IN SITU VALIDATION OF SATELLITE SEA SURFACE TEMPERATURE MEASUREMENTS IN A NORTHERN CALIFORNIA UPWELLING AREA

A Thesis submitted to the faculty of
San Francisco State University
In partial fulfillment of
the requirements for
the Degree

Master of Arts

In

Geography

by

Pamela van der Leeden

San Francisco, California

May 2015

CERTIFICATION OF APPROVAL

I certify that I have read *In situ* validation of satellite sea surface temperature measurements in a northern California upwelling area by Pamela van der Leeden, and that in my opinion this work meets the criteria for approving a thesis submitted in partial fulfillment of the requirement for the degree Master of Arts in Geography at San Francisco State University.

Andrew Oliphant, Ph.D.
Professor

Leonhard Blesius Ph.D.
Associate Professor

John Largier Ph.D.
Professor
University of California, Davis
Department of Environmental
Science and Policy

IN SITU VALIDATION OF SATELLITE SEA SURFACE TEMPERATURE MEASUREMENTS IN A NORTHERN CALIFORNIA UPWELLING AREA

Pamela van der Leeden
San Francisco, California
2015

This study validates Advanced Very High Resolution Radiometer (AVHRR) satellite sea surface temperature (SST) data with *in situ* data from oceanographic buoys in the Point Arena upwelling area, offshore of northern California. Sea surface temperature in this area is often colder, and more spatially and temporally variable than nearby deep ocean waters. Global AVHRR data is validated at the 25 km scale. It is possible that SST in coastal upwelling areas may not resolve well at this scale. This study uses a 1 km scale to explore AVHRR error under conditions where satellite SST data has been shown to vary in previous studies: day/night, season, and under low, moderate and high wind speeds. A data set of 25,000 matchups over 21 years was analyzed. An overall systematic error of 0.47 °C (MBE) was found. Analysis of matchup subsets found higher error during daytime, low wind speeds and the relaxation season. Several systematic errors and possible causes are discussed. An adjustment factor is provided that can be applied to future AVHRR SST to account for the bias found in this study.

I certify that the Abstract is a correct representation of the content of this thesis.

_____ _____

Chair, Thesis Committee Date

ACKNOWLEGEMENT

Thank you to those who made this study possible: my committee, especially Andrew Oliphant; Dave Foley, formerly with NOAA's CoastWatch Program, for making the satellite SST data set available to me in bulk; Seth Hiatt at San Francisco State University for his help with Python scripting; and Dan Howard with NOAA's Cordell Bank National Marine Sanctuary, for supporting my coursework in satellite oceanography. Satellite data courtesy of the NOAA CoastWatch Program and the NOAA NWS Monterey Regional Forecast Office.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

Figures Page

Introduction

Sea surface temperature (SST) is a basic measure of ocean surface conditions. SST measurements are a primary data input into applications such as climate research and modeling, weather forecasting including storm prediction and tracking, phenomena such El Niño and La Niña that affect weather and fisheries, and biological research on species and ecosystem health and productivity. SST usually refers to water temperature measured within the upper 10 meters of the ocean. SST of waters over the northern California continental shelf usually varies within a 10 °C range seasonally. In these waters, after variation due to seasonal changes, the oceanographic process called upwelling is a major cause of sea surface temperature variation. Changes in SST during upwelling occur on the scale of 1-2 °C and may be as large as 5 °C in a few days, triggering changes in ocean conditions for marine life. Upwelling brings colder, nutrient-rich waters from the ocean depths to the surface, where they support primary production and plankton blooms that benefit fisheries, sea birds and marine mammals (U.S. Department of Commerce 2007, 2014d).

Continental shelf waters off of northern California are part of the California Current System and are characterized by strong and variable horizontal gradients in temperature, due in part to the complex wind and current systems and the underlying topography of the continental shelf (Palacios *et al.* 2006). This means that the ocean surface temperature in this area is often colder, and more spatially and temporally variable than offshore ocean waters.

Sea surface temperatures can be measured by using *in situ* instruments, such as oceanographic buoys, or by remote sensing, using sensors deployed on satellites or other platforms such as aircraft. Buoy and satellite instruments measure different layers of the ocean surface, called bulk and skin layers respectively. The buoy bulk temperature is measured at approximately 0.5 meter depth, whereas the satellite radiometric

measurement of the skin temperature reaches only the top 11 micrometers of the sea surface.

The satellite-based SST sensor is a useful tool because, in cloud-free conditions, it provides large swaths of readily available measurements of the temperature of the world's vast oceans. Remotely-sensed SST measurements made by satellite are a standard data source for oceanographic and biological research, climate change research and weather forecasting. Applications include sea-ice analysis, describing El Niño/La Niña events, tracking trends in global ocean temperatures, and forecasting tropical cyclones.

The drawback to satellite SST is that it is an indirect, or interpreted measure of SST, and as such, potentially more subject to error than direct, local measurements made by *in situ* instruments. To derive SST, satellite sensors measure electromagnetic radiation (EMR) which is emitted from the point of observation, in this case, the sea surface. EMR is recorded by satellite sensor bands sensitive to thermal infrared (TIR) wavelengths. The radiance values are translated to SST values using a nonlinear sea surface temperature algorithm, which is a theoretical equation based on the Planck function, discussed in more detail in the methods section of this paper. In addition, visible and near-infrared wavelengths are used in the production of satellite-derived SST to identify clouds and water vapor. Buoy instruments, in contrast, are in direct contact with the sea water they measure.

AVHRR or "Advanced Very High Resolution Radiometer" is a sensor deployed on National Oceanic and Atmospheric Administration (NOAA) satellites. Several early studies on evaluating error in AVHRR SST focused on matching satellite with *in situ* observations as closely as possible in space and time (Kilpatrick *et al*. 2001, Li *et al.* 2001). Kilpatrick *et al*. stated that because the SST algorithms without correction for *in*

situ conditions assume average atmospheric conditions, the performance of the algorithm depends on how close the actual conditions are to the mean. Because of this, NOAA's AVHRR SST product, the global satellite data set used in this study, is routinely validated with *in situ* data and the SST algorithms adjusted (Martin 2004).

Several studies focused on validating satellite-derived SST data with *in situ* data have described biases that differed during day/night, seasonally and under varying wind conditions. A comparison study done on NOAA-12 and NOAA-14 AVHRR satellites focused on three coastal regions, the Gulf of Mexico, and the coastal Northeast and Southeast of the United States (Li *et al.* 2001). AVHRR SST was compared with NOAA moored buoy SST matched within 1 hour and 1 km. Li *et al.* described a positive (warm) bias averaging 0.2 °C for all three regions. Daytime SST bias was two to three times that of nighttime. Donlon *et al.* (2002) compared AVHRR SST with buoy data by removing AVHRR SST measured during wind speeds below 6 m s^{-1} because of potential bias from daytime water column stratification. With the lower wind speed data removed, day and night AVHRR SST showed a cool bias of -0.17 °C. Gentemann *et al.* (2004) validated passive microwave-derived SST with buoys in the tropics and noted a discrepancy between satellite-derived and buoy temperature measurements on warm, calm days (winds less than 6 m s^{-1}) due to ocean surface heating. After removing data associated with low wind speeds, they found a cool bias of -0.08 °C. Parekh *et al.* (2007) compared passive microwave-derived satellite SST with buoys in the Indian Ocean, daily and seasonally. They found that night-time satellite data had a large bias compared to daytime data, -1.20 to 0.00 respectively. They also found a larger bias during the pre-monsoon season than the monsoon season, 1.54 to -0.79 respectively.

This study investigates the relationship between satellite-based SST measurements with *in situ* data from oceanographic buoys in an upwelling area, where *in situ* buoy data are often used in tandem with satellite SST data for oceanographic and biological research.

The objective of this study is to explore how error varies in satellite-derived SST, using AVHRR as the representative satellite SST product and *in situ* SST from oceanographic buoys. This study explores error under conditions where it has been shown to vary in previous studies: day/night, upwelling season, and under low, moderate and high wind speeds, in order to evaluate how true the satellite-derived SST product is to buoy locations in the study area. More broadly, this study may shed light on how the AVHRR SST data performs in coastal upwelling areas characterized by variable, strong and localized SST gradients. If a consistent bias is found, an adjustment factor will be recommended in order to improve future AVHRR SST estimations in the study region.

Study Area

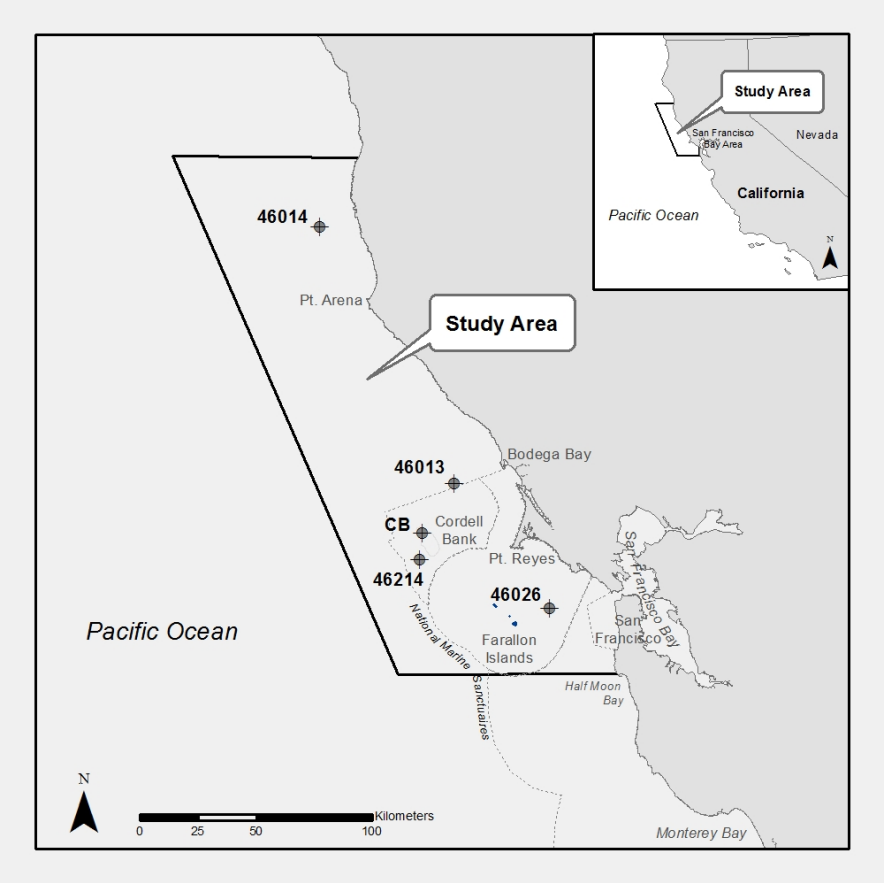


Figure1**.** Location map of study area and buoys selected for the current study.

This study's area of focus is the Point Arena upwelling center, or cell, located in the waters offshore of northern California. Point Arena is approximately 165 km north of San Francisco. The Point Arena upwelling cell commonly refers to the continental shelf waters between Point Arena and Point Año Nuevo near Half Moon Bay, an approximate area of 6,600 km^2. In the study area, the shelf break is approximately 40 km from shore. The study area extends south from the upwelling center at Point Arena; wind-driven currents carry upwelled waters southwards past Cordell Bank and Point Reyes (Halle and Largier 2011) and into the Gulf of the Farallones, which forms the southern extent of the study area, approximately 30 km south of San Francisco.

