, the magnitudes of satellite SST biases are of 0.5°C to 1°C, thus better than 0.5°C SST accuracy are achieved under the recommended buoy density (2 per 10° grid box). For example, simulations for biases with a global maximum of 1°C resulted in a residual bias of 0.32°C with a buoy density of 2 per 10° grid box. To achieve a better accuracy (e.g., 0.2°C), more buoys are needed or reprocessing of the satellite data are necessary to reduce satellite biases.

Observations from ships and buoys are combined according to their random noise levels. Because real time error information is usually not available from routine ship and buoy observations, the typical errors from statistical studies (section 1) are used. Because ship observations are noisier (estimated random error of 1.3°C) than buoy observations (estimated random error of 0.5°C), roughly 7 ship observations are required to have the same accuracy of one buoy observation (from $\frac{1.3}{\sqrt{n}} = 0.5$). Therefore an equivalent-buoy-density (*EBD*) is defined as

$$EBD = n_b + \frac{n_s}{7} \tag{7}$$

where n_b and n_s are the number of observations from buoys and ships in a 10° box, respectively.

The *EBD* was defined for each month, and then was averaged seasonally for operational buoy deployment. An example is shown in Fig. 7 for October – December 2003. As the focus is now on the open ocean, boxes poleward of 60°N and 60°S were not shown along with boxes with less than 50% ocean by area and boxes in Hudson Bay and the Mediterranean Sea. Color shading is used in the figure to indicate where and how many additional buoys are needed to reach the initial target of 2 buoys per 10°x°10 box. Note that this is a nowcast system. A forecast system would require reliable information on ocean surface currents. After more buoys are deployed, especially in the Southern Ocean, future improvements could include predictions of drifting-buoy tracks from statistical models or from dynamic models employing data assimilation.

The number of additional buoys that would be needed to reach an initial *EBD* of 2 for all shaded boxes in Fig. 7 has been computed and is plotted in Fig. 8 for four latitudinal bands. The number of buoys needed in the middle latitude Southern Hemisphere (60°S-20°S) shows a rapid drop with time in the mid 1990s due to the increase in the number of buoys deployed then (Reynolds et al. 2002). For the three-month average shown in Fig. 7, 189 additional buoys are needed between 60°N-60°S, of which 102 are needed between 60°S-20°S, 65 between 20°S-20°N, and 22 between 20°N-60°N.

Using the results shown in Figs. 6b and 7, it is possible to obtain the performance of the in situ observational system for SST. This is done using the PSBE as a function of BD_{10} as shown in Fig 6b (the solid line), where BD_{10} is defined by monthly values of EBD, similar to the seasonal values shown in Fig. 7. The individual PSBE values can then be averaged spatially. Fig. 9 shows the PSBE for four zonal bands from January

1990 to December 2003. Again the impact of additional drifting buoys in the mid 1990s shows a drop in the bias error with time, especially in the middle latitude Southern Hemisphere Ocean (20°S-60°S). In the middle latitude Northern Hemisphere Ocean (60°N -20°N), the EBD from the current in situ network is typically dense enough to correct potential satellite SST biases to an initial accuracy of 0.5^{o}C. The global (60°S-60°N) and tropical (20°S-20°N) averages of the potential satellite bias error are roughly 0.6^{o}C at the end of 2003.

6. Summary and discussions

The results presented here provide an objective method to determine the minimum in situ network required for SST analyses for climate in the satellite era. For this network it has been assumed that satellite data are available and that these data provide adequate coverage of the ocean on 5° spatial and monthly time scales. The purpose of the in situ network is to allow large satellite biases to be corrected to a required accuracy. Because satellite biases cannot be predicted, satellite biases in the NOAA/U.S. Navy operational AVHRR SST over the last 20 years were examined to determine bias magnitudes and typical spatial patterns. A worse case scenario was defined to have a global maximum bias of 2°C. The required in situ network is to reduce biases of this magnitude to below 0.5^{o}C, an initial target accuracy.

Through simulations it was found that the residual potential satellite bias error (PSBE) and in situ data density have a near exponential relationship. Thus once certainty accuracy is achieved, considerably more in situ data are needed to achieve even a small

improvement in accuracy. Fig. 6b shows that a buoy density of more than 3 in a 10° box does not reduce the bias error significantly for the biases presented here.

The simulations also showed that at least two buoys or buoy equivalents are needed in 10° boxes to reduce a 2^oC satellite bias to below 0.5^oC. The required buoy density of 2 is near the end of the rapid error reduction range, thus it may be considered as an optimal buoy density. It is important to point out that under normal conditions (e.g., without a major volcanic eruption or a satellite equipment failure), the magnitudes of satellite biases are of 0.5^oC to 1^oC, and the residual bias error can be reduced to about 0.3^oC with the buoy density of 2 per 10° box.

An equivalent buoy density (EBD) has been defined to combine ship and buoy observations according to their typical observational random errors. Once this is done, the spatial maps of the EBD were computed to determine where additional buoys would be needed to bring the EBD to 2 as an initial requirement. The simulations also allowed a PSBE to be defined as a function of EBD. This relationship allows an average PSBE to be computed with time to monitor the accuracy of the current in situ network for SST.

The EBD assumes typical ship observational random error. Recent work by Kent and Taylor (2004) and Kent and Challenor (2004) have begun to identify better ship SST observations. Once this work is completed, the better ship observations could be separated from the rest ship observations and then a new version of (7) could be developed to redefine the EBD.

The past in situ observation network was not specifically designed for climate SST and thus it was not necessarily the most efficient network for climate SST in the satellite era. For example, the EBDs exceeded 5 in most of the North Atlantic Ocean (see

Fig. 7), while the EBDs are less than 2 in a large number of boxes in the Southern Oceans. For climate SST purposes alone, a new in situ data distribution should be achieved.

