ADVANCES IN ATMOSPHERIC SCIENCES, VOL. 23, NO. 3, 2006, 474–486

Derivation of Regression Coefficients for Sea Surface Temperature Retrieval over East Asia

Myoung-Hwan AHN*[1], Eun-Ha SOHN[1], Byong-Jun HWANG[1], Chu-Yong CHUNG[1], and Xiangqian WU[2]

[1] *Remote Sensing Research Laboratory/Meteorological Research Institute, Seoul, Korea*

[2] *Office of Research and Applications, National Environmental Satellite, Data, and Information Service, National Oceanic and Atmospheric Administration*

(Received 20 July 2005; revised 16 November 2005)

ABSTRACT

Among the regression-based algorithms for deriving SST from satellite measurements, regionally optimized algorithms normally perform better than the corresponding global algorithm. In this paper, three algorithms are considered for SST retrieval over the East Asia region (15°–55°N, 105°–170°E), including the multi-channel algorithm (MCSST), the quadratic algorithm (QSST), and the Pathfinder algorithm (PFSST). All algorithms are derived and validated using collocated buoy and Geostationary Meteorological Satellite (GMS-5) observations from 1997 to 2001. An important part of the derivation and validation of the algorithms is the quality control procedure for the buoy SST data and an improved cloud screening method for the satellite brightness temperature measurements. The regionally optimized MCSST algorithm shows an overall improvement over the global algorithm, removing the bias of about −0.13°C and reducing the root-mean-square difference (rmsd) from 1.36°C to 1.26°C. The QSST is only slightly better than the MCSST. For both algorithms, a seasonal dependence of the remaining error statistics is still evident. The Pathfinder approach for deriving a season-specific set of coefficients, one for August to October and one for the rest of the year, provides the smallest rmsd overall that is also stable over time.

Key words: Regional SST algorithm, GMS-5, Data quality control, cloud detection

doi: 10.1007/s00376-006-0474-7

1. Introduction

Satellite estimation of sea surface temperature (SST) has greatly improved the spatial coverage and resolution provided by in-situ observations such as that from ship or buoy. Such estimation must overcome the interference by the atmosphere between the ocean surface and satellite, primarily the water vapor absorption of the upwelling radiation, even in the relatively "clean" atmospheric window spectrum. Following Saunders's early experiment (Saunders, 1967), McClain et al. (1985) introduced the Multi-Channel SST (MCSST) algorithm that assumes a linear relationship between the temperature difference at two window channels (hereafter referred to as DT) and the total amount of water vapor. Walton (1988) and Walton et al. (1998) suggested Cross Product SST (CPSST) and Non-Linear SST (NLSST) to account for the nonlinearity between DT and water vapor. Wu et al. (1999) adopted a quadratic formula to account for the nonlinearity effect in deriving an SST algorithm (QSST) for the Geostationary Operational Environmental Satellite (GOES).

It has also been found (e.g., Barton, 1995) that a locally optimized regression algorithm is superior to the corresponding global algorithm for specific regions. For example, Shenoi (1999) showed that the use of regional regression coefficients with the NLSST algorithm for the northern Indian Ocean reduces the root-mean-square difference (rmsd) by 50% compared to the global Pathfinder SST (PFSST, Kilpatrick et al., 2001). A theoretical study (Minnett, 1990) suggested that the global algorithm had a higher error, up to 1 K, than a regional algorithm in the Greenland-Iceland-Norwegian area. The improvement is primarily due to the adaptation of the algorithm to the regional atmospheric conditions.

Most of the space-based SST estimates are derived from polar orbiting satellites such as the National Oceanic and Atmospheric Administration (NOAA) series and European Remote Sensing (ERS) satellites. Compared to the geostationary platform, the polar or-

*E-mail: mhahn@kma.go.kr

biting satellites are much closer to the target. This affords a higher signal to noise ratio and a higher spatial resolution, which translate into higher radiometric accuracy and better cloud detection and avoidance, both of which are critical to SST retrieval. Nevertheless, the geostationary platform offers much more frequent observations over selected regions, 24 times a day in the case of the Geostationary Meteorological Satellite (GMS-5) compared to twice a day by a polar orbiting satellite. It is therefore much more likely for a geostationary platform to find a clear moment for SST retrieval (Bates and Smith, 1985; Legeckis and Zhu, 1997). Furthermore, Wu et al. (1999) introduced an algorithm that makes use of frequent observations by GOES to help with cloud detection. Their GOES SST product shows more complete SST maps in cloudy regions and the diurnal variation of SST in clear regions.

A series of GMS satellites has been operated by the Japan Meteorological Agency (JMA) since 1977, covering most of the Western Pacific Ocean and part of the Indian Ocean. The latest one, GMS-5 that was launched in 1995, has the split window channels (10.5–11.5 μm and 11.5–12.5 μm) that enable the derivation of SST. Shirakawa (1996) derived the MCSST regression coefficients for the full disk coverage of GMS-5, which has a bias of -0.4 K and an rmsd of 1.5 K. Later, Yasuda and Shirakawa (1999) improved the cloud screening method using the objectively analyzed climate SST, which reduced the rmsd by 0.3 K (to 1.2 K). It should be noted that GMS-5 has a relatively high NEDT (Noise Equivalent Delta Temperature), 0.35 K at 300 K (MSC, 1997), and lacks a shortwave infrared channel at 3.9 μm.

The main purpose of this paper is to derive a regionally optimized SST retrieval algorithm using the GMS-5 data over the East Asia region, from 15°N to 55°N and from 105° E to 170°E, which covers the domain of the Korea Meteorological Administration (KMA) regional weather prediction model. Based on previous studies, regression coefficients for three algorithms (MCSST, QSST, and PFSST) are derived using the collocated observations by buoy and the GMS-5 from 1997 to 1999, and these are then validated using similar data from 2000 to 2001. After this introduction, section 2 describes the data used in the current study, including the quality control procedure for the buoy data and the cloud screening procedure for the GMS-5 data. The regression coefficients for all algorithms are derived and compared in section 3. Validation results are presented in section 4, and the paper concludes in section 5 with a summary.

2. Collection and processing of data

This study is based on the analysis of collocated buoy and GMS-5 data. It is critical to understand and document how the data were collected and processed, as well as the general characteristics of the final datasets.

2.1 *Data collection and collocation*

The buoy data used in this paper are obtained from the drifting and moored buoys via the Global Telecommunication System. The GMS-5 data are acquired from the Division of Meteorological Satellite of the KMA. The satellite data are navigated with a coastline match-up process using the window channel brightness temperature and topographic data, and by manual adjustment when necessary. The navigation accuracy is better than one pixel (5 km). An array of 3×3 pixels is collocated with each buoy report within 30 minutes. Each collocated dataset consists of date, time, and location of the buoy report, the buoy SST, the reflectance and the split-window brightness temperatures (Tb_{11} at 11 μm, and Tb_{12} at 12 μm) from the GMS-5, and the satellite zenith angle of the measurement. All the available buoy and GMS-5 data within the region (15°–55°N, 105°–170°E) and time period (1997–2001) are collected. The total number of the collocated observations before any quality control and cloud screening procedures is 143980.