The study area is of interest to oceanographers and biologists, being at the center of one of four major upwelling regions globally, and including areas designated as national marine sanctuaries. (Largier 2007, Largier *et al.* 2006, U.S. Department of Commerce 2007). Upwelling in the study area is characteristic of the California current system, and is a hotspot of biological productivity (Palacios et al. 2006). The other three major upwelling systems are as follows: the Peru current system off the west coast of South America, the Canary current system off the west coast of North Africa, and the Benguela current system off the west coast of South Africa.

Upwelling is a wind-driven, seasonal phenomenon in the study area, usually occurring at irregular intervals, or pulses, between March and July. Although upwelling pulses can occur at any time of year, they are most persistent and sustained during the upwelling season. The upwelling season is characterized by strong alongshore winds, and strong alongshore currents over the continental shelf producing upwelling jets and cold sea surface temperatures (Largier *et al.*1993).

The upwelling season is followed by the relaxation season, a season of weak winds and warm SST in the fall, which is followed by a storms season characterized by rain and

southerly wind events in winter. The relaxation season is characterized by a stratified water column. During the winter, the water column is cold and well mixed (Largier *et al*. 1993).

Method

This study compares 21 years of SST measurements from five moored oceanographic buoys with matching Advanced Very High Resolution Radiometer (AVHRR) satellite SST measurements in the Pt. Arena coastal upwelling region.

Buoy data

Five oceanographic buoys were selected to provide *in situ* sea surface temperature data for comparison with the satellite data. Three buoys are owned and maintained by NOAA's National Weather Service National Data Buoy Center (NDBC), one by Scripps Institution of Oceanography (SIO/UCSD) for the Coastal Data Information Program (CDIP), and one by Bodega Marine Lab (BML/UCD) for the Cordell Bank National Marine Sanctuary (CBNMS). The buoys were selected because of their geographical location, rich historical, quality-controlled data sets and excellent metadata. Although the Cordell Bank buoy has a limited data set and is now discontinued (U.C. Davis Bodega Marine Laboratory 2014), it was included because of its location in the middle of this national marine sanctuary. Although Cordell Bank itself has granitic pinnacles that rise high above the sea floor, the CB buoy was moored at a depth comparable to that of the other buoys.

Other moored buoys were considered but were either too far inshore, such as the BML buoy located one mile off the coast, or were too far offshore, such as NDBC buoy 46059

located 400 miles offshore in deep ocean. The addition of drifting buoys to the *in situ* data set was considered but rejected for practical reasons. Drifting buoys are most often deployed in the open ocean and hence less likely to gather data on the continental shelf. Only a limited number of observations within the study area would be available, eliminating a long time series for any individual drifter. The increased complexity inherent in matching-up drifting buoy SST measurement time and location with satellite data, together with the limited amount of data that would be gained, argued against including drifters in this study.

Each buoy is moored on the continental shelf within the Point Arena upwelling cell. See Figure 1 for a map of buoy locations and Table 1 for buoy information. The distance between the furthest buoys, Point Arena and San Francisco is 191 kilometers (103 nautical miles). The closest buoys are Point Reyes and Cordell Bank which are 11 kilometers (6 nautical miles) apart.

The resolution of the buoy SST temperature sensors ranges from 0.0001 to 0.1 °C. and accuracy ranges from +/- 0.002 to 1.0 °C (Scripps 2014b, U.S. Department of Commerce 2014a, U.C. Davis Bodega Marine Laboratory 2011). Accuracy of the three NDBC buoys may be closer to 0.08 ° C, significantly better than the stated 1.0 °C (U.S. Department of Commerce 2009). This is because the +/- 1.0 degree accuracy stated by the NDBC represents a conservative estimate of system accuracy rather than individual instrument accuracy. NDBC system accuracy is a combination of sensor accuracy, errors known to be produced by the buoy and platform, and how well the NDBC judges it is able to monitor and control the measurement quality in the field. Analysis of SST field comparisons using duplicate sensors resulted in 0.08 °C accuracy (U.S. Department of Commerce 2009). For information on individual buoys, see Table 1.

Table 1. Information about buoys selected for the current study.

Buoy	Name	Location	Sensor Depth (m)	Sensor Type	Instrumentation Manufacturer	Resolution (deg C)	Accuracy (+/- deg C)
46014	Pt. Arena	39.235, -123.97389	0.6	thermistor	PSI, Baytech, Magnavox, or other	0.1	1 (0.08 in field tests)
46013	Bodega Bay	38.24194, -123.30056	0.6	thermistor	PSI, Baytech, Magnavox, or other	0.1	1 (0.08 in field tests)
46214	Pt. Reyes	37.94645, -123,46747	0.45	NTC thermistor	Datawell	0.05	0.2
CB	Cordell Bank	38.04948, -123.45725	2.5	thermistor	SeaBird Electronics (MicroCAT)	0.0001	0.002
46026	San Francisco	37.75889, -122.83333	0.6	thermistor	PSI, Baytech, Magnavox, or other	0.1	1 (0.08 in field tests)

Buoy	Acquisition Time (min)	Buoy Type	Manufacturer	Hull Width (m)	Water Depth (m)	Watch Circle (m)	Owned/Maintained By
46014	8	discus	NDBC	3	256	434	NOAA's NWS National Data Buoy Center
46013	8	discus	NDBC	3	116	191	NOAA's NWS National Data Buoy Center
46214	30	WaveRider directional	Datawell	0.9	550	700	Scripps Institution of Oceanography
CB	1 (averaged to 8)	spherical	Wet Tech Energy	0.75	85	122	Bodega Marine Lab (U.C. Davis)
46026	8	discus	NDBC	3	53	116	NOAA's NWS National Data Buoy Center

Most buoys record and transmit only a subset of data for each hour, in order to conserve battery power. This study's buoy data are reported at the one hour resolution, but represent an 8 or 30 minute average of that hour, depending on the buoy. Wind speed, used in this study for comparing bias error under different wind conditions, was recorded by the buoy anemometers at a height of five meters, at the same time as the SST observation. See Table 1 for individual buoy detail. The Cordell Bank buoy's SST data was the exception. This buoy's data was provided at one-minute resolution, so the first eight minutes of each hour were averaged in order to better match the other buoy data.

NDBC buoy data undergoes both automated and manual quality checks. The first level of quality check is performed by the National Weather Service (NWS) on real-time data, and removes gross errors only. These are values that are almost certain to be false, due to communication transmission errors or total sensor failure. A second level of automated quality check algorithms checks data against set parameters. Data considered erroneous receives a "hard" flag and is deleted. Data flagged as suspect receives a "soft flag" and is released in real-time but is manually reviewed within 24 hours.

Archived data undergo further, stricter quality review using a combination of automated and manual review. For SST data this includes range-limit and time-continuity checks (U.S. Department of Commerce 2009). Data that does not pass the quality review are removed before the data are archived. This study used archived data only.

The Point Reyes buoy data undergoes quality checking through Scripps's CDIP. SST data are checked against min/max values and rate of change between subsequent values; bad data are removed before archiving (Scripps 2014a). The Cordell Bank buoy data are graphed by BML to check visually for suspicious or bad data (U.C. Davis Bodega Marine Laboratory 2011). Data for this study was provided with flagged data removed.

The NDBC buoy sensors are calibrated prior to deployment and replaced with newly calibrated instruments every two years. The old instruments undergo a calibration analysis after being swapped out (NDBC 2014a). The CDIP buoy is swapped out every two to three years with a freshly calibrated buoy and the swapped out buoy sensors undergo calibration analysis (CDIP 2014b). The Cordell Bank buoy temperature sensor is replaced every six weeks with a freshly calibrated sensor (U.C. Davis Bodega Marine Laboratory 2011).

Moored buoys are connected to their anchor on the ocean floor by a tether with additional length built into the tether to accommodate wind, waves and currents. The floating buoy moves within a circle on the ocean surface called a watch circle; the size of this circle is determined by the depth of the mooring and length of the tether. Because the latitude/longitude of the buoy used to extract satellite SST values is a fixed point location, this could introduce some spatial uncertainty.

Data for the four buoys was downloaded from NDBC and CDIP for the years 1992 to 2012, the years for which the satellite data are available. Cordell Bank buoy data were available only from 2007.

Satellite data

Satellite data options were limited by several requirements of this study, specifically, a long time series of data that was readily available, fully processed, and quality-controlled, at a spatial resolution appropriate for local area coverage (LAC), and with a temporal resolution of one hour. The NOAA/OSU Heritage SST data set, a joint product of NOAA CoastWatch and Oregon State University and the only data set meeting all criteria, was used in this study. The NOAA/OSU local-area coverage data set was available back to 1992, with 2 to 8 passes every 24 hours. Matching hourly buoy data was available for the same time period. The NOAA/OSU Heritage SST satellite data set was provided by

NOAA's CoastWatch. The data set and metadata may be found here: http://coastwatch.pfeg.noaa.gov/erddap/info/erdAOsstdhday/index.html

The Polar Orbiting Operational Environmental Spacecraft (POES) satellites are outfitted with AVHRR sensors and are owned and operated by NOAA. These satellites orbit the Earth's poles, as opposed to other types of satellites that orbit the equator, or are geostationary. This study's data set was collected by several different POES satellites over the years, sometimes with two operational at the same time. POES-19 is the current POES satellite in operation (U.S. Department of Commerce 2014b). The AVHRR instrument is a multiband sensor with both visible and infrared radiance channels. The visible and near-infrared bands are used to identify clouds and water vapor. Table 2 provides a summary of band wavelength and SST-related use.

Table 2. AVHRR band wavelength and SST use.
Sources: U.S. Department of Commerce 2014c and Martin 2004.

Channel Number	Wavelength (um)	Spectrum subregion	SST Use
1	0.58 - 0.68	Visible	Daytime cloud mapping
3B	3.55 - 3.93	NIR (Near Infrared)	Night cloud mapping, SST
4	10.30 - 11.30	TIR (Thermal Infrared)	Night cloud mapping, SST
5	11.50 - 12.50	TIR (Thermal Infrared)	SST

The POES satellite AVHRR measures SST at 1.1km resolution (0.5 nautical mile), equivalent to 0.0125 degrees latitude by 0.0125 degrees longitude. SST accuracy is +/- 0.7 °C (U.S. Department of Commerce 2014e.) This represents nominal accuracy as reported by NOAA. Two to four images per satellite were acquired daily of the study area. In years with two operational satellites, four to eight images per 24 hours were available.

The POES satellite data for the west coast of North America downloads in High Resolution Picture Transmission (HRPT) format to a ground receiving station run by the NOAA National Weather Service in Monterey, California. Raw satellite data are recorded as a digital number (DN) and processed by NOAA's Environmental Satellite, Data, and Information Service (NESDIS) through several steps that include converting the DN to radiance, converting radiance to SST, removing data contaminated by clouds, quality checking and validating the data.

Radiance is expressed as mW m^{-2} sr^{-1} cm^{-1} (units of milliWatts per area per steradian per wavelength), by the thermal infrared (TIR) bands on the AVHRR instrument on board the POES satellites, and can then be converted to temperature values. The raw satellite data are translated to SST using a non-linear sea surface temperature algorithm (NLSST) (U.S. Department of Commerce 2014e.) Development and operational use of the algorithm is described by Walton *et al.* (1998). Calibration and validation of the AVHRR SST data for continuity and algorithm integrity is described in Li *et al*. (2001).

Radiance received at the satellite sensor is the sum of attenuated surface radiance and atmospheric path radiance. Translation of radiance to SST values is based on a theoretical equation where radiance received by the satellite sensor is shown as blackbody temperature T_1using the Planck function. Some of the evolution of the various components of the equations used for the POES AVHRR product, as well as sample equations, are presented below. All equations and explanations that follow are summarized from Martin (2004).