These results have already had an influence on future buoy deployments. The NOAA Atlantic Oceanographic and Meteorological Laboratory (AOML) is now using seasonal figures like Fig. 7 to guide surface drifting buoy deployments. The presently designed buoy need network is a nowcast system. It will require reliable information on the global ocean surface circulation to make it a forecast system. It is hoped that this study is a step toward objectively defining requirements for an integrated global ocean observing system. Although this study is specific for SST, the methodology presented here can be adopted for other parameters that are measured from both in situ and satellite networks. Such parameters may include sea-surface wind, sea-level pressure, sea-surface current, and in the future sea-surface salinity. The integrated observing system should also consider an integrated sensor system for multiple parameters. Future improvements of the work presented here could include farther study of the bias patterns associated with other satellite SSTs (TMI, ATSR, AMSR-E and MODIS) and reprocessed satellite SST, such as the most recent version of the Pathfinder AVHRR SST.

ACKNOWLEDGEMENTS:

The NOAA Office of Global Programs (OGP) and the National Climatic Data Center provided support for this project. Feedback from Michael Johnson at OGP, Ed Harrison at NOAA Pacific Marine Environmental Laboratory and Bob Molinari at AOML greatly improved the design of the network. Careful reading and many suggestions from Drs. Tom Peterson, Bomin Sun, Brian Nelson, David Parker, and Sharon LeDuc helped improve the manuscript. We thank the two anonymous reviewers

whose comments made our presentation more precise. Some of the graphs were plotted using the Grid Analysis and Display System (GrADS) developed at the Center for Ocean-Land-Atmosphere Studies (http://grads.iges.org/grads).

REFERENCES:

Davis, R. E., 1976: Predictability of sea surface temperature and sea level pressure anomalies over the North Pacific Ocean. *J. Phys. Oceanogr*., **6**, 249-266.

Emery, W. J., D. J. Baldwin, P. Schlüssel, and R. W. Reynolds, 2001: Accuracy of in situ sea surface temperatures used to calibrate infrared satellite measurements. *J. Geophys. Res.,* **106**, 2387–2405.

Esaias, W. E., and Coauthors, 1998: An overview of MODIS capabilities for ocean science observations. *IEEE Trans. Geosci. Remote Sens.,* **36**, 1250–1265.

Haywood, J. M., P. N. Francis, M. D. Glew, and J. P. Taylor, 2001: Optical properties and direct radiative effect of Saharan dust: A case study of two Saharan dust outbreaks using aircraft data. *J. Geophys. Res.,* **106**, 18417–18430.

Kagan, R. L., 1979: *Averaging Meteorological Fields*. Gidrometeoizdat, 212 pp. (in Russian).

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 Tech*., **10**, 591 - 608.

Kent, E. C., 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 Tech*., **16**, 905-914.

Kent, E. C., and P. K. Taylor, 2004: Towards estimating climatic trends in SST, Part I: Methods of measurements. *J. Atmos. & Oceanic Tech*. (submitted).

Kent, E. C., and P. G. Challenor, 2004: Towards estimating climatic trends in SST, Part II: Random errors. *J. Atmos. & Oceanic Tech*. (submitted).

Kilpatrick, K. A., G. P. Podestá, 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**, 9179-9197.

Koblinksky, C. J., and N. R. Smith (Editors), 2001: *Observing the Oceans in the 21st Century*, Published by 2001 GADOE Project Office, Bureau of Meteorology of Australia, 604pp.

Kummerow, C., W. Barnes, T. Kozu, J. Shiue, and J. Simpson, 1998: The Tropical Rainfall Measuring Mission (TRMM) sensor package. *J. Atmos. Oceanic Technol.,* **15**, 809–817.

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, 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,** 11587–11601.

Mutlow, C. T., A. M. Zavody, I. J. Barton, and D. T. Llewellyn-Jones, 1994: Sea surface temperature measurements by the along-track scanning radiometer on the *ERS-1* satellite: Early results. *J. Geophys. Res.,* **99**, 22575–22588.

Needler, G., N. Smith and A. Villwock, 1999: The action plan for GOOS/GCOS and

sustained observations for CLIVAR, *Ocean Obs '99 Proceedings*, 18-22 Oct. 1999 - Saint Raphael, France. (Paper available at http://www.bom.gov.au/OceanObs99/Papers/Needler.pdf.)

Parker, D. E., C. F. Folland, and M. Jackson, 1995: Marine surface temperature: Observed variations and data requirements. *Climatic Change*, **31**, 559-600.

Reynolds, R. W., 1993: Impact of Mount Pinatubo aerosols on satellite-derived sea surface temperatures. *J. Climate*, **6**, 768-774.

Reynolds, R. W. and T. M. Smith, 1994: Improved global sea surface temperature analyses using optimum interpolation. *J. Climate*, **7**, 929-948.

Reynolds, R. W., N. A. Rayner, T. M. Smith, D. C. Stokes and W. Wang, 2002: An improved in situ and satellite SST analysis for climate. *J. Climate*, **15**, 1609-1625

Reynolds, R.W., C. L. Gentemann, and F. Wentz, 2004: Impact of TRMM SSTs on a climate-scale SST analysis. *J. Climate*, **17**, 2938-2952.

Smith, T. M., R. W. Reynolds, C. F. Ropelewski, 1994: Optimal averaging of seasonal sea surface temperatures and associated confidence intervals (1860–1989). *J. Climate*, **7**(No. 6), 949–964.

Wentz, F. J., C. L. Gentemann, D. K. Smith, and D. B. Chelton, 2000: Satellite measurements of sea-surface temperature through clouds. *Science,* **288**, 847–850.

