Comparison of Sea Surface Temperature Collection Methods at Sea

LCDR JACOB C. HINZ, USN

October 14, 2004

Graduate Student,
Departments of Meteorology and Oceanography,
Naval Postgraduate School,
Monterey, California

Table of Contents

Comparison of Sea Surface Temperature Collection Methods at Sea

LCDR JACOB C. HINZ, USN

Abstract

A comparison of three methods of Sea Surface Temperature measurements was conducted using both bulk and skin sensing methods. Statistics were generated to define bias and mean standard deviation for comparison purposes. Analysis of environmental factors impacting the bias and variability of these various sensors was also conducted, including the impacts of wind, air temperature and down welling long wave radiation.

---

1. Introduction

The tactical and operational interest in sea surface temperature (SST) spans many different Navy applications. One of which is the ability to characterize the near surface environment to predict effective radio frequency ranges for ships and aircraft operating in a given region. These atmospheric conditions, driven in large part by SST, are critically important when it comes to offensive and defensive military operations with regard to communications, radar detection, and electronic warfare. For example, Eckard (2001) found that in some environment stability conditions, that SST differences of less than 0.5ºC was enough to produce significant differences in evaporative duct height, one of the most significant atmospheric features affecting radar and radio propagation.

The primary motivation for this study is the Navy's interest in automated SST collection efforts as a larger move toward decreasing the manning requirements for US vessels. Eckard (2001) describes the range of methods currently in the US fleet today:

> "The inputs for sea surface temperature come from either the seawater injection measurements, taken meters below the surface, or an installed infrared (IR) system that measures skin temperatures. Other projects, such as SMOOS, use float-sondes that measure the sea surface temperature in the top one-centimeter.

All of these methods are currently deployed in the fleet and all have different error mechanisms, levels of accuracy and human intervention."

By quantitatively measuring SST in this study utilizing a variety of available sensors and methods, an accurate assessment of the precision and error for each method can be made and possibly applied to the Navy's decision-making process for an automated SST collection system.

Previous studies of SST methodology involved comparisons of shipboard SST methodologies to each other and to satellite derived SSTs. These included Carino (1998), Strauhs (1999), Eckard (2001), and Gahard (2003). This study considers only shipboard SST measurements as the significant stratus coverage over the operations area tends to contaminate satellite-derived data and, as such it is not considered.

For this study, the research vessel (R/V) Pt Sur was used as the investigation platform for all SST measurements regardless of method. The cruise lasted ten (10) days and was designed to gather data at and between recognized CALCOFI stations. The stations coincided with CALCOFI lines 67, 77, and 83 in the East-West direction and North-South alongshore transects, at or near station 70 between these CALCOFI lines.

SST data collection methods were of two types, "skin" and "bulk" SST measurements. Skin measurements are those in which the temperature measured is that of the top 50μm to 1mm of the sea surface. Bulk measurements are those in which the temperature sensed is greater than 1mm below the surface and is more representative of what is assumed to be the well-mixed upper layer of the sea. It is primarily this difference in methodology that leads to differing values in measurement and the potential for erroneous tactical application of the SSTs measured. To be diligent in this investigation, comparisons are made of methods in the same category (e.g. methods of the 'skin' category) and in differing categories (e.g. 'skin' vs. 'bulk') to determine if there is significant deviation between similar methods.

2. Data Sources and Collection Methods

Three (3) data collection methods were used during this study and are outlined in Table 1 below.

'Skin' Methods	'Bulk' Methods
Automated Mast IR Probe (radiometer)	Underway Data Acquisition System (Thermistor)
	Automated Boom Probe (Thermistor)

Table 1. SST Collection Methods by 'Skin' or 'Bulk' Category

a. Method One - Automated Mast IR Probe

The only 'Skin' SST measurement method available on this cruise is the automated, mast-mounted, IR radiometer. This device is mounted on the superstructure of the vessel at an approximate height of 10m above the waterline. It samples sea surface radiance in a 2 steradian field of view and reports sea surface temperature to 0.01ºC once each second. A true SST calculation is derived from measured radiance values using equation (1):

$$T_{sea} = \left[\frac{T^4_{meas} - (1-\varepsilon)T^4_{sky}}{\varepsilon} \right]^{1/4} \quad (1)$$

where epsilon (ε), the emissivity of seawater is 0.98. This correction is reportedly done automatically by the system based upon down welling long wave (LW) radiance as measured form a point near the IR sensor.(Gahard, 2003)

b. Method Two –Automated Underway Data Acquisition System

The first 'Bulk' method is the thermistor sensor in the Underway Data Acquisition System (UDAS). Using this method, seawater is drawn from a seacock located on the ships hull at approximately 2.5 meters depth. The temperature is measured prior to any significant modification by the ships data collection systems (oceanographic sampling only) and recorded automatically to 0.01ºC.

