Click Here for Full Article

Evaluations of SST climatologies in the tropical Pacific Ocean

A. B. Kara,[1] C. N. Barron,[1] and T. P. Boyer[2]

[1] Accuracies of a total of six sea surface temperature (SST) climatologies are investigated in the tropical Pacific Ocean, spanning latitudes from 30°N to 30°S. Observation-based climatologies are directly obtained from their original sources and typically formed from satellite measurements, in situ data, or a combination of both. Other SST climatologies are constructed from reanalyses of a numerical weather prediction center. All products have different features, each with its own data sources, grid resolutions, interpolation methods, sampling periods, and biases. Even though time periods during which climatologies were formed differ, all products have good agreement with each other, with basin-averaged mean bias values of almost zero and RMS SST differences typically <0.3°C over the annual cycle. Nondimensional skill scores between the pairs of SST products are close to 1, indicating almost perfect agreement everywhere except in the western equatorial Pacific. Comparisons against an ensemble climatology based on the average of all SST climatologies further reveal the weaknesses and strengths of each product. Additional validations are performed by forming climatologies using SSTs from moored buoys of the Tropical Atmosphere Ocean (TAO) array during 1986–2005. Overall, very small cold biases of ≈0.2°C and near-perfect skill scores are found for all SST climatologies when comparing them with 408-month-long SST time series from 34 buoys. In general, all SST products are generally accurate enough to be used for various climate applications.

1. Introduction

[2] Accurate data sets for climatological mean SSTs are highly desirable to investigate many processes, especially in tropical Pacific Ocean. For example, tropical Pacific SST plays a central role in atmosphere-ocean heat exchange processes, with various implications for climate change [e.g., *Schneider et al.*, 2002; *Thum et al.*, 2002]. Reliable SSTs in the tropical Pacific are necessary to monitor climate changes [*Webster*, 1995] and El Niño Southern Oscillation (ENSO) events [*Diaz and Markgraf*, 2000]. The atmospheric internal variability over the North Pacific is also closely related to the tropical Pacific SST anomalies [*Chen and van den Dool*, 1997; *Zhang and McPhaden*, 2006].

[3] There are several observation-based SST products, providing monthly mean SSTs over the global ocean [e.g., *Casey and Cornillon*, 1999]. An observation-based climatology can be based on historical measurements, such as SSTs from ship measurements, drifting buoys, moored buoys and high-resolution Conductivity, Temperature, Depth (CTD) or eXpendable BathyThermograph (XBT) profiles

[*Stephens et al.*, 2002]. SSTs from satellites, from in situ sources or from both can also be used in forming the product climatologies [*Reynolds et al.*, 2002]. Additionally, there are also numerical weather prediction (NWP) model products which provide high temporal resolution (e.g., 6 hourly) SSTs from which monthly climatologies can be formed. An example NWP product is the European Centre for Medium-Range Weather Forecasts (ECMWF) 40-year Reanalysis (ERA-40) [*Uppala et al.*, 2005].

[4] There is not a concise and consistent statements of accuracy for each SST product. All SST climatologies have their unique biases. For example, different biases may emerge where the climatologies have been constructed over different time intervals. SSTs from NWP products could be too coarse to resolve many small features, especially near the coastal boundaries.

[5] The major goal of this paper is to quantify accuracy of climatological SSTs from various products which have their unique features and biases. Our purpose is not to fully identify one climatology as superior to the others. Rather, we examine multiple climatologies, each with its own data sources, methods, sampling periods and biases. Where the climatologies agree, we will have increased confidence in the consensus results, particularly since we arrive at the consensus using different approaches. Where agreement is less, we will have a better estimate of the uncertainty in the climatological fields. Variations that fall within this uncertainty are less significant than those that fall outside.

[1]Oceanography Division, Naval Research Laboratory, Stennis Space Center, Mississippi, USA.
[2]NOAA National Oceanographic Data Center, Silver Spring, Maryland, USA.

Table 1. Abbreviations Used for SST Products Throughout the Text

Acronym	Name of the SST Product
PATH	Pathfinder
NOAA	National Oceanic and Atmospheric Administration
MODAS	Modular Ocean Data Assimilation System
WOA01	World Ocean Atlas, 2001
ERA15	ECMWF 15-year reanalyses
ERA40	ECMWF 40-year reanalyses

[6] Besides the evaluations of various SST climatologies through comprehensive statistical analyses, one of the other focuses of this study is to provide a useful contribution to the tropical Pacific oceanography. Thus, in this paper we put together various existing but dispersed SST data sets and make their detailed assessment. This can assist in studying various climate-related topics mentioned above, and in developing ocean prediction systems where the accuracy of the database and the models is consistent with the needs of the user. Such evaluations can also help in establishing whether the grid resolution and data sources used for constructing climatologies are sufficient.

2. SST Climatologies

[7] Monthly mean SST climatologies are obtained or formed from various sources. A list of these SST products is given in Table 1, along with the product abbreviations used throughout the text. Sampling period and data sources used for constructing the monthly means vary (Table 2). There are mainly two categories: (1) observation-based SST climatologies (PATH, NOAA, MODAS and WOA01) and (2) SST climatologies formed from ECMWF reanalyses (ERA15 and ERA40).

[8] Monthly mean SST climatologies for observation-based products are directly obtained from their original sources. However, we form monthly mean SST climatologies for NWP products on the basis of their 6 hourly outputs. The domain considered in this study for evaluating SST products in Table 1 covers the tropical Pacific Ocean spanning the latitudes from 30°S to 30°N. On the basis of climatological means averaged over all months from the PATH data set, general depictions of the spatial variations of SST reveal distinct features in this region (Figure 1). Particularly, there are cold (warm) temperature values in the eastern (western) part of the domain.

2.1. Observation-Based SST Products

[9] The satellite-based 4 km resolution PATH climatology is derived from Advanced Very High Resolution Radiometer (AVHRR) Pathfinder SST data during 1985–2001. The data processing was done using SSTs from NOAA polar orbiting satellites. This climatology uses most of the same techniques presented by *Casey and Cornillon* [1999], and has a resolution of 4 km. Key improvements in the new PATH climatology include a more accurate, consistent land mask, resulting in improvements near the coastal regions. Both daytime and nighttime daily fields are included in each monthly average. The monthly climatologies were derived using data only with a quality flag of 7. In general, the Pathfinder algorithm combines the multiple observations as long as they have same quality level.

[10] The NOAA climatology is based solely on observational SST data [*Reynolds et al.*, 2002]. In particular, the climatology incorporates in situ SSTs from ships and buoys (e.g., moored and drifting buoys) along with SSTs from AVHRR satellite retrievals. The NOAA SST product was built from two intermediate climatologies: a 2° SST climatology developed from in situ data for the period of 1971–2000 and a 1° SST climatology for the period of 1982–2000 derived from the second version optimum interpolation (OI.v2) SST analysis. In situ SSTs are used for bias correction in forming the climatology.

[11] MODAS SST is a satellite-based product that includes no in situ SSTs [*Barron and Kara*, 2006]. Here, the climatology is formed over the time period 1993–2005. It is produced on a uniform 1/8° (latitude, longitude) grid by an OI of AVHRR nonlinear SST (NLSST) observations processed by the Naval Oceanographic Office (NAVOCEANO). All operational global AVHRR data from 1993 to the present have been used in the MODAS analysis, reflecting on any given day the collected global coverage data from one to three of the NOAA polar orbiters. The MODAS analysis uses a Gaussian error covariance with 60-h time and 20-km length scales, smaller than those used in the NOAA climatology. These scales were determined subjectively to balance fidelity in representing fronts with mitigation of spurious gradients around data-sparse regions.

[12] The climatology from WOA01 is formed from in situ observational SSTs only. Most of the moored SST measurements in the tropical Pacific are included in the climatology. Observational data used in WOA01 were averaged by 1° × 1° grids for input to the objective analysis [*Boyer et al.*, 2002]. The initial objective analyses usually contained some large-scale gradients over a small area or bulls eyes. These unrealistic features generally occurred because of the difficulty in identifying nonrepresentative values in data sparse areas.

2.2. ECMWF Reanalyses

[13] There are two reanalyses from ECMWF used for forming SST climatologies. The ERA15 (ERA40) reanalysis

Table 2. Time Intervals and Original Grid Resolutions for the SST Products[a]

Data Set	Interval	Grid Resolution	Time	Reference	Data Source
PATH	1985–2001	0.040° × 0.040°	variable	*Casey and Cornillon* [1999]	satellite only
NOAA	1971–2000	1.000° × 1.000°	variable	*Reynolds et al.* [2002]	satellite, in situ
MODAS	1993–2005	0.125° × 0.125°	daily	*Barron and Kara* [2006]	satellite only
WOA01	1773–2001	1.000° × 1.000°	variable	*Stephens et al.* [2002]	in situ only
ERA15	1979–1993	1.125° × 1.125°	6 hourly	*Gibson et al.* [1999]	NWP-based
ERA40	1957–2002	1.125° × 1.125°	6 hourly	*Uppala et al.* [2005]	NWP-based

[a]The reader is referred to given references for further details about each product. The availability of in situ and satellite observations can vary from one day to another one; thus, observation-based SST climatologies (i.e., PATH, NOAA, MODAS, and WOA01) are not based on continuous observations in time and space.
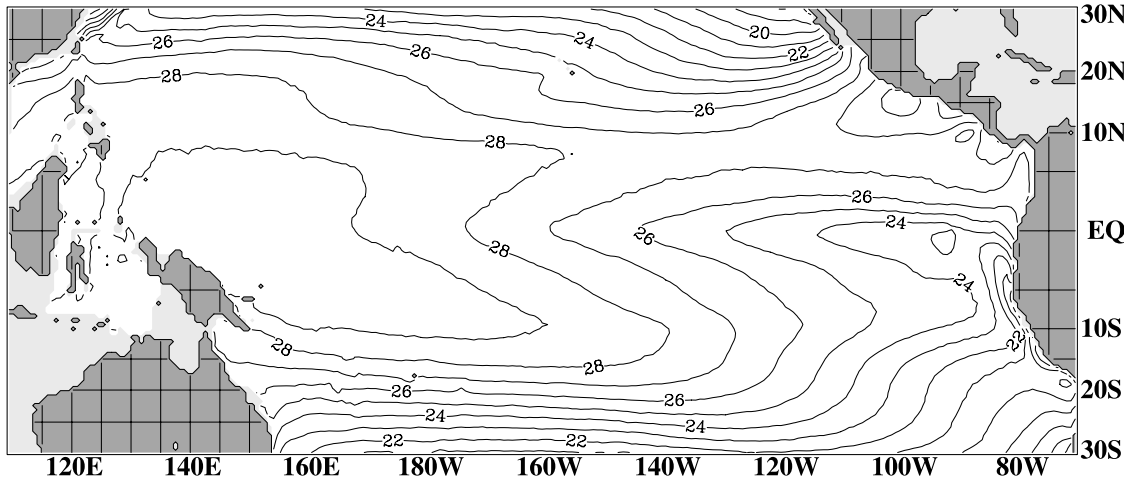
Figure 1. Climatological mean SST in °C from the ≈4 km resolution Pathfinder data set.

is available for 1979–1993 (1957–2002). The ERA-40 project applies a modern variational data assimilation technique for the past conventional and satellite observations. The model physics and the surface parameterization have been upgraded and improved since ERA-15. Note that the SSTs from ECMWF reanalyses are prescribed. More specifically, ERA15 and ERA40 come from the Hadley Centre Sea Ice and Sea Surface Temperature (HADISST) and National Center Environmental Center (NCEP) 2D-Var data sets, depending on the period considered. Thus, ECMWF climatologies examined here are the analyses of SST observations prescribed as boundary conditions. For this reason, they are also observation-based SST products. On the other hand, compared to the standard ones discussed in section 2.1, ECMWF reanalyses have higher temporal resolution and coarser spatial resolution.

