

**Satellite and Skin Layer Effects on the Accuracy of Sea Surface Temperature
Measurements from the GOES Satellites**

Gary A. Wick¹, John J. Bates¹, and Donna J. Scott²

¹NOAA Environmental Technology Laboratory, Boulder, CO

²CIRES/ETL, University of Colorado, Boulder

Corresponding author address: Gary A. Wick, NOAA Environmental Technology
Laboratory, 325 Broadway, R/ET1, Boulder, CO 80305.

E-mail: Gary.A.Wick@noaa.gov

ABSTRACT

The latest Geostationary Operational Environmental Satellites (GOES) have facilitated significant improvements in our ability to measure sea surface temperature (SST) from geostationary satellites. Nonetheless, difficulties associated with sensor calibration and oceanic near-surface temperature gradients affect the accuracy of the measurements and our ability to estimate and interpret the diurnal cycle of the bulk SST. Overall, measurements of SST from the GOES Imagers on the GOES 8-10 satellites are shown to have very small bias (< 0.02 K) and rms differences of between 0.6 and 0.9 K relative to buoy observations. Separate consideration of individual measurement times, however, demonstrates systematic bias variations of over 0.6 K with measurement hour. These bias variations significantly affect both the amplitude and shape of estimates of the diurnal SST cycle. Modeled estimates of the temperature difference across the oceanic cool skin and diurnal thermocline show that bias variations up to 0.3 K can result from variability in the near-surface layer. Oceanic near-surface layer and known “satellite midnight” calibration effects, however, explain only a portion of the observed bias variations, suggesting other possible calibration concerns. Methods of explicitly incorporating skin layer and diurnal thermocline effects in satellite bulk SST measurements were explored in an effort to further improve the measurement accuracy. While the approaches contain more complete physics, they do not yet significantly improve the accuracy of bulk SST measurements due to remaining uncertainties in the temperature difference across the near-surface layer.

1. Introduction

Sea surface temperature (SST) is a fundamental geophysical variable with important applications to many problems of interest. Satellite measurements have been used extensively to produce detailed maps of SST on scales ranging from regional to global. Regional maps have important applications to fisheries and studies of circulation and frontal features (e.g., Njoku et al. 1985; Ullman and Cornillon 1999) while global maps are used for climate monitoring, as inputs to atmospheric models, and as a diagnostic tool for oceanographic models (e.g. Reynolds and Smith 1994). Polar orbiting satellite sensors including the Advanced Very High Resolution Radiometer (AVHRR) on the NOAA series of satellites and the Along Track Scanning Radiometer (ATSR) on the European Remote Sensing Satellites currently provide global infrared measurements of SST at high resolution (~ 1 km) and high accuracy (0.3 – 0.6 K rms) approximately twice daily (more frequently at high latitudes) (see, e.g., Barton et al. 1995; Smith et al. 1996; May et al. 1998; Merchant and Harris 1999). Geostationary satellites provide the added potential for measuring SST over more limited areas at a much greater frequency. The primary motivation for measuring SST from geostationary satellites is that these measurements can provide a greater number of cloud-free observations and complete resolution of the diurnal cycle of SST. The ability to accurately measure the diurnal cycle of SST is of specific importance for applications such as remotely measuring the air-sea heat flux at three-to-six hourly resolution (Clayson and Curry 1996).

The recent launch of the latest generation of Geostationary Operational Environmental Satellites (GOES) beginning with GOES-8 has facilitated significant improvements in the accuracy of SST measurements from geostationary satellites. The accuracy was limited in the past by coarse channel digitization levels, single infrared imager channels, poor spatial

resolution, and sensor acquisition schedules (Maul 1981; Bates and Smith 1985). Bates and Smith (1985) were able to obtain SST measurements with a scatter of between 0.8 and 1.0 K on spatial scales of 8-16 km. The GOES Imagers on the new GOES satellites have infrared imager channels centered at approximately 10.7 and 12.0 μm with 10-bit resolution and 4-km spatial resolution (Menzel and Purdom 1994). These channels provide a “split-window” capability like that commonly used in SST measurements from the AVHRR and ATSR. Various groups have used these data to produce improved maps of SST from the GOES satellites (Legeckis and Zhu 1997; May and Osterman 1998; Wu et al. 1999). The rms accuracy of these products ranges between 0.6 and 0.9 K depending on the satellite and algorithm.

An important problem affecting the interpretation and validation of satellite measurements of SST from both polar orbiting and geostationary satellites is the presence of the oceanic skin layer (see, e.g., Katsaros 1980; Robinson et al. 1984; and Schlüssel et al. 1990) and diurnal thermocline. While infrared radiometers at 10-12 μm measure the temperature of the ocean at depths only on the order of 10 μm , the buoy measurements commonly used to derive and validate the satellite algorithms are made at depths on the order of 1 m. Models and measurements of the near-surface temperature profile have demonstrated the potential for significant variability in the temperature difference between these depths especially during the daytime at low winds and with strong solar heating (Wick 1995, Fairall et al. 1996; Soloviev and Schlüssel 1996; Webster et al. 1996; Donlon et al. 1999). Due to interest in the upper ocean heat content and for consistency with historical SST measurements, the “bulk” temperature beneath the skin layer is frequently the temperature of interest rather than the “skin” temperature actually measured by the satellite. Algorithms that attempt to estimate the bulk temperature by directly regressing the satellite brightness temperature measurements against in situ bulk temperature

measurements are subject to errors resulting from variability in the difference between the skin and bulk temperatures. This problem is very important when measuring SST from geostationary satellites since measurements are collected regularly through periods when the difference can vary strongly. All the geostationary SST algorithms cited above were derived either directly or indirectly from comparisons with buoys and thus are subject to this source of error.

Other accuracy problems can result from measuring SST from geostationary satellites. These include the possibility of extreme satellite viewing angles, variable heating and solar illumination of the satellite sensor, and nonuniform sampling of all possible surface conditions. Since the satellite position relative to the earth's surface remains constant, the sun will illuminate and heat the satellite from many different directions depending on the season and time of day. Nonuniform sampling occurs because midlatitude conditions are sampled only over a small range of satellite viewing angles while the equatorial region is sampled at angles between 0° and near 55° .

In this paper, we examine how these factors affect the accuracy of SST measurements from the latest generation of GOES satellites. The various effects can lead to systematic biases with time of day that are not revealed clearly in statistics computed from all the measurements. Our primary purpose is not to present a new GOES SST algorithm, but rather to describe sources of error that may be present in all the products. Knowledge of these errors is important for proper use of the products, especially for interpretation of the diurnal SST cycle. We will examine errors related to the presence of the skin layer and diurnal thermocline and study whether consideration of the near-surface temperature profile in satellite measurement techniques can lead to improvements in the accuracy of bulk SST estimates.

We begin in section 2 by describing GOES bulk SST algorithms derived for the GOES-8, -9 and -10 satellites using traditional techniques of regression against buoy measurements. In section 3 we look in detail at variations in the measurement accuracy on both daily and seasonal scales. We further study the theoretical impact that the skin layer is expected to have on measurement accuracy and compare these results with the observed variability. In section 4 we then demonstrate the impact of these bias variations on estimates of the SST diurnal cycle. Finally, in section 5 we apply different techniques to evaluate whether consideration of the skin layer can improve the accuracy of SST estimates in practice.

2. GOES SST measurements

We derived new algorithms for computing SST from the GOES Imager on the GOES-8, -9, and -10 satellites. GOES-8 has functioned as the eastern satellite (positioned at 75°W) since its launch in 1994. GOES-9 operated as the western satellite (135°W) from its launch in 1995 until it failed in July 1998. Following the failure of GOES-9, GOES-10 was positioned at 135°W as the new western satellite. The NOAA Forecast Systems Laboratory (FSL) regularly provides us with portions of the full disk scan taken every three hours from each satellite. Our data subset from the western satellites extends from approximately 45°N to 25°S and 90° to 180°W . Our subset from GOES-8 covers from 45°N to 25°S and 25° to 110°W . While the original resolution of the GOES infrared data is 4 km, our data were subsampled at 8-km resolution. We have data from GOES-9 from June 1997 through July 1998, and GOES-8 and -10 from August 1998 through the present. The characteristics of the GOES Imager are summarized in Table 1 (Komajda and McKenzie 1994, Ellrod et al. 1998).

The data were provided in the retransmitted GOES VARiable (GVAR) format. The data stream contains scaled counts that we then transformed to temperatures and geolocated. Our

conversion from scaled counts to radiance and temperature followed the recommended procedure (Weinreb et al. 1997) and did not introduce any new information. Conversion of the raw satellite data into the GVAR format by the NOAA National Environmental Satellite Data and Information Service (NESDIS) takes into account known effects on calibration such as variations in the scan mirror emissivity with viewing angle.

a. Cloud detection

Before SST estimates can be derived from infrared satellite data, the data must be carefully screened for cloud contamination. Our method evolved from those used by Olesen and Grassl (1985), Saunders and Kriebel (1988), and May and Ostermann (1998). We use a combination of several single-channel and channel difference threshold tests and visible and infrared spatial uniformity tests. Separate tests are applied during the day and night depending on the effectiveness of the visible and 3.9- μm channels. During twilight (solar zenith angles between 75° and 90°) we rejected all the data due to the possible ineffectiveness of both the visible and 3.9- μm channels for adequate cloud screening. The effects of solar radiance on the 3.9- μm channel at solar zenith angles slightly greater than 90° were found to be negligible. The thresholds were all chosen to exclude extra cloud at the risk of excluding valid surface data. A flow chart summarizing our procedure is shown in Figure 1.

An initial threshold test is made against the 11- μm brightness temperature to flag pixels with temperatures too cold for realistic SST values. A value of 12°C was chosen for our geographic subset though this may exclude a few valid measurements at the northern boundary. A second test is applied to the difference between the 11- and 12- μm brightness temperatures with both upper and lower thresholds. The upper threshold test is based on the approach of Saunders and Kriebel (1988) and is used to identify thin cirrus. The lower threshold was

introduced to mimic the tests with the 3.9- μm channel described below and was found to be effective in areas of uniform cloud cover. The final threshold values were obtained through visual examination of the effectiveness of the cloud removal.

During the daytime visible data are applied to threshold and spatial uniformity tests to exclude reflectances too large for the sea surface and large variability typically associated with clouds. The reflectance of the sea surface is very small and uniform at the moderate viewing angles considered. The maximum albedo of 6.5% was derived using a theoretical model of sea surface reflectance (Payne 1972) and the uniformity limit was obtained through visual inspection. During the nighttime upper and lower thresholds are applied to the 3.9–11 μm and 3.9–12 μm brightness temperature differences. These tests were taken from the schemes of Olesen and Grassl (1985) and Saunders and Kriebel (1988). Lower thresholds are used to detect fog and low stratus and the upper thresholds help identify sub-pixel, semi-transparent, and most medium- and high-level clouds. The threshold values were again modified based on visual inspection of the resulting images. All remaining pixels during both the day and night are then tested for spatial uniformity in the 11- μm brightness temperature since large temperature gradients most commonly correspond to cloudy regions or cloud boundaries.