Wentz, F. J., C. L. Gentemann, and P. Ashcroft, 2003: On-orbit calibration of AMSR-E and the retrieval of ocean products. The 83rd Annual Meeting of the American Meteorological Society, 9-13 February 2003, Long Beach, California (available at http://ams.confex.com/ams/pdfpapers/56760.pdf)

Zhang, H.-M., R. W. Reynolds, and T. M. Smith, 2004: Bias characteristics in the

AVHRR sea surface temperature. *Geophys. Res. Lett.*, **31**, L01307, doi:10.1029/2003GL018804.

FIGURE CAPTIONS:

Fig. 1. Mean (upper panel) and standard deviation (lower panel) of the satellite AVHRR SST biases (in °C) for January 1990 through December 2002. The bias is defined as the difference between the OI SST without bias correction minus the OI SST with bias correction. Contour intervals are 0.2°C. Positive values increase towards coastlines in (a).

Fig. 2. EOF Modes 1 to 6 (top to bottom panels, respectively) of the AVHRR SST biases for January 1990 through December 2002. Spatial patterns are on the left. Time series are on the right with x-axis labeling the years. Multiplying the spatial pattern by the corresponding time amplitude results in bias in °C. The time series indicate that mode 1 represents biases associated with volcano aerosols. Mode 2 and 3 represents two different seasonal variations. Mode 4 represents a trend. Modes 5 and 6 represent interannual variations.

Fig. 3. Percent variance of the 1990-2002 AVHRR SST bias EOF modes. Only the first 25 modes are shown. The buoy need network is designed using the first six EOF modes, which account for 52.7% of the total variance (shaded).

Fig. 4. (a) Absolute values of EOF mode 1 (scaled for a maximum of 2°C). (b) Potential Satellite Bias Error (PSBE) for EOF mode 1 with a buoy grid resolution of 20°. By defintion the buoy density is one per 20^ox20^o grid box. The PSBE is the remaining bias that cannot be removed with the specified buoy grid resolution. (c) Same as (b) but for buoy grid resolution of 7°. Note different grey color scales are used in (a) than (b) and (c) for clarity.

Fig. 5. Adjusted Potential Satellite Bias Error (PSBE) as a function of buoy grid resolution for each bias regime represented by each of the first six EOF spatial patterns. As designed in the simulations, each of the scaled EOF modes represents a typical bias regime that could occur independently of each other. Thus each one has been scaled to have a global maximum of 2°C to represent a total bias magnitude. The buoy need network is designed to correct any of the biases to a required accuracy of 0.5°C or better. The results show that a buoy grid resolution of 7° or smaller is required. Modes 1 to 3 are easier to correct due to their larger spatial scales. Requirements for the regimes of modes 4 to 6 are more stringent and similar.

Fig. 6. Average of the Potential Satellite Bias Error (PSBE) for modes 4-6 of Fig. 5. (a) (left) shows the averaged PSBE as a function of the buoy grid resolution as in Fig. 5; the dashed line is a linear fit to the data. (b) (right) shows the averaged PSBE as a function of the buoy density (BD) for a 10°x10^{o} spatial grid. The dashed line is the model fit by Eq. (5). Note that the PSBE decreases rapidly as BD increases from 0 to 3, but the error reduction levels off beyond a BD of 3. The thin vertical dash line indicates where BD=2, which is the recommended minimum buoy density to reduce a 2°C bias to below 0.5°C.

Fig. 7. Seasonally (October – December 2003) averaged monthly equivalent buoy density (EBD) on a 10°x10^{o} grid. EBD includes contributions from both buoys and ships, accounting for their typical random errors. Green shading indicates where $EBD \geq 2$ which satisfies the initial requirement. Red shading indicates critical regions where $EBD < 1$ and two more buoys are needed. Yellow shading indicates $1 \leq EBD < 2$ and at least one more buoy is needed.

Fig. 8. Number of additional buoys would have been needed to bring the equivalent buoy density (EBD, Fig. 7) to 2 or more per 10° box for four zonal bands. The effect of large buoy deployment in the mid 1990s is indicated by a drop in the number of buoys needed.

Fig. 9. Simulated potential satellite bias errors (PSBEs). Shown are zonal averages over four latitude bands. The simulated PSBEs are for simulated biases with a global maximum of 2°C. The PSBE is the residual bias error that cannot be removed by the existing in situ observing network. Because the actual biases are often smaller than 2°C, the residual bias errors may be smaller than those shown.