[14] We form climatologies of monthly SST from ERA15 (1979–1993) and ERA40 (1957–2002) using 6 hourly outputs. These reanalyses use a consistent data assimilation scheme. Therefore, they are unaffected by operational model upgrades.

3. Statistical Metrics

[15] The finest resolution (4 km) PATH is taken as the reference SST climatology. Other products (i.e., NOAA, MODAS, WOA01, ERA15, and ERA40) are evaluated with respect to PATH. Comparisons will also be performed among all products. All climatologies are interpolated to a common grid of 0.5° × 0.5°, so that they can be evaluated at common grid locations. While this resolution may inflate the information content of products which have resolution lower than this, it is subjectively chosen to represent all products.

[16] There are always advantages (or disadvantages) of using a fine resolution SST product (e.g., PATH) in evaluating a coarse resolution SST product (e.g., NOAA). For example, spatial variability of SST may be better represented by a fine resolution one, but a climate modeler may prefer a coarse resolution product because of its computational efficiency since global climate models have typically coarse resolution to reduce computational expense on super-computers. When interpolating a coarse resolution SST product to a finer grid resolution, there are definitely

problems in the accuracy of the resulting field. This is particularly true for NWP-based SST climatologies as demonstrated by *Kara et al.* [2008]. In that study, SSTs from various climatologies are interpolated to a common grid of 4 km grid, the finest resolution of the products considered (i.e., PATH), resulting in some coastal SST differences. In this paper, our target is a wider audience, not only dealing with coastal and regional SSTs but also global SSTs for climate studies. Therefore, the resolution of 0.5°, that is more/less representative of all products is chosen for product comparisons.

[17] Evaluations presented here are based on monthly mean time series of SST at each ocean grid point. Let X_i ($i = 1, 2, , n$) be the set of n reference (i.e., PATH) SSTs, and let Y_i ($i = 1, 2, , n$) be the set of SSTs for any one of the other products (OTHER), referring to any one of the climatologies (i.e., MODAS, NOAA, WOA01, ERA15, and ERA40). Also let $\overline{X}$ ($\overline{Y}$) and σ_X (σ_Y) be the mean and standard deviations of the PATH (OTHER) SSTs, respectively. Because we would like to evaluate SSTs on the climatological timescales, in the time series comparisons n is equal to 12. Following *Murphy* [1995], the statistical metrics are mean error (ME), root mean square (RMS), correlation coefficient (R) and non-dimensional skill score (SS). They are expressed as follows:

$$ME = \overline{Y} - \overline{X}, \tag{1}$$

$$RMS = \left[\frac{1}{n} \sum_{i=1}^{n} (Y_i - X_i)^2 \right]^{1/2}, \tag{2}$$

$$R = \frac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y}) / (\sigma_X \, \sigma_Y), \tag{3}$$

$$SS = R^2 - \underbrace{[R - (\sigma_Y/\sigma_X)]^2}_{B_{COND}} - \underbrace{[(\overline{Y} - \overline{X})/\sigma_X]^2}_{B_{UNCOND}}. \tag{4}$$

In brief, ME is the mean SST difference, and RMS is absolute measure of the distance between the PATH and OTHER SST
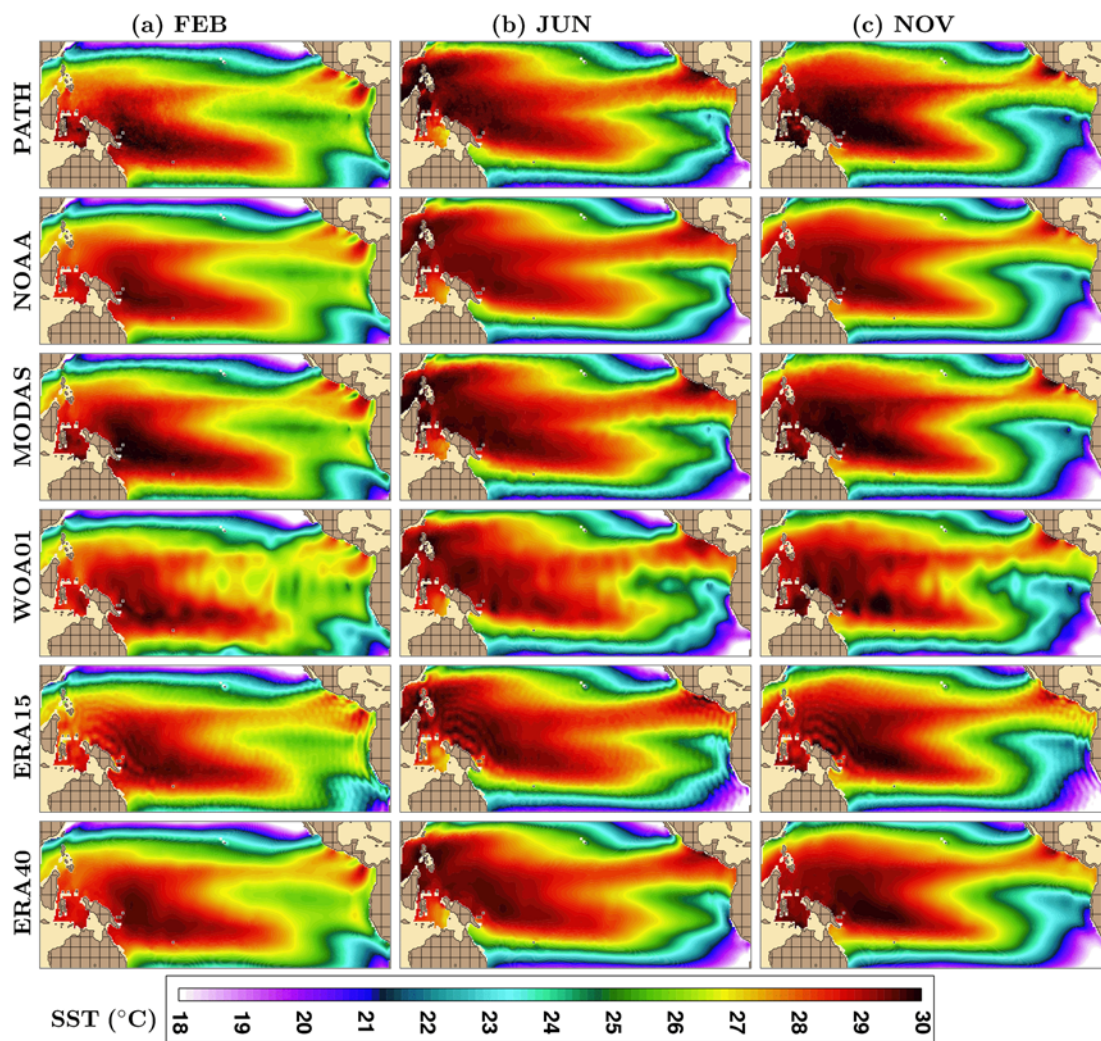
Figure 2. Monthly mean SST climatologies in (a) February, (b) June, and (c) November. They are shown for all SST products as described in section 2, in detail.

time series over the annual cycle. R is a measure of the degree of linear association between the two time series.

[18] The SS value in equation (4) is the fraction of variance explained by PATH minus two nondimensional biases (conditional bias, B_{COND}, and unconditional bias, B_{UNCOND}) that are not taken into account in the R formulation (see equation (3)). B_{UNCOND} is also called systematic bias and is a measure of the difference between the means of PATH and OTHER SST time series. B_{COND} is a measure of the relative amplitude of the variability (standard deviation) in the two data sets. R^2 is equal to SS only when B_{COND} and B_{UNCOND} are zero. Note that SS is 1.0 if PATH and any given climatology product agree with each other perfectly. Positive skill is usually considered to represent a minimal level of acceptable performance. Negative SS indicates poor agreement for a given SST climatology in comparison to PATH.

4. Evaluations of SST Climatologies

[19] The most obvious feature of the climatological long-term annual mean SST is the cold tongue region in the eastern equatorial Pacific (Figure 1). Annual mean of SST in this region can be as cold as 22°C, as much as 8°C cooler than other regions of the tropical Pacific. Vertical mixing and entrainment are critical for understanding the seasonal cycle of SST in the eastern Pacific cold tongue [e.g., *Wang and McPhaden*, 1999]. Cold SST in this region is typically a result of equatorial upwelling [e.g., *Fiedler and Talley*, 2006]. SST variability in the cold tongue can mainly be controlled by the changes in net solar radiation at the sea surface, and annual variations of surface winds can also significantly influence the SST annual cycle in the eastern Pacific as well [e.g., *Kessler et al.*, 1998].