The Planck function for SST can be expressed as:

$$fp(T_i, \lambda_i) = fp(T_s, \lambda_i)t_i + fp\left(\overline{T}, \lambda_i\right)(1 - t_i) \quad (1)$$

where fp is the Planck function,T_i is the blackbody temperature corresponding to the radiance received at the satellite, $\lambda_i$is an individual band, T_s is the water skin temperature, T_i is the brightness temperature of a band, $\overline{T}$= mean temperature of lower troposphere, t_i=transmittance, and *V*= columnar water vapor, so that t_i= ƒ($\overline{T}$,*V*,λ_i). The term on the left is radiance received at the satellite, the first term on the right is attenuated surface radiance, and the second term on the right is atmospheric path radiance. The $1 - t_i$ factor represents atmospheric emissivity. This theoretical equation is applied to the data collected by the AVHRR sensor bands to create the AVHRR operational algorithms. The AVHRR POES SST algorithm uses bands 4 and 5 for daytime SST; the nighttime SST algorithm use bands 3, 4 and 5.

AVHRR operational algorithm

The theoretical equation was refined by linearizing the Planck function, and accounting for the dependence of transmittance on water vapor as a function of wavelength. The following equation written for AVHRR sensor bands 4 and 5 was derived, called the "split window" algorithm. The split window algorithm is typically used to estimate surface temperature for AVHRR (Li *et al.* 2001, Martin *et al.* 2001):

$$T_s = T_4 + \Gamma(T_4 - T_5),\ where\ \Gamma = (1 - t_4)/(t_4 - t_5) \qquad (2)$$

where T_s is the skin temperature, T_4 and T_5 are temperatures based on the radiance received at bands 4 and 5, and t_4 and $t_5$are transmittances due to the effect of columnar water vapor on bands 4 and 5.

Equation (2) is the basis of the AVHRR operational algorithm but results in large cool biases in satellite temperatures relative to a comparison *in situ* data set (McClain *et al*. 1985). Therefore, constants are added to reduce the bias, resulting in equation (3), called the multi-channel SST, or MCSST:

$$\text{SST} = C_1T_4 + C_2(T_4 - T_5) + C_3 \qquad (3)$$

where SST is the temperature estimated from satellite data as corrected by fitting to buoy observations, and C_1, C_2, C_3 are coefficients obtained from a least-squares regression of SSTs calculated from satellite T_4 and T_5 against buoy matchup temperatures.
Equation (3), the MCSST, gives good results within a limited subset of measurements, where the sun angle is small (close to overhead) and water vapor values are low. Other equations are used to cover higher sun angle and water vapor values.

The water vapor SST (WVSST) equation was used to create an improved MCSST. It relies directly on water vapor values and sun angle and incorporates water vapor values from radiosonde or passive microwave observations. The WVSST value of Γ is substituted into the split-algorithm equation (2):

$$\Gamma = [m_4/(m_5 - m_4)][1 + (m_5 V \sec\theta)/2 + \cdots] \quad (4)$$

where m_4 and m_5 are passive microwave channels, V is the water vapor value, and θ is the sun angle.

Equation (3) was revised to include larger sun angle values, but not large water vapor values. Larger sun angle values mean increased path length for the radiance to travel. The resulting equation (5) was called the NLSST.

$$\text{SST} = C_1T_4 + C_2(T_4 - T_5) + C_3(T_4 - T_5)(\sec\theta - 1) + C_4 \quad (5)$$

Finally, a revised NLSST algorithm produced a better match with buoy temperatures:

$$\text{SST} = C_1T_4 + C_2T_{sfc}(T_4 - T_5) + C_3(T_4 - T_5)(\sec\theta - 1) + C_4 \quad (6)$$

where T_{sfc} is the surface temperature from a climatological lookup table or the MCSST equation (3).

A sample of the daytime nonlinear operational equation NLSST with constants used for the NOAA-14 satellite is:

$$\text{SST} = 0.9336T_4 + 0.079T_{sfc}(T_4 - T_5) + 0.77(T_4 - T_5)(\sec\theta - 1) - 253.69 \quad (7)$$

where SST and T_{sfc} are in °C and T_4 and T_5 are in Kelvin.

AVHRR POES nighttime SST employs a "triple window" version of the NLSST algorithm using all three infrared bands. This uses the difference between bands 3 and 5 to remove water vapor effects. A sample NLSST "triple window" algorithm used for the NOAA-14 satellite nighttime passes is:

$$\text{SST} = 0.980064T_4 + 0.031889T_{sfc}\,(T_3 - T_5) + 1.817861(sec\,\theta - 1) - 266.186 \quad (8)$$

where T_4 is the main SST estimate based on band 4. The other terms provide corrections and convert the units from *K* to *C*. Finally, for all algorithms described here, constants are adjusted depending on which satellite collected the data.

To minimize cloud contamination, SST data are routinely classified for cloud-coverage using the Clouds from AVHRR-Phase 1 (CLAVR-1) method described by Stowe et al. (1999), so that radiance values from clouds can be excluded. This algorithm makes use of the fact that clouds are usually colder and more reflective than the surfaces below them. Several variations of CLAVR-1 are used, optimized for daytime ocean and nighttime ocean use. During the daytime, both solar reflectance and infrared emittance are used to identify clouds; at nighttime only infrared emittance is used (Stowe *et al*. 1999). See Table 2 for a description of AVHRR bands and their uses in cloud detection.

CLAVR-1 is a sequential, decision-tree, multispectral threshold algorithm that uses a series of reflectance and emission tests to identify cloud pixels. CLAVR-1 was designed to be conservative in its cloud classification, the goal being a data set of entirely clear pixels. If a pixel is flagged by any one of the cloud tests, meaning it is identified as cloudy, that pixel is classified as cloudy. The various tests compare contrast signatures with threshold values, spectral signatures between two bands, and spatial signatures for variability. Each 2 X 2 pixel array is classified as clear, mixed or cloudy. Mixed and

cloudy are considered cloud-contaminated. If only 1 of the 4 pixels in the array is classified as cloudy, the entire array is classified mixed. The conservative design of CLAVR-1 results in many more false positives (clear pixels classified as cloudy), than false negatives (missed cloudy pixels) (Stowe et al. 1999).

Martin (2004) suggested that spatial variability tests used in cloud-detection algorithms may be more error-prone in coastal ocean areas, due to the strong SST gradients from upwelling, possibly resulting in clear ocean pixels being incorrectly classified as cloudy.

In general, clouds drastically reduce the amount of AVHRR SST data available for matchup with *in situ* data. Matchup data summaries for AVHRR Pathfinder show that for the years 1995-1999, the average matchup rate for NDBC buoys with AVHRR data was 16%. The matchup rate for the same years for all buoys within latitude band 25-50 °N was also 16% (NASA 2001).This study's rate of cloud-free matchups of 19% falls within a similar range. The slightly higher rate of matchups may be explained by the higher resolution of the satellite data used. The NDBC uses 25 km resolution SST data for its global matchup database, as described in the validation discussion below. The smaller pixel size of 1 km compared with the 25 km used by the NDBC, would likely produce more matchups in areas of patchy cloudiness.

Unlike cloud contamination, aerosol affects are not detected and removed in AVHRR data processing, but they can cause bias. Volcanic aerosols from the 1992 Mt. Pinatubo eruption caused a positive bias in nighttime SST measurements lasting more than six months (Li *et al*. 2001). For this reason, AVHRR SST values need to be frequently calibrated with *in situ* SST data (Li *et al.* 2001). SST measurements from moored and drifting buoys are periodically used to adjust the satellite SST data. The data set is validated monthly by NOAA's CoastWatch West Coast Regional Node (WCRN) using

buoy data from the NOAA's NDBC, one of the data sources for this study's buoy SST data.

Validation of the SST data set is conducted for the entire global data set using both moored and drifting buoys. The moored buoys number approximately 100, and include buoys from NOAA's NDBC off the east and west coasts of the United States, the Gulfs of Mexico and Alaska, and Hawaii, Japan, the north Atlantic, and equatorial moored buoys. Drifting buoys are managed by the Global Drifter program and number approximately 1,250.

SST data from buoys is matched with satellite data measured within the same 4 hour interval, and within 25 kilometers, and added to a matchup database. Standard deviation and bias error for geographic regions and day/night is calculated from the matchups. The statistical results are assessed weekly, monthly and annually, and used to adjust the SST algorithm if necessary (Martin 2004).In other words, the regression algorithm is adjusted with new coefficients.

The NOAA/OSU SST product is adjusted with buoy measurements from the three moored buoys owned and maintained by NDBC, and any drifting buoys that pass through. Because the satellite SST data set consists of satellite data frequently adjusted with buoy data, it is essentially a hybrid of satellite and buoy measurements (Martin 2004).

Buoy and satellite SST differences

A major difference between SST measurements of buoy and satellite is the depth at which the ocean temperature is measured. SST instruments on the moored buoys in this study are at depths between 0.45 and 2.5 meters below sea surface. Drifter buoy temperature sensors are at 0.2 to 0.3 m. The AVHRR sensor on the satellite measures

only the very thin surface layer of the ocean, and this is an indirect measurement, within the top 11 micrometers, or one-thousandth of a centimeter of the sea surface (Martin 2004). SST measured at the buoy instrument depth is called bulk temperature; the satellite surface measurement is called skin temperature.

The difference between bulk and skin temperatures can cause differences between buoy and satellite temperature readings in the same geographical location because they are sampling different spectrums of the water column. Sea water is one thousand times denser than air, and sunlight reaching the water surface is absorbed and scattered before it can travel very far (Martin 2004). If the water column is well-mixed, temperature readings taken from different depths will agree. If the water column is not well-mixed but stratified into different temperature layers, bulk and skin temperatures will diverge.

The main cause of mixing near the ocean surface is wind and the main reason for a stratified water column in the study area is solar heating. Because this study's data are all cloud-free, sun angle will have the largest impact on relative solar heating. Calm windless days with a high sun angle would be associated with the most stratification and therefore larger differences, and windy days would be associated with the most mixing, and therefore less difference in bulk and skin temperatures.

In addition to bulk and skin temperature, buoy and satellite measurements differ in several important ways. Buoy measurements represent a single point on the ocean surface, albeit moving within a fixed watch circle, recording data around the clock, whereas satellite measurements are made in kilometer-wide swaths within seconds as the satellite passes overhead in its orbit around the earth. Table 3 summarizes the differences between buoy and satellite instrument measurements.

Because this study matches buoy and satellite data, and buoy SST measurements represent only a single point on the ocean surface compared to the satellite swath, the highest spatial resolution satellite data available that also met other study objectives was used. This is the 1 km AVHRR SST data set. This is a much higher spatial resolution than is typically used in buoy-satellite matchup comparisons, for example a representative study by Kilpatrick *et al.* (2001) used 16 km. The buoy matchup data set used by NESDIS for ongoing global validation for the POES/AVHRR SST product uses 25 km. In other studies, Gentemann *et al.* (2004) matched buoy and satellite data within 50 km and Parekh *et al*. (2007) used a 75 km area. An exception to these examples is an early study using *in situ* data to evaluate error in AVHRR SST by Li *et al.* (2001), who used 1km.

Table 3. Comparison of buoy and satellite instrument measurements used in the current study.