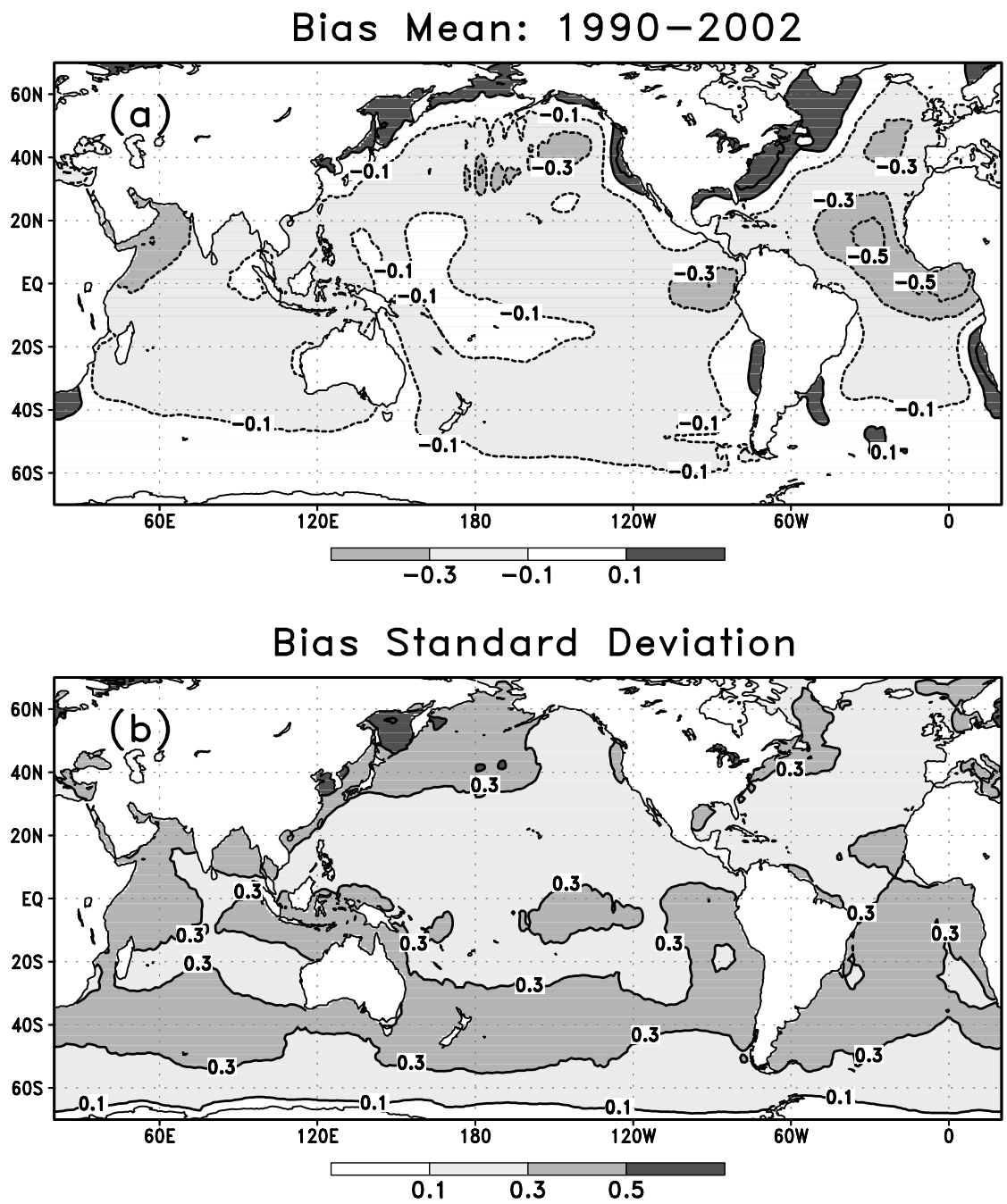
Bias Mean: 1990–2002
(a)
60N
40N
20N
EQ
20S
40S
60S
60E
120E
180
120W
60W
0
−0.1
−0.3
−0.5
0.1
−0.3
−0.1
0.1
Bias Standard Deviation
(b)
0.1
0.3
0.5

Figure 1

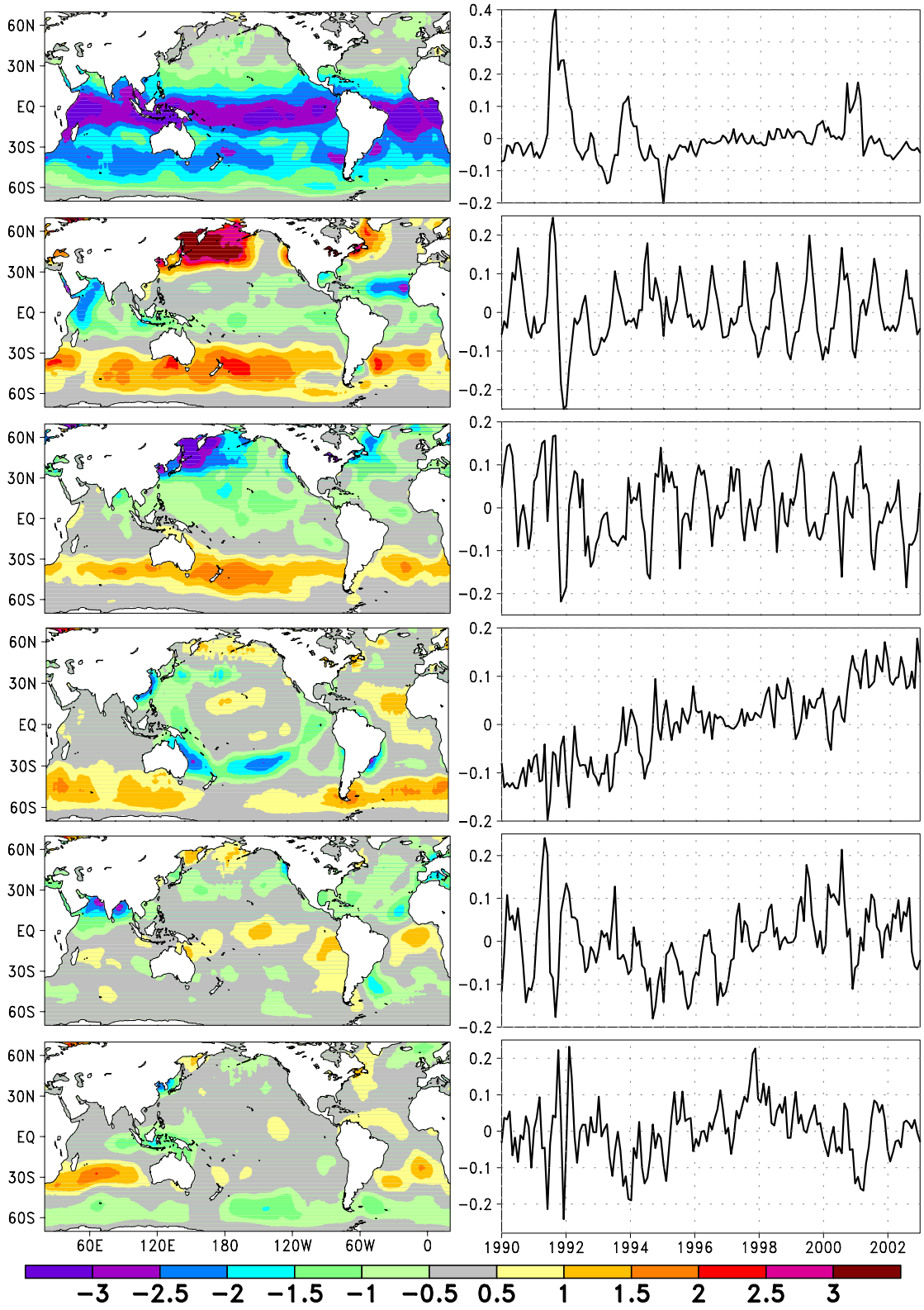
60N
30N
EQ
30S
60S
60E
120E
180
120W
60W
0
1990
1992
1994
1996
1998
2000
2002
−3
−2.5
−2
−1.5
−1
−0.5
0.5
1
1.5
2
2.5
3