Following these tests some obviously cloud contaminated pixels occasionally remained especially in regions of frequent stratus clouds. Two additional tests were performed to detect these cases. A local histogram was constructed from the 11- μm brightness temperatures within a 100x100 pixel array in an attempt to find pixels markedly cooler than their surroundings. If a local minimum was observed below an upper peak corresponding to clear surface values, brightness temperatures below this minimum were rejected. Finally, computed SST values were compared against a continuously updated field of recent clear nighttime SST retrievals. Values

significantly cooler than recent observations were considered influenced by cloud. The threshold was selected so as to allow for normal seasonal cooling. All threshold values used are shown in Figure 1.

b. Algorithm derivation

The primary function of an infrared satellite SST algorithm is to correct for absorption and reemission effects of atmospheric water vapor. Absorption due to other atmospheric gases (including carbon dioxide and ozone) and aerosols also hinder the retrieval of SST. The effects of aerosols, however, are not well accounted for in most traditional algorithms. The majority of the SST algorithms have been based on the split-window multichannel technique where the atmospheric correction is obtained using measurements subject to different amounts of absorption at different frequencies (Anding and Kauth 1970; Deschamps and Phulpin 1980; McMillin and Crosby 1984). The original operational SST product of NOAA using the AVHRR was called the Multi-channel SST or MCSST (McClain et al. 1985) and used a linear combination of the 11- and 12- μm brightness temperatures. Since then, several refinements to the algorithm have been explored including nonlinear terms and explicit incorporation of water vapor measurements (Emery et al. 1994; Barton 1995; Walton et al. 1998). Further improvements to the atmospheric corrections were possible for the ATSR using both multiple frequencies and two different viewing geometries (e.g. Mutlow et al. 1994). Specific attempts to better account for aerosols with the ATSR were introduced by Brown et al. (1997). The algorithm we used closely resembles the current operational MCSST form (May et al. 1998). This form was chosen for its simplicity and consistency with other existing GOES SST algorithms (Legeckis and Zhu 1997; May and Osterman 1998; Wu et al. 1999).

The algorithm coefficients can be derived in a variety of ways depending on the desired SST product. Procedures to estimate the subsurface or bulk temperature commonly obtain the coefficients through regression against direct measurements of the bulk temperature by drifting buoys. Skin SST algorithms, in contrast, are normally derived through the regression of SST values against brightness temperatures simulated using a forward radiative transfer model. The use of models in the generation of skin SST products has been necessary because of a lack of continuous, global *in situ* skin SST measurements. Both skin and bulk SST products are useful and the product choice depends on the application. In this work we produce a bulk SST product for consistency with the existing GOES SST algorithms and to evaluate the impact of the skin layer on bulk SST estimates. Our initial coefficients were obtained through direct regression against drifting and moored buoy measurements.

For each satellite, matches were constructed between satellite and buoy measurements that coincided within 1.5 hr and 25 km. These limits were chosen for consistency with those used to evaluate the operational MCSST algorithms (4 hr and 25 km) (May et al. 1998). The finer time limit was used to ensure that a buoy measurement was matched with only one GOES image. No attempt was made to exclude regions of strong SST gradients and spatial sampling differences remain a problem in interpreting the matches.

The drifting and moored buoy SST measurements were obtained from the NOAA National Center for Environmental Prediction (NCEP). We performed additional quality control on the buoy data including removal of periods exhibiting excessive noise in the data and values outside the range of physically expected temperatures. The unrealistic values could have resulted from sensor malfunctions or data transmission errors. Where multiple clear satellite pixels coincided with the same buoy measurement, we constructed one match using the median

value of the clear satellite brightness temperatures rather than considering several matches with the same buoy. This helped to reduce any noise in the brightness temperature measurements and the possible effects of undetected cloud contamination.

We compiled matches between June 1997 and July 1998 for GOES-9 and between August 1998 and December 1999 for GOES-8 and GOES-10. Accumulation of matches over these periods and our entire GOES coverage region enabled us to obtain matches over all seasons and a wide range of satellite viewing angles. Each set of matches was then split into independent training and testing sets consisting of every other match.

Significant scan-to-scan variations in the brightness temperature were occasionally observed from each satellite. This problem has been noted previously (Baucom and Weinreb 1996; Ellrod et al. 1998) and its effect on SST retrievals has been discussed (May and Osterman 1998; Legeckis and Zhu 1997). The variations are especially important for full resolution GOES data where different infrared detectors are used on alternating scan lines. To account for these variations, the brightness temperatures were averaged over a 5-line by 3-pixel array (40 km x 24 km) prior to the computation of the SST. The rms errors in the SST measurements increased by 0.1 – 0.2 K if individual brightness temperatures were used instead but the biases were not significantly affected. Comparisons with and without the averaging suggested that the averaging did not influence the conclusions of this paper.

We explored several different algorithms based on the MCSST form. Our best overall results were obtained when the GOES data were regressed to the buoy measurements using an algorithm of the form:

$$\text{SST} = a T_{11} + b (T_{11} - T_{12}) + c (T_{11} - T_{12}) (\sec \theta - 1) + d, \quad (1)$$

where T_{11} and T_{12} are the GOES brightness temperatures in Kelvin at 10.7 and 12.0 μm respectively, θ is the satellite zenith angle, and a , b , c , and d are the regression coefficients. The coefficients for the best fit to the training data are shown in Table 2.

c. Overall accuracy assessment

The overall statistics for the application of our GOES SST algorithms to the independent testing sets are shown in Table 3. These results were obtained by considering all the available measurements in the respective sets. For GOES-9 and -10 the rms accuracy is approximately 0.6 K relative to the buoy measurements. The results are slightly worse for GOES-8 with an rms accuracy of near 0.8 K. This is consistent with the higher noise level of the GOES-8 imager (May and Osterman 1998). For all satellites, the bias (satellite – buoy) in the measurements is very small.

Failure to perform the additional buoy quality control resulted in rms differences that were approximately 0.1 to 0.2 K higher than those in the table. This illustrates the potentially significant contribution of the buoy accuracy to the apparent satellite measurement accuracy. This result is consistent with recent findings by Emery et al. (2000) that indicate that the buoy measurement uncertainty may be as much as 0.3 – 0.5 K if no additional quality control is applied. The larger values observed by Emery et al. (2000) suggest that uncertainty in the buoy measurements is still an important factor in the buoy-satellite differences even with the quality control.

The statistics computed using all the data are encouraging and are consistent with the accuracy obtained by May and Osterman (1998) and Wu et al. (1999). Inclusion of a quadratic term for $T_{11} - T_{12}$ such as that of Wu et al. (1999) did not lead to an improvement in the overall

accuracy. Dropping the $T_{11} - T_{12}$ multiplier in the $\sec \theta$ term as did May and Osterman (1998) led to an increase in the rms differences on the order of hundredths of a degree.

3. Measured SST bias variations

a. Seasonal variations

More detailed studies, however, suggest important limitations on the measurement accuracy that are not readily apparent from the statistics computed from all the data. We first determined if there were any seasonal variations in the SST accuracy by computing the monthly mean bias and rms difference for each satellite for several individual months. These results are shown in Table 4. No data were available for GOES-10 in May 1999 due to a relocation of the NOAA FSL receiving antennas.

While there is no prominent seasonal variation in the GOES-9 and -10 results, the GOES-8 results exhibit notable seasonal changes during the period analyzed. The SST values measured by GOES-8 tend to be biased warm in the boreal winter and cold in the boreal summer. This is likely related to the sampling of GOES-8 where a significant majority of the ocean surface pixels are located in the Northern Hemisphere. The algorithm may not reproduce the full range of seasonal SST variability and cloudiness variations might also have an effect. There was a small increase in the amount of detected cloud from GOES-8 during May through July of 1999 and a similar increase in undetected clouds would lead to a cold bias in the satellite measurements. An increase in undetected cloud would also contribute to the increase in the rms difference observed at the same times. While additional seasonal variations could result from changing solar illumination geometries, this does not appear to be significant factor because there is not a large seasonal change in other effects known to be related to solar illumination of the spacecraft nor are there similar variations for the other satellites. Both the bias and rms difference varied little

over the period of study for GOES-9. The SST values measured by GOES-10 show some slow bias variations over time, but they do not appear to be correlated with season.

b. Hourly variations

The most notable accuracy concern becomes apparent when the statistics are computed separately for each time of day at which the satellite measurements are obtained. The average SST bias and rms difference relative to the buoys are plotted as a function of local solar time at the satellite subpoint for all the satellites in Figure 2 and the statistics are summarized in Table 5. Midnight (0000) LST corresponds to 0600 UTC for GOES-8 and 0900 UTC for GOES-9 and -10. All satellites are seen to exhibit strong diurnal variations in the bias with maximum peak-to-peak amplitudes greater than 0.6 K. While the rms differences are as large, the bias variations are repeatable between satellites and over different averaging periods. Variations of this magnitude will clearly affect the application of the SST product, especially for estimating the diurnal SST cycle.

There are many possible causes of the bias variations. Variations with time of day are possible due to the position of the sun relative to both the measurement point on the earth and to the satellite itself. It is necessary to distinguish the measurement time from the local surface solar time at the measurement location. Because of the large region viewed by the geostationary satellites, a large range of local surface solar times is viewed at each measurement time. A grouping of measurements at a given local surface solar time will be comprised of measurements taken at several different UTC times. Individual treatment of satellite measurement time and local surface solar time at each pixel reveals that bias variations occur with both the measurement time and local solar time. Some bias variation with local time of day is to be expected from the presence of the skin layer and diurnal thermocline. Diurnal variations in

cloudiness and the atmosphere can also cause changes with the time of day. Bias variations with satellite measurement time can result from changes in the solar illumination of the satellite independent from processes at the earth's surface. In the remainder of this section we explore the possible contributions of these different effects.

1) INSTRUMENTAL EFFECTS

The most prominent feature for all satellites is a strong negative bias near the local satellite midnight. Satellite midnight refers to local solar midnight at the satellite subpoint. At this time the satellite is on the opposite side of the earth from the sun. Depending on the season, the sun may directly illuminate the front, imaging side of the satellite and instrument surfaces (near the solstices) or may be eclipsed by the earth leaving the satellite in darkness (near the equinoxes). While the earth's surface in view will be in darkness, the period around the satellite midnight is the only time when sunlight can directly illuminate the imaging side of the satellite.