2.2 *Quality control of the buoy data*

The buoy data are prone to error arising from the malfunctioning of the buoy itself or in the process of data communication. The malfunctioning can be caused by a drained buoy battery, accidental ship traction, or the grounding of the buoy that exposes the sensor to the sun. A rigorous quality control procedure is required to reject the erroneous buoy data, and the procedure must be objective to maintain uniform criteria for selection while processing large amounts of data. The procedure adopted for this study tests short-term and long-term variations of the measured SST (Hansen and Poulain, 1996). The basic strategy is to retain only the highly reliable data while sacrificing a few "good" data points.

Figure 1 shows the overall flowchart of the quality control procedure for the buoy data. First, a buoy is disqualified if the total number of data points is less than 20, which simply removes infrequently reporting buoy stations. The next test is the short-term variation, which can be caused by the passage of a thermal front or an eddy and the diurnal variation due to solar insolation (Stramma et al., 1986; Webster et al., 1996). The insolation-induced diurnal variation of SST is normally less than 3 K per day (Wu et al., 1999; Webster et al., 1996). The SST variation caused by frontal passage is larger and can be estimated from its spatial gradient and movement speed. While the temperature gradient is typically on the order of 0.1–0.2°C km^{-1} in the East Sea (Park, 1996), a more intensive gradient

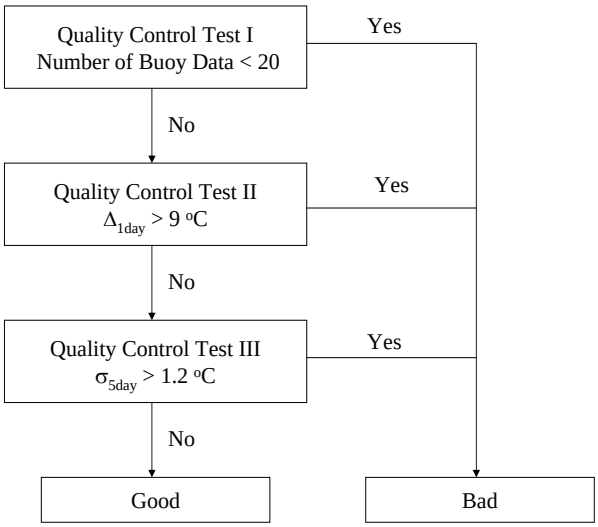

Fig. 1. Flow chart for the quality control of the buoy SST. $\Delta_{1\mathrm{day}}$ and $\sigma_{5\mathrm{day}}$ represent the temperature gradient for one day and the standard deviation of a given buoy for five days, respectively.

(up to 1.5°C km^{-1}) has been reported for very strong thermal fronts (Rodon, 1975). The typical speed of oceanic eddies or fronts is 9 km per day (Gill, 1982). With all these considerations, we believe that the temperature variation by natural processes is less than 9°C d^{-1}; the buoy data with short-term variation larger than this threshold value are discarded.

The final test is the longer-term variation of the buoy SST. During a 5-day period and in the absence of a repeated passage of strong thermal eddies, the SST fluctuation is mainly caused by its diurnal variation that ranges from 0.5°C to 3.5°C (Legeckis and Zhu, 1997; Webster et al, 1996; Stramma et al., 1986). Assuming the maximum SST diurnal variation to be less than 4°C, the expected standard deviation of SST during a 5-day period is less than 1.2°C. Thus, if the standard deviation of 5-day SST is more than 1.2°C, all data in the period are discarded.

2.3 *Cloud screening*

Shirakawa (1996) suggested a cloud screening method for the operational retrieval of the GMS-5 SST, which includes a thermal gross test, thermal spatial coherence test, split window test, albedo test, and climatology test. Yasuda and Shirakawa (1999) later modified the threshold values to improve the cloud screening performance. While it may be appropriate for operations, Ahn et al. (2001) found that some cloudy pixels escape the Yasuda and Shirakawa tests, especially at night due to the lack of the visible and 3.9 μm data.

For this study, we developed a cloud screening procedure that is similar in concept to that of Yasuda and Shirakawa (1999) but with more stringent threshold values, as shown in Fig. 2. The first test removes the collocated data whose Tb_{11} is more than 15°C cooler than the buoy SST. This threshold value is based on a model simulation of the water vapor effect on Tb_{11}. The second test removes the collocated data whose split window channel difference is too large or negative. The latter test detects instrument irregularities and heavy dust loading (Ahn et al., 2003b). The third test ensures thermal uniformity around the buoy location, which is often destroyed by broken cloud. Based on previous theoretical estimates (Ahn et al., 2003a), the collocated dataset is assumed to be cloud contaminated if the standard deviation of the 3×3 pixel array of the GMS-5 Tb_{11} data is more than 0.8°C. During the daytime, additional tests are applied to the visible data. If the averaged albedo of the nine pixels is larger than 0.05, or the standard deviation is larger than 0.03, we assume the pixel is contaminated by clouds. This test is effective for all clouds, thick or thin, uniform or broken, high or low. Finally, we delete the collocated data whose estimated SST using the Yasuda and Shirakawa algorithm is more than 4°C cooler than the buoy SST. This threshold value is sufficiently larger than the reported accuracy (1.2°C) and removes the

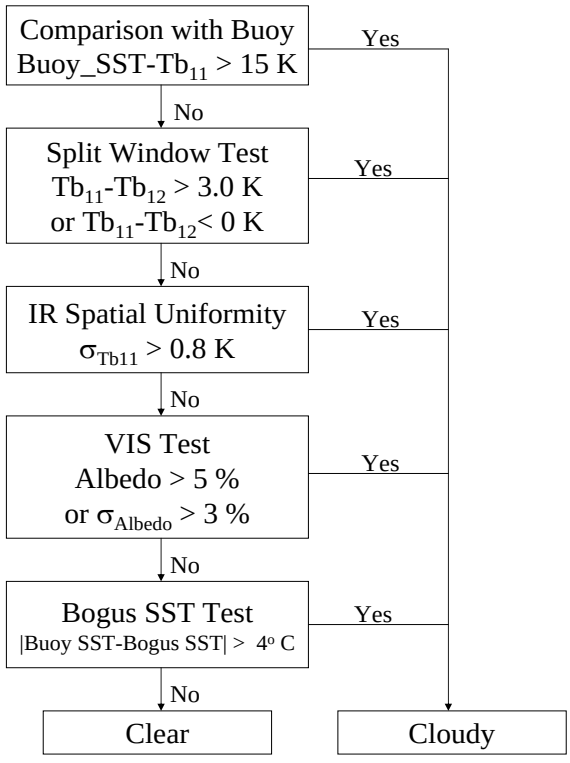

Fig. 2. Flow chart for the cloud screening procedures. The brightness temperature is given at 11 μm (Tb_{11}) and 12 μm (Tb_{12}), and albedo is the averaged value for 9 pixels near the buoy position, while σ_{Tb11} and σ_{Albedo} are the standard deviations of Tb_{11} and albedo, respectively. The global SST is obtained from the global MCSST algorithm (see text).