Figure 2

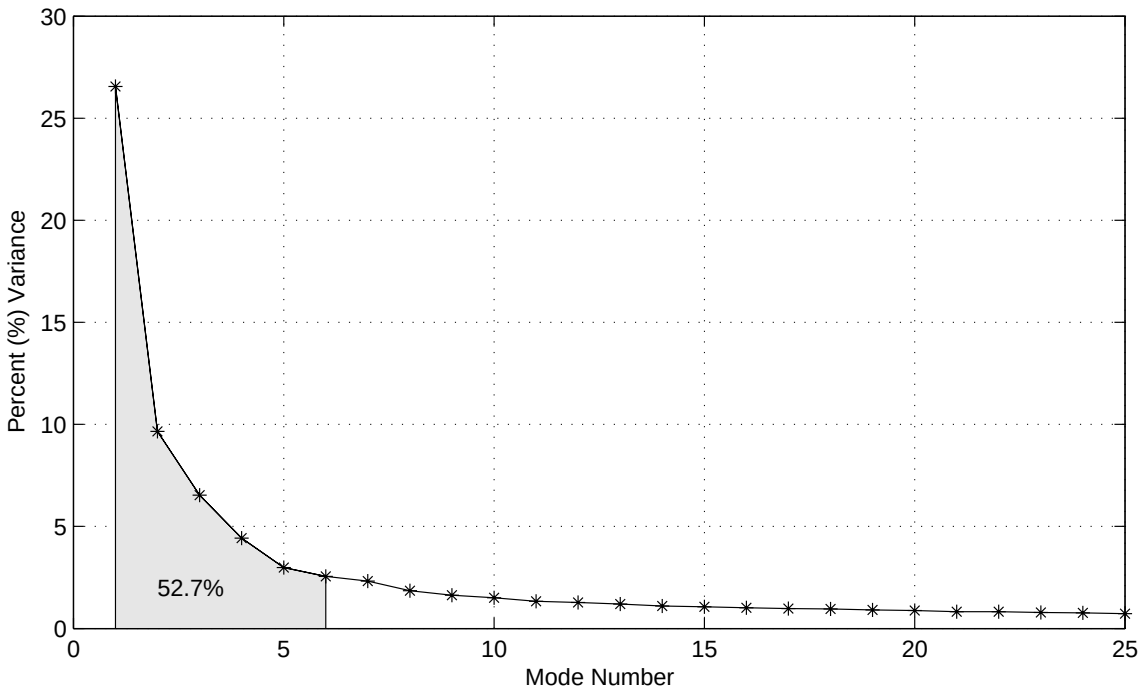
30
25
20
15
10
5
0
Percent (%) Variance
52.7%
0
5
10
15
20
25
Mode Number

Figure 3

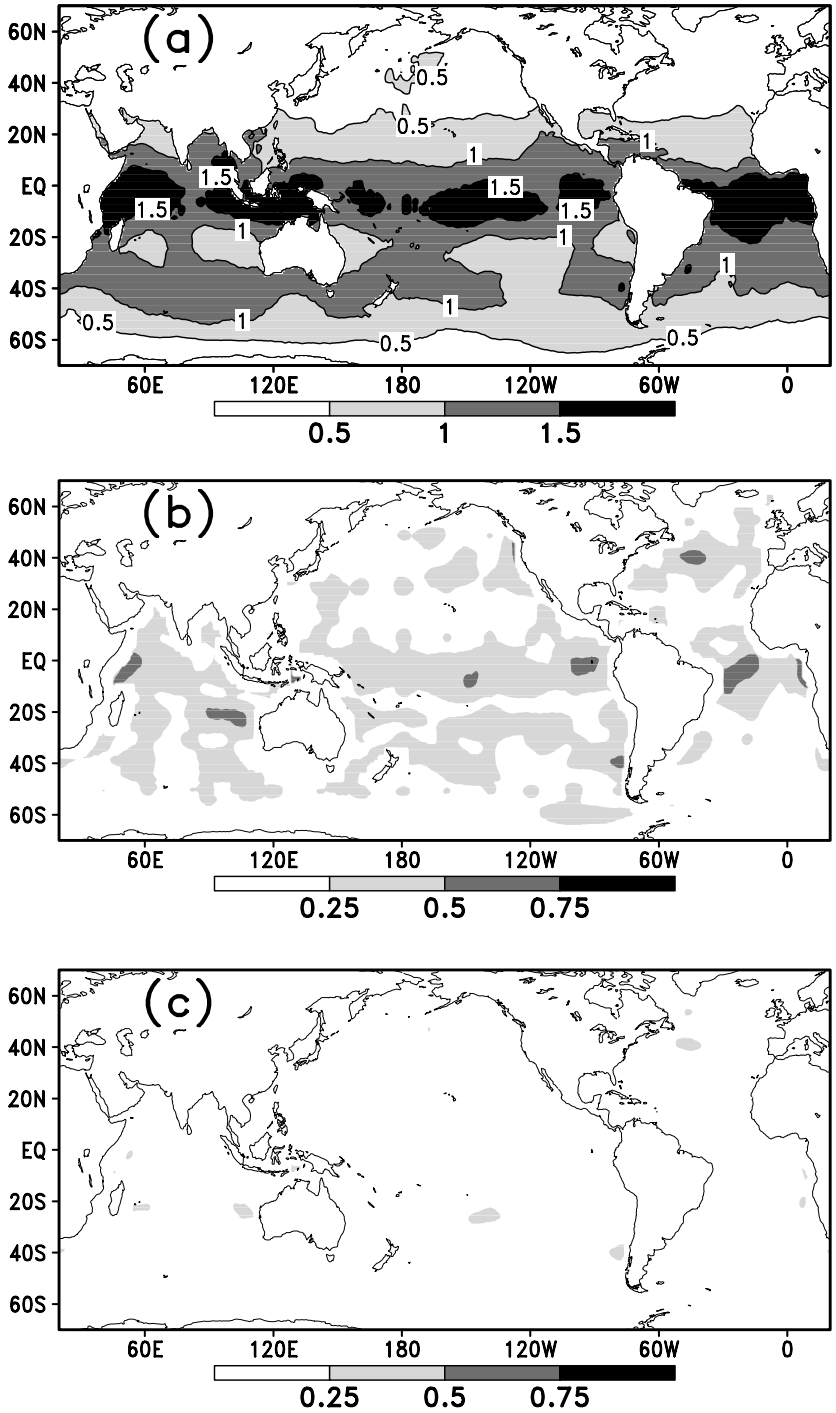
(a)
60N
40N
20N
EQ
20S
40S
60S
60E
120E
180
120W
60W
0
0.5
1
1.5
(b)
0.25
0.5
0.75
(c)
0.25
0.5
0.75