The strong negative bias appears to be caused by a known measurement problem termed the "midnight effect" resulting from solar heating of the instrument surfaces (Johnson and Weinreb 1996). Excessive radiation from these surfaces falls on the detector and affects the calibration. The problem was first found to affect GOES-8 measurements within the eight hours around the satellite midnight. The effect was believed to be strongest in the boreal summer. Effects on GOES-9 and later satellites were predicted to be smaller. Corrections were proposed (Johnson and Weinreb 1996) but have not been implemented in the normal processing sequence.

Our data suggest that the problem may be more significant than originally believed. The observed brightness temperature changes and corresponding bias variations for each satellite closely resemble the behavior described by Johnson and Weinreb (1996). The bias variations are not any weaker for GOES-9 and -10. The hourly bias variations change only slightly with

season and are significant throughout the year. There is no season during which the variations are clearly more prominent. Moreover, our results suggest that such changes in bias are not confined only to times near the local satellite midnight and can instead affect all measurements throughout the day.

Some specific arguments support this assertion that instrumental effects influence more than just the measurements within four hours on either side of the satellite midnight. Of note are the increase in positive bias between 0600 and 0900 LST observed for all the satellites and the maximum positive bias at 0900 for GOES-9 and -10. This increase is largely independent of the local solar time in the different portions of the scene suggesting that it is also independent from diurnal surface and cloud processes. Further, the apparent warming of the satellite relative to the buoys occurs just after dawn and results primarily from increases only in the brightness temperature at 12 μm . Physical warming of the ocean surface layer is not expected to be this large this early in the day and would be accompanied by an increase in the 11- μm brightness temperature as well. Changes in the amount of undetected clouds would also likely result in a change in the 11- μm brightness temperature. To further verify the larger impact of satellite calibration effects, it is necessary to examine some of the other factors that could influence the bias variations and demonstrate that they do not likely explain the remaining variability.

Other instrumental factors that can influence the satellite and buoy differences include satellite scan position and zenith angle effects and differences in satellite and buoy sampling. Variations in scan mirror emissivity with incidence angle have been shown to affect measurements from the GOES Imagers (Weinreb 1996). Changes in bias with scan position were observed in our data but the mean effects were removed prior to displaying the hourly bias variations. Changes in satellite zenith angle can also affect the results through differences in

atmospheric path length and sea surface emissivity. By removing the mean variability with scan position, however, we effectively precluded any attempt to isolate the effects of satellite zenith angle.

Some of the differences between the satellite and buoy SST measurements are due to the fact that buoys provide a point measurement while satellites provide a measurement averaged over a larger area. While this difference is potentially significant, it is not believed to affect the bias variations observed here. Assuming that small-scale variability in SST may differ with time of day, the differences in the spatial extent of the measurements would have a greater impact on variations in the rms differences than on changes in the biases. By considering the mean bias of many measurements at the same local solar time, the effect of any instantaneous differences between point and area-averaged SST is reduced.

2) CLOUDS AND CLOUD DETECTION

Diurnal variations in cloud cover and the efficiency of the cloud detection routines also have a potential impact on the observed bias variations. An increased amount of undetected clouds would cause a tendency for negative biases. Such an increase could be caused by an increase in the total number of clouds (with constant cloud screening efficiency) or a decrease in the effectiveness of the cloud detection routines. The overall effect of cloud, however, is expected to be small because the cloud detection thresholds were chosen to error on the side of excluding too much data. The mean percentage of data rejected as cloud contaminated was over 90% which is significantly higher than typical cloud fraction amounts.

Since separate cloud detection procedures were used during the day and night and visible data improves the ability to detect clouds over the water, it is possible that our cloud screening was less effective during the night. To address this, separate processing of the entire set of

GOES-9 data was performed using simplified cloud detection tests that could be applied in the same manner to both day and nighttime data. The tests consisted of the 11- μm simple threshold test, 11–12 μm difference threshold test, 11- μm spatial uniformity test, 11- μm dynamic threshold test, and SST field test as shown in Figure 1. A comparison of the hourly bias variations obtained using the different techniques is shown in Figure 3. Removing differences in the cloud detection method did not remove or significantly reduce the bias variations between night and day. This suggests that the observed bias variations are not likely due to changes in the cloud detection efficiency. The overall accuracy of the SST measurements was, however, reduced by the lack of additional information. Additional tests using more restrictive thresholds on our normal cloud detection procedure also did not significantly change the bias variations.

Diurnal variations of cloud cover over the oceans derived from satellite data in the International Satellite Cloud Climatology Project (ISCCP) showed the peak high-cloud fraction over convective locations to occur near 1700 LST and the peak low-cloud fraction over nonconvective locations to occur near 0400 LST (Bergman and Salby 1996). These variations, however, were typically smaller in amplitude than those over land except for low clouds just off the west coast of South America. A plot of our average percentage of cloud-flagged pixels as a function of satellite measurement time is shown for a subset of the GOES-8 and -10 measurements in Figure 4. The results for both satellites show peaks in the number of rejected pixels at 0600 and between 1500 and 1800 UTC. Though these peaks are also influenced by increased data rejection in periods of twilight, they generally agree with the results of Bergman and Salby (1996). At these periods of possible peak cloudiness, the mean bias between the satellite and buoy measurements is near 0.0 K and the rms differences are near minimum values suggesting little impact of clouds on the differences. While it is impossible to completely rule

out the possible influence of clouds on the observed bias variations, these results seem to argue against a large contribution.

Independent investigation of the NESDIS prototype operational GOES SST product (E. Maturi, personal communication) also revealed similar hourly bias variations. This result further implies that the variations are not due to specific processing techniques including cloud clearing and averaging of multiple scan lines.

3) THEORETICAL SKIN AND NEAR-SURFACE LAYER EFFECTS

Some diurnal bias variations relative to buoy measurements are expected to result from variability in the temperature difference across the skin layer and diurnal thermocline. While infrared measurements give the temperature near the top of the skin layer, the buoy measurements at depths near 1 m are beneath the skin layer and can be within or beneath a diurnal warm layer if present. A schematic illustration of representative near-surface temperature profiles identifying the key elements for both daytime and nighttime is shown in Figure 5. While the skin temperature is usually cooler than that at the base of the skin layer (only a small fraction of solar radiation is absorbed within the skin layer), the skin temperature can be greater or less than the bulk temperature depending on solar heating and the presence of a diurnal thermocline. Under low wind speed conditions and especially during the daytime, it is possible for the temperature difference between the skin and the depth of the buoy measurement (ΔT) to differ significantly from its mean value (Wick 1995; Fairall et al. 1996). In this section, we calculate the theoretically predicted variability in the near-surface temperature profile and explore its contribution to the measured hourly bias variations and overall GOES SST measurement accuracy.

A detailed model for the temperature variations across the oceanic skin and near-surface layer was used to predict the value of the temperature difference throughout our GOES sampling domain during four one-month periods. The model couples a parameterization for the temperature change across the skin layer (Wick et al. 1996) with a complete one-dimensional, second-moment turbulence closure mixed layer model (Kantha and Clayson 1994). The parameterization is used to estimate the temperature difference between the ocean surface and the base of the skin layer, while the mixed layer model predicts the temperature profile from the base of the skin layer to the depth of the buoy measurement. A complete description of the coupled model and illustration of its ability to reproduce the temperature difference under different conditions is provided by Wick (1995). The model is very similar to that employed by Webster et al. (1996).

To predict ΔT , the model requires an estimate of the initial temperature and salinity and time series of the wind speed and net and solar heat flux. The temperature field was obtained from the Reynolds optimum interpolation SST analysis (Reynolds and Smith 1994). To obtain regular flux estimates throughout the domain, we used the outputs of numerical weather prediction models. For convenience, we used the four-times-daily wind speed and heat flux components from the NCEP/NCAR 40-year reanalysis project (Kalnay et al. 1996). These fields are expected to have greater uncertainties than direct measurements but they should enable realistic estimates of the range of possible temperature differences. Alternatively, one could use heat flux maps derived entirely from satellite data (e.g., Curry et al. 1999), but such products are not readily available at all times. At this time, there are not a large enough number of *in situ* observations to perform the analysis solely with direct flux measurements.

The model was used to simulate the temperature differences during September and October 1997 to correspond with GOES-9, and December 1998 to correspond with GOES-8 and -10. Hourly variations in the mean temperature difference averaged over all the available satellite-buoy match locations are shown in Figure 6 along with the corresponding standard deviations. Here, ΔT is defined as the temperature at the ocean skin minus that at the buoy measurement depth so that a negative ΔT corresponds to a cooler skin. The results show that during the nighttime ΔT has an average value between roughly -0.3 and -0.35 K and exhibits little variability. During the daytime, however, the variability between individual measurements increases and the mean difference decreases in magnitude to around -0.1 K due to warming of the near-surface layer. These modeled values are consistent with previous direct measurements of skin and bulk temperatures (e.g., Wick et al. 1996; Donlon et al. 1999). The results are similar for each satellite and period. The decrease in the average ΔT at GOES-8 match locations at 1200 LST is due to changes in the locations of the matches. An average of ΔT computed over the entire measurement domain has its largest value at 1200 LST.

The simulations suggest that average satellite SST bias variations of as much as 0.3 K between day and night can result from diurnal warming and the presence of the skin layer. It is important to emphasize that this is an average over all wind speeds and daily values can be much larger at low winds. Mean bias variations in GOES SST products could theoretically be reduced by up to this amount through explicit treatment of the skin layer and diurnal thermocline. The theoretical mean bias variations resulting from the near-surface temperature profile are compared with the observed variations in Figure 7 for GOES-10. The values of the simulated ΔT have been converted to a mean of zero to facilitate comparison. The near-surface layer variations do appear to explain a portion of the positive bias observed between 0600 and 1500 LST.

Excluding the measurements near the local satellite midnight (2100 – 0300 LST), the 0.4 K variation in observed bias is only slightly greater than the ΔT variations. Nonetheless, there is still a difference in the timing of the peak observed bias and the maximum warming of the near-surface layer. The observed bias increases sharply and peaks at 0900 LST. While some modeling errors are to be expected due to the forcing data, which is not ideal, it is unlikely that the skin layer and diurnal thermocline could completely explain the large observed bias increase early in the morning.

The results suggest that variability in the near-surface temperature profile and satellite midnight effects both contribute to the observed bias variations but still do not fully explain the observations. Variations in cloudiness and sampling characteristics of the buoys do not appear to explain the remaining bias variations. It seems that calibration problems related to the known satellite midnight effects influence a larger range of measurements than originally believed.

4. Impact on measurements of the diurnal SST cycle

The near-surface temperature profile and measured hourly bias variations have an important effect on estimates of the shape and amplitude of the diurnal SST cycle. When speaking of the diurnal SST cycle, it is essential to specify whether one is referring to the diurnal cycle at the ocean skin or at depths representative of the bulk measurements. Since the majority of algorithms used to date (including that discussed in sections 2 and 3) have attempted to provide the bulk SST, we will focus on estimating the diurnal variability of the bulk SST.