At this sensors depth, the system is sampling water beneath the hull and may be considered mixed by the ship's hull and bow wave.

c. Method Three – Automated Boom Probe

The third 'Bulk' method is the automated boom probe (Boom). This method employs a basic thermistor, towed approximately 1.5 meter off the port quarter of the vessel. While making way, the boom skims the surface of the water, and when the vessel heaves to on a CTD station, it sinks and measures at a depth of approximately 1 meter. During both of these situations, the sea temperature is measured at a depth shallower than the UDAS system, but still not truly a skin temperature. Since it is assumed in this method that the upper few inches of the sea are well mixed that this is considered an accurate representation of the SST. The boom thermistor records automatically, continuously and at a precision of 0.01ºC.

The boom, like the bucket, is in the influence of the vessels wake and therefore the SST measured may be considered a well-mixed sample of the upper meter of the sea. There are no warm water discharges forward of the sensor, and little influence by other discharges is expected from this vessel.

3. Analysis and Results

In order to compare results with those obtained in Eckard (2001), a similar procedure was employed. The time marker of each piece of data recorded was converted to a decimal Julian day and compared.

The first available data was the UDAS injection temperature and the Boom probe. Each of these were collected in the UDAS system over a 10.4 Julian day period resulting in a data set of more than 42000 data points.

The next available dataset was derived from the IR sensor mounted on the ships superstructure. This dataset covered the length of the cruise and encompassed nearly 650,000 data points over a 9-day period.

Lastly, UDAS data was available that provided air temperature (ºC), LW radiance (W/m^2, a proxy for cloud cover), and wind speed (m/s) and direction (ºT) data that may provide insight into variation of SST measuring techniques.

To best compare the three SST datasets and the atmospheric data sets (wind and air temperature), they had to be somehow normalized. Since each of the datasets was collected at varying sampling rates (of less than 2 minutes), and with varying starting and ending dates, a new dataset was created for this analysis. All datasets were interpolated using MATLAB's 'interp1' command to a new time vector that ran from day 218 (August 5) to 223 (August 10) with equal time steps of approximately 2 minutes. This served to place all datasets on the same sampling rate, with minimal distortion of the data statistics.

a. General SST Method Comparison

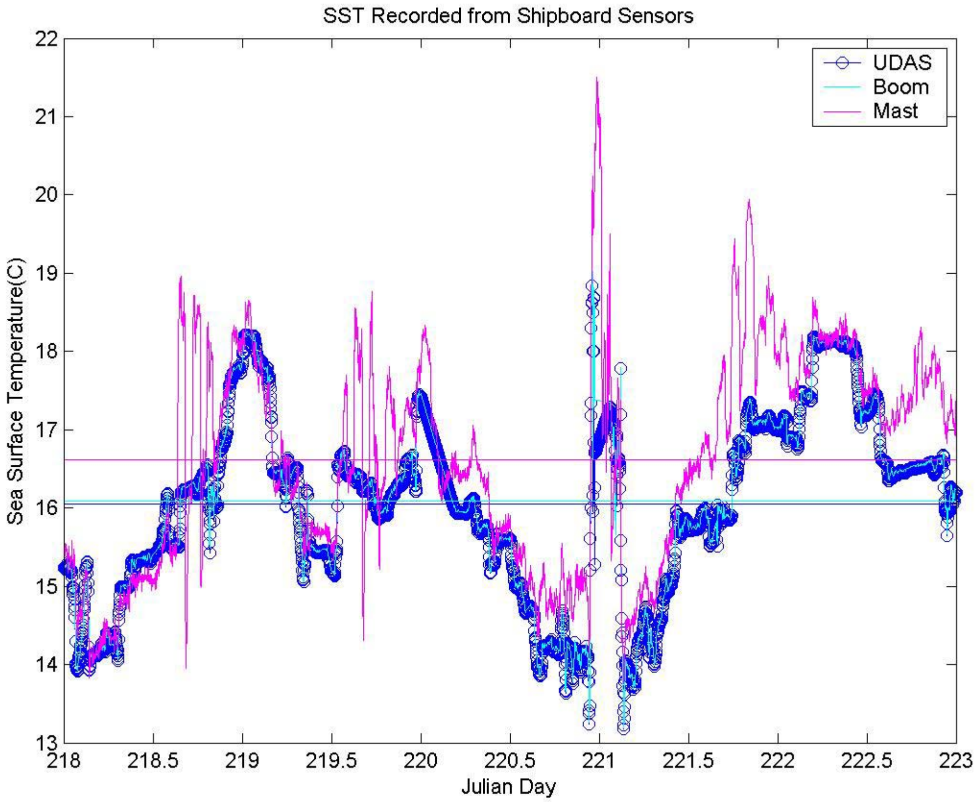

Figure 1. SST Recorded from Shipboard Sensors

Figure 1 contains a plot of the three sensors analyzed form shipboard collected data during the OC3570 cruise. A horizontal line of the same plotting color indicates the mean temperatures for each compared method.

General comparisons show that the Boom and UDAS systems measure values very close to each other, including the variations evident in the time series. The Mast appears to have far more variation than the other sensors and in general seems to have a warm bias as compared to each of the other sensors. The large peaks in the Mast as compared to the Boom/UDAS SSTs near days 218, 220, 221.5, 222, and 223 will be explored later in this study. The dramatic warming in all three sensors near day 221 can be explained by the vessel's transit into Port San Luis. This will be discussed in more detail later in this study as well.