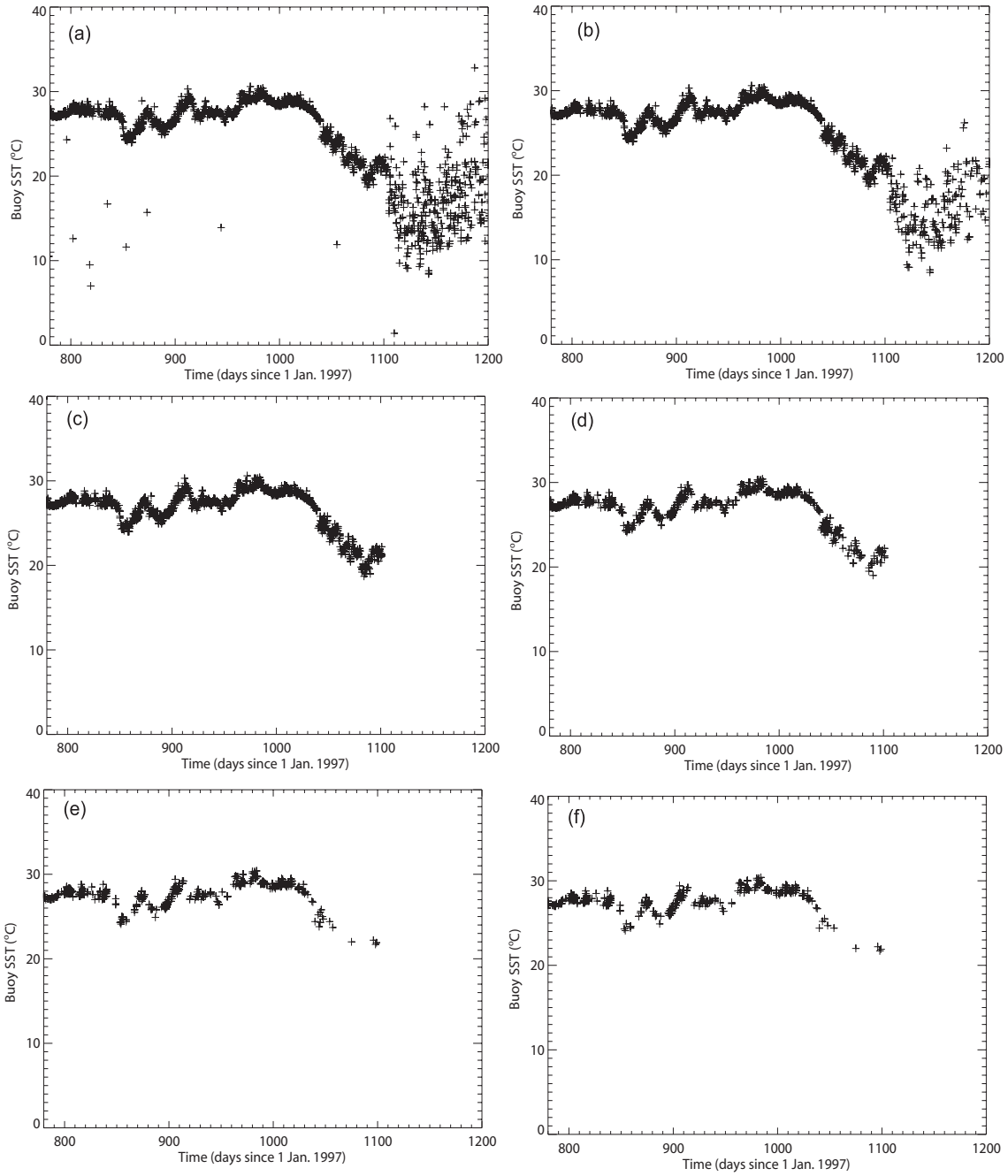

Fig. 3. Time series of the buoy SST obtained from GTS for WMO drift buoy number 21612. (a) The original time series shows the many randomly distributed peaks with the especially high-varied peaks at the end of series indicating error caused by malfunction. (b) The distribution of the buoy SST after quality control II, (c) after quality control III, (d) comparison with buoy SST under the cloud screening procedure, (e) split window test and IR (infrared) spatial uniformity test, and (f) VIS (visible) test and Global SST test.

remaining erroneous data (especially at night).

2.4 *Example*

Figure 3 shows the results of the sequence of tests described above when applied to the observed SST by drifting buoy number 21612. The early part of the original time series (Fig. 3a) varies smoothly with occasional anomalies, whereas the later part of the time series is quite noisy. The first quality control test for the buoy data does not apply to this particular case, but the second test that limits the daily SST variation to less than 9°C removes virtually all the abnormal

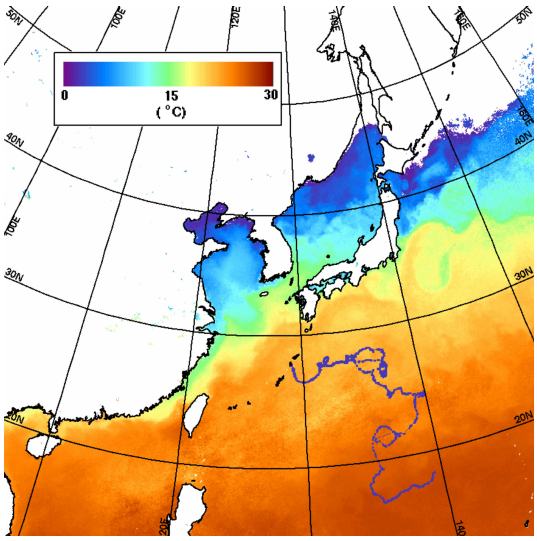

Fig. 4. Trajectory of drift buoy number 21612, which ran into an island. The background is the spatial distribution of the monthly mean SST in January 2000.