For each satellite, we computed the average diurnal SST cycle over a selected region for a period of one month. The regions extended from 85° to 100°W and 0° to 25°N for GOES-8 and 135° to 150°W and 0° to 25°N for GOES-9 and -10. The diurnal cycle was averaged over the extended region to reduce the effect of noise in individual satellite scenes. In Figure 8, we

plot the diurnal cycle both from the direct measurements and after the hourly bias variations shown in Figure 2 were subtracted from the measurements. The measured bias variations incorporate both instrumental and skin layer effects on the SST values. Subtracting the biases should provide an improved estimate of the average diurnal bulk SST cycle. For all the cases the variability in the hour averages is greater than the diurnal amplitude. There is large variability in the satellite measurements and the amount of warming over the averaging period but the adjustments shown are as they would be for any estimate of the diurnal cycle.

The impact of the bias variations on estimates of the diurnal cycle is very apparent. The effect is most important for GOES-9 where the shape of the diurnal cycle is strongly distorted. The sharp increase in the bias observed between 0300 and 0600 LST leads to a too rapid onset of perceived diurnal warming and an apparent dual peak in the diurnal cycle. The amplitude of the diurnal warming is also overestimated if the bias variations are not accounted for. While the bias variations for GOES-8 and GOES-10 do not similarly alter the shape of the estimated diurnal cycle, they do affect the amplitude of the diurnal warming. It remains unclear why the bias variations differ between the satellites.

Clearly, direct application of Eq. (1) can lead to misinterpretation of the diurnal bulk SST cycle. While the use of a single SST algorithm at all times may be desirable for a continuous time series, some compensation for the bias variations is necessary. Even in the absence of sensor bias, a single “all-hour” satellite bulk SST algorithm derived through regression against buoy observations is not expected to directly give the bulk SST diurnal cycle. Changes in the measured satellite brightness temperatures are representative of changes in the skin temperature and might not be accurately mapped into diurnal bulk SST variations. Sensor errors add more complications.

Subtraction of mean bias variations may not be sufficient for accurate estimates of instantaneous diurnal variations. The mean bias variations appear to adequately reflect most satellite effects, but skin layer and diurnal thermocline effects are more variable. The small variations in the rms errors in the satellite measurements between measurement times suggests that the satellite effects on the bias are relatively constant and can be removed with a bias correction. The small observed seasonal variations further suggest that removal of a single annual mean bias is sufficient. Only the measurements taken at the single time nearest to the local satellite midnight might not be adequately corrected. Further modeling of the near-surface temperature profile may be necessary to correct for effects of the skin layer and diurnal thermocline. The magnitude of ΔT is very sensitive to instantaneous values of the wind speed and heat flux. The impact of explicitly incorporating near-surface layer effects in satellite bulk SST measurements is explored further in the following section.

5. Explicit incorporation of the near-surface temperature profile in SST measurements

The previous sections suggested that the accuracy of a GOES satellite bulk SST product could theoretically be improved through explicitly accounting for the presence of the skin layer and diurnal thermocline. In this section we apply two different techniques to determine if these improvements can be achieved in practice. In the first we limit our derivation to high wind speed conditions where data has suggested that the bulk-skin temperature difference approaches a near-constant value. In the second, we apply modeled estimates of ΔT at the location of all buoy matches to convert between skin and bulk temperatures.

a. Limitation to high wind speed conditions

Recent measurements (Donlon 1999; P. Minnett, personal communication) have suggested that the temperature difference across the skin layer might approach a near constant

value of ~ -0.1 K at higher wind speeds (above roughly 6-8 m/s). This led some scientists to suggest that satellite skin SST algorithms might be derived using measurements from buoys if the comparisons are limited to such high wind speeds. This would also suggest that the accuracy of satellite bulk SST algorithms would improve at higher wind speeds due to reduced variability in the near-surface temperature profile. This is in contrast to the results of Wu et al. (1999) who found that GOES measurement errors increased at higher wind speeds.

A revised set of matches were constructed for GOES-8 and GOES-10 between August 1998 and October 1999 using only buoys that had both valid SST and wind speed measurements. From this dataset, algorithms of the form of Eq. (1) were derived first using all the available measurements and then using only those measurements obtained outside some critical wind speed. Since wind speed measurements were required, the buoys consisted primarily of the TOGA Tropical Atmosphere-Ocean (TAO) buoys in the equatorial Pacific and the National Data Buoy Center (NDBC) buoys along the coasts of the United States including Hawaii.

The results of these experiments are summarized in Table 6 for four different choices of the critical wind speed. The overall rms error in the measurements shows little improvement if the measurements are limited to the highest wind speeds. If, however, only the lowest wind speeds are considered, there is an increase in the rms error at least for GOES-10 which has a lower instrument noise level. This increase is likely due to increased surface layer variability. For GOES-8 the instrument noise and calibration errors apparently dominate over surface layer effects. These results support the hypothesis that satellite bulk SST products can be improved by excluding measurements at the lowest wind speeds but imply that the improvements may be small in the total rms statistics because of the relatively small number of points involved and the size of the effect compared to the other factors influencing the retrievals.

The results at the highest wind speeds also help to interpret the possible causes of the hourly bias variations. At these wind speeds, mixing is sufficient to prevent the formation of a diurnal thermocline during the daytime and little variation in ΔT with time of day would be expected. The mean hourly bias variations at wind speeds greater than 8 m/s shown in Figure 9, however, illustrate that there are still significant hourly bias variations in the measured SST. The statistics and number of points are also included in Table 7. The variability is still very similar to that seen in Figure 2 with a peak near midday. This provides strong evidence that satellite calibration effects cause a significant part of the bias variations during the daytime. Variability in the near-surface temperature profile appears to be responsible for only a small portion of the bias variations.

b. Direct simulation of ΔT at each satellite-buoy match

We next evaluated a measurement method where we simulated the temperature difference across the skin layer and diurnal thermocline at each buoy match location. In this procedure it is assumed that the satellite provides a measurement of the skin temperature and that an estimate of ΔT is necessary to compute a bulk SST value. The basic steps of this procedure can be summarized as follows: 1) At each match location to be used in the derivation of the algorithm, a skin temperature is computed from the buoy's measured bulk SST and the modeled ΔT value. 2) The satellite brightness temperatures are regressed against these skin temperature estimates to derive a satellite skin SST algorithm. 3) The satellite skin SST algorithm is used to compute skin temperature measurements throughout the coverage region in the normal way. 4) At independent match locations where the algorithm is to be validated, bulk SST estimates are computed from the satellite skin temperature measurements and the modeled ΔT values. 5) These bulk SST estimates are compared with the actual buoy SST measurements and the

accuracy of the procedure is determined. This technique represents the most explicit and direct consideration of the near-surface temperature profile.

This technique was applied to the same four one-month periods for which ΔT was computed throughout the satellite domain. For each period, the bulk SST estimates computed with the above-cited explicit procedure were compared with bulk SST measurements computed through the normal technique of directly regressing the satellite measurements against the bulk temperatures. A slightly modified technique was also employed where instead of using the simulated ΔT value at each match location, we used the mean value for that time of day. The bias and rms differences of the different procedures relative to independent buoy measurements are summarized in Table 8.

The results demonstrate that we obtain essentially the same bulk SST measurement accuracy regardless of whether we explicitly account for the skin layer and diurnal thermocline or just directly regress the satellite measurements against bulk SST measurements. Currently, any benefits obtained from incorporating the detailed skin layer physics are countered by inaccuracies in the near-surface layer model and the wind speed and heat flux forcing data used to run the model. When applying the modeled ΔT values at individual match points, the uncertainties in the modeled fluxes are critical. Applying the mean ΔT value at a given measurement time provides improved results over the use of the individual ΔT values but still does not produce a significant improvement over the direct bulk SST regression.

Despite the neutral result, the results are still encouraging. When computations of this type were first attempted, the results were significantly worse when the near-surface layer was explicitly included. The models and forcing data have improved in quality to the point that the accuracy is now the same. With further improvements to the models and the availability of

improved heat flux fields, it may be possible to improve the accuracy of satellite bulk SST measurements by explicitly treating the skin layer and diurnal thermocline.

6. Conclusions

The latest generation of GOES satellites has facilitated significant improvements in our ability to measure SST from geostationary satellites. The overall rms accuracy of bulk SST measurements relative to buoy measurements was shown to be between 0.6 and 0.9 K. This compares very favorably with traditional SST measurement accuracy from the polar orbiting AVHRR. More detailed analyses, however, demonstrate that sensor calibration and oceanic near-surface temperature profile effects have a significant impact on the accuracy and interpretation of the measurements.

The SST measurements exhibited large variations in their mean bias with season and time of day. Most notable was a mean bias variation of over 0.6 K with satellite measurement time. These bias effects were found to be dominated by satellite calibration problems related to the known satellite midnight effect. Bias changes of from 0.3 to 0.6 K between 2100 and 0300 LST are consistent with the known behavior of the satellite midnight effect. Modeled estimates of ΔT showed that bias variations up to 0.3 K of the remaining 0.3-0.4 K differences could result from variability in the skin layer and diurnal thermocline but the bias variations were also observed when the measurements were limited to high wind speed conditions where diurnal variations in ΔT are negligible. In addition, the peak positive bias occurred at 0900 LST when surface effects are still typically much smaller and most of the change was due solely to increases in the 12- μm brightness temperature. Diurnal cloudiness variations did not appear to significantly affect the hourly bias variations as observed peaks in cloudiness showed little correlation to the bias variations. Independent evaluation of a different GOES SST product showed similar hourly bias

variations despite different processing and cloud screening. These results suggest that additional satellite calibration problems contribute to the bias variations. Smaller bias variations were observed with season and due to scan mirror emissivity effects from the GOES-8 satellite. The seasonal bias variations were likely due to the majority of matches occurring in the Northern Hemisphere and changes in cloud cover.

The hourly bias variations have an important effect on attempts to measure the diurnal cycle of bulk SST. The variations led to errors in estimates of both the amplitude and shape of the diurnal cycle. Without some correction for satellite effects and consideration of the near-surface temperature profile, the amplitude of the diurnal bulk SST cycle can be overestimated by from 0.3 to nearly 1.0 K and the peak warming can be predicted to occur from 3-6 hours too early.

Given the theoretically predicted impact of the near-surface layer on bulk SST measurements, we evaluated a procedure that explicitly incorporated skin layer and diurnal thermocline effects to determine if the improvements could be obtained in practice. The technique resulted in essentially the same accuracy as that obtained through direct regression against bulk SST data, indicating that the improvements in physical accuracy were balanced by uncertainties in the near-surface layer models and forcing fields used to drive the models. Increased direct measurements of the skin temperature and further improvements in models for the skin layer and global flux fields are required to produce improvements in bulk SST measurements over those based on direct regressions of the satellite data against in situ bulk SST observations.