To determine the mathematical relationship between the three methods, the mean and standard deviation of each method was calculated.

Method	**Mean**	**Standard Deviation**
UDAS	16.05	1.14
BOOM	16.08	1.15
MAST	16.61	1.37

Table 2. Comparison of Mean and Standard Deviation for Varying Methods

b. 'Bulk' Method Comparisons

Based on mean and standard deviation alone (Table 2), it would appear that the UDAS and Boom were reporting within .03ºC of each other (the Boom slightly warmer) with nearly identical standard deviations. A further look at the correlation value of .996 between them indicates that they are indeed very closely matched. To further emphasize the accuracy of this high correlation, it is similar to those found by Strauhs (1999).

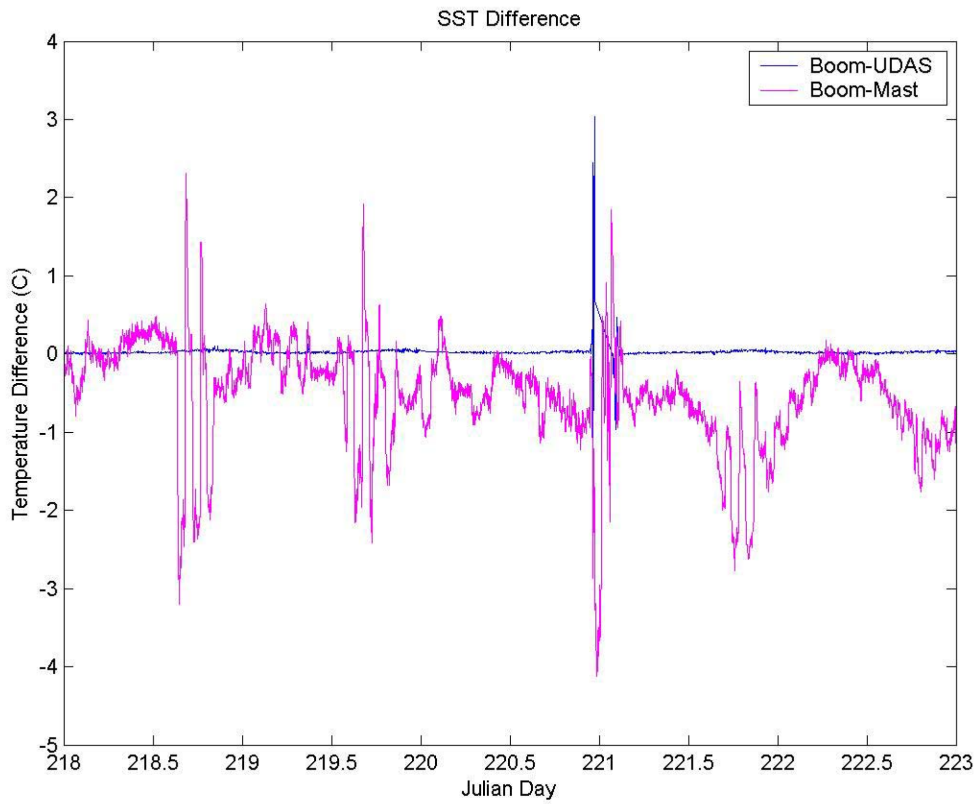

Figure 2. SST Differences Between Measuring Methods

This bias is confirmed by analyzing the differences between the Boom values and the UDAS (Figure 2, Blue). This mean value is again +.03 confirming that the Boom is warmer than the UDAS on average.

The differing locations of the sensors may provide some of the explanation for this bias. The deeper UDAS intake would presumably be colder, while the surface-floating Boom would be warmer due to solar radiation. This is a similar bias and explanation as cited by Strauhs (1999).

Also, the bias may be an electronic bias in the hardware or software of either system. But, whichever the solution, the bias can pretty well be factored into any use of the data due to its very consistent nature.

The close proximity of these two measured SST values for the majority of the time series indicates that there is little difference between the water that they measured for the majority of the cruise and they could very well be used interchangeably if they were properly calibrated to rid the data of the bias.

The only area of interest is the spike near day 221, the transit into Port San Luis. This is the reason for the large change in absolute temperature and of more importance, the change in the difference between the Boom and the UDAS.