	In situ	**Satellite**
Source	Moored buoys	Pathfinder AVHRR satellites
Spatial resolution (horizontal)	a point within a watch circle	1.1km (0.0125 degree)
Temporal resolution	hourly (data recorded for 8 minutes each hour and averaged) (NDBC buoys)	every 6 to 12 hours (each satellite pass takes about 1 minute to cross the study area)
Measurement resolution	0.1 °C (NDBC buoys)	0.0001 °C
Measurement accuracy	+/- 1 °C (NDBC buoys)	+/- 0.7 °C
Measurement water depth	bulk: 0.6 m below sea level (NDBC buoys)	skin: extremely thin layer at surface (11 micrometers)

Buoy data and satellite data occurring in the same hour were matched. The highest temporal resolution of satellite data available was one hour, so this determined the

temporal resolution of the matchup. This one hour temporal resolution is in line with other studies. Li *et al*. (2001) and Gentemann *et al.* (2004) used one hour for matchups, Parekh *et al.* (2007) used two hours, and the NESDIS matchup database uses one hour (Martin 2004).

Data pre-processing

The satellite data was provided for this study in Esri (manufacturers of ArcGIS mapping software) Ascii format, a raster format. Once the project data was received from NOAA's CoastWatch, all further processing was done using ArcGIS. The files were organized into batches for processing. Using Python scripts, the Ascii files were renamed with a nomenclature that would allow for efficient processing with ArcGIS geoprocessing tools, described later. Each satellite file was named with a date and time stamp and the hour of the satellite pass for easy identification and matchup later with the buoy data.

The files were converted to Imagine format files for smaller file size, and easier handling by the ArcGIS geoprocessing tools. Next, their geographic projection was defined to the native projection of the files, which was Geographic (latitude/longitude), using the WGS84 datum. The point data projection was adjusted to the raster data projection to avoid resampling the raster files which can alter the original pixel values. In other words, the buoy data was projected to the native satellite file projection and geoprocessing was done in WGS84.

Buoy measurements are assumed to be a true representation of field conditions, providing an accurate measurement of bulk sea surface temperature at their depth of measurement. Every effort was made to locate the buoys as precisely as possible for matchup with the satellite data. The latitude/longitude for the buoy location is recorded when it is moored and a mooring history for each buoy was pieced together for this study. The buoy with the most location changes was the Point Reyes buoy 46214, with 12 locations during the

time period of the study. The distance between first and subsequent locations was significant because the buoy was moved 36 km away. Subsequent locations were within a half kilometer of each other.

Watch circle radius for the buoys in this study ranges from 116 to 700 meters, with an average of 300 meters. In general then, a buoy will be within +/- 300 meters from the latitude/longitude of its mooring. Relative to the satellite data pixel size of 1 km, this is a fairly small area. The buoy with the largest watch circle, 46214 at 700 m, would be the most likely to incur this spatial error. Therefore, the maximum spatial mismatch between buoy and pixel would be one adjacent pixel error. In addition, results from this study show that Buoy 46214 does not have a greater error than the other buoys. It seems unlikely that the spatial uncertainty introduced by the watch circle is significant at the 1 km scale.

For each year of data an ArcGIS shape file was created with a point for each buoy location during that year. The satellite SST value was extracted at each of those points using the ArcGIS “Extract MultiValues to Points” geoprocessing tool. This tool extracts the pixel value at multiple locations from a raster image and exports the extracted values to a table. No interpolation was used, so that the value extracted is the SST value of the satellite pixel in which the buoy point location falls.

After extracting the satellite SST values at the buoy point locations all data manipulation and analysis was done in Microsoft Excel. Buoy and satellite values from the same hour were matched in Excel. Each satellite observation was joined to the corresponding buoy observation using the date/time stamp for that hour as the join field. Buoy and satellite observations were recorded in Coordinated Universal Time (UTC), and were converted to Pacific Standard Time (PST) and Pacific Daylight Time (PDT) for this project. Table 4 summarizes matchups by year and buoy.

Table 4. Summary of matches between cloud-free imagery and buoys in the current study, 1992 to 2012.

Year	Buoy 46013	Buoy 46014	Buoy 46026	Buoy 46214	Buoy CB	# of Matches
1992	270	224	271	0	0	765
1993	426	324	437	0	0	1,187
1994	403	385	389	0	0	1,177
1995	364	292	68	0	0	724
1996	431	453	359	27	0	1,270
1997	57	487	472	438	0	1,454
1998	299	309	195	337	0	1,140
1999	0	0	0	0	0	0
2000	116	111	121	106	0	454
2001	0	0	0	0	0	0
2002	0	0	0	0	0	0
2003	113	245	267	266	0	891
2004	0	0	0	0	0	0
2005	497	441	509	447	0	1,894
2006	423	291	419	383	0	1,516
2007	419	417	408	340	191	1,775
2008	485	426	460	409	193	1,973
2009	357	566	511	448	369	2,251
2010	363	369	378	358	0	1,468
2011	425	423	417	644	615	2,524
2012	500	405	462	397	432	2,196
Total Matches	**5,948**	**6,168**	**6,143**	**4,600**	**1,800**	**24,659**

The matchups were analyzed for bias using descriptive statistics, including histograms, mean bias error (MBE), standard deviation, skewness and kurtosis, root mean squared error (RMSE) and Nash Sutcliffe model effectiveness. The matchups were also filtered in Excel to create various subsets for different physical and measurement conditions; day/night, low/moderate/high wind speeds, and upwelling season subsets were explored. Results apply to cloud-free conditions between 1992 and 2012.

Results

Results show an overall warm bias of approximately 0.5 °C in the satellite measurement of SST as compared to buoy SST. Figure 2 shows results for all matchups without filtering. The R-squared value is 0.63, reflecting the modest fit of the regression line of satellite data, the dependent variable, to buoy data, the independent variable, see Figure 2a. A scattering of outliers is mostly confined to satellite temperatures biased towards colder temperatures, with a very few outliers with very warm satellite bias. The cold outliers can also be seen in Figure 2b and appear to be contributing to the strong negative skewness value of -2.0. The tail of negative values to the left of center together with the greater number of values appearing to the right of center are indicators of negative skewness. Data also show high kurtosis (value of 20.4), as evidenced by the high peak at 0 °C and 0.5 °C data values, extending far above a normal distribution curve.

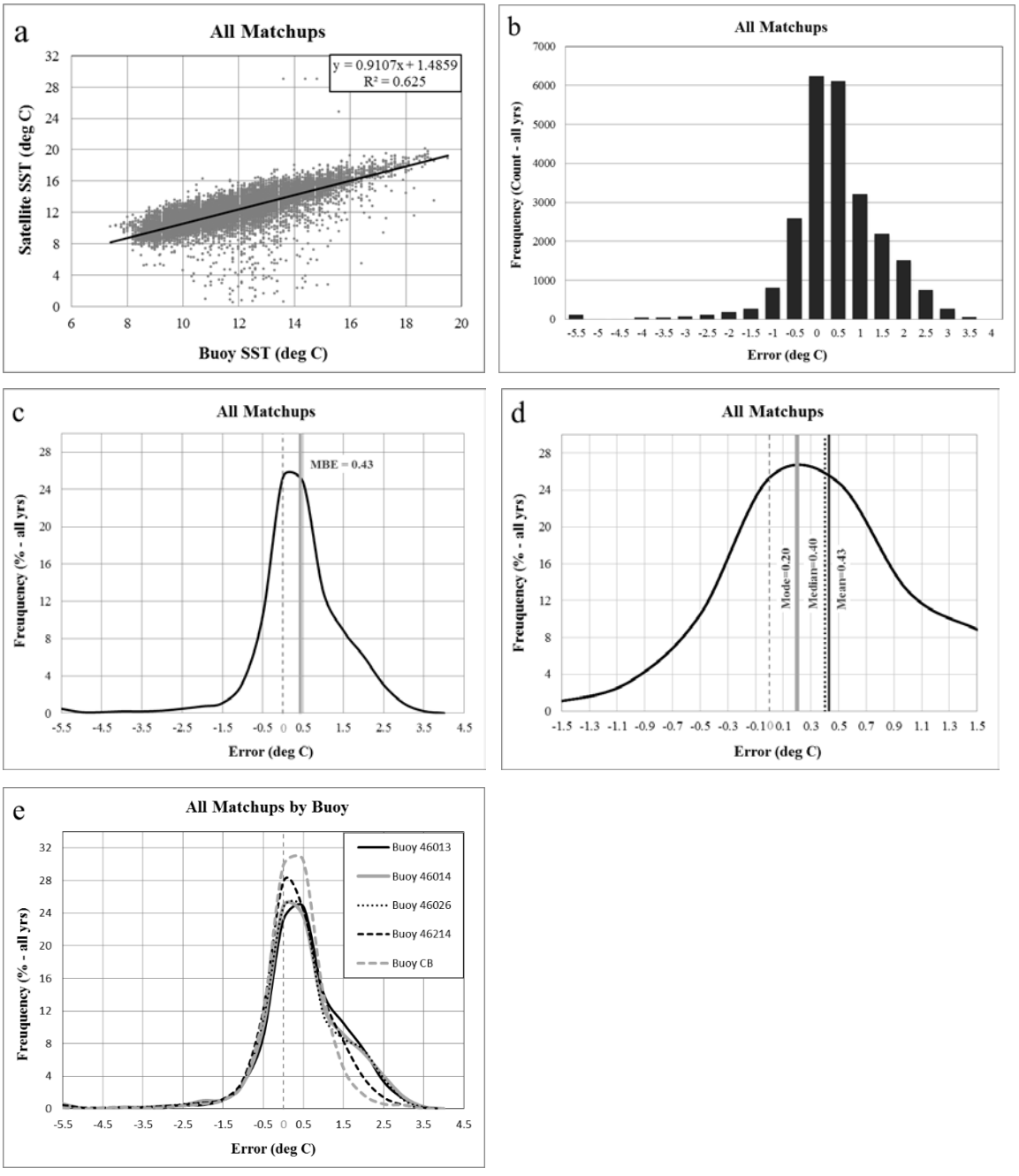


Figure 2. (a) Regression analysis of AVHRR data against buoy data, all matchups, (b) histogram of frequency count of error, (c) percent frequency of error and MBE, (d) mean, median and mode of bias error, and (e) individual buoys.

Mean difference between satellite SST and buoy SST, or mean bias error (MBE), is 0.43 °C, Figure 2c. In Figure 2c and subsequent figures, error is plotted using percent frequency of error to standardize the results in order to facilitate comparison between matchup subsets. The mean absolute value of the difference between satellite and buoy bias error, or mean absolute error (MAE), is 0.84 °C. Bias error ranges from 15.4 °C to 13.0 °C, with a standard deviation of 1.17 °C. Root mean squared error (RMSE), or the square root of the mean of the square of all the bias error is 1.25 °C. Nash Sutcliffe, which measures model efficiency similar to the regression r-squared value, is 0.43.

Nash Sutcliffe values are almost always lower than R-squared values because Nash Sutcliffe measures fit to the 1:1 line, as opposed to regression's best fit line. Nash Sutcliffe analysis values may range between -∞ and 1.0, with 1.0 representing a perfect fit. Nash Sutcliffe represents one minus the total squared differences between predicted and observed values adjusted for the variation of the observed values, as summarized from Krause *et al.* (2005).

Additional measures of "average" error are shown in Figure 2d, which gives a close-up view of the top of the distribution curve from Figure 2c. The relationship between mean, median and mode are shown, with mode at 0.20 °C and median at 0.40 °C relative to the mean of 0.43 °C. So the most common error value is 0.20 °C but there are enough larger error values to pull the median and mean to twice that.

The general pattern of error is quite similar between individual buoys, see Figure 2e. All five buoys show a warm bias, with MBE values from lowest to highest as follows: CB buoy at 0.29 °C, 46214 at 0.34 °C, 46026 at 0.49 °C, 46014 at 0.52 °C, and 46013 at 0.58 °C. Although the CB buoy had the lowest MBE, it also had the fewest matchups due to a limited time series, approximately one third the number of matchups of the other buoys.