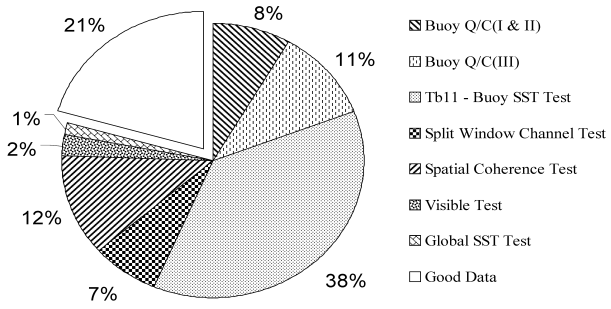

Fig. 5. Pie chart of the percentage of data removed by each buoy Q/C (quality control) and cloud screening procedure. The largest number of data points is removed by the simple comparison between Tb_{11} and the buoy SST test with a threshold of 15°C.

data from the early part of the time series (Fig. 3b). However, most of the noisy data at the end of the time series remain, which is often indicative of a drifter running aground or in shallow waters (Hansen, 2001; private communication). To confirm this hypothesis, we examined the buoy trajectory over the monthly mean SST field (Fig. 4), which shows that the buoy indeed became stuck near the island of Tokunoshima of Japan at the end of the time series. The third quality control test for the buoy data that limits the 5-day SST variation to less than 1.2°C removes most of the noisy data at the end of the time series (Fig. 3c), which is justified by Fig. 4. About 19% of the collocated data are removed by the quality control procedure for buoy data.

The remaining collocated datasets are further subject to cloud screening tests. Those that survived the buoy SST and Tb_{11} comparison test, the spatial uniformity test and split window test, and the bogus SST test are plotted in Figs. 3d, 3e, and 3f, respectively. A large portion of data, about 45% of the data remaining after the buoy Q/C procedure, is removed by the simple comparison between Tb_{11} and buoy SST. At the end, about 19% of the original collocated data pass all the tests and are considered highly reliable data (Fig. 3f).

2.5 *Characteristics of the final data*

The quality control procedure for the buoy data and the cloud screening procedure for the GMS-5 data are applied to all the collocated data. The portion that was removed by each test is summarized in Fig. 5. Of the original 143980 collocated data points, 25344 (or 21%) pass all the tests and are considered reliable. There are roughly equal number of collocations during daytime and nighttime (12917 vs. 12427, each is about 21% of the original daytime and nighttime data), suggesting that the overall data preparation procedures for nighttime might be as good as those for daytime.

Because the regression coefficients depend strongly on the latitudinal coverage of the collocated data (Emery et al., 2001) and on other parameters such as the observation geometry, the characteristics of the final data are examined. Figure 6 shows the geographical distribution of the collocated data from 1997 to 2000. While the data are sparse in some high latitude regions and in the upwelling coastal area of the East China Sea, and data are dense over the East Sea, and for the most part the collocated data are fairly uniformly distributed.

We further examined the data characteristics in terms of the month of observation, the satellite zenith

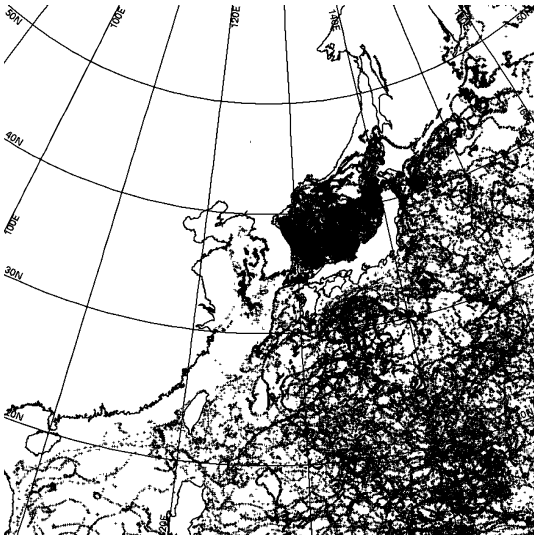

Fig. 6. Geographical distribution of the quality-controlled and cloud-screened collocated data for the East Asia region during 1997 to 2000.

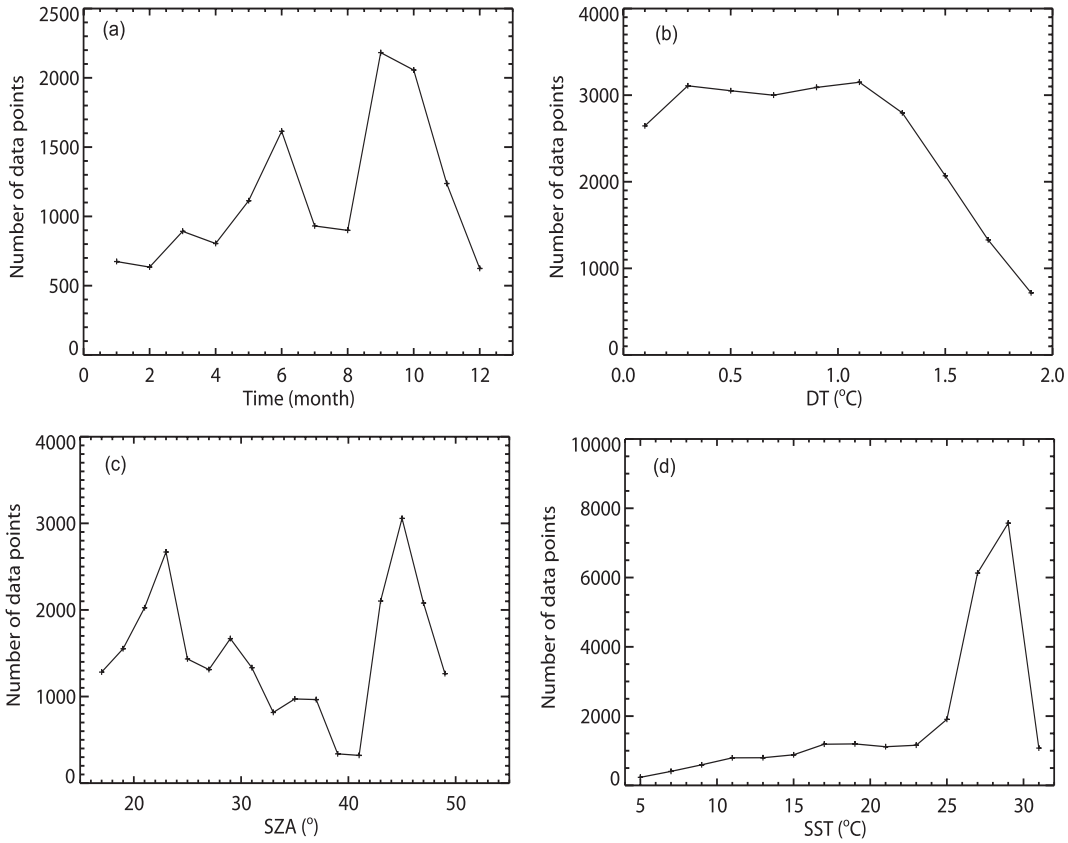

Fig. 7. Frequency distribution of the collocated satellite and buoy data as a function of month, DT (Tb_{11}–Tb_{12}) and SZA (Satellite Zenith Angle), and buoy SST for the period of 1997 to 1999 in the East Asia region.