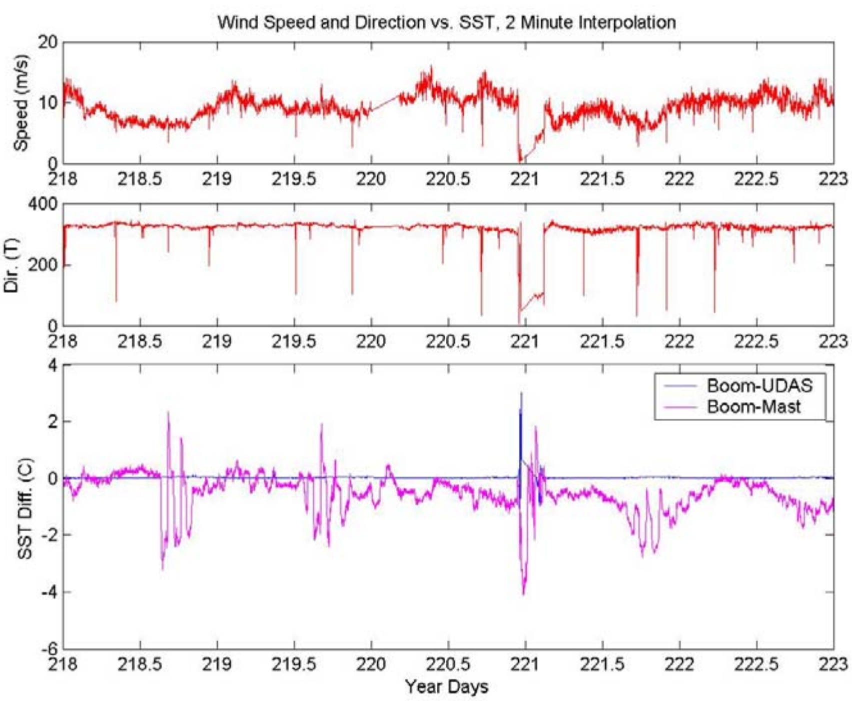

Figure 3. Wind Speed and Direction Vs. SST Differences

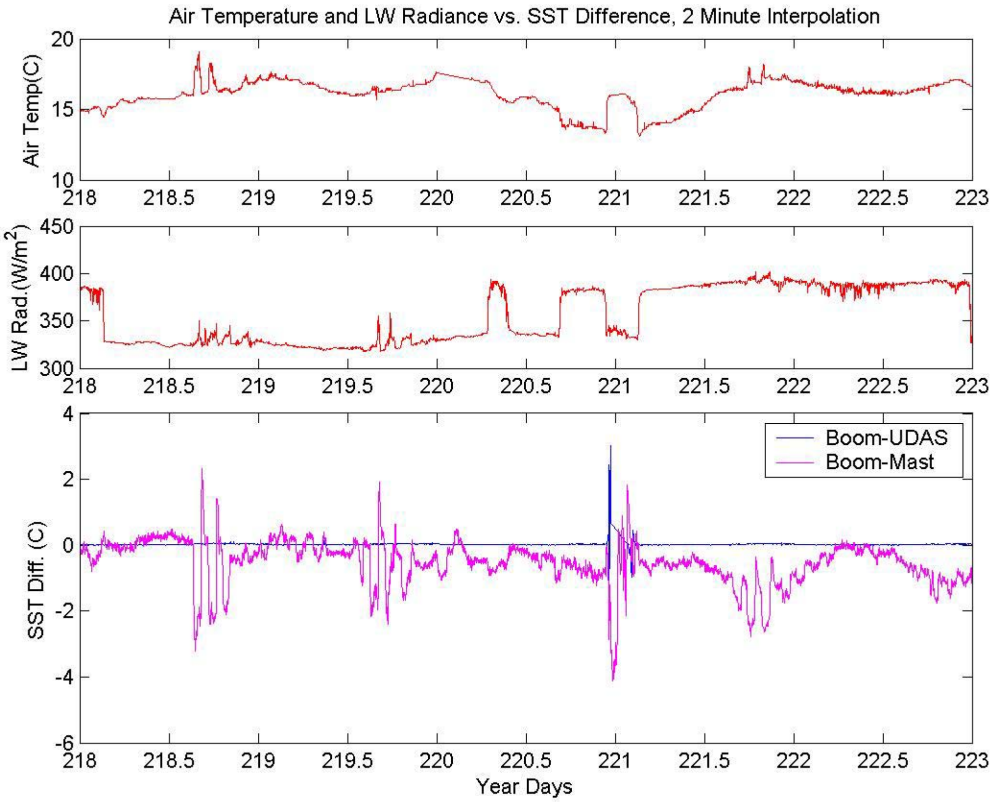

Figure 4. Air Temperature and LW Radiance vs. SST Differences

Figures 3 and 4 are plots of the Boom/UDAS difference (blue line) in measured SST aligned with the corresponding changes in wind, temperature, and down-welled LW radiation.

Evident in the figures is the significant decrease in wind speed as the predominant winds blowing from the northwest decreased to near zero meters per second and became southerly. Additionally, the significant warming and decrease in LW radiance that came with the clear skies that ere found in the lee of the mountainous region to the north and west of Port San Luis.

These events are related to the increased difference between the two bulk methods. The change in wind direction is merely an indication that he vessel came into the lee of the local terrain, but the decreased wind speed served to allow the sea surface to become smooth. This smoother seas surface allowed for less mixing, and more stratification at the upper two to three meters of the water column, where warmer waters were allowed to remain closer to the surface while cooler layers sank lower. This contributed to a more positive difference (Boom>UDAS).

Additionally, the clearing of clouds as indicated by the decreasing LW radiance and the associated increasing air temperature (due also to proximity to land) caused the surface of the water to be heated more than that at the increased depths. This too will contribute to the increased (positive) difference.