Acknowledgments. The GOES data used in this study were generously provided by the NOAA Forecast Systems Laboratory. The drifting and moored buoy data were supplied by Dr. Richard Reynolds of the NOAA National Center for Environmental Prediction via an archive at the NOAA Climate Diagnostics Center and by Daniel Olszewski of the Naval Oceanographic Office. We thank Dr. William Emery, Dr. Joseph Shaw, Dr. Jeffrey Hare, and four anonymous reviewers for their useful comments on the manuscript. This work was supported in part by Grant NAG5-6521 from NASA.

REFERENCES

- Anding, D., and R. Kauth, 1970: Estimation of sea surface temperature from space. *Remote Sens. Environ.*, **1**, 217-220.
- Barton, I. J., 1995: Satellite-derived sea surface temperatures: Current status. *J. Geophys. Res.*, **100**, 8777-8790.
- Bates, J. J., and W. L. Smith, 1985: Sea surface temperature: Observations from geostationary satellites. *J. Geophys. Res.*, **90**, 11,609-11,618.
- Baucom, J. G. and M. Weinreb, 1996: Characteristics of E/W stripes in infrared images from the GOES-8 imager. *GOES-8 and Beyond*, SPIE Proc. Vol. 2812, Denver, CO, The International Society for Optical Engineering, 587-595.
- Bergman, J. W., and M. L. Salby, 1996: Diurnal variations of cloud cover and their relationship to climatological conditions. *J. Climate*, **9**, 2802-2820.
- Brown, S. J., A. R. Harris, I. M. Mason, and A. M. Zavody, 1997: New aerosol robust sea surface temperature algorithms for the along-track scanning radiometer. *J. Geophys. Res.*, **102**, 27,973-27,989.

- Clayson, C. A., and J. A. Curry, 1996: Determination of surface turbulent fluxes for the Tropical Ocean-Global Atmosphere Coupled Ocean-Atmosphere Response Experiment: Comparison of satellite retrievals and in situ measurements. *J. Geophys. Res.*, **101**, 28,515-28,528.
- Curry, J. A., C. A. Clayson, W. B. Rossow, R. Reeder, Y.-C. Zhang, P. J. Webster, G. Liu, and R.-S. Sheu, 1999: High-resolution satellite-derived dataset of the surface fluxes of heat, freshwater, and momentum for the TOGA COARE IOP. *Bull. Amer. Meteor. Soc.*, **80**, 2059-2080.
- Deschamps, P. Y., and T. Phulpin, 1980: Atmospheric correction of infrared measurements of sea surface temperature using channels at 3.7, 11, and 12 μm . *Boundary-Layer Meteor.*, **18**, 131-143.
- Donlon, C. J., T. J. Nightingale, T. Sheasby, J. Turner, I. S. Robinson, and W. J. Emery, 1999: Implications of the oceanic thermal skin temperature deviation at high wind speed. *Geophys. Res. Lett.*, **26**, 2505.
- Ellrod, G. P., R. V. Achutuni, J. M. Daniels, E. M. Prins, and J. P. Nelson III, 1998: An assessment of GOES-8 imager data quality. *Bull. Amer. Meteor. Soc.*, **79**, 2509-2526.
- Emery, W. J., Y. Yu, G. A. Wick, P. Schlüssel, and R. W. Reynolds, 1994: Correcting infrared satellite estimates of sea surface temperature for atmospheric water vapor attenuation. *J. Geophys. Res.*, **99**, 5219-5236.
- Emery, W. J., D. J. Baldwin, P. Schlüssel, and R. W. Reynolds, 2000: Accuracy of in situ sea surface temperatures used to calibrate infrared satellite measurements, submitted to *J. Geophys. Res.*

- Fairall, C. W., E. F. Bradley, J. S. Godfrey, G. A. Wick, J. B. Edson, and G. S. Young, 1996: Cool skin and warm layer effects on sea surface temperature, *J. Geophys. Res.*, **101**, 1295-1308.
- Johnson, R. X., and M. Weinreb, 1996: GOES-8 imager midnight effects and slope correction. *GOES-8 and Beyond*, SPIE Proc. Vol. 2812, Denver, CO, The International Society for Optical Engineering, 596-607.
- Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, A. Leetmaa, R. Reynolds, R. Jenne, and D. Joseph, 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.*, **77**, 437-471.
- Kantha, L. H., and C. A. Clayson, 1994: An improved mixed layer model for geophysical applications. *J. Geophys. Res.*, **99**, 23,235-25,266.
- Katsaros, K. B., 1980: The aqueous thermal boundary layer. *Bound.-Layer Meteor.*, **18**, 107-127.
- Komajda, R. J., and K. McKenzie, 1994: An introduction to the GOES I-M imager and sounder instruments and the GVAR retransmission format. NOAA Tech. Rep. NESDIS 82, NOAA/NESDIS, Washington, DC, 56 pp. [Available from National Technical Information Service, U.S. Dept. of Commerce, Sills Bldg., 5285 Port Royal Road, Springfield, VA 22161.]
- Legeckis, R., and T. Zhu, 1997: Sea surface temperatures from the GOES-8 geostationary satellite. *Bull. Amer. Meteor. Soc.*, **78**, 1971-1983.
- Maul, G. A., 1981: Application of GOES visible-infrared data to quantifying mesoscale ocean surface temperatures. *J. Geophys. Res.*, **86**, 8007-8021.

- May, D. A., and W. O. Osterman, 1998: Satellite-derived sea surface temperatures: Evaluation of GOES-8 and GOES-9 multispectral imager retrieval accuracy. *J. Atmos. Ocean. Technol.*, **15**, 788-797.
- 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. Pichel, and C. Walton, 1985: Comparative performance of AVHRR-based multichannel sea surface temperatures. *J. Geophys. Res.*, **90**, 11,587-11,601.
- McMillin, L., and D. Crosby, 1984: Theory and validation of the multiple window sea surface temperature technique. *J. Geophys. Res.*, **89**, 3655-3661.
- Menzel, W. P., and J. F. W. Purdom, 1994: Introducing GOES-I: The first of a new generation of geostationary operational environmental satellites. *Bull. Amer. Meteor. Soc.*, **75**, 757-781.
- Merchant, C. J., and A. R. Harris, 1999: Toward the elimination of bias in satellite retrievals of sea surface temperature 2. Comparison with in situ measurements. *J. Geophys. Res.*, **104**, 23,579-23,590.
- Mutlow, C. T., A. M. Závody, 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**, 22,575-22,588.
- Njoku, E. G., T. P. Barnett, R. M. Laurs, and A. C. Vastano, 1985: Advances in satellite sea surface temperature measurement and oceanographic applications. *J. Geophys. Res.*, **90**, 11,573-11,586.
- Olesen, F.-S., and H. Grassl, 1985: Cloud detection and classification over oceans at night with NOAA-7. *Int. J. Remote Sens.*, **6**, 1435-1444.

- Payne, R. E., 1972: Albedo of the Sea Surface. *J. Atmos. Sci.*, **29**, 959-970.
- Reynolds, R. W., and T. M. Smith, 1994: Improved global sea surface temperature analyses using optimum interpolation. *J. Climate*, **7**, 929-948.
- Robinson, I. A., N. C. Wells, and H. Charnock, 1984: The sea surface thermal boundary layer and its relevance to the measurement of sea surface temperature by airborne and spaceborne radiometers. *Int. J. Remote Sens.*, **5**, 19-45.
- Saunders, R. W., and K. T. Kriebel, 1988: An improved method for detecting clear sky and cloudy radiances from AVHRR data. *Int. J. Remote Sens.*, **9**, 123-150.
- Schlüssel, P., W. J. Emery, H. Grassl, and T. Mammen, 1990: On the bulk-skin temperature difference and its impact on satellite remote sensing of sea surface temperature. *J. Geophys. Res.*, **95**, 13,341-13,356.
- Smith, E., J. Vazquez, A. Tran, and R. Sumagaysay, 1996: Satellite-derived sea surface temperature data available from the NOAA/NASA Pathfinder program. http://www.agu.org/eos_elec/95274e.html.
- Soloviev, A. V., and P. Schlüssel, 1996: Evolution of cool skin and direct air-sea gas transfer coefficient during daytime. *Bound.-Layer Meteor.*, **77**, 45-68.
- Ullman, D. S., and P. Cornillon, 1999: Satellite-derived sea surface temperature fronts on the continental shelf off the northeast U.S. coast. *J. Geophys. Res.*, **104**, 23,459-23,478.
- 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.

- Webster, P. J., C. A. Clayson, and J. A. Curry, 1996: Clouds, radiation, and the diurnal cycle of sea surface temperature in the tropical western Pacific. *J. Climate*, **9**, 1712-1730.
- Weinreb, M., 1996: Real-world calibration of GOES-8 and -9 sensors. *GOES-8 and Beyond*, SPIE Proc. Vol. 2812, Denver, CO, The International Society for Optical Engineering, 572-586.
- Weinreb, M. P., M. Jamison, N. Fulton, Y. Chen, J. X. Johnson, J. Bremer, C. Smith, and J. Baucom, 1997: Operational calibration of the imagers and sounders on the GOES-8 and -9 satellites. *Appl. Opt.*, **36**, 6895-6904.
- Wick, G. A., 1995: Evaluation of the variability and predictability of the bulk-skin sea surface temperature difference with application to satellite-measured sea surface temperature. Ph.D. thesis, University of Colorado, Boulder, 139 pp.
- Wick, G. A., W. J. Emery, L. H. Kantha, and P. Schlüssel, 1996: The behavior of the bulk-skin temperature difference under varying wind speed and heat flux. *J. Phys. Oceanogr.*, **26**, 1969-1988.
- Wu, X., W. P. Menzel, and G. S. Wade, 1999: Estimation of sea surface temperatures using GOES-8/9 radiance measurements. *Bull. Amer. Meteor. Soc.*, **80**, 1127-1138.

TABLE 1. Characteristics of the GOES Imager on GOES-8, -9, and -10.

Channel #	Description	Wavelength	FOV	NEΔT at 300 K		Bits
		(μm)	(km)	(K)		Transmitted
				GOES-8	GOES-9	
1	Visible	0.55-0.75	1			10
2	Shortwave IR	3.8-4.0	4	0.16	0.08	10
3	Moisture	6.5-7.0	8	0.27	0.15	10
				(230K)	(230K)	
4	Longwave IR 1	10.2-11.2	4	0.12	0.07	10
5	Longwave IR 2	11.5-12.5	4	0.2	0.14	10

TABLE 2. GOES SST algorithm coefficients.