angle of observation (SZA), the split window channels brightness DT, and the buoy SST. These are the parameters later used in the regression equation. The results are plotted in Fig. 7. The number of collocated data points is relatively small in winter and summer, peaking in June and particularly in September and October (Fig. 7a). The number of collocated data points is uniformly distributed for DT less than about 1.5°C, decreasing gradually with increasing DT above 1.5°C and dropping abruptly when DT is greater than about 2°C (Fig. 7b). There are more than 700 collocated data points for all values of SZA except for around 40 degrees (Fig. 7c). Finally, the number of collocated data points demonstrates a profound peak around 28°C while it is uniformly distributed for other values of buoy SST (Fig. 7d). The effect of these irregular distributions of data points is discussed later.

3. Derivation of the regression coefficients

Three types of regression algorithms are tested for suitability for SST retrieval using the GMS-5 data in the East Asia region. These include MCSST that assumes a linear relationship between the atmospheric absorption and the DT (McClain et al., 1985); QSST that accounts for the non-linear effect of water vapor absorption with a quadratic term of DT (Wu et al., 1999); and PFSST (Kilpatrick et al., 2001) that accounts for the different atmospheric conditions.

3.1 *Regional MCSST and QSST*

When the global MCSST algorithm of Yasuda and Shirakawa (1999) is applied to our dataset, the bias and rmsd are −0.13°C and 1.36°C, respectively, which are larger than the reported accuracy when evaluated over the full disk. The bias and rmsd are defined as:

$$\text{bias} = \frac{1}{n}\sum_{n}(\text{SST}_{\text{Satellite}} - \text{SST}_{\text{Buoy}}) \,,$$

$$\text{rmsd} = \sqrt{\frac{1}{n}\sum_{n}(\text{SST}_{\text{Satellite}} - \text{SST}_{\text{Buoy}})^2} \,.$$

Furthermore, Fig. 8 shows a clear dependence of the errors on certain parameters. The bias is negative from August to October while positive for the rest of the year; negative for DT greater than 0.7°C while positive when the atmosphere is dry; negative for small SZA (< 35°) while positive when the scene

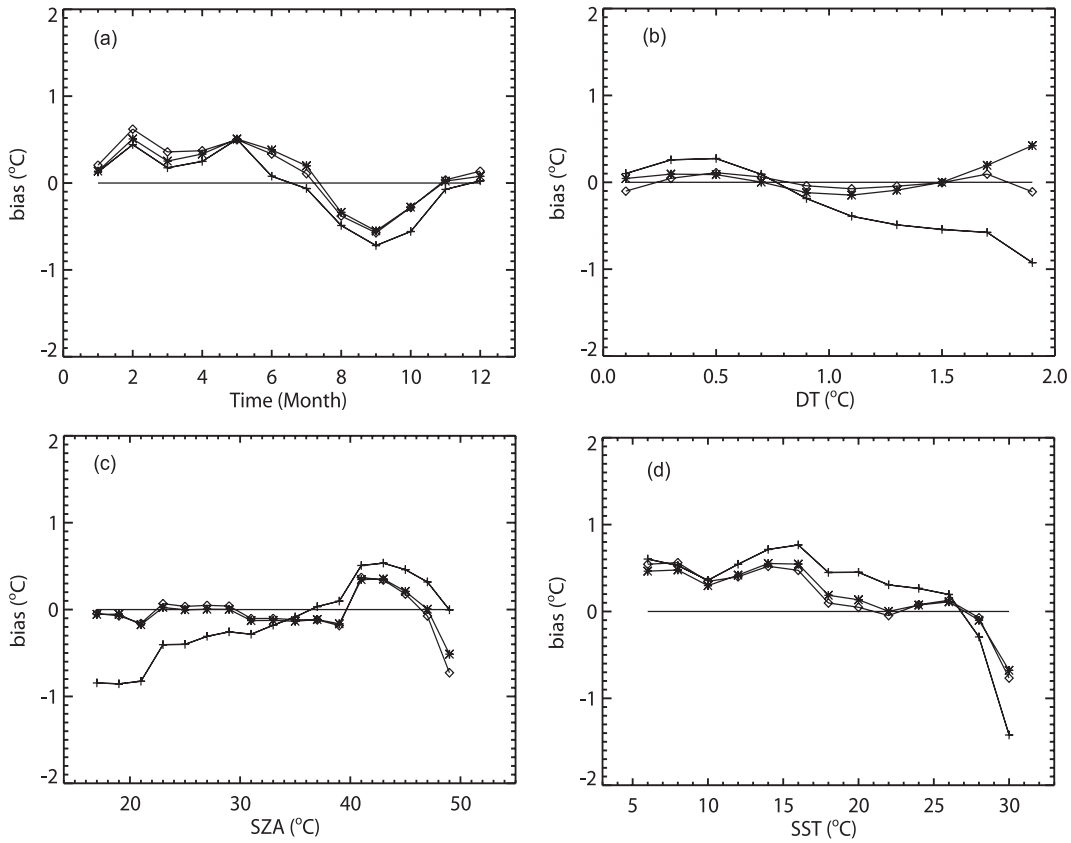

Fig. 8. Biases of global MCSST (GMCSST; +), regional MCSST (RMCSST; *), and regional QSST (RQSST; ♢) algorithms as functions of (a) month, (b) DT, (c) SZA, and (d) buoy SST.

Table 1. Regression coefficients, rmsd, and bias for global MCSST (GMCSST), regional MCSST (RMCSST) and QSST (RQSST), temporally-varying RMCSST (RTMCSST), and RQSST (RTQSST). For the RTMCSST and RTQSST, the regression coefficients are shown for the months of August, September, and October (Period 2) and for the other months (Period 1).

Algorithm		A	B	C	D	E	rmsd (°C)	bias (°C)
GMCSST		1.07177	2.31327	2.59312	−16.8281		1.36	−0.13
RMCSST		1.0480	3.2672	−0.9151	3.0144		1.26	0.0
RQSST		1.0170	3.5635	−1.5840	−0.2507	3.7818	1.25	0.0
RTMCSST	Period 1	1.0336	3.3583	−2.1301	3.0839		1.19	0.0
	Period 2	0.9180	3.1452	−1.8803	6.2805			
RTQSST	Period 1	0.9969	2.9302	−2.7186	−0.008	4.3860	1.16	0.0
	Period 2	0.7383	3.9528	−5.1217	−0.7299	10.6243		

is farther away from the nadir; and negative for warm SST (> 27°C) while positive otherwise. As will be shown later, similar dependences are also found for the rmsd. Although the causes are not clear, many of these dependencies can be reduced empirically.