One of the things not indicated in the plotted data is that the ship's motion slowed significantly, and became nearly stationary in Port San Luis. This minimized the mixing of the water column being sampled due to the hulls motion through the water.

Once the ship returned to the open sea, all of these events were reversed and the difference between the methods returned to the .03 mean difference.

This evidence indicates that the Boom and UDAS can be used interchangeably in turbulent, well-mixed waters driven by the ship's movement and/or wind-driven seas. It can also be stated that there is a large difference between them when used in stratified, warming waters under clear or partly cloudy sky.

c. 'Skin' vs. 'Bulk'

A similar analysis can be conducted looking at the Mast (skin method) and the Boom (bulk method). The Boom can be considered representative of the bulk methods due to its close correlation to the other bulk method. The difference between the Boom and the Mast can be analyzed using a mean difference and standard deviation. Table 3 contains these statistics.

	Mean Difference	Mean Standard Deviation
Boom - Mast	-0.53	0.69
Boom - UDAS	0.03	0.11

Table 3. Statistics for SST Sensor Differences.

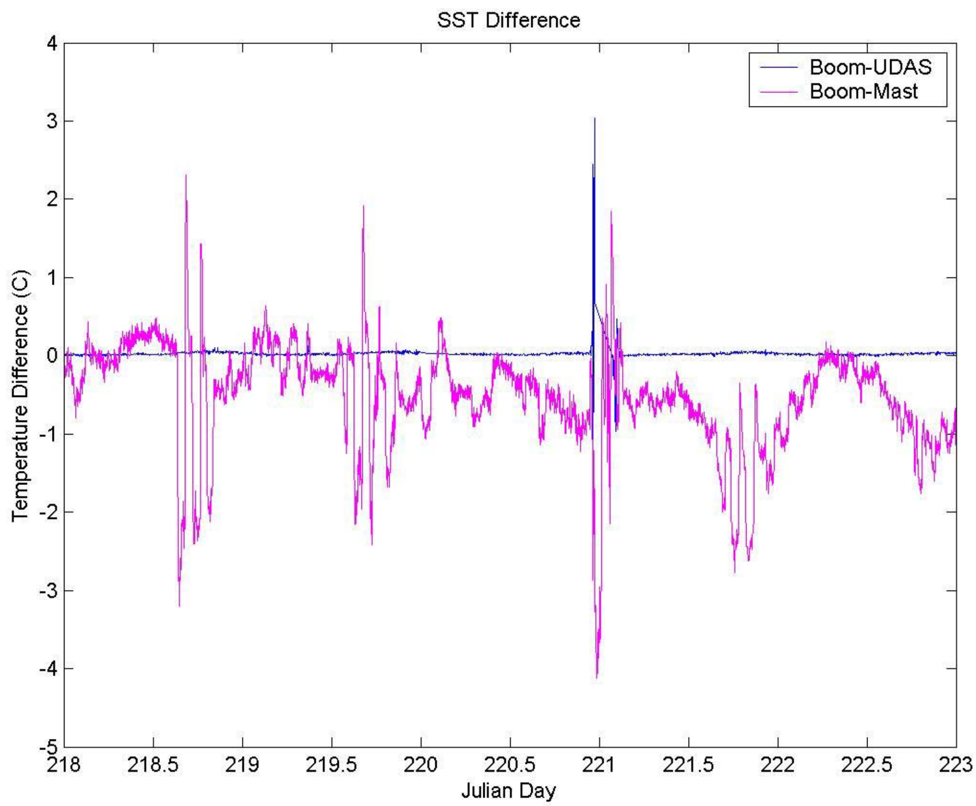

Figure 5. As in Figure 2.

The mean and standard deviation of the Mast difference are much higher as compared to the UDAS difference statistics. The large mean difference indicates a much larger bias than the UDAS had, but with a large standard deviation, this bias may fluctuate significantly over time and conditions. This adds to the argument that the Mast sensor is more sensitive to other environmental factors than the bulk methods. We will now explore some possible explanations for this bias and for this variability.