Satellite	a	b	c	d (K)
GOES-8	1.0466	2.0227	0.7741	-13.2965
GOES-9	1.0361	1.9132	0.8597	-10.0473
GOES-10	1.0914	1.7113	1.2030	-26.0475

TABLE 3. Overall accuracy statistics of the GOES SST algorithms.

Satellite	Bias (K)	RMS difference (K)	Number of points
GOES-8	0.00	0.81	11,436
GOES-9	-0.01	0.61	3469
GOES-10	0.01	0.62	20,981

TABLE 4. Seasonal evolution of SST bias and rms difference.

GOES-9			
Month	Bias (K)	rmsd (K)	# matches
6/97	-0.03	0.69	174
7/97	-0.09	0.65	245
8/97	0.07	0.62	339
9/97	0.00	0.67	364
10/97	0.02	0.54	270
11/97	0.11	0.63	309
12/97	-0.02	0.51	295
1/98	0.05	0.64	262
2/98	0.13	0.55	189
3/98	0.05	0.65	204
4/98	-0.07	0.56	225
5/98	-0.28	0.65	306
6/98	0.00	0.56	287
GOES-8			
	Bias (K)	rmsd (K)	# points
8/98	-0.24	0.92	130
9/98	-0.06	1.04	566
10/98	0.07	0.71	774
11/98	0.24	0.67	740
12/98	0.13	0.70	782
GOES-10			
	Bias (K)	rmsd (K)	# points
8/98	-0.08	0.83	470
9/98	0.10	0.68	1029
10/98	0.06	0.61	1236
11/98	-0.05	0.71	1199
12/98	-0.09	0.58	953

1/99	0.23	0.78	642	-0.15	0.59	1011
2/99	0.22	0.73	786	-0.20	0.63	744
3/99	0.04	0.68	827	-0.18	0.65	967
4/99	-0.18	0.77	619	-0.16	0.91	486
5/99	-0.45	1.11	181	N/A	N/A	0
6/99	-0.38	0.99	602	-0.02	0.51	959
7/99	-0.25	1.04	955	-0.02	0.57	1840
8/99	-0.22	0.92	1077	0.00	0.52	2421
9/99	0.24	0.73	811	-0.01	0.60	2138
10/99	0.08	0.74	905	0.11	0.60	2051
11/99	0.14	0.72	549	0.15	0.65	1616
12/99	0.20	0.65	490	0.13	0.58	1861

TABLE 5. Hourly variation of SST match statistics.

Hour (LST)	GOES-8			GOES-9			GOES-10		
	Bias (K)	rmsd (K)	# points	Bias (K)	rmsd (K)	# points	Bias (K)	rmsd (K)	# points
0000	-0.48	0.86	815	-0.34	0.74	78	-0.73	1.00	507
0300	-0.20	0.83	1788	-0.22	0.63	395	-0.23	0.60	2419
0600	0.01	0.63	644	0.10	0.54	613	0.07	0.52	3187
0900	0.13	0.73	1363	0.28	0.63	374	0.37	0.63	3315
1200	0.30	0.84	1855	0.09	0.59	442	0.31	0.68	2551
1500	0.11	0.74	1900	-0.04	0.61	925	0.00	0.50	3461
1800	-0.02	0.74	2159	-0.07	0.56	409	-0.02	0.54	2495
2100	-0.22	0.71	912	-0.12	0.61	232	-0.43	0.69	3046

TABLE 6. Measurement accuracy changes at high wind speeds.

Satellite	Wind speed range	Bias (K)	RMS error (K)	Number of points
GOES-8	All	-0.05	0.74	8157
	> 8 m/s	0.00	0.76	1288
	> 10 m/s	-0.01	0.73	173
	< 3 m/s	0.00	0.69	1147
	< 2 m/s	0.05	0.71	434
GOES-10	All	0.00	0.66	9462
	> 8 m/s	0.00	0.64	2386
	> 10 m/s	0.06	0.64	363
	< 3 m/s	0.00	0.82	730
	< 2 m/s	0.02	0.86	303

TABLE 7. Hourly variation of SST match statistics at wind speeds above 8 m/s.

GOES-8				GOES-10			
Hour (LST)	Bias (K)	rmsd (K)	# points	Hour (LST)	Bias (K)	rmsd (K)	# points
2345	-0.57	0.88	120	0000	-0.95	1.15	94
0245	-0.32	0.81	207	0300	-0.32	0.61	186
0545	0.17	0.54	114	0600	-0.06	0.55	251
0845	0.28	0.74	206	0900	0.24	0.55	509
1145	0.20	0.86	214	1200	0.26	0.65	640
1445	0.24	0.70	197	1500	-0.01	0.56	240
1745	-0.06	0.71	87	1800	-0.03	0.54	222
2045	-0.20	0.69	143	2100	-0.52	0.75	244

TABLE 8. Comparison of direct regression for bulk SST with explicit skin layer treatment.

Satellite	Bulk regression		Explicit skin layer (Local ΔT)		Explicit skin layer (Avg. hourly ΔT)	
	Bias (K)	RMS (K)	Bias (K)	RMS (K)	Bias (K)	RMS (K)
GOES-9 (Sep. 97)	-0.03	0.68	-0.03	0.71	-0.05	0.70
GOES-9 (Oct. 97)	-0.01	0.59	0.01	0.64	0.00	0.60
GOES-8 (Dec. 97)	0.00	0.74	-0.01	0.73	0.02	0.73
GOES-10 (Dec. 97)	0.01	0.53	0.02	0.53	0.02	0.50

Figure Captions

Figure 1. Schematic diagram of the cloud filtering technique used for the GOES data. In the inputs, $T_{3.9}$, T_{11} , and T_{12} refer to brightness temperatures at 3.9, 11, and 12 μm , respectively, albedo is derived from the GOES visible channel, SZA is the solar zenith angle, and Field is a running “best guess” of the nighttime SST at each location.

Figure 2. Hourly variations in the GOES SST measurement bias (satellite – buoy) for each satellite plotted as a function of the satellite measurement time. Mean variations with scan position were first removed and the measurements were averaged over all seasons. All the satellites demonstrate large changes in the average bias.

Figure 3. Comparison of hourly bias variations observed for GOES-9 using both the normal cloud detection procedure and one with no change in technique between day and night. The figure demonstrates that the bias variations do not likely result from a greater occurrence of undetected clouds during the night.

Figure 4. Illustration of the hourly variation of the fraction of pixels rejected in cloud screening for GOES-8 and -10. The overall cloud fractions are very high and the patterns show little correlation to the observed hourly bias variations.

Figure 5. Schematic illustration of the elements of the oceanic near-surface temperature profile during day and night. The drawings are not to scale and depths have purposely been excluded.

The thickness of the skin layer is typically assumed to be a millimeter or less and the buoy measurements are usually at depths on the order of a meter. The depth of the diurnal warm layer, if present, is highly variable and can extend to depths greater than a meter. In such a case, the buoy measurement could also be within the warm layer. Temperature differences across the skin layer typically vary between 0.1 and 0.3 K while the differences across the diurnal thermocline vary significantly with wind speed and can reach more than a degree.

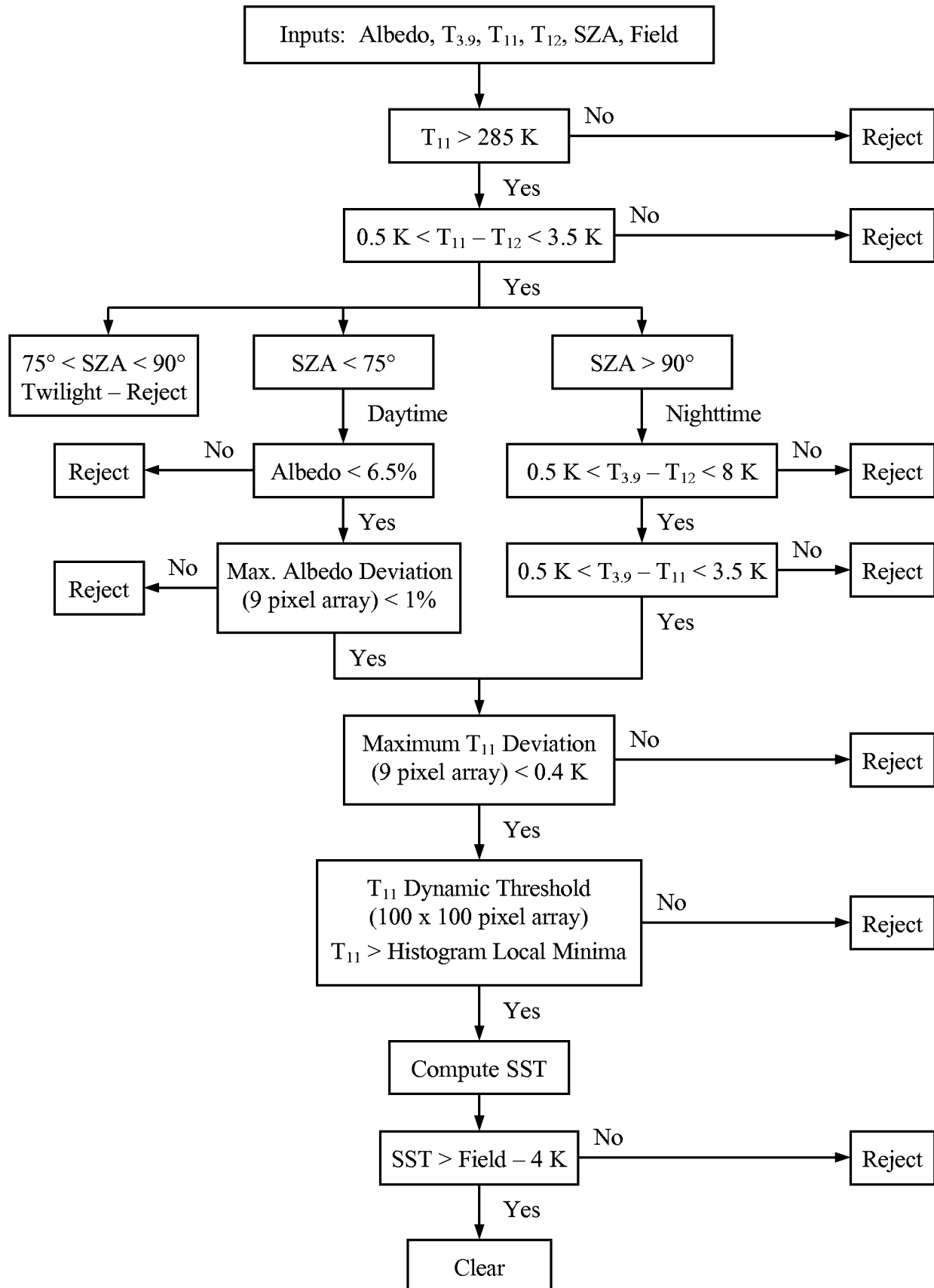
Figure 6. Hourly variation of the modeled ΔT values at the locations of the matches between satellite and buoy data. Measurements were grouped by satellite measurement time and then plotted against the equivalent local solar time at the satellite subpoint.