[20] There are no major changes in SST by month in the tropical Pacific (Figure 2). For simplicity, we only show climatological SSTs in February, June, and November. Spatial variations of SSTs from all SST products are generally similar, while some differences arise especially in the equatorial Pacific cold tongue. SST magnitudes from all climatologies also fairly resemble each other. For example, SST values of ≈25°C in February become ≈23°C in June and even colder in November in the eastern equatorial Pacific. Climatological SST varies slightly by month in the
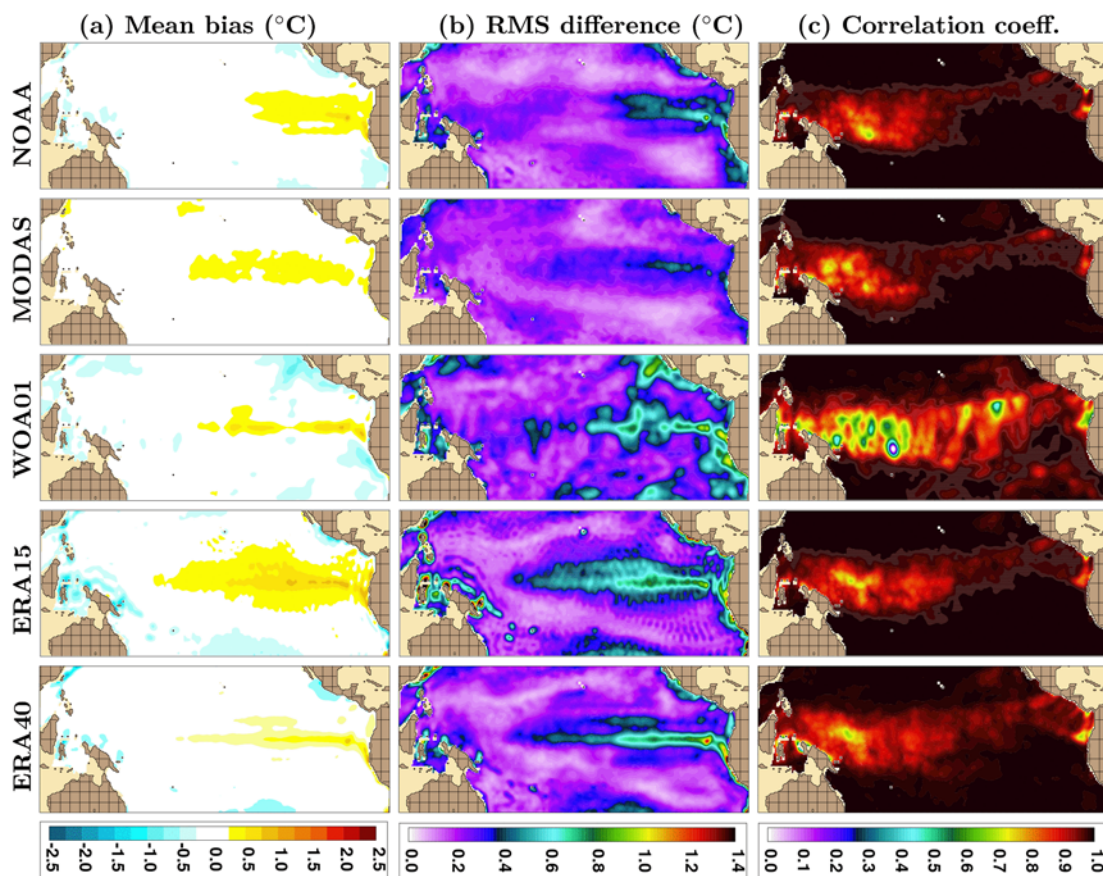
Figure 3. Spatial fields of (a) mean bias, (b) RMS difference, and (c) correlation coefficient. Computations are performed at each grid point on the basis of 12 monthly mean time series over the annual cycle between SST products and PATH climatology.

western Pacific warm pool, where SST is ≈30°C in February, June, and November.

4.1. SST Bias and Correlation

[21] We now take PATH as the reference data set and evaluate other climatologies on the basis of various statistical metrics of ME, RMS, R, B_{COND}, B_{UNCOND}, R and SS. Statistical metrics (see section 3) are computed at each grid point over the annual cycle and then mapped over the basin. Figure 3 presents the resulting spatial fields of ME, RMS SST difference and R. The white color in Figure 3a represents bias values between −0.25°C and 0.25°C, indicating that annual mean SST bias for each product is negligibly small over most of the region. However, there are some biases in the eastern and central tropical Pacific. In particular, all SST products are typically ≈0.5°C warmer than PATH, with WOA01 having warming confined to a very small region in comparison to other products.

[22] Similar to mean bias, RMS values are generally small (<0.3°C) over the majority of the region, and this is true for almost all products (Figure 3b). Relatively large RMS SST values ranging between ≈0.5°C and ≈1.0°C are noted in the eastern and central equatorial Pacific. The satellite-based MODAS SST product is an exception, in that RMS values with respect to PATH are lowest (<0.3°C) in the same region.

[23] While WOA01 shows almost no SST bias in the eastern and central tropical Pacific, there are relatively large RMS values (≈0.8°C) in comparison to PATH. The same is also true for ERA15 and ERA40. As mentioned earlier, a difference between the two ECMWF products is that the SST data set produced by the Hadley Centre and NOAA has been used in ERA40.

[24] As to the annual cycle of SST products, R values are close to 1, indicating nearly perfect agreement with the PATH climatology (Figure 3c). The only region where R values are relatively low for WOA01 is the western equatorial Pacific warm pool. Very little variability in SST is the main reason for having low correlations in these regions as will be further discussed in the following section. Correlations are weak in some areas, for example, in the western equatorial Pacific warm pool.

4.2. SST Skill Score

[25] SST standard deviation varies from one location to another over the annual cycle. For example, standard deviations can be quite large (e.g., >2°C) in the eastern equatorial Pacific, while they are very small (e.g., <0.5°C) in the western equatorial Pacific warm pool (Figure 4). As a result, mean bias or RMS values are expected to be small (large) in the western (eastern) tropical Pacific. This implies that the RMS SST difference does not provide fair evaluation about the accuracy of SST products in comparison to the PATH
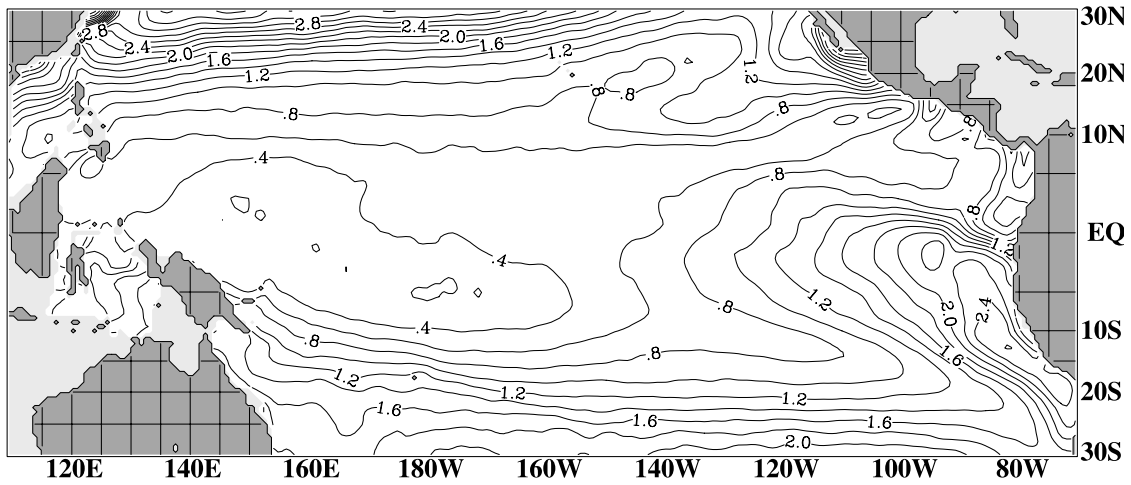
Figure 4. Standard deviation of SST from the Pathfinder climatology calculated on the basis of 12 monthly mean time series over the annual cycle.

climatology at those locations where SST standard deviations vary greatly. A normalization is therefore essential, which can be accomplished by using a skill score.

[26] As described in section 3, the nondimensional skill score (equation (4)) accounts for both conditional and unconditional biases. SS is equal to 1 (i.e., perfect relationship with respect to PATH) only when RMS is 0. Since SS

includes both RMS and standard deviation of the reference (i.e., PATH) SSTs, assessing the performance of a given SST product at two different locations which have quite different SST annual cycles can be accomplished in a normalized sense.

[27] Figure 5 includes two nondimensional biases (B_{COND} and B_{UNCOND}) and the resulting skill values when SST products are validated with respect to PATH climatology
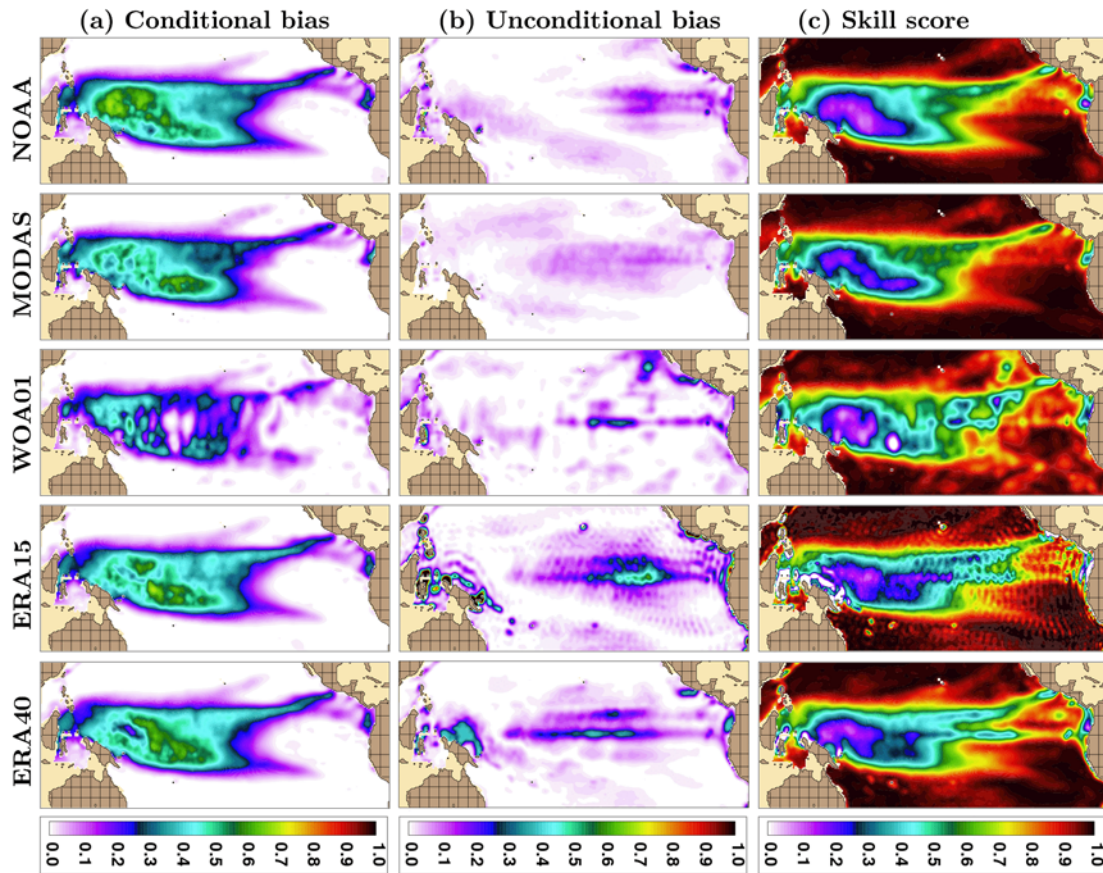