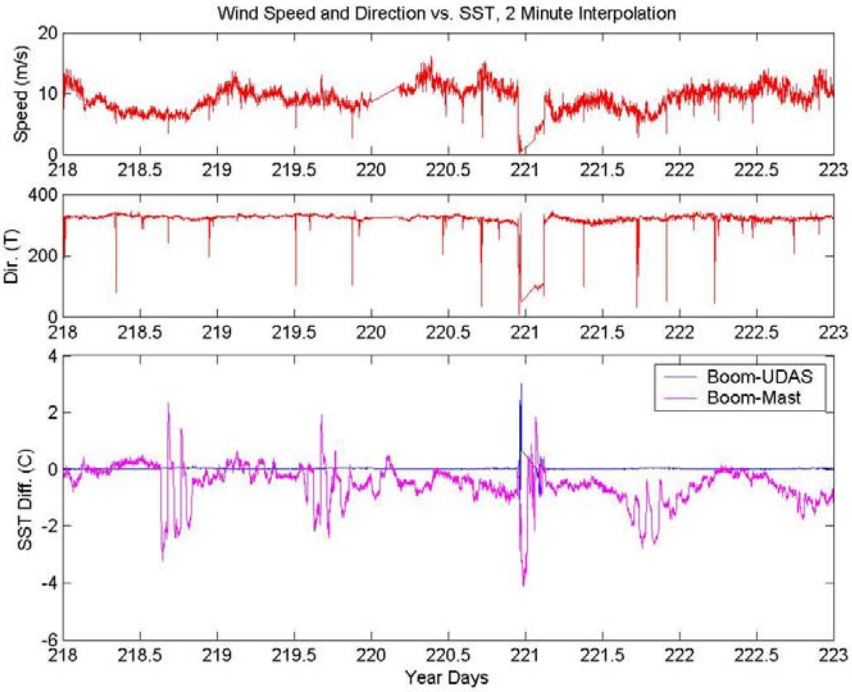

Figure 6. As in Figure 3, Now Focusing on the Boom-Mast Difference (Magenta).

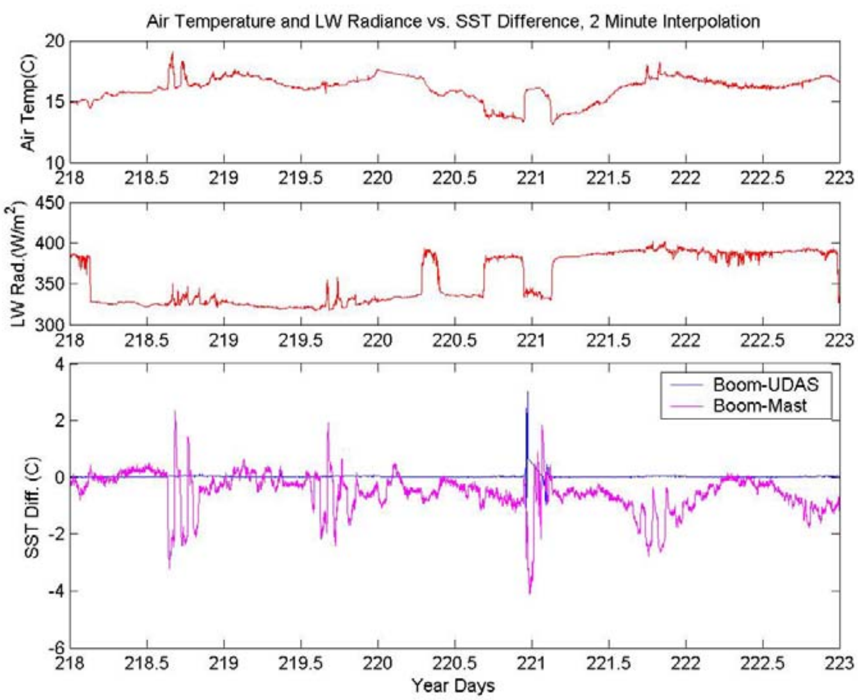

Figure 7. As in Figure 4, Now Focusing on Boom-Mast Difference (Magenta).

In general the difference between the two sensors can be demonstrated through the statistics in Table 2, above. The mean of the values reported from the Mast are consistently higher than the bulk methods (resulting in the negative differences in Figures 5 and 6(magenta)). The standard deviation is much higher than the bulk methods as well indicating that the Mast maybe more sensitive to natural and mechanical (man-made) environmental factors than the bulk method sensors.

The first, and easiest to describe, large difference is the spike on day 221 related to the vessel's visit to Port San Luis. In this case, the factors mentioned previously are involved in warming the sea surface more than at depth. Due to the fact that the Mast sensor measures the very surface (upper few millimeters) as compared to the Boom, which tends to sink about a half meter when the vessel is not moving, the difference between the two sensors is even more significant. The very surface of the water is significantly warmer than the water just below the surface and this leads to a much larger difference than that seen between the Boom and the UDAS.

The next phenomena noticed in the Mast difference are manifested between days 218.5 and 219, 219.5 and 220, and again between 221.5 and 222. These major deviations appear not to be related to wind speeds and directions as much as they are to air temperatures and LW radiance. In all of these events, increased differences (greater negative values in the Boom-Mast difference) are observed coincidently with sharp spikes in LW radiance and often with sharp spikes in air temperature. All of these indicate that there was a variable pattern in cloud thickness and or height and a relative change in air temperature as a result of the varying cloud conditions. This variation occurred each time just prior to a large increase in SST, as if the vessel was crossing a ocean frontal boundary, and in the daylight hours between 0600 and 1600 local.

To confirm the cloud variability and the possibility of crossing frontal boundaries, hourly cloud observations and a plot of SST along the voyage plan were consulted.

For each of the periods of consideration, hourly observations of cloud cover and fog showed significant variation over the periods of interest. Cloud cover was described as stratocumulus varying from 1/8 to 3/8 throughout the first period (218.5-219) with varying amounts of fog in the distance. The second period of interest (219.5 - 220) had reports indicating stratus clouds covering 1/8 to 2/8 of the sky. Again, as in the first period, there were indications of increasing and decreasing cloud amounts/development throughout the period. The third period (221.5-222), much as the first two, there were observed variations in the cloud cover, but this time it was between 7/8 and 8/8 coverage in stratocumulus.