Comparisons to other buoys should be made with caution, given the smaller sample size for CB.

Cold outliers were explored further to see if it made sense to exclude them from the analysis data set. The histogram in Figure 2b shows that focusing on values below approximately -4.0 would capture the extreme cold outliers. Selecting matchups that are greater than 4 standard deviations (+/- 4.68 °C) from the mean bias error captures these outliers. This subset contained 154 data points, or 0.006 percent of all matchups. Of these, 148 showed extremely cold bias, and the remaining 6, extremely warm bias. Satellite temperatures for matchups with extremely cold bias error ranged from 0.5 °C to 13.5 °C; corresponding buoy matchup SST for cold bias outliers ranged from 10 to 19 °C. For comparison, buoy temperatures for all matchups in this study ranged from 7.4 °C to 19.5 °C. The cold outlier values are either so far below the range of expected SST values in the study area that they cannot be accurate measurements, or are within expected SST range but are 5° C or more below buoy measurements. Given that they represent only 0.006 percent of data and that their accuracy is doubtful, the outliers were excluded from the analysis data set to prevent skewing the results and the analyses were re-run.

All Matchups without Outliers

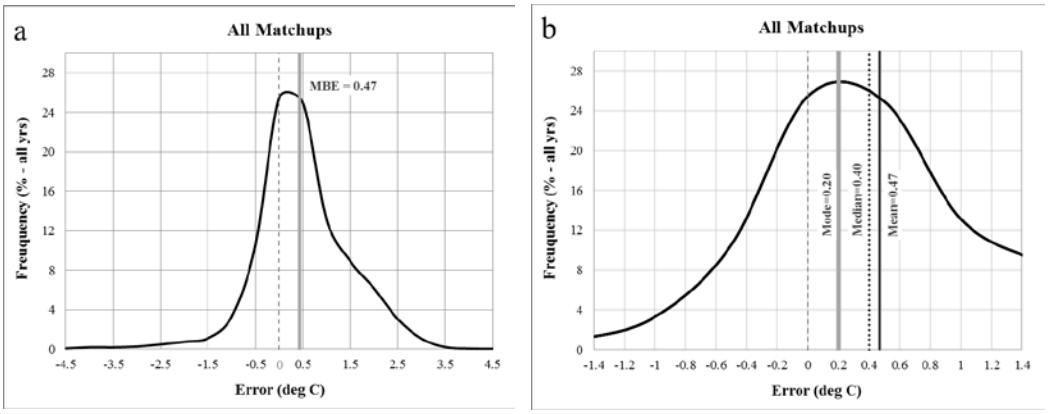


Figure 3. (a) MBE for matchups with outliers excluded and (b) mean, median and mode of bias error.

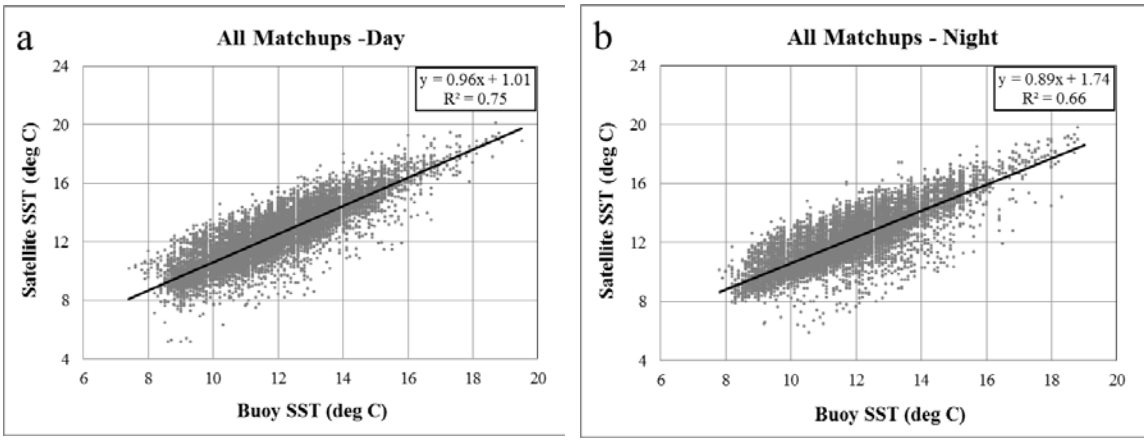
After removing outliers, the warm bias increased from 0.43 °C to 0.47 °C, see Figure 3a, and all other statistical measures behaved as would be expected after removing a small set of very cold outliers. Mode and median stayed the same, Figure 3b. Skewness went from -2.01 to -0.22 and kurtosis fell to 2.19. R-squared and Nash Sutcliffe values improved. The data set without outliers was used in all subsequent analyses. Differences in the performance of the satellite SST estimate under different physical and measurement conditions is discussed in the following sections and summarized in Table 5.

Table 5. Summary of statistical analysis results for all matchups and matchup subsets. N and r2 values refer to matchups between satellite and *in situ* SST. All other statistics refer to satellite SST error (difference between matched satellite and *in situ* SST).

	All (unfiltered)	All	Day	Night	Winter	Upwelling	Relaxation	Wind Speed 0-2 ms	Wind Speed >10 ms	Wind Speed 8-11 ms
N	24,659	24,505	14,036	10,469	8,480	9,781	6,244	2,013	3,620	3,905
R^2	0.63	0.71	0.75	0.66	0.64	0.63	0.67	0.66	0.68	0.70
MAE (°C)	0.84	0.80	0.78	0.82	0.63	0.88	0.89	0.89	0.84	0.82
MBE (°C)	0.43	0.47	0.53	0.40	0.29	0.55	0.61	0.44	0.62	0.59
Median (°C)	0.40	0.40	0.40	0.30	0.20	0.40	0.60	0.40	0.45	0.40
Mode (°C)	0.20	0.20	0.40	0.00	0.30	0.20	0.50	0.40	0.10	0.20
max BE (°C)	15.40	4.40	4.40	4.40	4.40	4.40	4.25	3.90	4.40	4.25
min BE (°C)	-13.00	-4.65	-4.50	-4.65	-4.60	-4.64	-4.65	-4.50	-4.30	-4.30
STD BE (°C)	1.17	0.98	0.91	1.06	0.85	1.05	0.99	1.09	0.96	0.96
Skewness	-2.01	-0.22	0.03	-0.37	-0.30	-0.11	-0.59	-0.40	0.24	0.16
Kurtosis	20.41	2.19	1.86	2.17	3.36	1.35	2.99	1.50	1.19	1.34
RMSE (°C)	1.25	1.09	1.05	1.13	0.90	1.18	1.16	1.17	1.14	1.13
Nash Sutcliffe	0.43	0.56	0.59	0.53	0.38	0.41	0.48	0.44	0.41	0.47
regression equation	y = 0.9107x + 1.4859	y = 0.93x + 1.32	y = 0.96x + 1.01	y = 0.89x + 1.74	y = 1.00x + 0.29	y = 0.88x + 1.85	y = 0.86x + 2.37	y = 0.9796x + 0.6928	y = 0.9478x + 1.1817	y = 0.9371x + 1.308

Day/Night comparison

Day and night matchups were compared. Results for the day/night comparison are shown in Figure 4 and are summarized in Table 5.