Figure 5. Spatial fields (a) conditional bias, (b) unconditional bias, and (c) skill score computed at each grid point on the basis of monthly time series over the annual cycle between SST products and PATH climatology. Results are shown for each SST product.
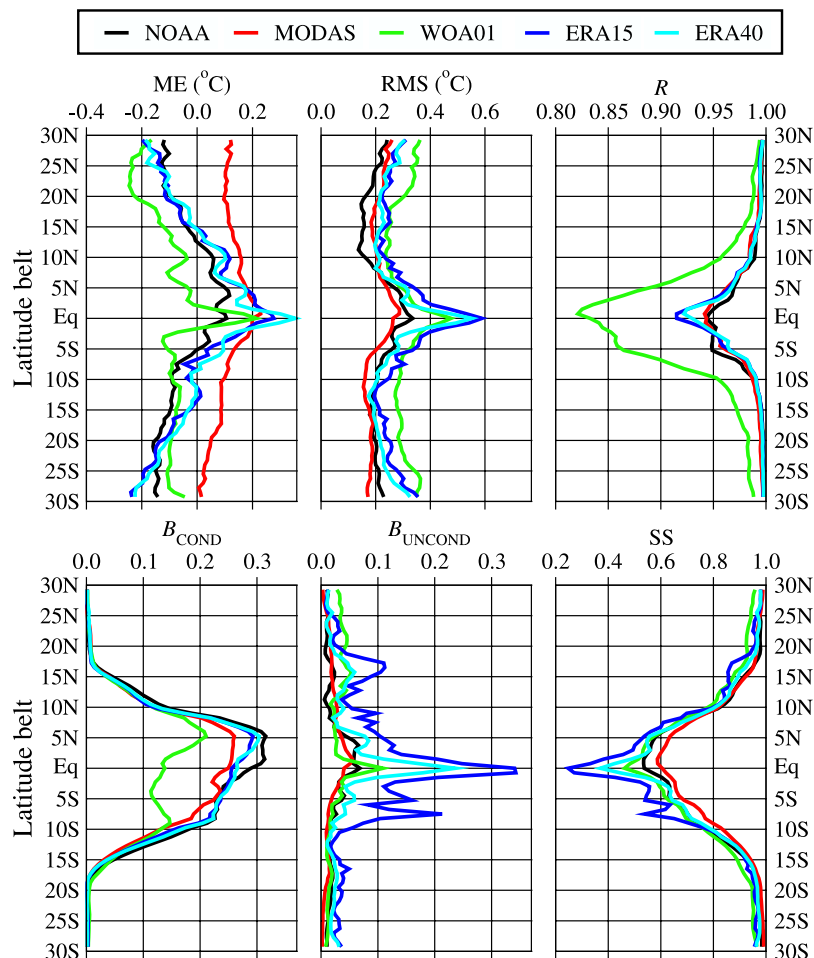
Figure 6. Zonal averages of statistical metrics shown in Figures 3 and 5. Zonal averaging is performed at each 0.3° latitude belt.

over the annual cycle. There are very high SST skills which are almost perfect with values of close to 1 for all climatologies everywhere except in the western equatorial Pacific warm pool. Previously, *Kara and Barron* [2007] demonstrated that cloud cover can affect accuracy of satellite-based daily MODAS SST in this region, and addition of in situ SST is typically an advantage for improving the accuracy. However, in the case of forming a climatology, that is not a concern since daily SSTs from many years contribute to the MODAS climatology, thereby yielding results similar to other climatological products.

[28] Given the high SST skills, all SST products agree with PATH well in the eastern tropical Pacific (Figure 5c). On the other hand, previously we showed in Figure 3b that RMS is very high in the eastern equatorial Pacific but is relatively low in the western equatorial Pacific, implying that all products agree with PATH better in the western equatorial Pacific. RMS errors in the western equatorial Pacific are low because SST standard deviations are also low (i.e., small SST variability over the annual cycle). This does not necessarily mean that there is good agreement between any one of the SST products and PATH since SST skill scores reveal otherwise. However, using the skill score in measuring the western equatorial Pacific may also appear to be less reliable since it is inflated because of the low variance in the region.

[29] Nondimensional biases in SS explain the source of differences between the SST products and PATH (Figures 5a and 5b). Low skill values in the western tropical Pacific are mainly due to conditional biases which are relatively larger than unconditional biases in this region. On the basis of the definition of B_{COND} (see section 3), such large values reveal that biases resulting in low skill are due mainly to differences in standard deviations between any one of the SST products and PATH. On the contrary, B_{COND} are very small and very close to zero except in the central tropical Pacific for SSTs from ERA15 and ERA40, explaining that biases due to the differences in mean SST in comparison to PATH are generally negligible.

[30] Zonal averages of all statistical metrics shown in Figures 3 and 5 are computed in order to investigate similarities and differences for each SST product depending on the latitudinal region (Figure 6). In comparison to the PATH climatology, all SST products generally yield similar patterns, having relatively poor agreement near the equatorial regions, mostly affected by the results in the western equatorial Pacific warm pool. The agreement with respect to the PATH climatology improves significantly away from the equatorial regions. Overall, the largest conditional and unconditional values are noted along the equatorial region. This contributes

Table 3. Areal Averages of Statistical Metrics for Each Observation-Based SST Climatology in the Tropical Pacific[a]

	PATH	NOAA	MODAS	WOA01	ERA15	ERA40
			PATH			
ME(°C)	0.00	−0.04	0.12	−0.09	0.00	0.00
RMS(°C)	0.00	0.21	0.21	0.31	0.29	0.24
B_{cond}	0.00	0.04	0.02	0.05	0.04	0.03
B_{uncond}	0.00	0.08	0.06	0.07	0.14	0.10
R	1.00	0.97	0.97	0.92	0.97	0.97
SS	1.00	0.83	0.87	0.74	0.76	0.82
			NOAA			
ME(°C)	0.04	0.00	0.16	−0.05	0.03	0.03
RMS(°C)	0.21	0.00	0.24	0.26	0.20	0.14
B_{cond}	0.01	0.00	0.00	0.01	0.00	0.00
B_{uncond}	0.06	0.00	0.12	0.05	0.11	0.05
R	0.97	1.00	0.98	0.93	0.99	0.99
SS	0.88	1.00	0.82	0.81	0.86	0.92
			MODAS			
ME(°C)	−0.12	−0.16	0.00	−0.21	−0.12	−0.09
RMS(°C)	0.22	0.24	0.00	0.35	0.30	0.26
B_{cond}	0.00	0.02	0.00	0.02	0.02	0.01
B_{uncond}	0.04	0.14	0.00	0.13	0.11	0.09
R	0.97	0.98	1.00	0.92	0.97	0.98
SS	0.90	0.80	1.00	0.71	0.81	0.85
			WOA01			
ME(°C)	0.09	0.05	0.21	0.00	0.08	0.08
RMS(°C)	0.31	0.26	0.35	0.00	0.33	0.29
B_{cond}	0.02	0.05	0.04	0.00	0.05	0.04
B_{uncond}	0.06	0.05	0.13	0.00	0.15	0.08
R	0.92	0.93	0.92	1.00	0.93	0.94
SS	0.77	0.77	0.69	1.00	0.68	0.76

[a]The domain for the tropical Pacific is shown in Figure 1. See text for descriptions of statistical metrics.

to relatively poor agreement in the equatorial SST, also evident from the relatively low skill values in these regions.

[31] Finally, areal averages of statistical metrics over the whole domain are further computed to examine relationship for all pairs of the SST products. This is done for the observation-based climatologies (Table 3) and other climatologies (Table 4), separately. For example, in Table 3, first the PATH is taken as the reference data set, and areal averages of each statistical metric are computed relative to all other climatologies as shown in the statistical maps of Figures 3 and 5. Then, the NOAA climatology is taken as the reference data set and evaluated against all other products. Similar computations are repeated for the other products.

[32] Evaluations provided in Tables 3 and 4 further confirm the close relationship among all data products. Areal averages of mean bias values are almost zero, and those of RMS SST differences rarely exceed 0.3°C regardless which product pairs are compared. There are also remarkable agreements for the SST annual cycle, as evident from high correlation coefficients, close to 1, calculated over the annual cycle. Similarly, relatively small conditional biases in comparison to unconditional biases indicate that biases are mainly due to differences in the means of SSTs rather than standard deviations.

5. Ensemble SST Climatology

[33] In sections 4.1 and 4.2, the PATH climatology is mainly taken as a reference product and compared to each one of the other climatologies. Although PATH is taken as a reference, we have not determined whether or not it is the best climatology. In addition, analysis in terms of differences from PATH provide very little information about the shortcomings of it. For example, PATH is affected by cloud contamination [e.g., *McClain et al.*, 1985; *Luo et al.*, 1995]. While its higher spatial resolution is an advantage, this does not necessarily imply that it is the best product. Similar to PATH, other SST products have also their own advantages and disadvantages.

[34] For the reasons mentioned in the preceding paragraph we produce a new SST climatology for validating all SST products rather than using any single product as the reference. To accomplish that goal, we form means of SSTs from all the products (PATH, NOAA, MODAS, WOA01, ERA15 and ERA40) at each grid point, and then obtain spatial variations over the tropical Pacific. Forming a mean climatology is performed month by month. For example, January climatology is constructed by computing mean January SSTs from all six SST climatologies. Thus, we obtain a reference monthly mean climatology, incorporating SST information from all products. This new product would average out errors in any individual SST climatology. In addition, it does not imply a judgment about which product is best. Hereafter, we will call this new reference product the ensemble climatology.

[35] Mean SST biases for each SST product with respect to the ensemble climatology are shown in Figure 7. No systematic bias is noted and differences are within ±0.1°C for all climatologies over most of the tropical Pacific Ocean. Relatively large biases are generally noted in the eastern and central equatorial Pacific for PATH and ERA40. In particular, PATH (ERA40) is typically colder (warmer) than the ensemble product by ≈0.3°C in the eastern equatorial Pacific. *Reynolds et al.* [2007] noted a similar overall cold PATH bias of 0.1–0.3°C, with nighttime SSTs showing relatively larger cold bias. Basin averages of annual mean SST biases are near zero for all products except for MODAS which has a warm bias of ≈0.12°C.