In all of these cases, the most common factor was not the amount of cloud cover, but the variations in increasing or dissolving cloud amounts. These increases and decreases in cloud cover varied quickly on the order of one hour's time. When compared to the more consistent SST difference periods before and after the periods of interest, there were cloud cover transitions that occurred, but they occurred gradually over much longer time periods. In other consistent SST difference cases, there were no transitions at all and cloud cover remained fairly constant.

This finding would suggest that the rapid variation in cloud cover changes the LW radiance as seen in the figures rapidly and that the Mast SST sensor does not handle the variation well and over corrects. The fact that these SST difference variations occurred only during daylight in not surprising as they may be more related to solar impacts and would not happen as much at night as there would be less incoming LW and shortwave radiation from the sun.

Another related possibility is that the LW radiometer may be looking in an area affected by clouds while the SST sensor is looking in an area affected by clear skies, or vice versa. This would produce poor corrections. This is thought to be unlikely though, as a similar scenario might exist during the longer-term cloud cover transitions or on the edges of cloud or fog banks, where SST differences were noted to be consistent.

What this finding does suggest is that one should take great care when selecting a sensor like the Mast sensor for daytime and highly variable cloud cover regimes and might be better suited to use the Boom and UDAS type sensors for SST measurements.

Another possible contributor to these SST difference variations may have been the ship's movement. At each of the first two points of interest, there was at least one point at which the ship slowed to do a CTD cast. This may provide some explanation for the variations, but is considered unlikely as the ship conducted CTD casts at many other periods when the SST difference was more consistent.

The SST plot in Appendix A shows crossing of warm frontal features near the points of variation in the Mast SST difference and the corresponding LW radiance.

Year Days of Period of Interest	**LAT / LON of Warm Feature**
218.5 to 219	35.5N / 123W
219.5 to 220	34.25N / 122.2W
221.5 to 222	33.0N / 121.5W

Table 4. Locations of Significant Warm Frontal Features Associated with SST Difference Variations

These frontal features (and the dynamic atmospheric boundary layer conditions they may create) may a good explanation for why the cloud cover (and specifically the marine stratus common to the coast of California) may have varied in such a short distance.

Each of these phenomena by themselves or some combination of them may be the reasons for these variations in SST difference and they do warrant further, focused study as to their impact on IR-type skin SST measuring systems.

The last phenomena does not have a strong signature nor a long time series by which to definitively draw conclusions, but is based on physical properties and points to a recommended area of further study. If one looks at figure 7 (magenta line), it is apparent that there is an increasingly negative trend along the Boom-Mast difference. This trend appears to be related in some way to the increase in LW radiance over the same period (while other environmental parameters remain fairly constant over the long term average).

By breaking the plots in figure 7 into two, one from day 218.5 to 220 and the other from day 221.5 to 223, to isolate the low and high LW radiance, respectively, Figures 8 and 9 are created.

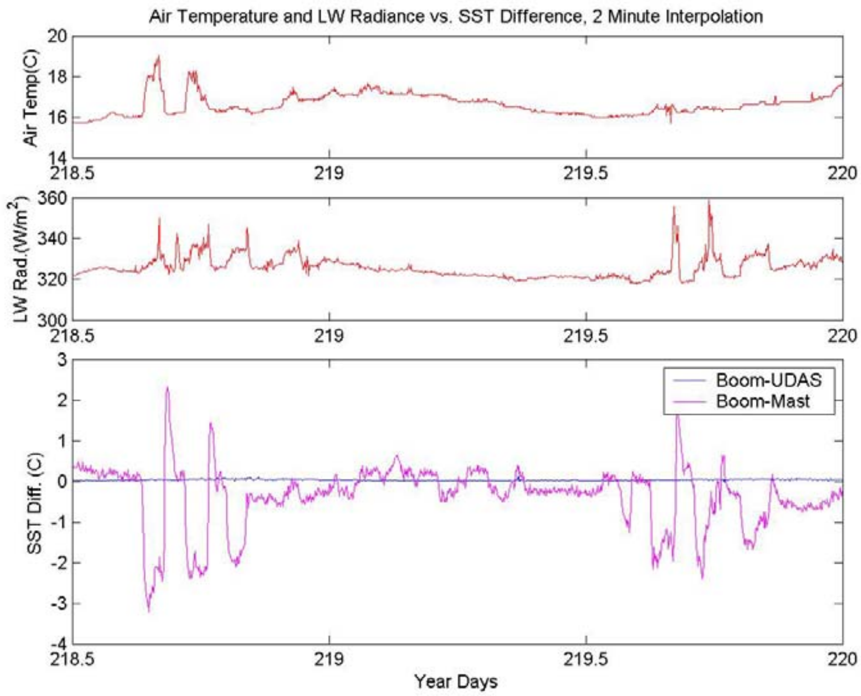

Figure 8. As in Figure 7, Now From Day 218.5 to 220.