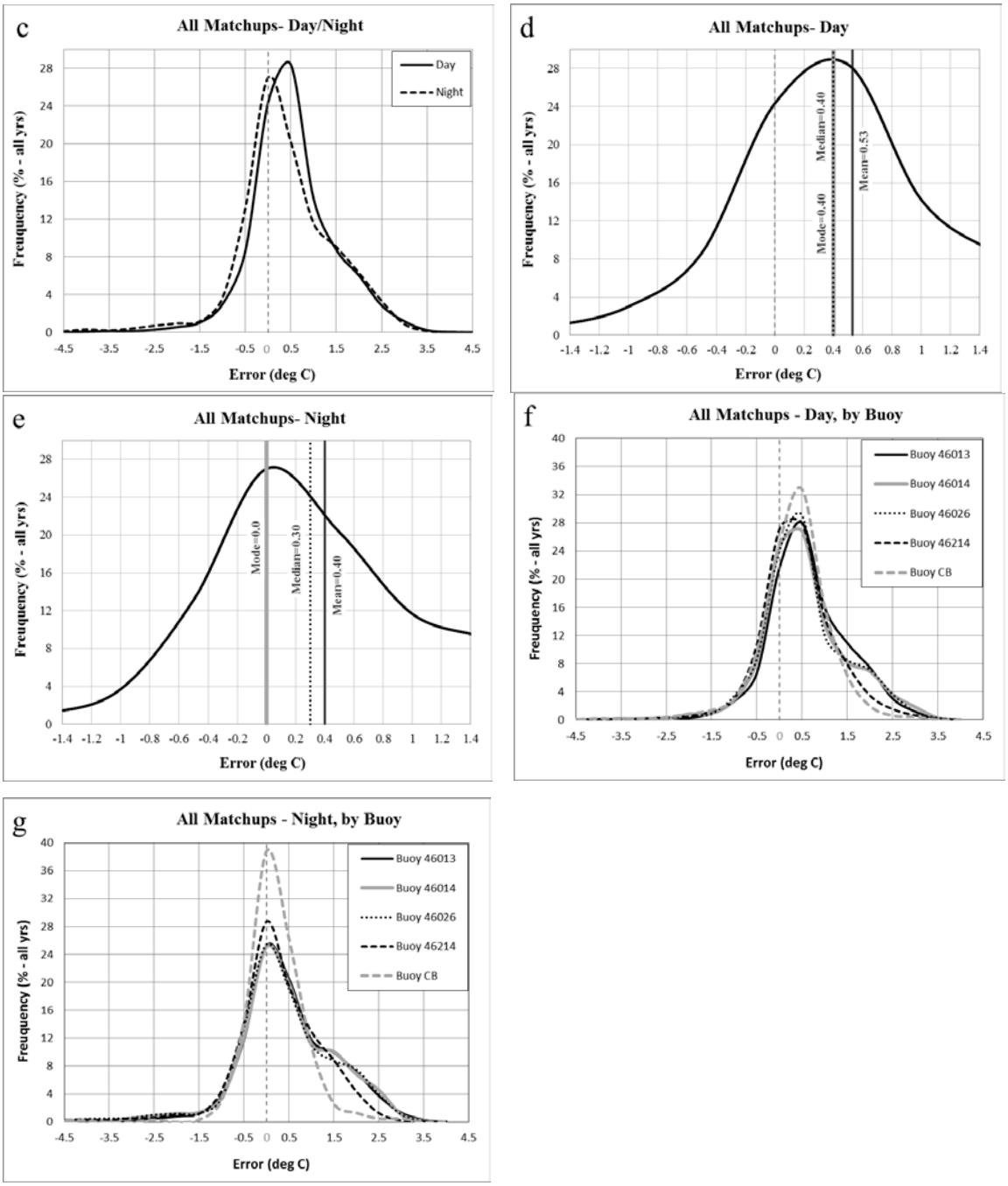


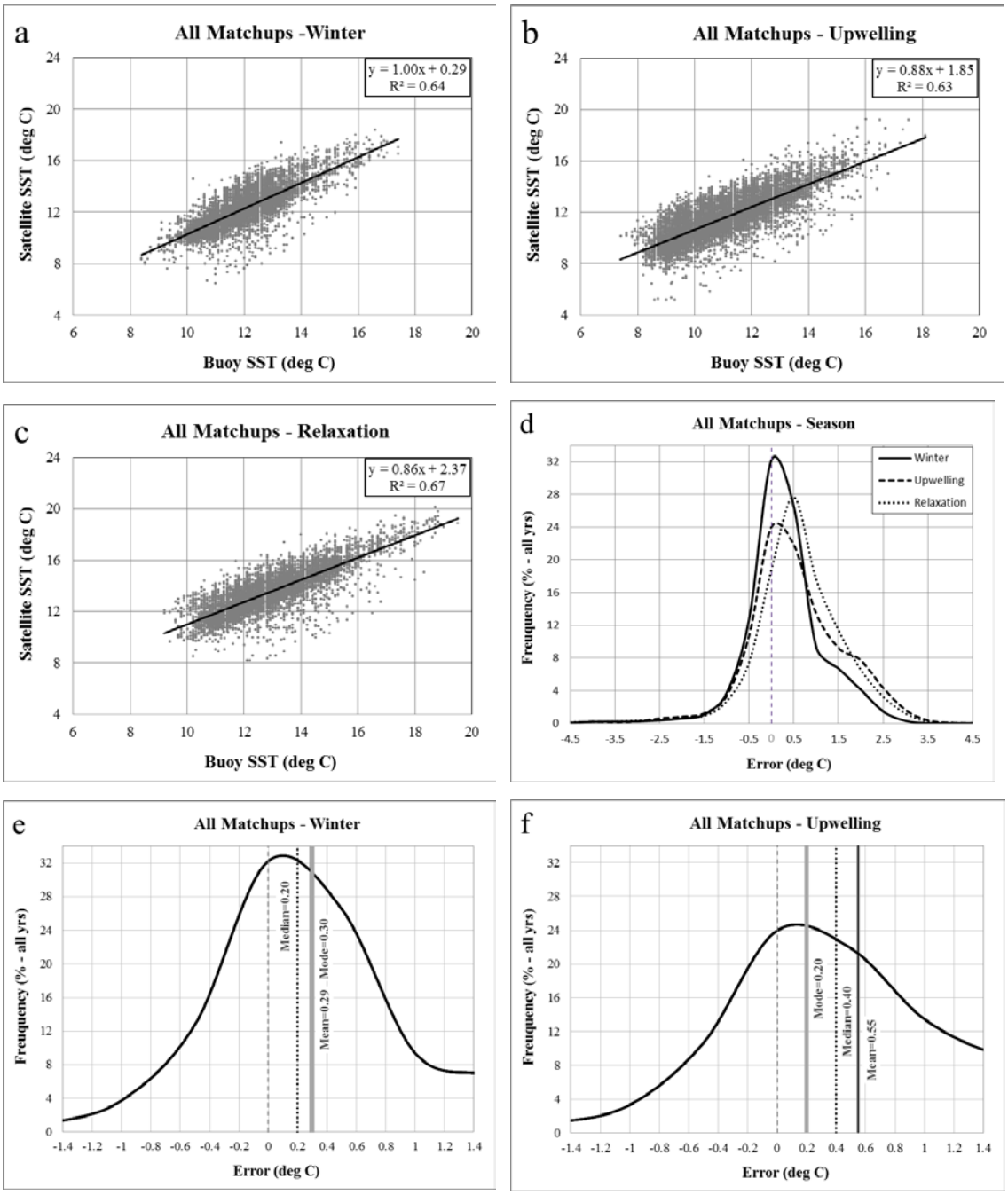
Figure 4. (a) (b) Regression analysis, day and night matchups, (c) percent frequency of error, (d) (e) mean, median and mode of bias error, and (f) (g) individual buoys.

Figures 4a and 4b show day and night regression lines, indicating more scatter at night. Night error values from regression plots before outliers were removed showed considerably more scatter below the regression line than day values, and had twice the skewness of day values, indicating that cold outliers occurred predominately at night. With outliers removed, both day and night continued to show an overall warm bias, with more bias during the day (0.53 °C MBE) than night (0.40 °C), Figures 4d and 4e. A comparison of daytime mean, median and mode show a strong warm bias centering around 0.5 °C. Nighttime mean, median and mode are more stratified. Mode is at zero, however median and mean are higher, indicating enough high positive values to pull those measures to the right. The error distribution curve may be far enough from a normal distribution that it is better described by mode rather than mean. Using mode, daytime error is four times that of night. Daytime values show relatively less scatter and relatively more bias; nighttime values show more scatter and less bias.

A comparison of individual buoys in Figure 4f and 4g shows that they follow a similar pattern, although the night-time satellite SST shows the closest agreement with the CB buoy, with a percent frequency a good ten points higher than the other buoys.

Seasonal comparison

Matchups were compared by season. Figures 5 and 6 show results for the seasonal comparison.


a
All Matchups -Winter
y = 1.00x + 0.29
R² = 0.64
Satellite SST (deg C)
Buoy SST (deg C)
b
All Matchups - Upwelling
y = 0.88x + 1.85
R² = 0.63
Satellite SST (deg C)
Buoy SST (deg C)
c
All Matchups - Relaxation
y = 0.86x + 2.37
R² = 0.67
Satellite SST (deg C)
Buoy SST (deg C)
d
All Matchups - Season
Winter
Upwelling
Relaxation
Freuqency (% - all yrs)
Error (deg C)
e
All Matchups - Winter
Median=0.20
Mode=0.30
Mean=0.29
Freuqency (% - all yrs)
Error (deg C)
f
All Matchups - Upwelling
Mode=0.20
Median=0.40
Mean=0.55
Freuqency (% - all yrs)
Error (deg C)

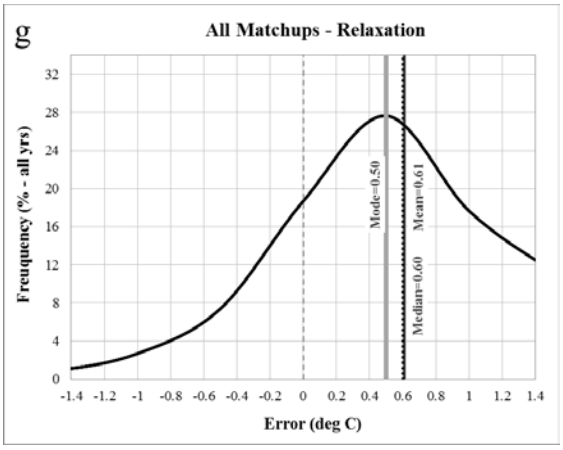


Figure 5. (a) (b) (c) Regression analysis for matchups by season, (d) percent frequency of error, and (e) (f) (g) mean, median and mode of bias error.

The regression model for each season in Figures 5a, 5b, and 5c is similar, explaining about two thirds of the scatter. All seasons show a warm bias, see Figure 5d, with interesting differences in measures of central tendency shown in Figures 5e, 5f, and 5g. Mean bias for upwelling and relaxation seasons is about twice that of winter. Median error for upwelling season is midway between that of winter and relaxation, while mode is lowest for upwelling and highest for relaxation.

Mean, median and mode for winter are grouped around values of 0.20 °C to .30 °C, as compared to the relaxation season at 0.50 °C to 0.60 °C; the relaxation season also has the largest skewness value. Interestingly, the upwelling season seems to be a hybrid of the winter and relaxation seasons, with a more divergent mean, median and mode. The mode at 0.20 °C indicates many low error values similar to the winter season, however the mean at 0.55 °C, indicates some fairly high bias values similar to the relaxation season.

A buoy by buoy comparison (Figure 6) shows some differences, but general agreement between the buoys. For the upwelling season in figure 6b, buoy 46013 appears to have higher error.