[36] Similar to mean bias, we also compute RMS SST difference, correlation and skill score values for each SST products in comparison to the ensemble climatology over the seasonal cycle (Figure 8). The most striking feature of RMS differences is that both NOAA and MODAS have much smaller biases in the cold tongue region than other products (Figure 8a). Overall, all products have RMS values of <0.2°C over majority of the tropical Pacific Ocean. The NOAA

Table 4. Same as Table 3 but for the Climatologies From ECMWF Reanalyses

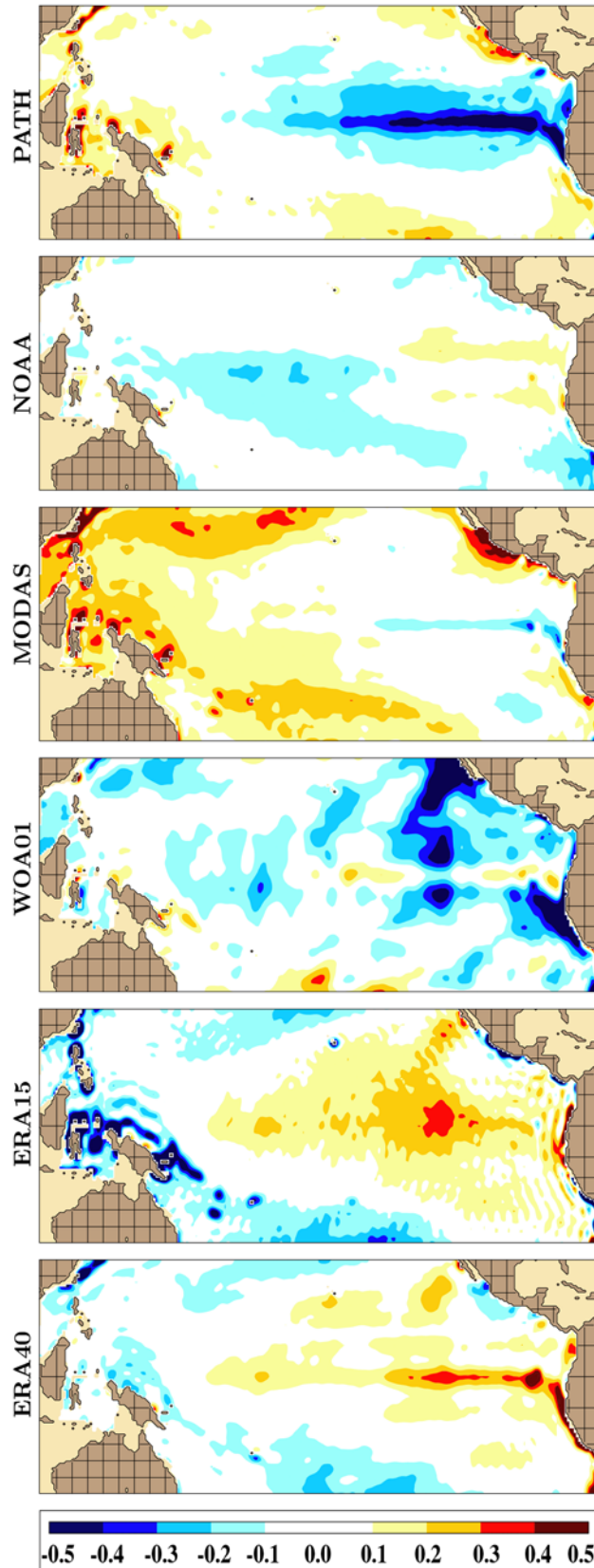|	PATH	NOAA	MODAS	WOA01	ERA15	ERA40
			ERA15			
ME(°C)	0.00	−0.04	0.12	−0.08	0.00	0.00
RMS(°C)	0.29	0.20	0.30	0.33	0.00	0.12
B_{cond}	0.00	0.00	0.01	0.01	0.00	0.00
B_{uncond}	0.11	0.11	0.11	0.13	0.00	0.04
R	0.97	0.99	0.97	0.93	1.00	0.99
SS	0.82	0.85	0.82	0.73	1.00	0.94
			ERA40			
ME(°C)	0.00	−0.04	0.12	−0.09	0.00	0.00
RMS(°C)	0.26	0.16	0.28	0.30	0.12	0.00
B_{cond}	0.00	0.00	0.01	0.01	0.00	0.00
B_{uncond}	0.09	0.06	0.10	0.09	0.04	0.00
R	0.97	0.99	0.97	0.93	0.99	1.00
SS	0.85	0.91	0.83	0.77	0.94	1.00

Figure 7. Spatial fields of mean biases for SST climatologies with respect to the ensemble climatology. The bias is computed over the seasonal cycle.

climatology has the lowest basin-averaged RMS value of 0.11°C. However, basin-averaged RMS values for other observation-based climatologies are similar: 0.15°C for PATH, 0.16°C for MODAS and 0.21°C for WOA01. Similarly, NWP-based climatologies have RMS values of 0.18°C and 0.15°C for ERA15 and ERA40. All products also capture SST seasonal cycle in a manner consistent with the ensemble climatology as evident from correlation values close to 1 (Figure 8b). Finally, all climatologies yield near-perfect skill values (Figure 8c). Relatively low SST skill with respect to ensemble climatology is seen in the western equatorial Pacific. Basin-averaged SST skill values ranges from 0.86 for WOA01 being lowest to 0.95 for NOAA being highest.

[37] We also examine SST differences for each product along with the standard deviation of the ensemble climatology. Our purpose is to identify the regions of high or low variability among the climatologies and indicate where biases of one product are significantly different from the ensemble variation. The differences reveal patterns that may indicate modes of variability in tropical Pacific SST.

[38] Standard deviation of the ensemble climatological SST is computed each month using six mean SSTs, one from each contributing product (i.e., PATH, NOAA, MODAS, WOA01, ERA15, and ERA40). For simplicity, the resulting fields of SST standard deviations are given for three selected months of February, June, and November (Figure 9a). The common feature evident from all months is that standard deviation values are <0.5° over majority of the tropical Pacific, with some larger variation in parts of the eastern equatorial Pacific. All fields reveal noticeable spatial variations. In particular, the standard deviations indicate at least two regions of high variability: a zone centered on the equator with higher variability in the eastern Pacific and a tropical rim zone with higher variability in the far western Pacific and the subtropical western Pacific near ±30° latitude.

[39] Monthly mean SST differences relative to the ensemble climatology are shown in Figure 9b (i.e., PATH ensemble, NOAA ensemble, etc.). This is done for each month separately. Scaling these by the standard deviation of the ensemble climatology overemphasizes some of the low-amplitude variations and obscures some of the variability patterns. However, SST differences are considered relative to the variation patterns identified in the standard deviation (Figure 9a). In general, absolute values of SST biases from each product are within the range of SST standard deviations given in Figure 9a. There are exceptions. For example, standard deviation of SST biases for WOA01 are generally >0.5°C in the eastern tropical Pacific, especially in June. Similar features are also evident from ERA15 and PATH in the central tropical Pacific, along the higher-variability portions of the equatorial zone. ERA15, ERA40 and MODAS all have higher deviations along the tropical rim, with cold bias in the ECMWF products and warm bias in MODAS. This may reflect a warming in the western Pacific between the 1979–1993 period of ERA15 climatology and the 1993–2005 MODAS period.

6. Climatological SST Evaluation at Moored Buoys

[40] In section 4, SST climatologies are evaluated relative to the finest resolution PATH product. Further comparisons
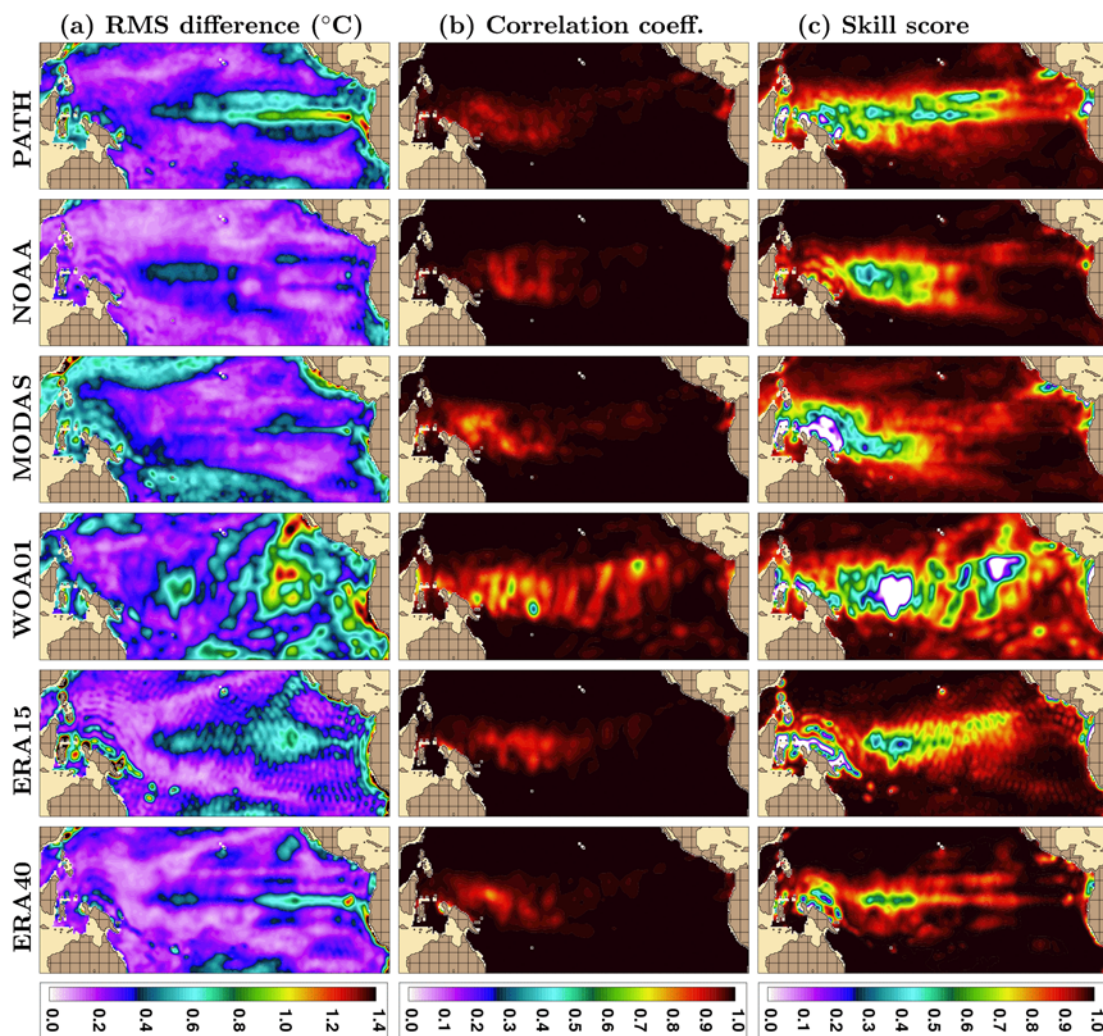
Figure 8. Spatial fields of (a) RMS difference, (b) correlation coefficient, and (c) skill score for SST climatologies in comparison to the ensemble SST climatology. Statistical metrics are computed at each grid point on the basis of 12 monthly mean time series over the annual cycle.

are also presented when choosing each one of SST climatology as a reference, separately. In addition, evaluations are performed relative to the ensemble climatology in section 5.

[41] The preceding analysis procedure is helpful for providing information about differences among the products in the tropical Pacific Ocean, but do not provide an independent assessment procedure since no validation was performed on the basis of observed SSTs. In this section, additional evaluations on the accuracy of climatological products are obtained on the basis of the SSTs measured from moored buoys. To evaluate accuracy of each climatological product, we use SSTs from many moored buoys located in the tropical Pacific Ocean (Figure 10). These buoys, maintained by the Tropical Atmosphere Ocean (TAO) array project office, regularly reports high-resolution (e.g., daily) SSTs [*McPhaden et al.*, 1998]. Although the spatial sampling is sparse and there can be missing daily SST measurements, a data set like this provides a good opportunity for evaluations of climatological products. This is because SSTs from the TAO array are available since

1986, providing a time series of observations sufficient for constructing local SST climatologies at many buoys.