Regionally optimized MCSST and QSST algorithms are found by deriving regression coefficients for the following equations using the 1997–1999 data from our dataset:

$$\text{MCSST} = A \times \text{Tb}_{11} + B \times \text{DT} + C \times \left[\frac{1}{\cos(\text{SZA})} - 1\right] \times \text{DT} + D\,, \quad (1)$$

$$\text{QSST} = A \times \text{Tb}_{11} + B \times \text{DT} + C \times \left[\frac{1}{\cos(\text{SZA})} - 1\right] + D \times \text{DT}^2 + E\,, \quad (2)$$

where A, B, C, D, and E are the regression coefficients. In Eqs. (1) and (2), the first terms on the right hand sides represent the close relationship between

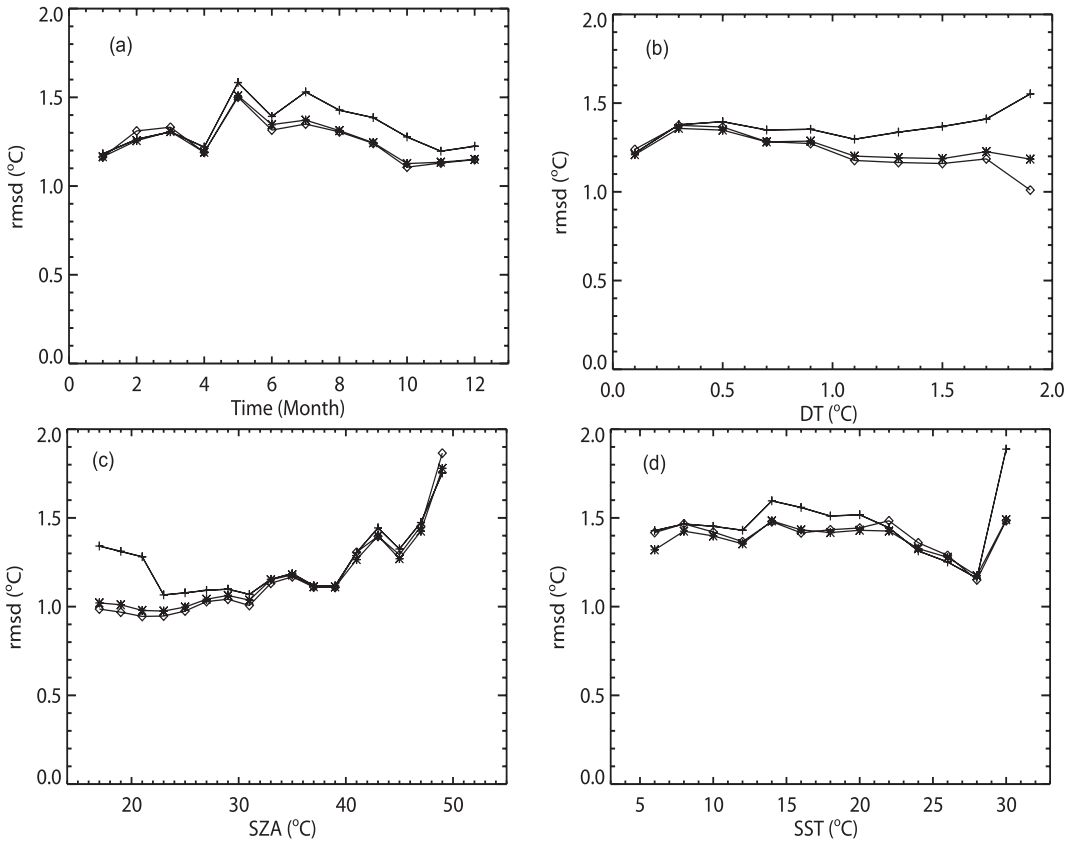

Fig. 9. Same as Fig. 8 but for rmsd.

Tb_{11} and the SST; the second terms account for the linear relationship between the atmospheric effect and DT; and the third terms correct for the optical path difference when viewed at different angles. Additionally, the fourth term in Eq. (2) represents the nonlinear correction for the atmospheric effect. The results are summarized in Table 1. The rmsd for the regionally optimized algorithm RMCSST is 1.26°C, similar to that of the RQSST (1.25°C) but lower than the global algorithm GMCSST (1.36°C). This is as expected, as a regionally optimized algorithm should represent the regional characteristics better. The noise amplification factor (NAF; Barton, 1995; Pearce et al., 1989) is also smaller in the regional algorithms, which means a smaller random error in the derived SST.

Analysis of bias characteristics further verifies the impact of the regionally optimized algorithm in more detail. Figure 8b shows that the regionally optimized algorithms have basically eliminated the dependence of the bias on the brightness temperature difference of the split window channels (DT). The dependence of the bias on the satellite zenith angle and on the buoy SST has also been significantly reduced (Figs. 9c and 9d), particularly for smaller viewing angles ($< 40°$) and warmer SST (> 15°C). Compared to the bias, Fig. 9 shows that, except for the case of a large viewing angle, the rmsd is generally less dependent on season, DT, SZA, and SST. And similar to the case of bias, the regionally optimized algorithms reduce such dependence. However, the dependence of the bias on season (Fig. 8a), while reduced, largely remains even when the algorithm has been regionally optimized.

There are several possible reasons for the different bias characteristics for August to October compared to the rest of the year. Without rigorous analysis, we can still reduce this difference empirically by deriving regression coefficients that are optimal for a specific period of time, i.e., one set for August to October and one set for the other months. To smooth the transition period, we weighted the two derived SST values in terms of time for the first and last weeks of the beginning and ending of each month. The coefficients and error statistics for these algorithms, called RTMCSST and RTQSST, are also shown in Table 1. The rmsds of the RTMCSST and RTQSST are 1.19°C and 1.16°C, respectively, which is an improvement of about 0.1°C over the corresponding algorithms for all seasons. The first of the regression coefficients for the second time domain of the RTQSST is much smaller than unity, which is compensated by the square term. Thus the expected NAF of the RTQSST algorithm is expected to be larger than that of the RTMCSST algorithm.