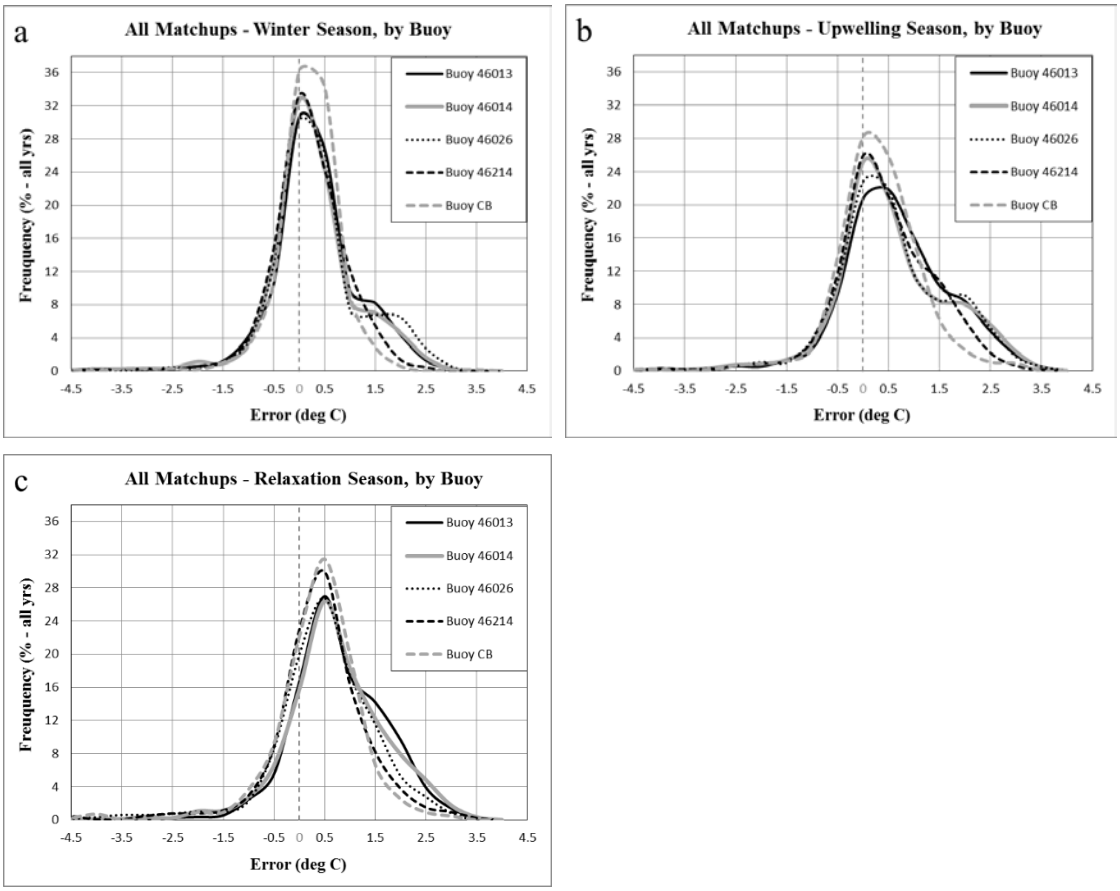


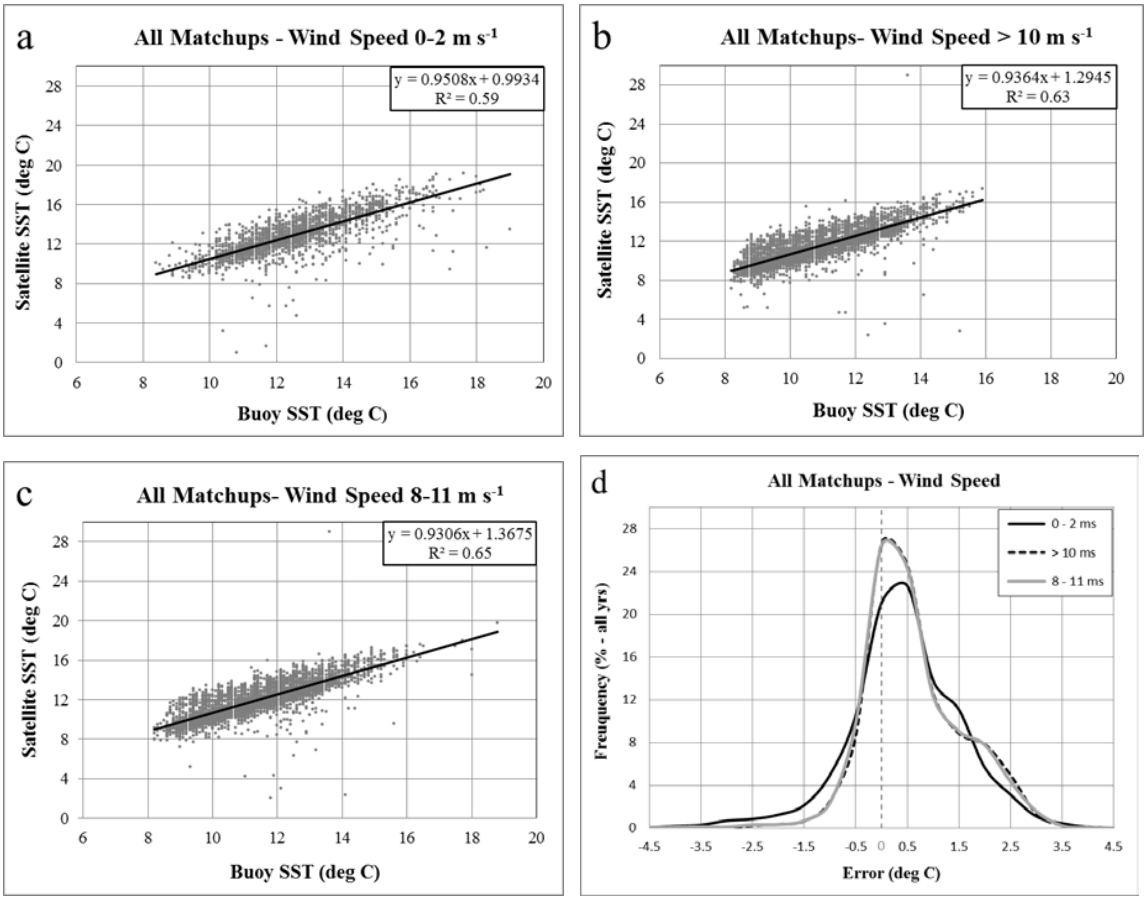
Figure 6. (a) (b) (c) Percent frequency of error, seasonal comparison by buoy.

Interestingly, Figure 6 clearly shows a second, smaller bump in error distribution to the right of the main distribution peak, limited to three of the five buoys (buoys 46013, 46014 and 46026). These three are coincidentally the ones with wind instrumentation. Winter and especially upwelling seasons show the secondary peak most clearly.

Wind speed comparison

Matchups were compared in three wind speed categories. Wind speed data was only available for three of the five buoys in this study because buoys 46214 and CB did not have wind instrumentation. Wind speed categories of low, moderate, and high winds

were selected as most likely to show differences in error due to associated sea surface conditions and were arrived at after preliminary data exploration; the categories are not mutually exclusive. The moderate wind category (8-11 m $s^{-1)}$ represents the wind speed at which there are moderate waves and many white caps at the sea surface. Figures 7 and 8 show results of the wind speed comparison.



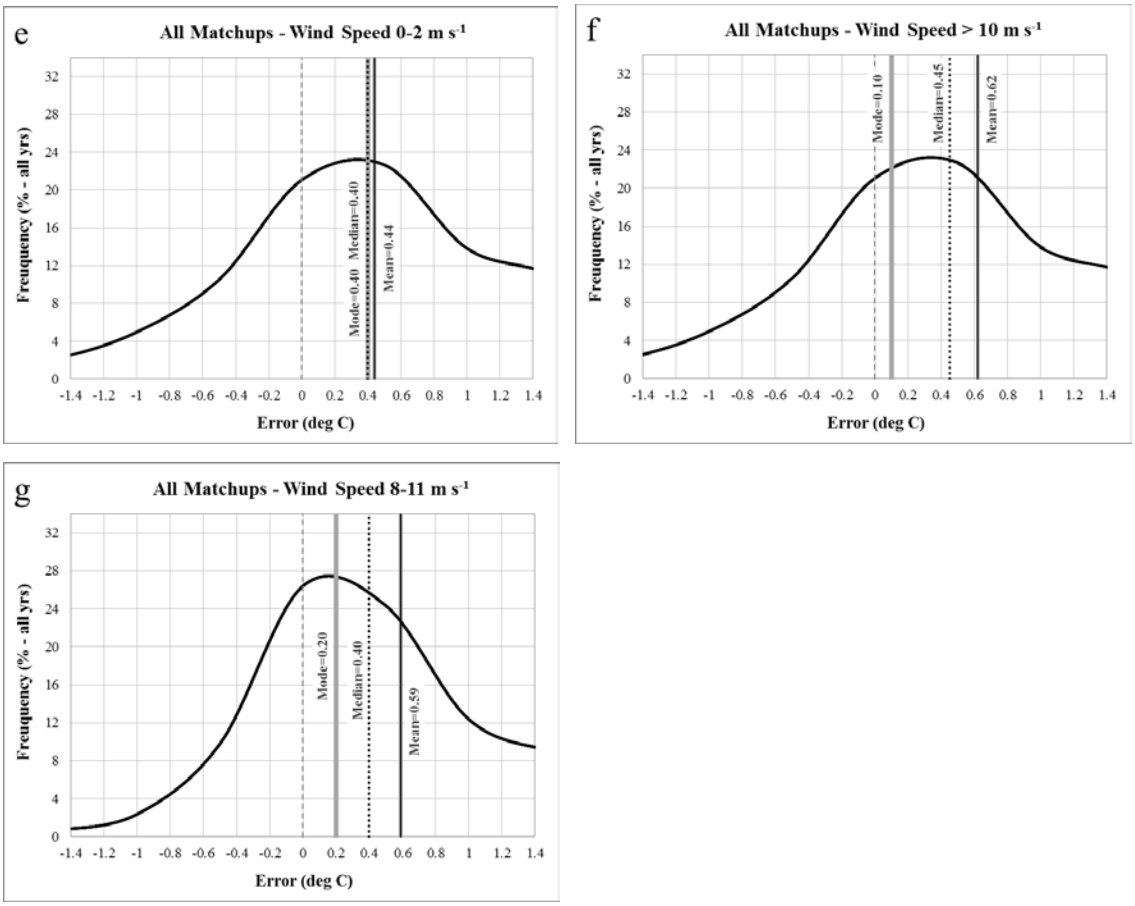


Figure 7. (a) (b) (c) Regression analysis for matchups by wind speed category, (d) percent frequency of error, and (e) (f) (g) mean, median and mode of bias error.

R-squared values for the three wind speed categories are similar, with the regression line explaining about two-thirds of scatter, see Figures 7a, 7b, and 7c. Wind speeds of 0-2 m s^{-1} show more scatter than higher wind speeds and are more negatively skewed, at approximately -1.7, as opposed to higher wind speeds at approximately -0.4. A secondary peak in warm bias similar to that seen in Figure 6 is moderate at low wind speeds and larger at higher wind speeds (Figure 7d).

All three categories show a warm bias, but lowest MBE (0.44 °C) occurs at low wind speeds. High and moderate wind speeds had MBEs of 0.62 °C and 0.59 °C respectively. Using mode, the highest error is in the low wind category (value of 0.40 °C) when water column stratification would be most likely to occur, two to four times that of the modes of moderate and high wind categories, where we would expect more water column mixing.

A buoy by buoy comparison (Figure 8) shows general agreement between the buoys. In the >10 m s^{-1} wind speed category in Figure 8b, buoy 46013 shows higher error.

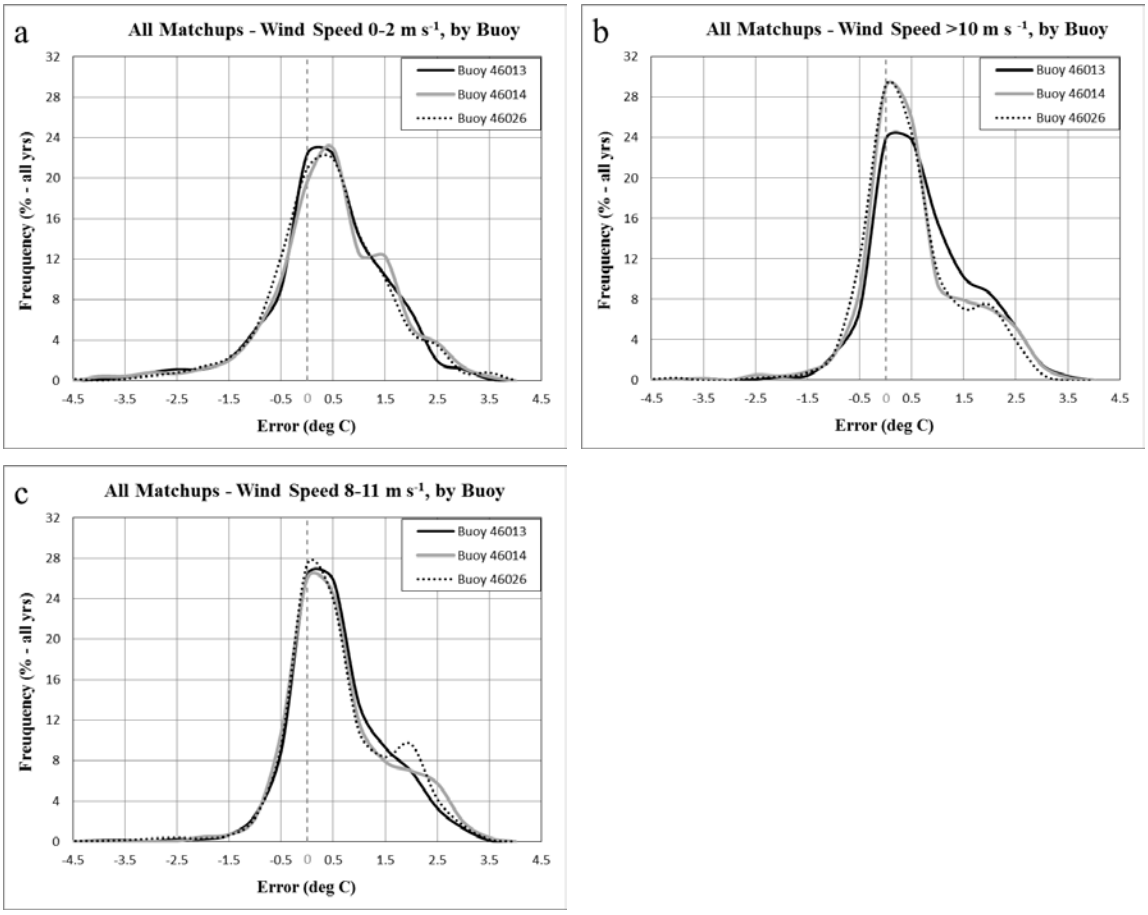


Figure 8. (a) (b) (c) Percent frequency of error, wind speed comparison by buoy.

Wind speed category graphs in Figures 8b and 8c show a bi-modal distribution similar to that of Figures 6 and 7, pointing to an additional source of systematic error.

Discussion and Conclusions

This study set out to explore error in satellite-derived SST in a coastal upwelling area, both overall and under conditions where differences in error might be expected to occur: day/night, upwelling season, low/moderate/high wind speeds. A data set of 25,000 matchups over 21 years was analyzed. For comparison purposes, the Li *et al.* (2001) study cited earlier analyzed 1,800 matchups from three coastal locations over one year. An overall systematic error of approximately 0.5°C was found for the entire matchup data set. Interesting differences in error were found in the subsets examined.