[42] For our validation purposes, daily averaged SST from each buoy is obtained, and monthly mean climatology is formed depending on the availability of SSTs from 1986 to 2005. Each buoy has variable number of year-long daily SST time series within this time frame. We first form monthly mean SST at each buoy location for a given year. If a buoy has fewer than 30 days of observed SST in a month, then this month is ignored in the local analysis. Monthly mean SSTs at a buoy location is constructed only when at least 10-year-long monthly mean SST time series are available. All TAO buoys are marked in Figure 10, but only buoys with 10 or more years of daily long SST were used. As shown, there are a total of 34 buoys where there are at least 10-year-long time series of monthly SSTs, and the number of years during which climatologies are constructed are given in Figure 10. For example, SST climatology at (0°N, 110°W), located in the eastern equatorial Pacific, is formed on the basis of daily SSTs from a total of 15 years of data. The number of daily SSTs used in obtaining the monthly climatologies is given in
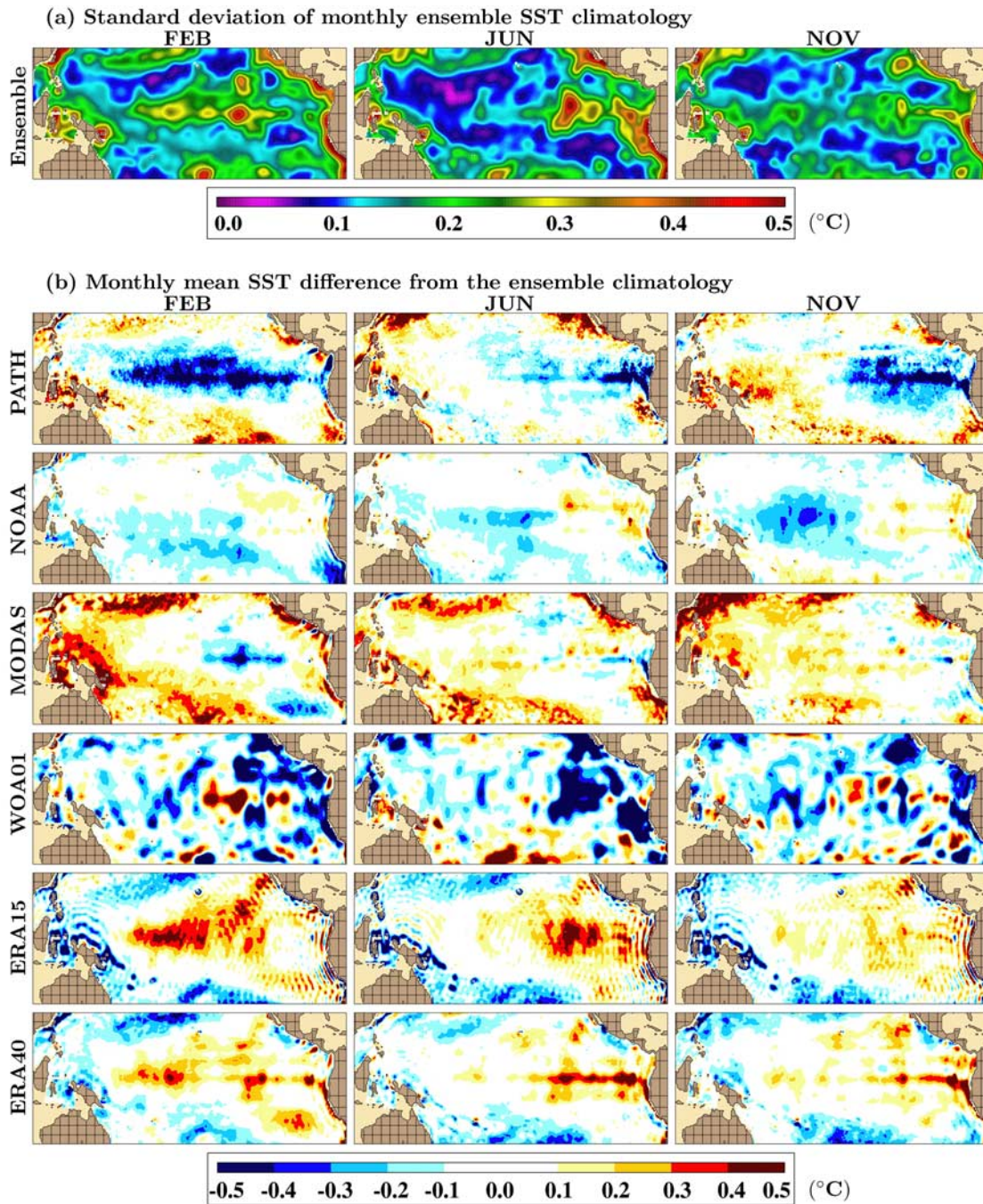
Figure 9. (a) SST standard deviation of the ensemble climatology computed using monthly mean values from PATH, NOAA, MODAS, WOA01, ERA15, and ERA40 for selected months of February, June, and August. (b) Monthly mean SST difference for each climatological SST product in comparison to the ensemble climatology. The difference is computed by subtracting the ensemble climatological SST from each product as explained in the text.

Figure 11, and may vary depending on the buoy; for example, there are relatively fewer daily SST at the other TAO location, (0°N, 165°E), because of missing buoy SST measurements in some months.

[43] Time series of climatological SST at (0°N, 110°W) from all products are compared in Figures 12a and 12b, respectively. Sampling errors based on interannual variability are computed at (0°N, 110°W) on the basis of TAO SST. This

is done to put differences among the climatologies in context and examine whether the monthly differences for each product in comparison to the buoy SST exceeds these sampling errors. In the analysis we assume that each year is independent. An estimate of sampling error is determined as follows: (1) The monthly mean SST from the TAO buoy is calculated for each year. (2) Standard deviation of these monthly means are then computed. This is the standard
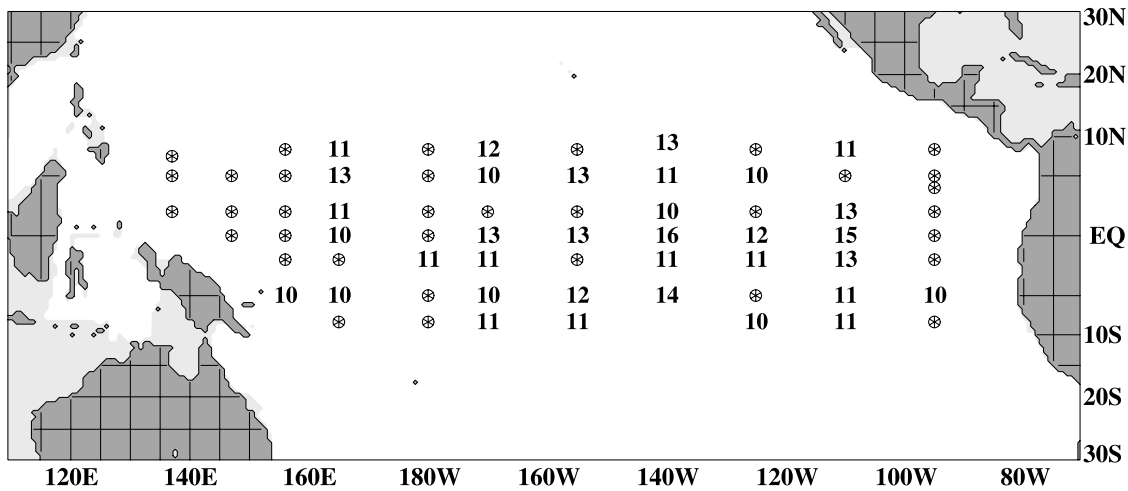
Figure 10. Locations of the moored buoys from the TAO array where monthly mean SST climatologies were constructed to compare with climatological SSTs from the six products described in section 2. The number of years which are used for forming monthly climatologies is indicated by a two-digit number.

deviation associated with interannual variability. (3) The standard error of the mean is obtained for each month by dividing standard deviation by $\sqrt{n}$ where n is the number of means for each month in the interannual series As mentioned earlier, n is set to 15 years at (0°N, 110°W), and n is set to 10 years at (0°N, 165°E). Standard error of the mean at both locations is given in Figure 13, demonstrating relatively large values at (0°N, 110°W) since SST standard deviation is higher in comparison to those at (0°N, 165°E). (4) Finally, the confidence limits at 95% are obtained. For this, standard error of the mean SST is multiplied by 1.96 and is added/subtracted from the climatological mean SST for each month.

[44] SSTs from all products generally agree with those from the buoy as seen from large positive skill scores

(Table 5). RMS values for ERA15 and ERA40 are relatively larger ($\approx$0.6°C) than those for the observation-based climatologies (<0.4°C) over the 12-month period. As seen from Figure 12, monthly SSTs from all products are generally within the upper and lower limits of observational buoy SST. Exceptions occur in February and April when SSTs from WOA01 and ERA40 fall outside of TAO uncertainty bands. ERA15 is also outside of the uncertainty range in April but it is just within the limits for February. NOAA, ERA15 and ERA40 are close but nearly inside the upper SST limit in January.

[45] SST skill is the largest for MODAS and NOAA climatologies at the two TAO buoy locations (Table 5). In the former, the climatology was formed using daily
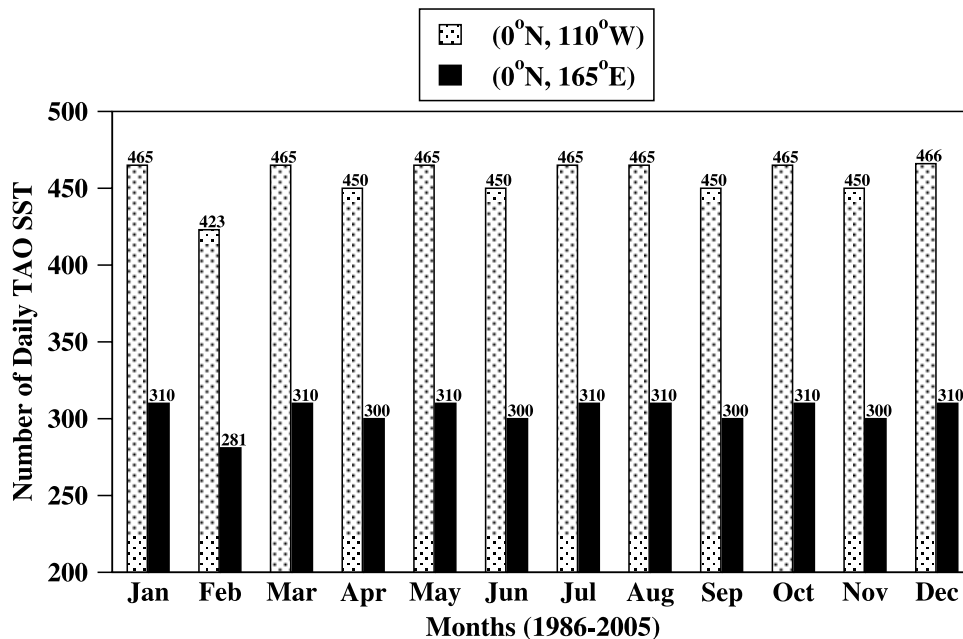