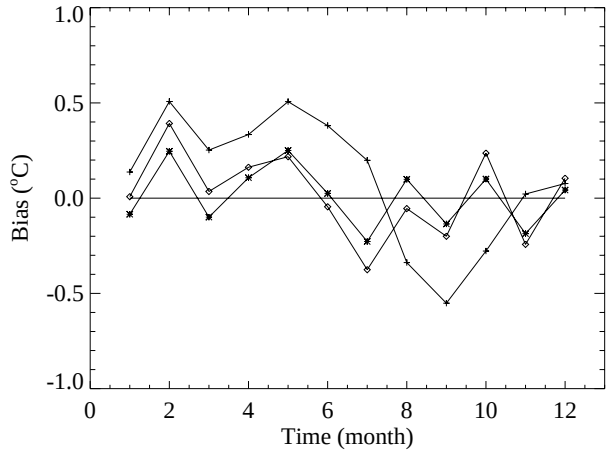

Fig. 10. Biases of RTMCSST (*), RTQSST (◇), and RMCSST (+) as functions of time (month).

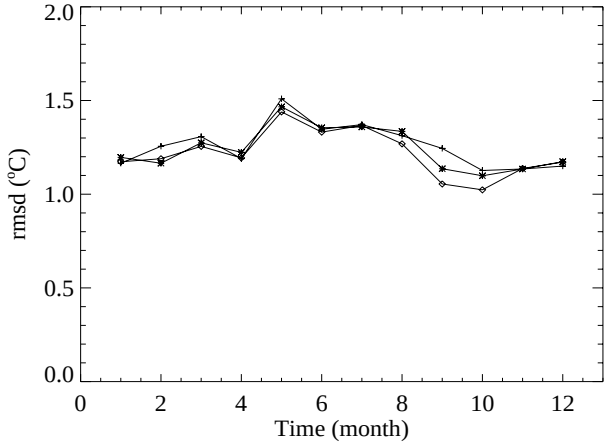

Fig. 11. The rmsd values of RTMCSST (□), RTQSST (Δ), and RMCSST(+) as functions of time (month).

Close examination of bias and rmsd as functions of time further confirms the positive impact of the regionally and temporally optimized algorithms. The seasonal dependence of bias, which is clear in RMCSST, is mostly absent in RTMCSST and RTQSST (Fig. 10). In addition to the overall improvement, both algorithms seem to perform much better for the August to October period, and the RTMCSST algorithm seems slightly better than the RTQSST algorithm. A similar conclusion can be made for the rmsd (Fig. 11), except that the RTQSST algorithm seems slightly better than the RTMCSST algorithm. Though not shown, the independence of bias and rmsd for the RTMCSST and RTQSST algorithms on DT, SZA, and SST is similar or better compared to the RMCSST and RQSST algorithms. In summary, the RTMCSST and RTQSST are comparable, and both perform better than the ones without temporal or regional optimization.

4. Validation

During the derivation of regression coefficients, a question is raised, viz. whether a sufficient number of collocated data points has been used to derive the regression coefficients when the original dataset is divided, especially for the August to October period. To answer that question, we performed the regression repeatedly with randomly selected collocated data and plotted the first regression coefficient and rmsd from these regression analyses as a function of sample size in Fig. 12. As expected, the first regression coefficient and rmsd fluctuate with sample size when the sample size is small but become stable as the sample size increases. The first regression coefficient varies by less than 0.01 when the sample size is larger than 2500, and the rmsd stabilizes when the sample size reaches about 4500. Since the sample size in all our regression is larger than 5000, we conclude that the derived regression coefficients are valid.

For the validation of the derived various SST retrieval algorithms, a total of 17431 collocated data

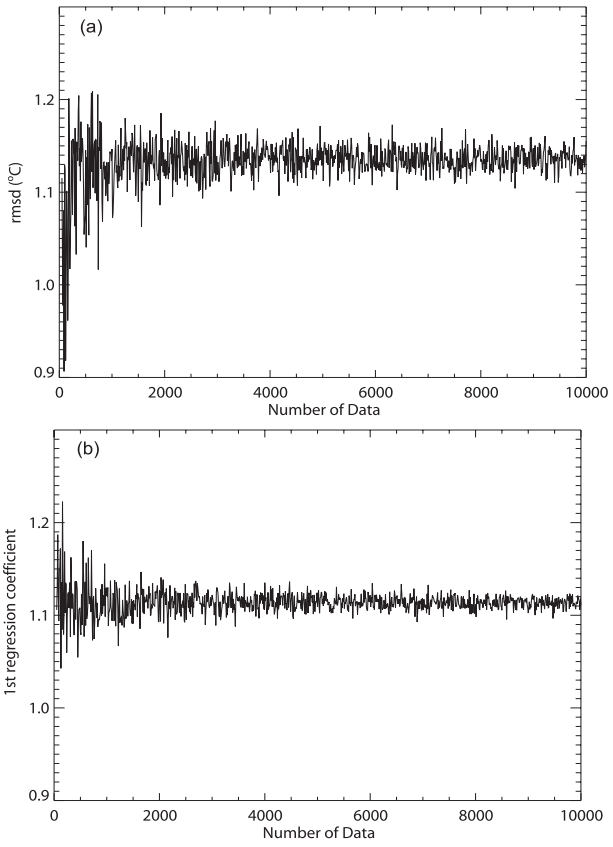

Fig. 12. Variation of rmsd and the first regression coefficient of MCSST as functions of the number of data points used for the coefficient derivation.