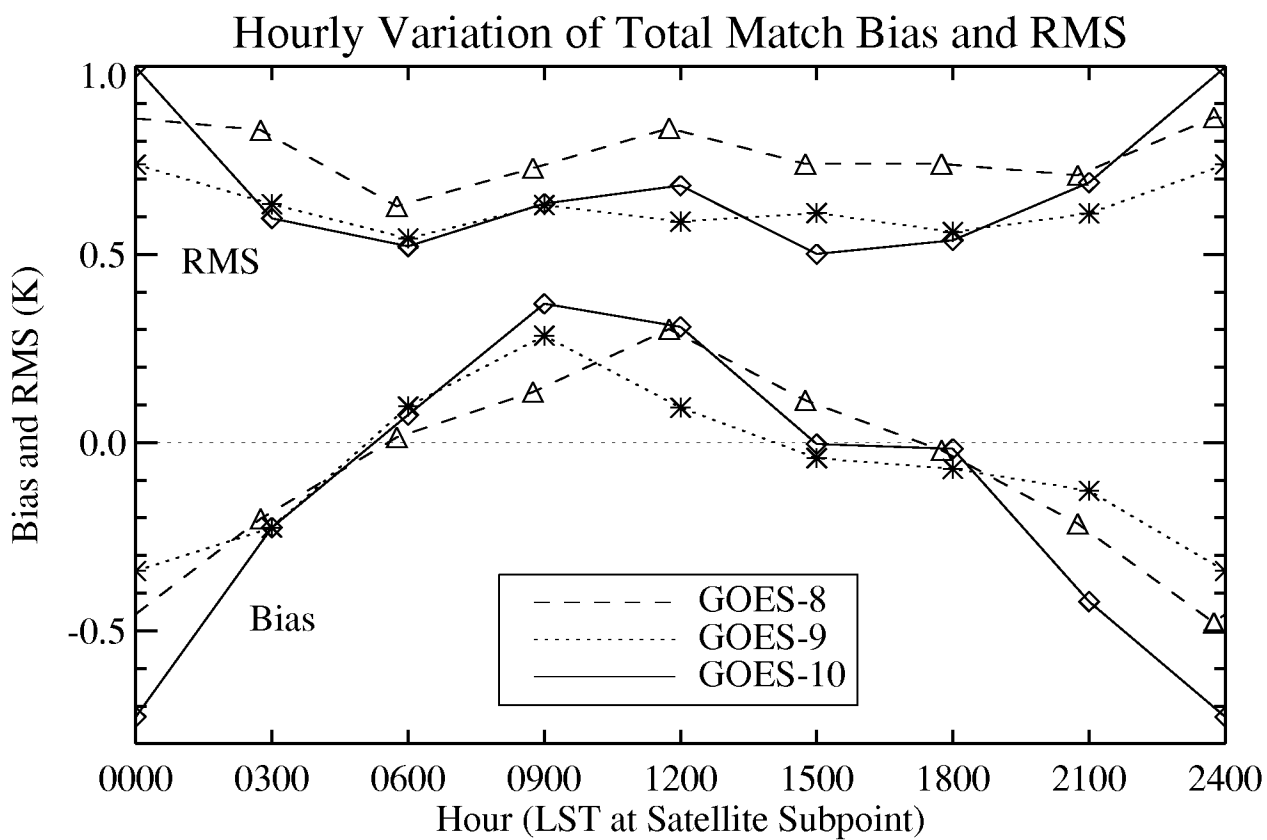
Figure 7. Comparison of the actual measured hourly bias variation for GOES-10 with the simulated variations due to the presence of the near-surface temperature profile. The simulated ΔT variations were shifted to have a mean of 0.0 and inverted so that a cool skin now corresponds to a negative ΔT . The skin layer variability is seen to explain only a portion of the observed bias variations. The rms differences corresponding to the bias variations are as shown in Figure 2.

Figure 8. Impact of the observed bias variations on estimates of the shape and magnitude of the bulk SST diurnal cycle. The dotted lines labeled “Raw” correspond to the diurnal cycle inferred directly from the satellite SST measurements. The solid lines show the corresponding cycle if the mean observed hourly bias variations are first subtracted from the measurements. The

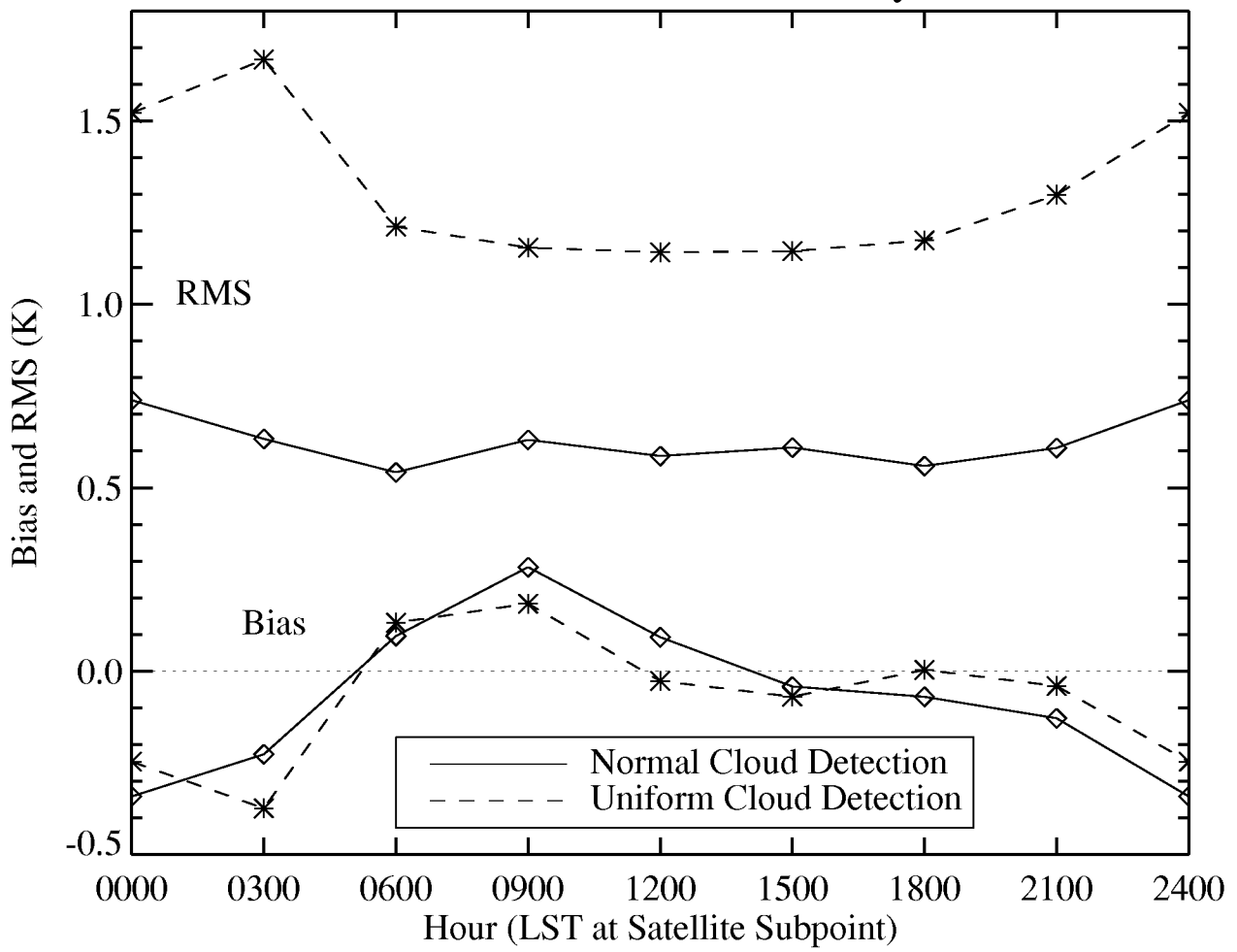
diurnal cycle estimates appear much more realistic after correction for the bias variations. The measurements were averaged over a $15^\circ \times 25^\circ$ region for each satellite.

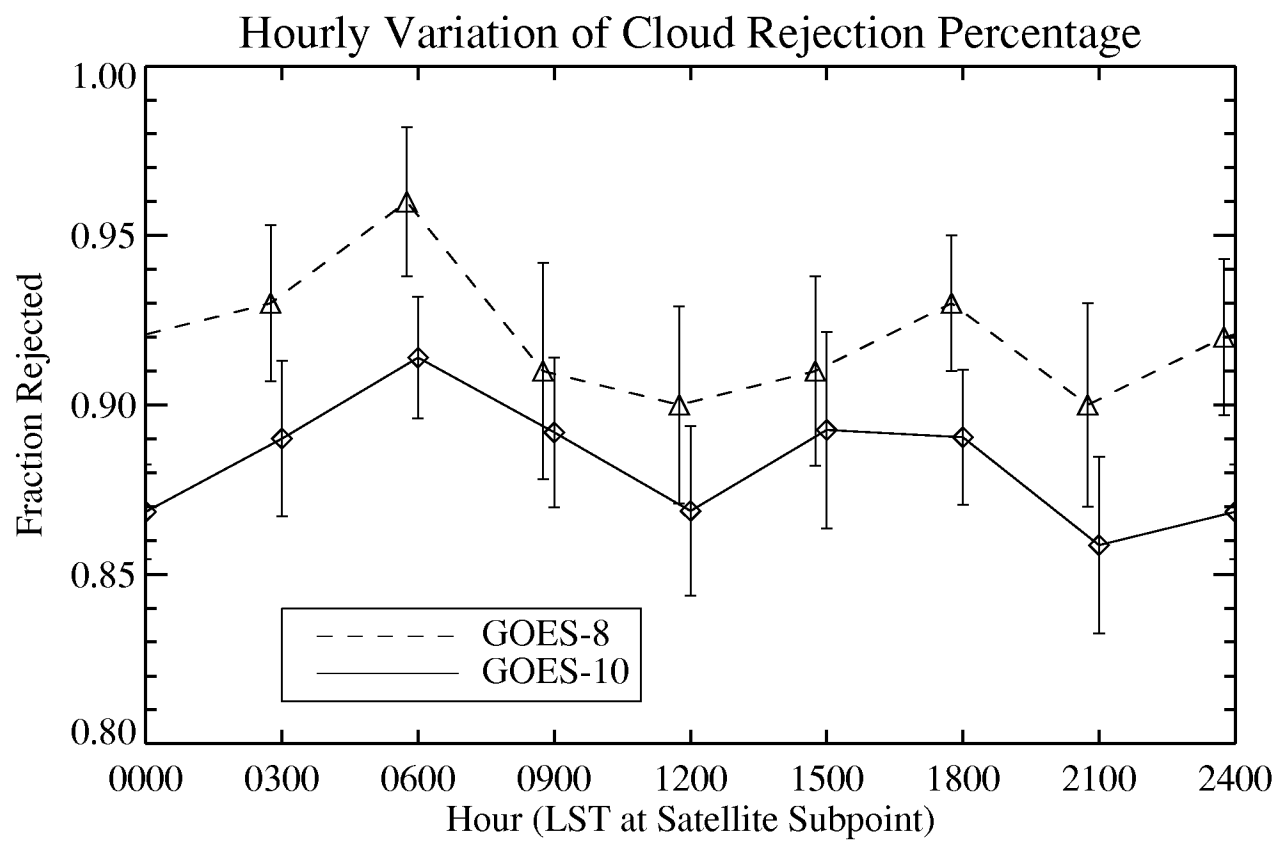
Figure 9. Hourly variations in the GOES SST measurement bias at wind speeds greater than 8 ms^{-1} for GOES-8 and -10 plotted as a function of the satellite measurement time. Hourly bias variations are observed even at high wind speeds when no diurnal thermocline is expected.