Figure 11. Total number of daily SST measurements from two TAO buoys by month during 1986–2005. These SSTs are used for forming monthly mean climatology at each buoy location.
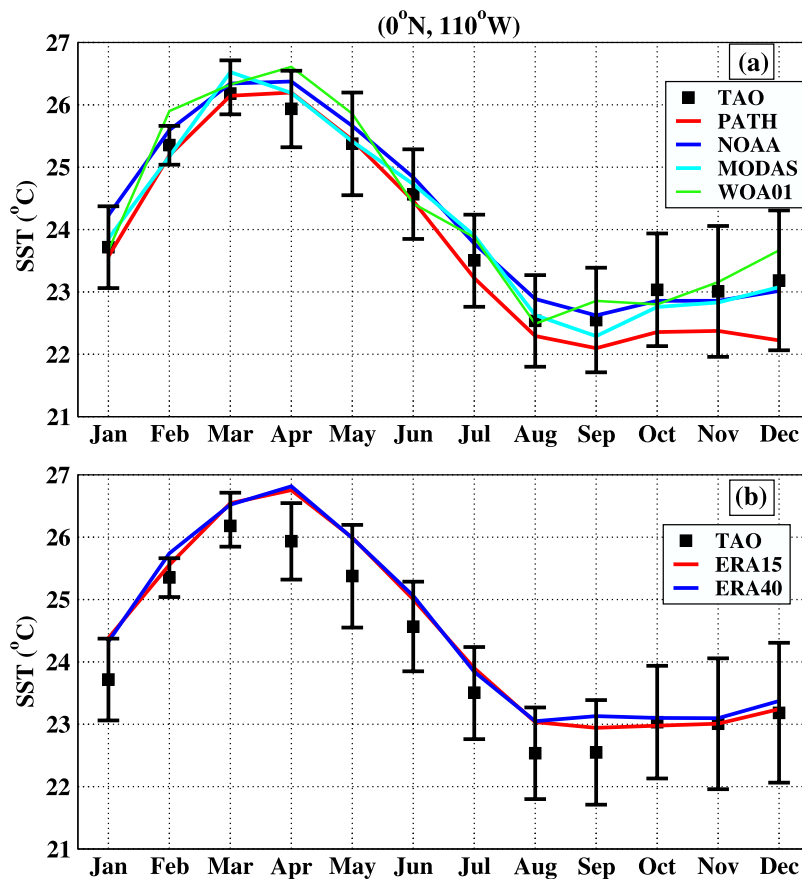
Figure 12. Climatological mean of SST time series at (0°N, 110°W) located in the eastern equatorial Pacific. Monthly SST climatology from the TAO buoy is also shown for comparison purposes. Also shown are the confidence limits at 95% based on the standard error of the SSTs from the TAO buoy.
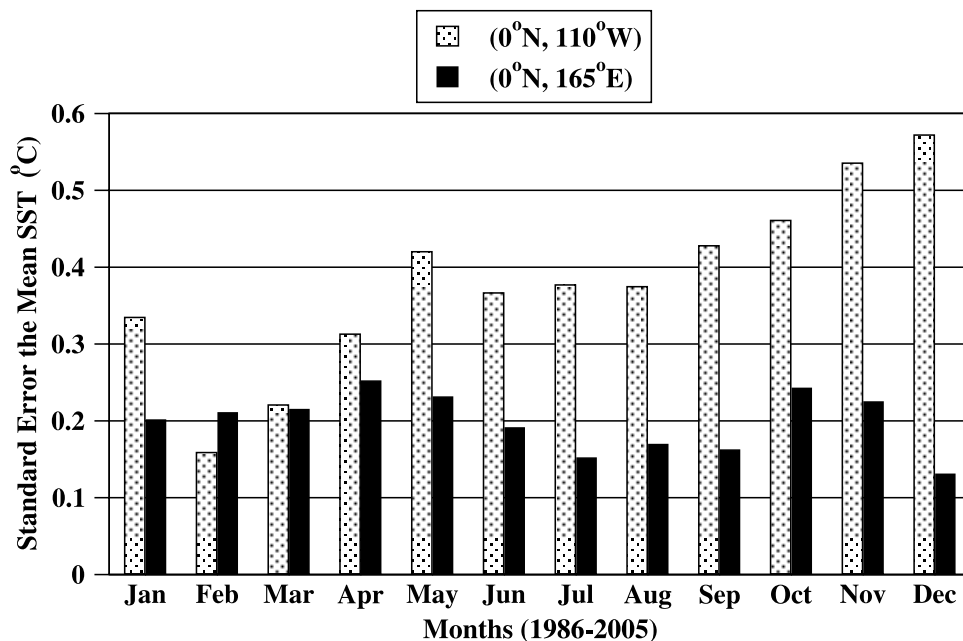

Figure 13. The standard error of the mean SST for each month. As described in the text, computations are performed on the basis of daily SSTs at two locations.

Table 5. Validation of Climatological SST at Two TAO Buoy Locations[a]

Climatology	ME (°C)	RMS (°C)	σ_X (°C)	σ_Y (°C)	R	SS
		Buoy (0°N, 110°W)				
PATH	−0.28	0.43	1.28	1.54	0.99	0.89
NOAA	0.18	0.28	1.28	1.39	0.99	0.95
MODAS	0.04	0.23	1.28	1.40	0.99	0.97
WOA01	0.22	0.36	1.28	1.42	0.98	0.92
ERA15	0.47	0.53	1.28	1.41	0.99	0.83
ERA40	0.52	0.58	1.28	1.40	0.99	0.79
		Buoy (0°N, 165°E)				
PATH	−0.23	0.29	0.14	0.24	0.71	<0
NOAA	−0.33	0.34	0.14	0.10	0.75	<0
MODAS	−0.08	0.15	0.14	0.18	0.75	<0
WOA01	−0.26	0.31	0.14	0.16	0.37	<0
ERA15	−0.01	0.09	0.14	0.10	0.74	0.55
ERA40	−0.02	0.06	0.14	0.13	0.85	0.71

[a]Buoys are located at eastern and western equatorial Pacific. Statistical values are computed on the basis of 12 monthly SSTs. Here, σ_X (σ_Y) refers to standard deviation of SSTs from buoys (products) over the annual cycle.

AVHRR-only SST time series from 1993 to 2005 (see Table 2), a time period including the one during which the monthly climatology from the TAO buoy was constructed. This indicates the consistency in producing a climatology from MODAS and TAO buoys, resulting in slightly better agreement in comparison to other products. NOAA SST climatology incorporates the buoy SST in the analysis, contributing to good agreement with buoy SSTs. Overall, even the agreement for SST climatologies from the ECMWF reanalyses in comparison those from buoys at TAO locations is about as good as the for the observation-based climatologies.

[46] One region where SST skill is relatively low for climatological products in comparison to PATH involves the western equatorial warm pool as was already indicated in Figure 5. Hence we examine the accuracy of SSTs from all climatologies relative to those from a TAO buoy located within the warm pool at (0°N, 165°E). The most obvious feature of the SST at this location is that there is very small SST variability over the 12-month period (Figure 14). Standard deviation of SST is 0.14°C, which is much smaller than 1.28°C observed at the buoy located in the eastern equatorial Pacific (0°N, 110°W). However, SST skill is also small at the former location. In fact, the SST standard deviations are so small that even though RMS values are low, skill calculations result in negative values (Table 5).

[47] Despite the low skill values at (0°N, 165°E), the monthly means from the various climatologies usually fall within the 95% confidence range for the TAO means (Figure 14). The only exceptions are that NOAA and/or WOA01 are outside the confidence limits of buoy SST in July, November, and December. Overall, results are similar to those seen at (0°N, 110°W), in that most of the monthly SSTs are within the range of observational TAO values.

[48] Finally, we combine SSTs from all 34 TAO buoys at which SST climatologies were formed using at least 10 years of monthly SSTs. The corresponding SST values at each location are extracted from each climatology, and combined
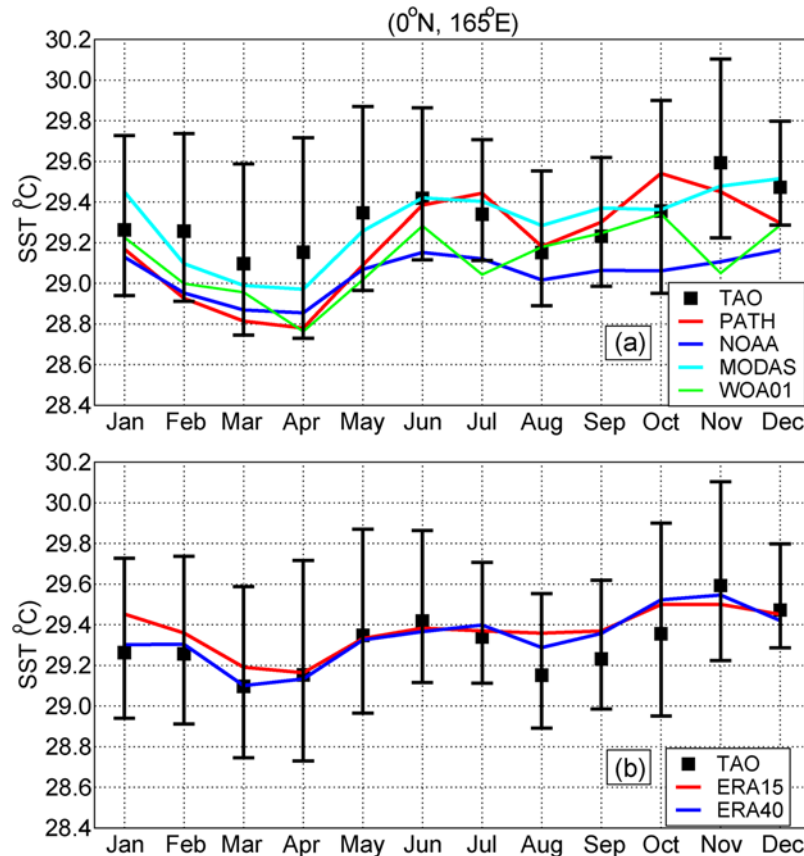

Figure 14. The same as Figure 12 but at (0°N, 165°E) located in the western equatorial Pacific.
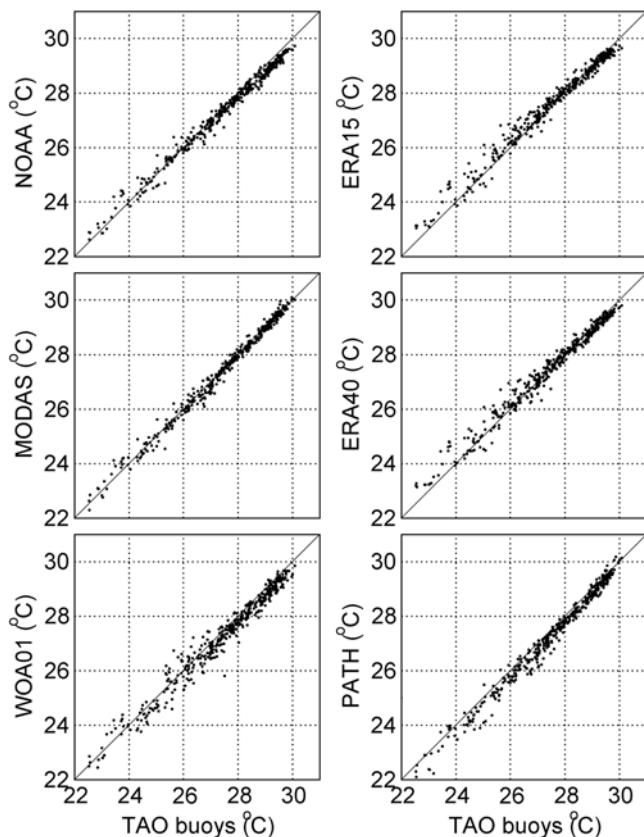
Figure 15. Scatterplots of climatological monthly mean SSTs from all 34 TAO buoys versus those from various products.

together as in the buoy SSTs. A scatterplot is then produced for SST climatologies between the buoy versus each product (Figure 15).