Figure 4

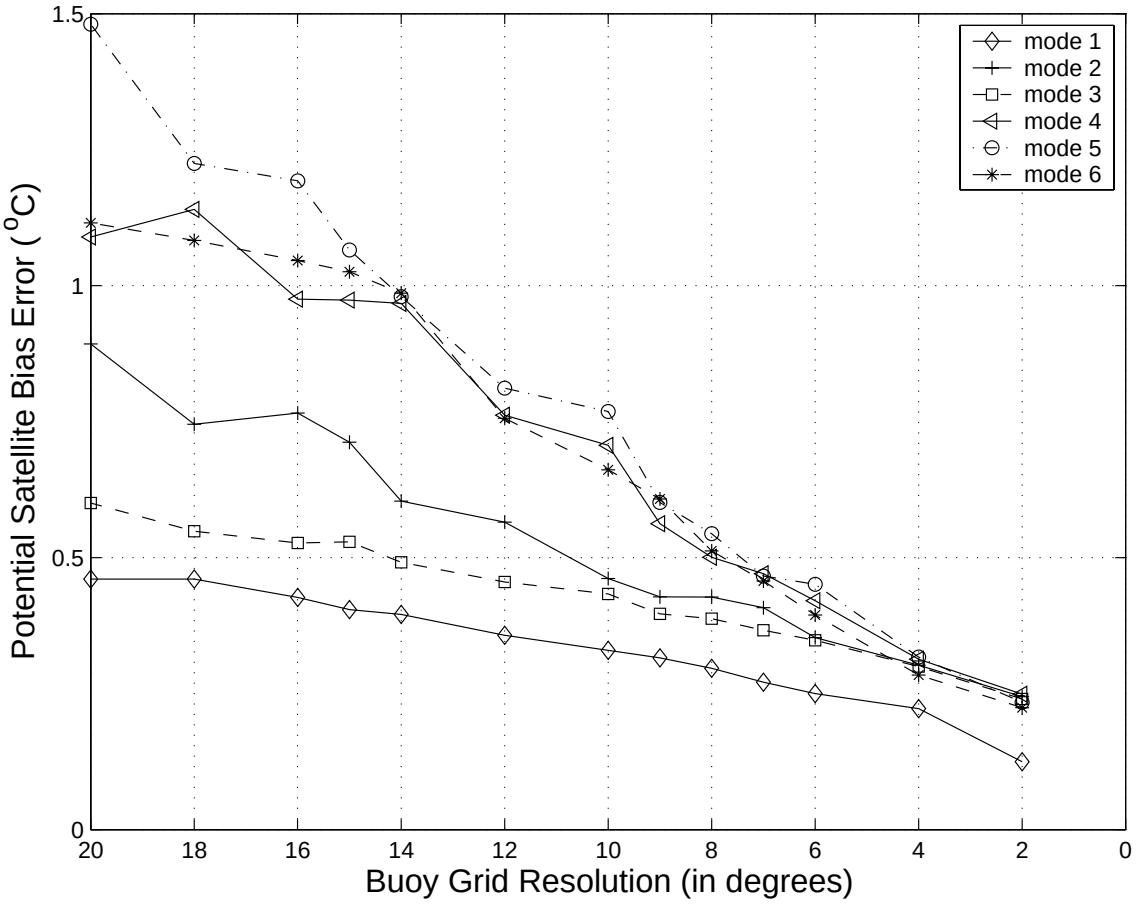
mode 1
mode 2
mode 3
mode 4
mode 5
mode 6
Potential Satellite Bias Error (^{o}C)
Buoy Grid Resolution (in degrees)
1.5
1
0.5
0
20
18
16
14
12
10
8
6
4
2
0