Using mode, daytime error is four times that of night. This is consistent with Li *et al.* (2001) who found that daytime bias was two to three times that of nighttime. Daytime values show relatively less scatter and relatively more bias; nighttime values show more scatter and less bias. Comparing matchups by season using mode showed that the upwelling season had the lowest error, followed by the winter season. The relaxation season had two and a half times the error of the upwelling season and the most scatter. The low error for winter season SST may be explained by generally cooler temperatures and winter storms producing a better-mixed water column. The relaxation season, characterized by warmer temperatures and lower winds, showed the highest error as might be expected.

For wind speed categories, the highest error using mode is in the low wind category, two to four times that of moderate and high wind categories. Low winds also had the most scatter. These results are consistent with expected lower error with more water column mixing, and more error expected when the water column is more stratified.

This study found both random and systematic error. A scattering of error in both cool and warm directions is evidence of random error. Increased scatter was found in the mid-temperature range, between approximately 10 °C and 16 °C where SST changes quickly. Rapid changes in sea surface conditions would be expected to cause random error as buoy bulk temperatures and satellite skin temperatures become out of step in both temperature directions. In addition to rapid temporal changes in SST, spatial variability could cause scatter if *in situ* data picks up on localized temperature differences that were not resolved at the coarser scale of satellite SST. Although random error cannot be corrected for, systematic error by definition occurs in a regular pattern and would cause bias that could be corrected for.

Three notable patterns in error distribution appear to point to several sources of systematic error. First, a small subset of very cold outliers was found, with values at least 5 °C below buoy matchup, or with values so low they cannot be accurate SST measurements. These cold outliers occurred predominantly at night and at low wind speeds. It is possible that some of these are instances where thin fog or clouds that would be expected to be much colder than sea surface, were mistaken for sea surface at nighttime when the visible spectrum is not available for cloud detection. Error caused by very cold outliers could be mitigated in the study area by excluding satellite measurements below 6 °C which would remove many of them. This would allow an extra 1.4 °C below the lowest buoy temperature (7.4 °C) in the matchup data set, on the assumption that there may be some colder buoy SST values that were not matched with satellite data due to cloudy or foggy conditions.

Second, a warm bias of approximately 0.5 °C is the most prominent systematic error found. This is consistent under the various conditions examined, except nighttime if mode is used, and is most notable during daytime, low wind, and the relaxation season, when we would expect water column stratification to occur. De-coupling of skin and bulk

temperatures due to water column stratification appears to be a major contributor to the warm bias. This warm bias was also consistent across the five buoys in the study area, despite differences in location and instrumentation. This result is in line with Li *et al's* (2001) finding of a warm bias of 0.2 °C in their analysis of SST in three coastal regions (Gulf of Mexico, US coastal Northeast and Southeast).

Third, an additional mode of warm bias is suggested by the second warm shoulder that appears most notably in the high wind category, as well as during moderate wind, winter and upwelling seasons. Interestingly, three of the five buoys show the second warm shoulder clearly, while the other two do not. In a cross-shelf transect, these three buoys are located mid-shelf, whereas the other two are located further from shore at the edge of the continental shelf. Because upwelling is associated with strong alongshore winds and currents, it is possible that the three mid-shelf buoys experience stronger winds and colder upwelled waters than the two buoys at the shelf edge. This secondary mode of warm bias error may also be related to sea surface conditions at higher wind speeds, such as increased reflectance from sea spray or foam. Satellite-derived SST may need to be adjusted to account for a warm bias in coastal upwelling areas.

Adjustment factor

Future satellite SST values in the study area can be adjusted to account for the systematic bias found in this study by using the regression equation for all matchups from Table 5, where y is the unadjusted satellite data and x is the buoy data,

$$y = 0.93x + 1.32 \quad (9)$$

Substituting a sample buoy SST value of 8 °C into equation 9, we get an unadjusted satellite SST value of 8.76 °C, a higher value than buoy SST, reflecting the warm bias error found by this study. If we take the inverse of the slope constant 0.93 from equation

9, and the negative of the intercept, we get an equation that can be used to adjust the satellite SST value for bias:

$$y = 1.0753x - 1.4194 \quad (10)$$

where the y value is the adjusted satellite SST and the x value is the unadjusted satellite SST. Substituting a sample unadjusted satellite SST value of 8 °C into equation 11, gives an adjusted satellite SST value of 7.183 °C, a lower temperature than 8 °C, and what we would expect when adjusting for warm bias. This adjustment factor can be applied to all SST values, as the only values excluded from the matchup data set analysis were the very small number of cold outliers. Adjustment factors for other matchup subsets examined in this study such as day/ night, can be obtained in the same way, using the regression equation for that subset from Table 5.

This study describes three types of systematic error in 1 km AVHRR SST measurements in an upwelling center, the most prominent being a 0.5 °C warm bias, which can be corrected for with the adjustment factor provided. At the scale of SST changes during upwelling, 1-2 ° C and up to 5 °C, this bias is significant. Given that global AVHRR data is validated at the 25 km scale, it seems likely that SST in coastal upwelling areas does not resolve well at this scale. SST measurements in other upwelling areas would likely benefit from additional validation with *in situ* data at a suitable spatial scale.

References

Donlon, C.J., Minnett, P.J., Gentemann, C., Nightingale, T.J., Barton, I.J., Ward, B., and Murray, M.J. 2002. Toward improved validation of satellite sea surface temperature measurements for climate research. *Journal of Climate* 15: 353-369.

Gentemann, C.L., Wentz, F.J., Mears, C.A., and Smith, D.K. 2004. *In situ* validation of Tropical Rainfall Measuring Mission microwave sea surface temperatures. *Journal of Geophysical Research* 109: C04021.

Halle, C.M. and Largier, J.L. 2011. Surface circulation downstream of the Point Arena upwelling center. *Continental Shelf Research* 31: 1260-1272.

Kilpatrick, K.A., Podestá, G. P., and Evans, R. 2001. Overview of the NOAA/NASA advanced very high resolution radiometer Pathfinder algorithm for sea surface temperature and associated matchup database. *Journal of Geophysical Research* 106: 9179-9197.

Krause, P., Boyle, D.P., and Bäse, F. 2005. Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences* 5: 89-97.

Largier, J.L. 2007. Cordell Bank in Upwelling Region and Offshore of San Francisco Bay. Slide presentation at the March 2007 meeting, Sanctuary Advisory Council of Cordell Bank National Marine Sanctuary, Olema, California.

Largier, J.L., Lawrence, C.A., Roughan, M., Kaplan, D.M., Dever, E.P., Dorman, C.E., Kudela, R.M., Bollens, S.M., Wilkerson, F.P., Dugdale, R.C., Botsford, L.W., Garfield, N., Cervantes, B. Kuebel, Koračin, D. 2006. WEST: A northern California study of the role of wind-driven transport in the productivity of coastal plankton communities: *Deep-Sea Research II* 53: 2833-2849.

Largier, J.L., Magnell, B.A., and Winant, C.D. 1993. Subtidal circulation over the northern California shelf. *Journal of Geophysical Research* 98: 18147-18179.

Li, X., Pichel, W., Clemente-Colon, P., Krasnopolsky, V. and Sapper, J. 2001. Validation of coastal sea and lake surface temperature measurements derived from NOAA/AVHRR Data, *International Journal of Remote Sensing* 22: 1285-1303.

Martin, S. 2004. *An Introduction to Ocean Remote Sensing*. Cambridge, Cambridge University Press, 426pp.

McClain, E. Paul, Pichel, William G., and Walton, Charles C. 1985. Comparative Performance of AVHRR-Based Multichannel Sea Surface Temperatures. *Journal of Geophysical Research* 90: 11,587-11,601.

National Aeronautics and Space Administration (NASA). 2001. AVHRR Pathfinder Oceans.http://yyy.rsmas.miami.edu/groups/rrsl/pathfinder/Matchups/description.html. (last accessed 7 June 2014).

Palacios, D.M., Bograd, S.J., Foley, D.G., and Schwing, F.B. 2006. Oceanographic characteristics of biological hot spots in the North Pacific: A remote sensing perspective. *Deep-Sea Research II* 53: 250-269.

Parekh, A., Sharma, R., and Sarkar, A. 2007. A comparative assessment of surface wind speed and sea surface temperature over the Indian Ocean by TMI, MSMR, and ERA-40.*Journal of Atmospheric and Oceanic Technology* 24: 1131-1142.

Scripps Institution of Oceanography. 2014a. Coastal Data Information Program.http://cdip.ucsd.edu/themes/cdip?pb=1&d2=p20&u3=tab:1:display:data_qc. (last accessed 7 June 2014).

Scripps Institution of Oceanography. 2014b. Coastal Data Information Program.

http://cdip.ucsd.edu/?sub=index&nav=documents&xitem=gauge. (last accessed 7 June 2014).

Stowe, L.L., Davis, P.A., and McClain, E.P. 1999. Scientific basis and initial evaluation of the CLAVR-1 global clear/cloud classification algorithm for the advanced very high resolution radiometer. *Journal of Atmospheric and Oceanic Technology* 16: 656-681.

U.C. Davis Bodega Marine Laboratory. 2014. Bodega Ocean Observing Node, Cordell Bank Buoy. http://bml.ucdavis.edu/boon/buoy_obs.html. (last accessed 30 April 2015).

U.C. Davis Bodega Marine Laboratory. 2011. BML/CBNMS CTFX Mooring Protocol. http://www.bml.ucdavis.edu/boon/pdf/BML_CBNMS_Moored_TSFX_QAQC_Methods.pdf (last accessed 30 April 2015).

U.S. Department of Commerce. 2014a. National Atmospheric and Oceanic Administration's National Data Buoy Center. http://www.ndbc.noaa.gov. (last accessed 7 June 2014).

U.S. Department of Commerce. 2014b. National Atmospheric and Oceanic Administration's National Environmental Satellite, Data, and Information Service- Office of Satellite and Product Operations. http://www.ospo.noaa.gov/Operations/POES/index.html. (last accessed 7 December 2014).

U.S. Department of Commerce. 2014c. National Atmospheric and Oceanic Administration's National Environmental Satellite, Data, and Information Service- Office of Satellite and Product Operations. http://noaasis.noaa.gov/NOAASIS/ml/avhrr.html. (last accessed 7 December 2014).

U.S. Department of Commerce. 2014d. National Atmospheric and Oceanic Administration's National Marine Sanctuaries, Sanctuary Integrated Monitoring Network.http://sanctuarysimon.org/cordell/sections/oceanography/overview.php?sec=0. (last accessed 6 December 2014).

U.S. Department of Commerce. 2014e. National Atmospheric and Oceanic Administration's CoastWatch Program and NWS Monterey Regional Forecast Office. http://coastwatch.pfel.noaa.gov/infog/AT_sstd_las.html. (last accessed 29 November 2014).

U.S. Department of Commerce. 2009. *Handbook of Automated Data Quality Control Checks and Procedures.* NDBC Technical Document 09-02, Stennis Space Center, Mississippi: National Data Buoy Center, National Oceanic and Atmospheric Administration, 78 pp.

U.S. Department of Commerce. 2007. *A Biogeographic Assessment off North/Central California: In Support of the National Marine Sanctuaries of Cordell Bank, Gulf of the Farallones and Monterey Bay. Phase II – Environmental Setting and Update to Marine Birds and Mammals.* NOAA Technical Memorandum NOS NCCOS 40. Silver Spring: NOAA National Centers for Coastal Ocean Science. 240 pp.

Walton, C.C., Pichel, W.G., Sapper, J.F., May, D.A. 1998. The development and operational application of nonlinear algorithms for the measurement of sea surface temperatures with the NOAA polar-orbiting environmental satellites. *Journal of Geophysical Research* 103: 27999-28012.