[49] In comparison to buoy SSTs, remarkable agreement is evident for all SST products as evident from the small scatter. In general, differences among all climatologies can be considered negligibly small as noted from positive and large skill values close to 1 (Table 6). SST skill drops to a smaller value when the same statistics are calculated only at the six buoys located within the western equatorial Pacific warm pool. Again, one should note that small RMS errors in the warm pool region due to low SST variability do not really indicate that all products have good skill in representing observed SST changes. For example, NOAA SST has lower RMS value of 0.22°C in the warm pool region, but it has also much lower skill. *Reynolds et al.* [2007] further provided a discussion of PATH and in situ SST differences, affecting these climate biases. Typically, PATH tends to be biased cool because of cloud contamination. For in situ data, ships tend to be biased warm relative to buoys.

7. Conclusions

[50] Studies of climate change require that baseline climate be accurately identified. The tropical Pacific is a key region that drives global heat budgets, and its influence on climate variability extends far beyond its local vicinity. Accurate assessment of mean SST, its variability and its uncertainty is

necessary to discern the significance on spatial and temporal timescales and thereby give increased confidence in analyzing various topics on climate change.

[51] Through statistical analyses, we demonstrate that the agreement among all SST climatologies considered in this paper is quite remarkable. However, they typically have low skill in matching the buoy climatology in the western Pacific warm pool; in fact, only the ERA15 and ERA40 climatologies have positive skill. Yet the mean and RMS errors are fairly small. The low skill is a consequence of the very low standard deviation in the TAO observations, which amplifies the impact of the small RMS errors. On the contrary, RMS errors are larger for the comparisons at 110°W, but correlations and skill score are much higher here, approaching the perfect value of 1. The other climatologies indicate similarly low standard deviation at (0°N, 165°E) with the TAO SST approximately the median of the values. SST standard deviations in this region are close to the precision of the measurements themselves. Thus, this paper also provides a sound basis for interpreting SST correlations and skill score.

[52] The significance of differences among SST products elsewhere should be judged on the basis of the intended applications. SST differences of a few tenths of a degree are unimportant for daily weather forecasts but may be crucial for estimates of global climate change over multiple decades. The variability among climatologies suggests that a multiyear temperature change of this magnitude might be difficult to discern in the eastern equatorial Pacific but an enormous event in the western Pacific warm pool. Those using the climatologies can use the information presented here to estimate the effects of climatological uncertainty on their own model simulations or other calculations. Identifying the size of the differences necessary to produce significantly different model outcomes is beyond the scope of this paper, but is a critical issue for climatological applications.

[53] Finally, we expect that the spatial and temporal variability displayed by the SST climatologies in the tropical Pacific, as presented in this paper, extend to other regions. This is very likely for those data sets which do not themselves

Table 6. SST Error Statistics Between Climatological Products and Buoys at TAO Locations[a]

Climatology	ME (°C)	RMS (°C)	σ_X (°C)	σ_Y (°C)	R	SS
Statistics in the Whole Pacific Domain						
PATH	−0.23	0.32	1.72	1.78	0.99	0.97
NOAA	−0.13	0.26	1.72	1.59	0.99	0.98
MODAS	−0.03	0.18	1.72	1.71	0.99	0.99
WOA01	−0.19	0.35	1.72	1.66	0.98	0.96
ERA15	0.07	0.27	1.72	1.59	0.99	0.97
ERA40	0.02	0.26	1.72	1.59	0.99	0.97
Statistics in the Western Equatorial Pacific						
PATH	−0.13	0.22	0.41	0.49	0.94	0.72
NOAA	−0.23	0.27	0.41	0.33	0.95	0.57
MODAS	−0.01	0.12	0.41	0.41	0.96	0.91
WOA01	−0.21	0.28	0.41	0.38	0.90	0.53
ERA15	−0.13	0.19	0.41	0.36	0.94	0.78
ERA40	−0.11	0.16	0.41	0.39	0.97	0.84

[a]Statistics are computed on the basis of monthly mean time series from all 34 TAO buoys (34 × 12 = 408 monthly SST values) and from 6 buoys (6 × 12 = 72 monthly SST values) located west of 180°W, separately. Here, σ_X refers to standard deviation of buoy SSTs, and σ_Y refers to standard deviation of climatological SST product.

incorporate the TAO observations. Long-term observational efforts similar to the TAO effort would be useful for answering these questions and identifying the uncertainties in predictions of climate change.

References

Barron, C. N., and A. B. Kara (2006), Satellite-based daily SSTs over the global ocean, *Geophys. Res. Lett.*, *33*, L15603, doi:10.1029/2006GL026356.

Boyer, T. P., C. Stephens, J. I. Antonov, M. E. Conkright, R. A. Locarnini, T. D. O'Brien, and H. E. Garcia (2002), *World Ocean Atlas 2001*, vol. 2, *Salinity, NOAA Atlas NESDIS 50*, 176 pp., edited by S. Levitus, U.S. Gov. Print. Off., Washington, D. C.

Casey, K. S., and P. Cornillon (1999), A comparison of satellite and in situ based sea surface temperature climatologies, *J. Clim.*, *12*, 1848–1863.

Chen, W. Y., and H. M. van den Dool (1997), Asymmetric impact of tropical SST anomalies on atmospheric internal variability over the North Pacific, *J. Atmos. Sci.*, *54*, 725–740.

Diaz, H. F., and V. Markgraf (2000), *El Niño and the Southern Oscillation: Multiscale Variability and Global and Regional Impacts*, 496 pp., Cambridge Univ. Press, New York.

Fiedler, P. C., and L. D. Talley (2006), Hydrography of the eastern tropical Pacific: A review, *Prog. Oceanogr.*, *69*, 143–180.

Gibson, J. K., P. Kållberg, S. Uppala, A. Hernandez, A. Nomura, and E. Serrano (1999), ERA description (version 2), *ECMWF Re-Anal. Proj. Rep. Ser. 1*, 74 pp., Reading, U.K.

Kara, A. B., and C. N. Barron (2007), Fine-resolution satellite-based daily sea surface temperatures over the global ocean, *J. Geophys. Res.*, *112*, C05041, doi:10.1029/2006JC004021.

Kara, A. B., C. N. Barron, A. J. Wallcraft, T. Oguz, and K. S. Casey (2008), Advantages of fine resolution SSTs for small ocean basins: Evaluation in the Black Sea, *J. Geophys. Res.*, *113*, C08013, doi:10.1029/2007JC004569.

Kessler, W. S., L. M. Rothstein, and D. Chen (1998), The annual cycle of SST in the eastern tropical Pacific, diagnosed in an ocean GCM, *J. Clim.*, *11*, 777–799.

Luo, G., P. A. Davis, L. B. Stowe, and E. P. McClain (1995), A pixel-scale algorithm of cloud type, layer and amount for AVHRR data. Part I: Nighttime, *J. Atmos. Oceanic Technol.*, *12*, 1013–1037.

McClain, E. P., W. G. Pichel, and C. C. Walton (1985), Comparative performance of AVHRR-based multichannel sea surface temperatures, *J. Geophys. Res.*, *90*, 11,587–11,601.

McPhaden, M. J., et al. (1998), The Tropical Ocean-Global Atmosphere (TOGA) observing system: A decade of progress, *J. Geophys. Res.*, *103*, 14,169–14,240.

Murphy, A. H. (1995), The coefficients of correlation and determination as measures of performance in forecast verification, *Weather Forecasting*, *10*, 681–688.

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

Reynolds, R. W., T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax (2007), Daily high-resolution-blended analyses for sea surface temperature, *J. Clim.*, *20*, 5473–5496.

Schneider, N., A. J. Miller, and D. W. Pierce (2002), Anatomy of North Pacific decadal variability, *J. Clim.*, *15*, 586–605.

Stephens, C., J. I. Antonov, T. P. Boyer, M. E. Conkright, R. A. Locarnini, T. D. O'Brien, and H. E. Garcia (2002), *World Ocean Atlas 2001*, vol. 1, *Temperature* [CD-ROM]*, NOAA Atlas NESDIS 49*, 167 pp., edited by S. Levitus, U.S. Gov. Print. Off., Washington, D. C.

Thum, N., S. K. Esbensen, D. B. Chelton, and M. J. McPhaden (2002), Air-sea heat exchange along the Northern sea surface temperature front in the eastern tropical Pacific, *J. Clim.*, *15*, 3361–3378.

Uppala, S., et al. (2005), The ERA-40 re-analysis, *Q. J. R. Meteorol. Soc*, *131*, 2961–3012, doi:10.1256/qj.04.176.

Wang, W., and M. J. McPhaden (1999), The surface layer heat balance in the equatorial Pacific Ocean, part I: Mean seasonal cycle, *J. Phys. Oceanogr.*, *29*, 1812–1831.

Webster, P. J. (1995), The annual cycle and the predictability of the tropical coupled ocean-atmosphere system, *Meteorol. Atmos. Phys.*, *56*, 33–55.

Zhang, D., and M. J. McPhaden (2006), Decadal variability of the shallow Pacific meridional overturning circulation: Relation to tropical sea surface temperatures in observations and climate change models, *Ocean Modell.*, *15*, 250–273.

————————————

C. N. Barron and A. B. Kara, Oceanography Division, Naval Research Laboratory, Code 7320, Building 1009, Stennis Space Center, MS 39529, USA. (birol.kara@nrlssc.navy.mil)

T. P. Boyer, NOAA National Oceanographic Data Center, 1315 East-West Highway, Silver Spring, MD 20910–3282, USA.