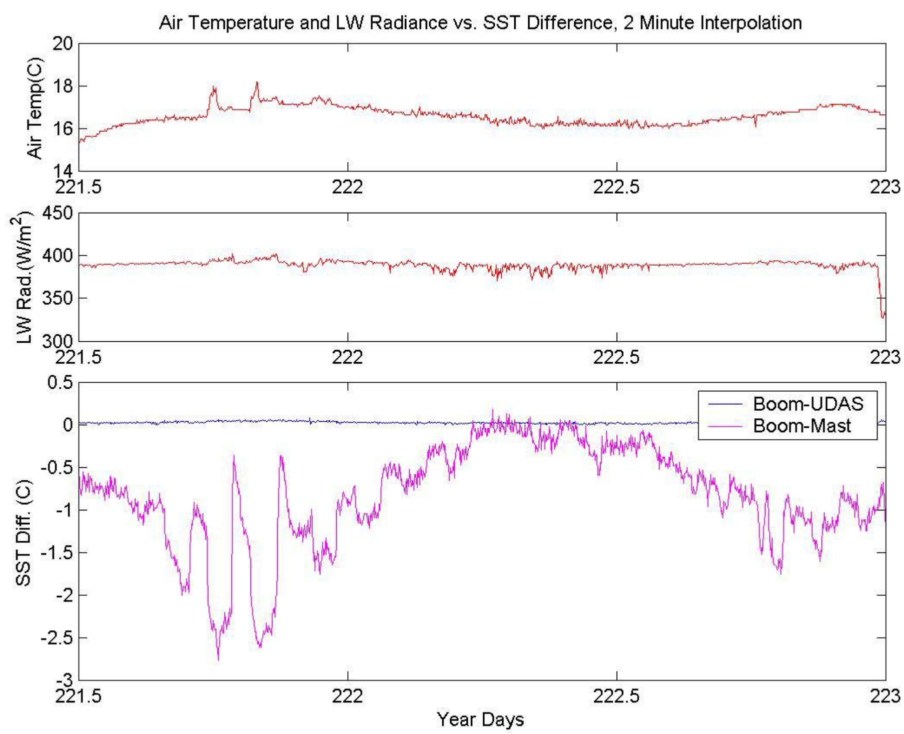

Figure 9. As in Figure 8, now from day 221.5 to 223.

When the average of these differences are calculated, the effect becomes pronounce. The average of Figure 8 and 9 differences are -.36ºC and -.84ºC respectively. With an increase in LW radiance of nearly 50 W/m^2 increased the SST difference by more than double.

Considering this event and the others related to LW radiance mentioned above, LW radiance is an issue worth much further study. These may be explained by a poor Mast sensor calibration or a miscalculation in the Mast sensor's algorithm for LW radiance correction. Irrespective of these or other explanations, this is a significant difference that brings to light an SST difference that exceeds that which will have significant impact on military operations as described by Eckard (2001).

4. Summary and Conclusions

This study found that the bulk methods, UDAS and Boom, although at different depths, provided very similar temperatures with only a slight bias of .03ºC that may be attributed to the differing location of the sensor in depth, and that this bias decreases when turbulence caused by vessel speed through the water or wind generated mixing is present.

This study also found that skin SST methods similar to the Mast sensor have very large biases on average and that under certain circumstances these biases can be significantly exaggerated. These include, still waters with little mixing, light or not wind regimes, slow vessel speeds, high cloud amounts, and increased air temperatures. During these events, differences from bulk methods can reach nearly 3ºC, but mostly average at a bias of greater than 0.5ºC.

Concrete conclusions can be made from this study but it is important to clarify the scope of these conclusions. Eckard (2001) demonstrated that variations of as little as 0.5ºC can have significant effects on radar and other EM propagation. With the demonstrated differences between IR skin measurement systems and bulk SST methods, great caution must be taken when using skin measurements in tactical decision aids and military planning. This said, bulk methods were used as the control group in this study and no statement about their accuracy was made against in-situ data, just a relative comparison to other methods.

This study should serve (as the ones before and after it) as a lesson that more detailed study is necessary before making the rash judgment that IR skin temperature is should be the preferred method for future automated environmental collection systems for use in military applications.

List of Acronyms

CALCOFI	California Cooperative Oceanic Fisheries Investigations
SMOOS	Shipboard Meteorological and Oceanographic Observation System
UDAS	Underway Data Acquisition System

List of References

Carino, D., “AVHRR SST Comparisons With Sail ‘In Situ’ Data”, September 14, 1998.

Eckerd, M., “Impact of Sea Surface Temperature Errors on Evaporative Duct Height”, Summer 2001.

Gahard, C., “Comparison of Sea Surface Temperature Measurement Methods”, March 18, 2003.

Strauhs, H., “AVHRR SST Comparisons with CTD, Sail and Boom Probe ‘In Situ’ Data”, March 1999.

Appendix A. SST Plotted Along OC3570/4270 Cruise Track (Source: Collins)