Figure 5

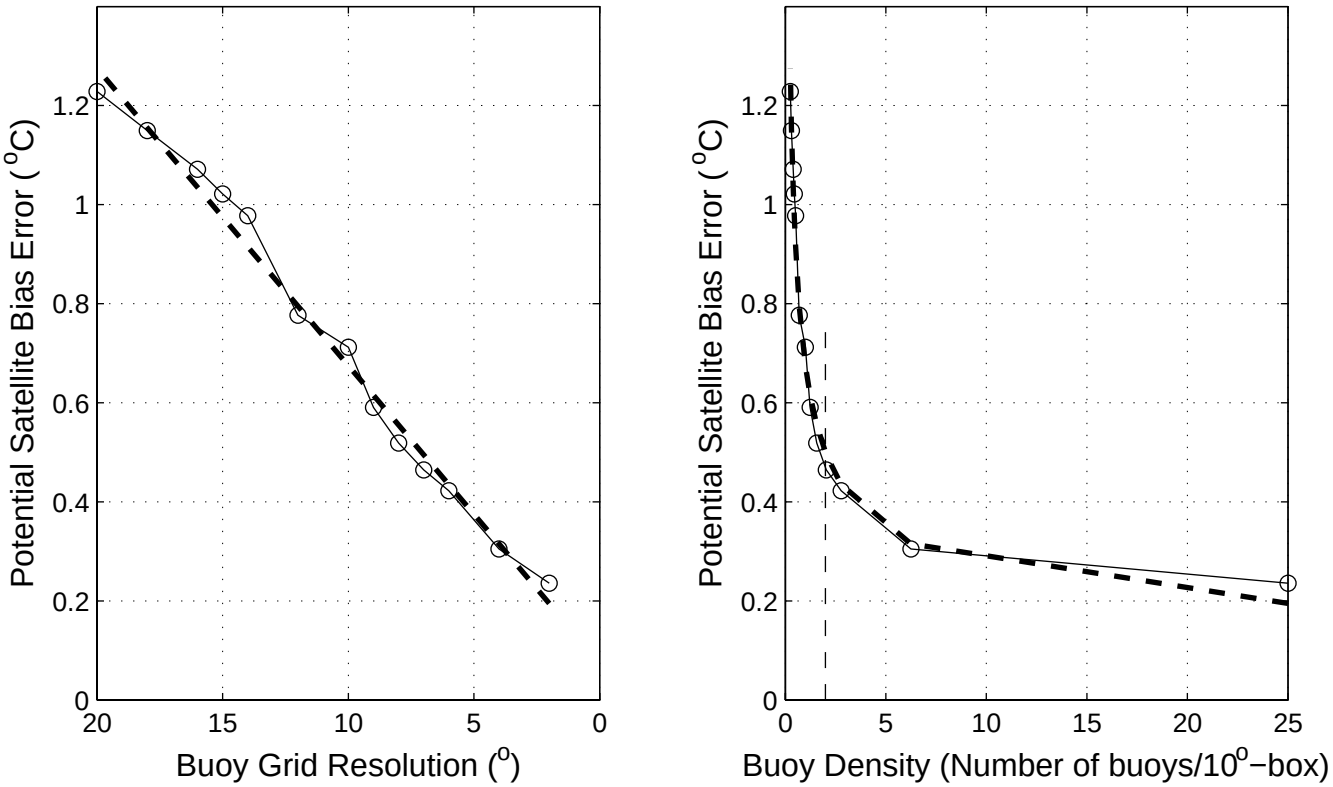
Potential Satellite Bias Error (^{o}C)
Buoy Grid Resolution (^{o})
Potential Satellite Bias Error (^{o}C)
Buoy Density (Number of buoys/10°−box)

Figure 6

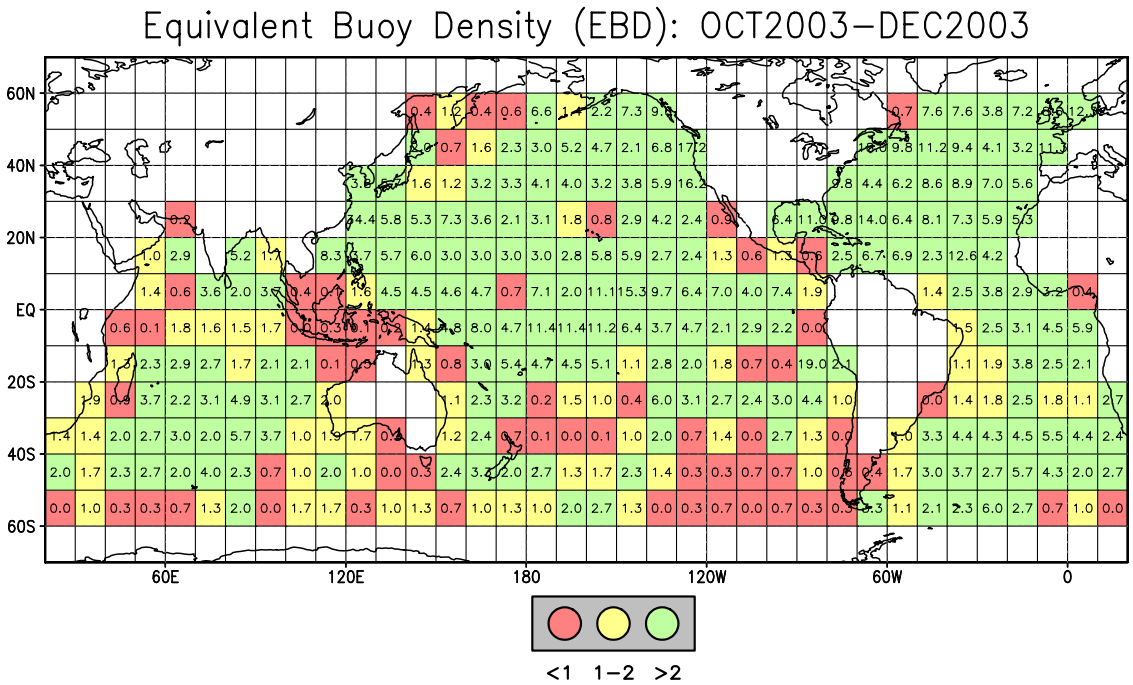
Equivalent Buoy Density (EBD): OCT2003–DEC2003
60N
40N
20N
EQ
20S
40S
60S
60E
120E
180
120W
60W
0
<1
1–2
>2