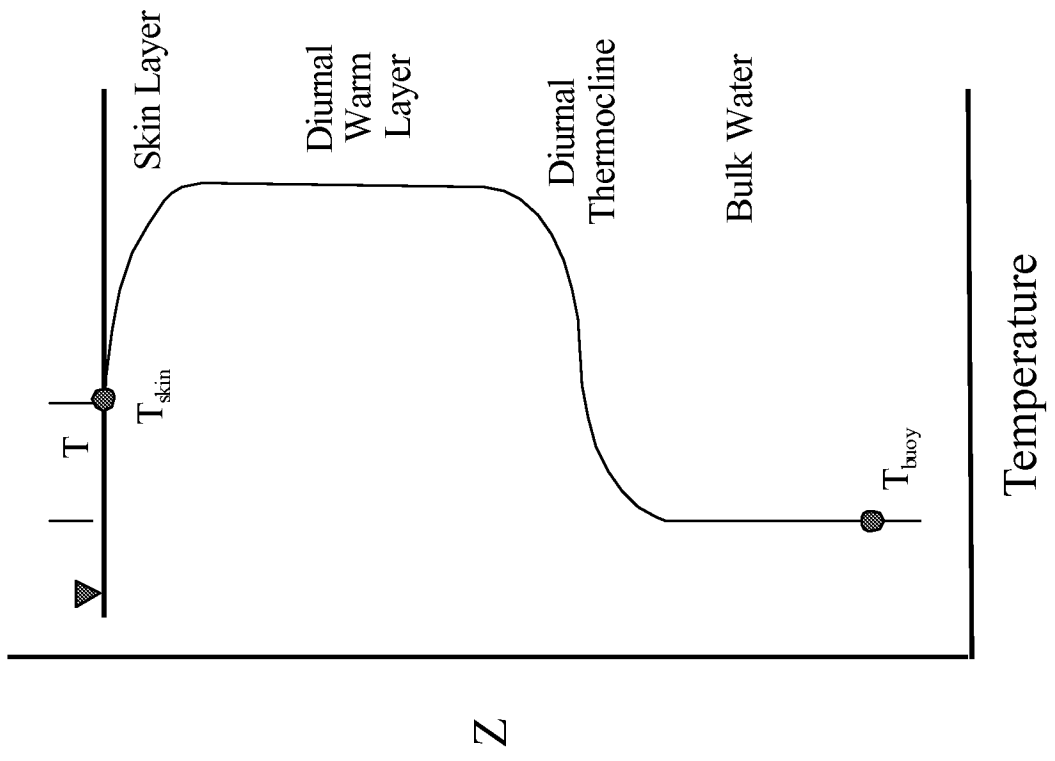


Cloud Detection Effects on GOES-9 Hourly Bias Variations

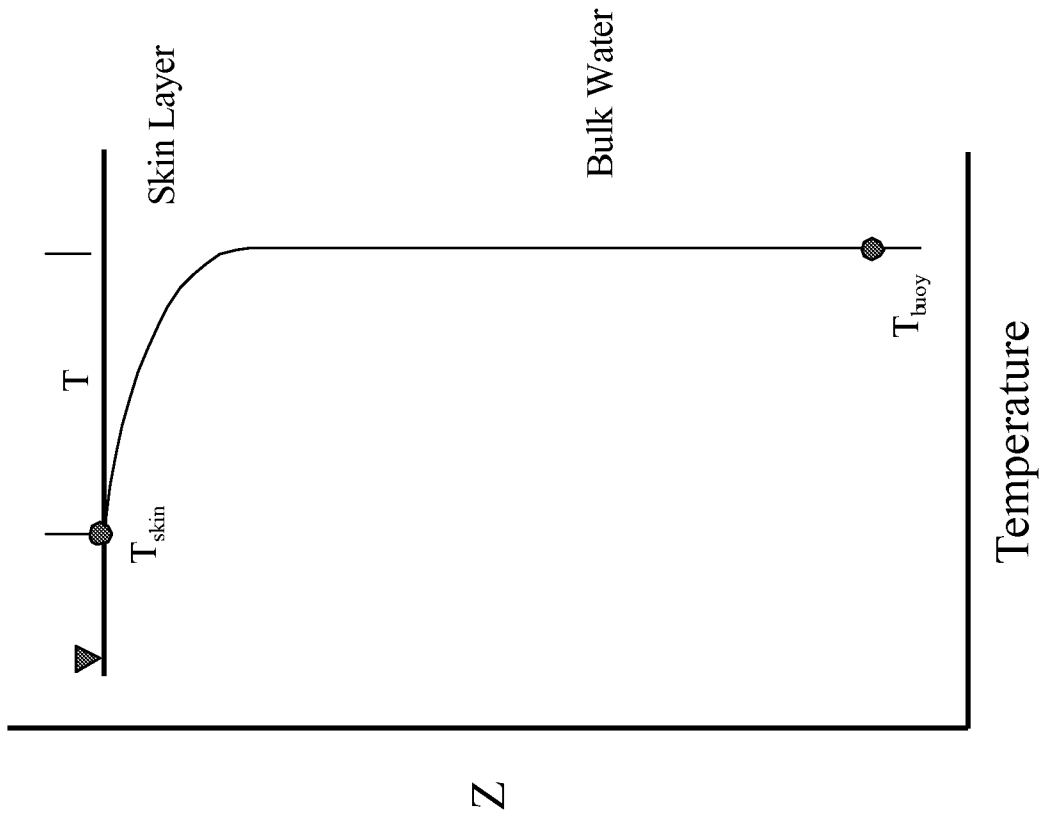




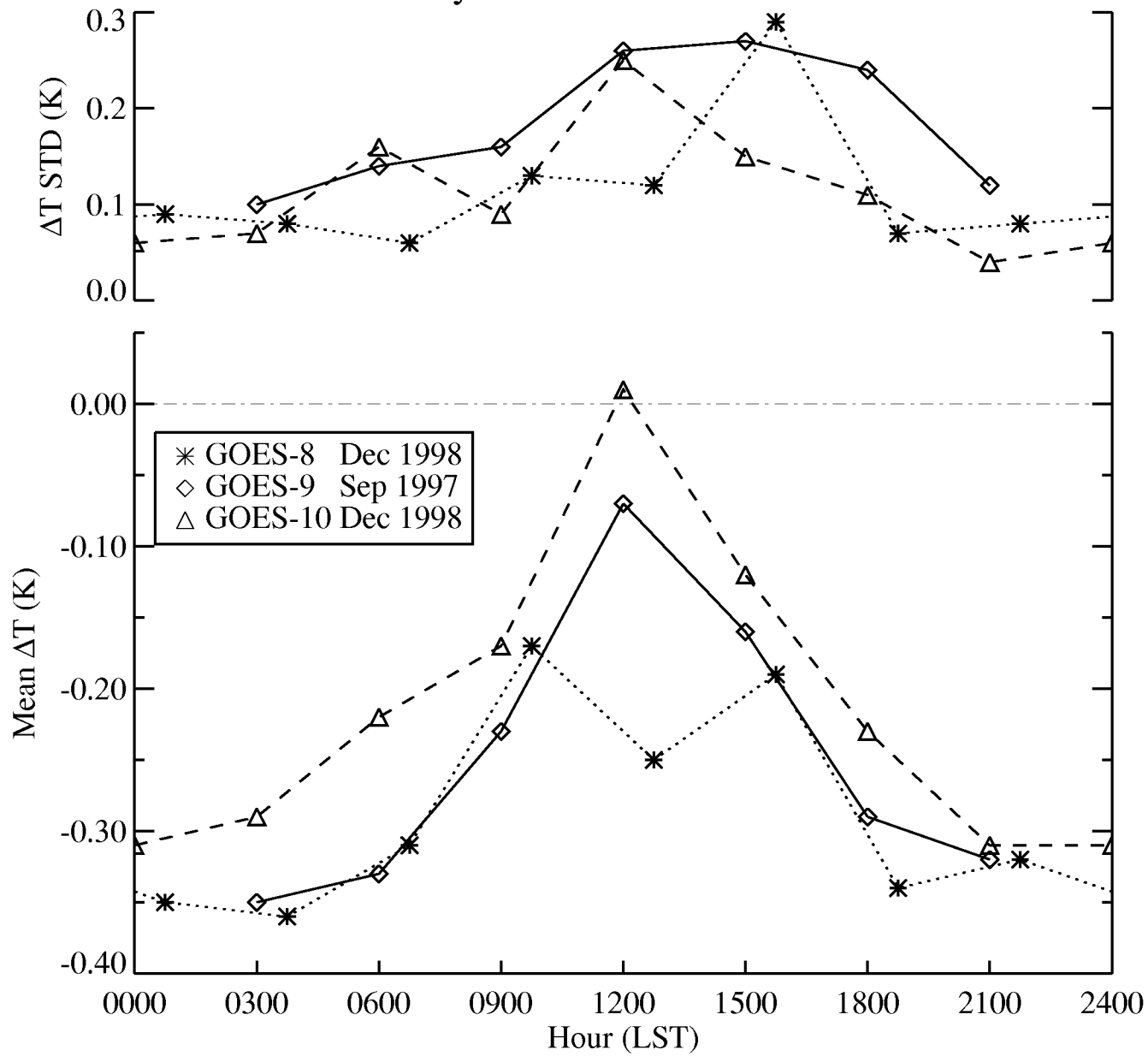
Day



Night



Hourly Variation of Modeled ΔT



Comparison of Observed GOES-10 Bias with Modeled ΔT Variation