Figure 7

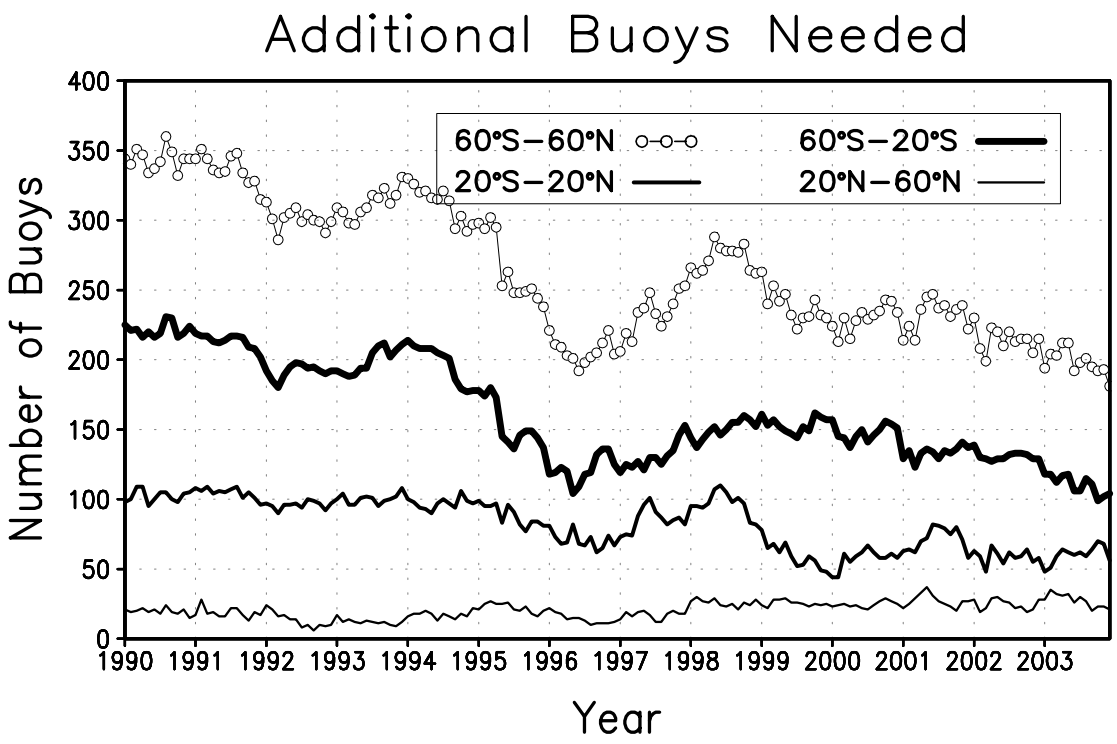
Additional Buoys Needed
60°S–60°N
60°S–20°S
20°S–20°N
20°N–60°N
Number of Buoys
400
350
300
250
200
150
100
50
0
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
Year

Figure 8

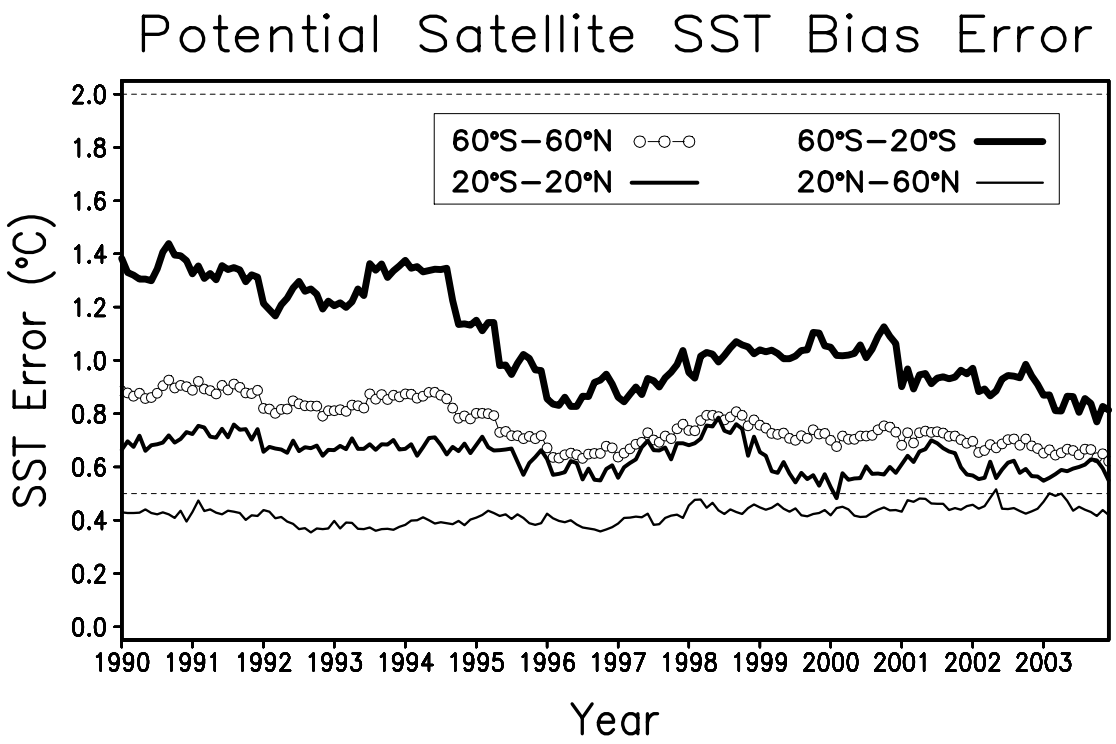
Potential Satellite SST Bias Error
60°S–60°N
60°S–20°S
20°S–20°N
20°N–60°N
SST Error (°C)
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003
Year

Figure 9