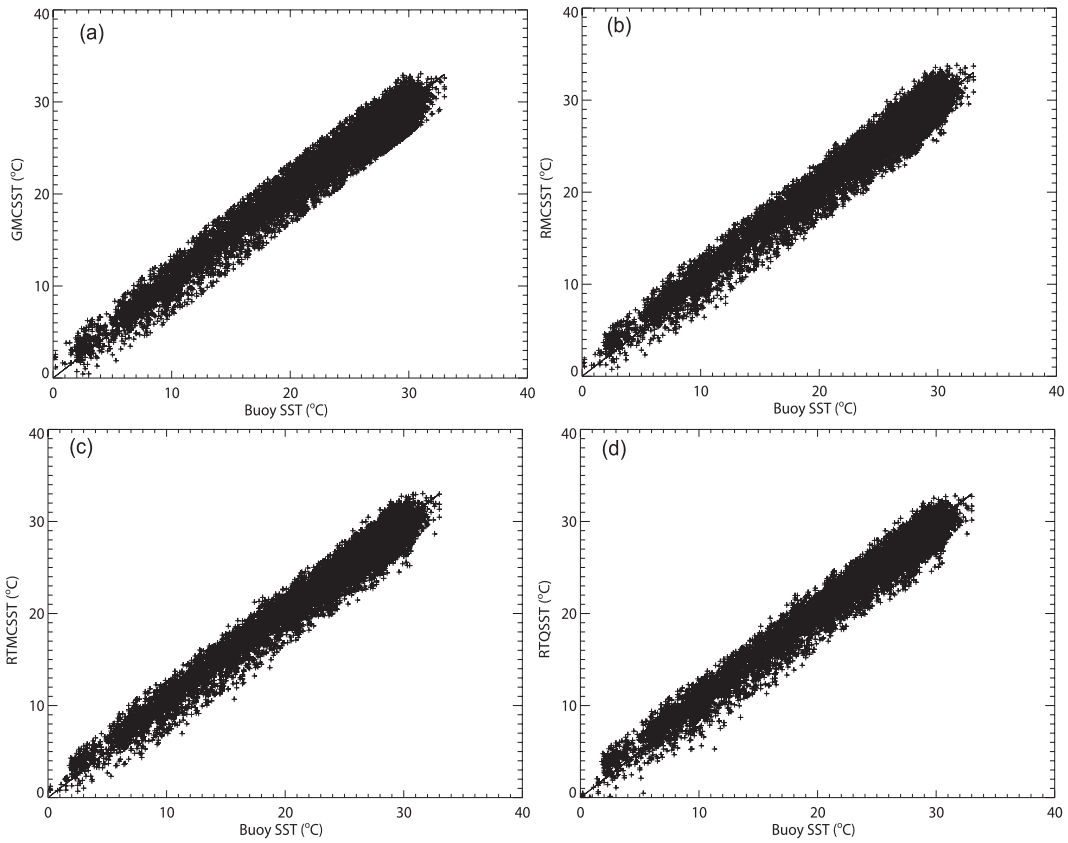

Fig. 13. Scatter diagram of the derived SST using the (a) GMCSST, (b) RMCSST, (c) RTMCSST, and (d) RTQSST algorithms for the time period 2000–2001. The rmsd values are 1.29, 1.22, 1.16, and 1.14°C, respectively.

points collected from January 2000 to October 2001 are used. Figure 13 shows the scatter plots of the buoy SST and GMS-5 SST using the GMCSST, RMCSST, RTMCSST, and RTQSST algorithms. Respectively, the biases for these algorithms are 0.18°C, 0.03°C, 0.05°C and 0.05°C, and the rmsd values are 1.29°C, 1.22°C, 1.16°C, and 1.14°C. These results are consistent with the error analysis during coefficient derivation that shows progressive improvement due to regional and temporal optimization.

The monthly mean bias and rmsd values of GMCSST, RMCSST, and RTMCSST as functions of the four different parameters are shown in Fig. 14 and Fig. 15, respectively. The characteristics for bias and rmsd are similar to the cases of the coefficient derivation. The time dependence of bias from RTMCSST is significantly reduced compared to other algorithms. The bias varies within about ±0.6°C for RMCSST and GMCSST, while it varies only within ±0.3°C for the RTMCSST. All algorithms show the similar dependency on DT, buoy SST and SZA to that of the coefficient derivation. In the case of the DT values, the two regional algorithms show a range of ±0.3°C, while the global algorithm shows values larger than 0.5°C for a higher DT value of about 1.5°C. For the different SZA and buoy SST, the regional algorithms also reduce the variability, especially for the lower values of SZA and buoy SST.

The rmsd of the RTMCSST and RMCSST algorithm is also smaller than that of GMCSST, while the RTMCSST shows the smallest rmsd of almost all cases, especially for the months of August and September and for different DT values. During August and September, rmsd shows an improvement of about 0.4°C in RTMCSST compared to the GMCSST algorithm. Also, there is significant improvement, by about 0.2°C to 0.3°C, in rmsd for the higher DT values for both RTMCSST and RMCSST.

5. Summary and conclusions

Algorithms exist to derive sea surface temperature from GMS-5 satellite measurements (Shirakawa, 1996; Yasuda and Shirakawa, 1999). When applied to the domain of the KMA regional weather prediction model in East Asia (15°–55°N, 105°–170°E), however, the retrieval algorithm was found to have a negative bias that also seems to be dependent on season, atmosph-

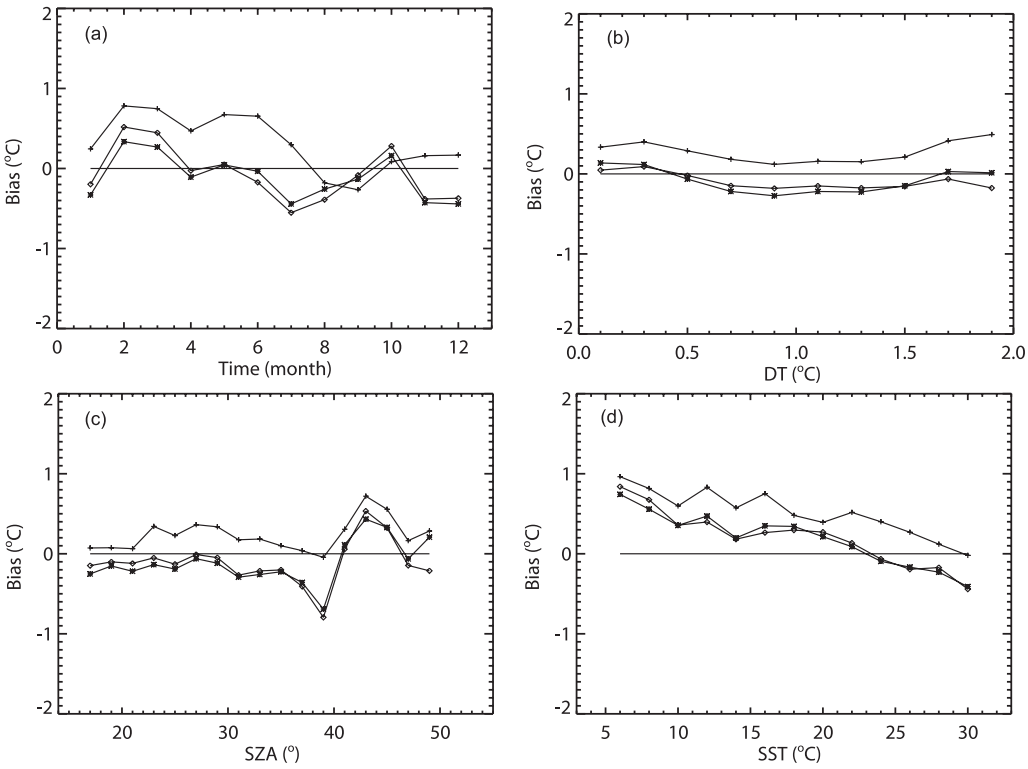

Fig. 14. Biases of GMCSST (+), RMCSST (*), and RTMCSST (◇) for the time period 2000–2001 as functions of (a) time, (b) DT, (c) SZA, and (d) SST.

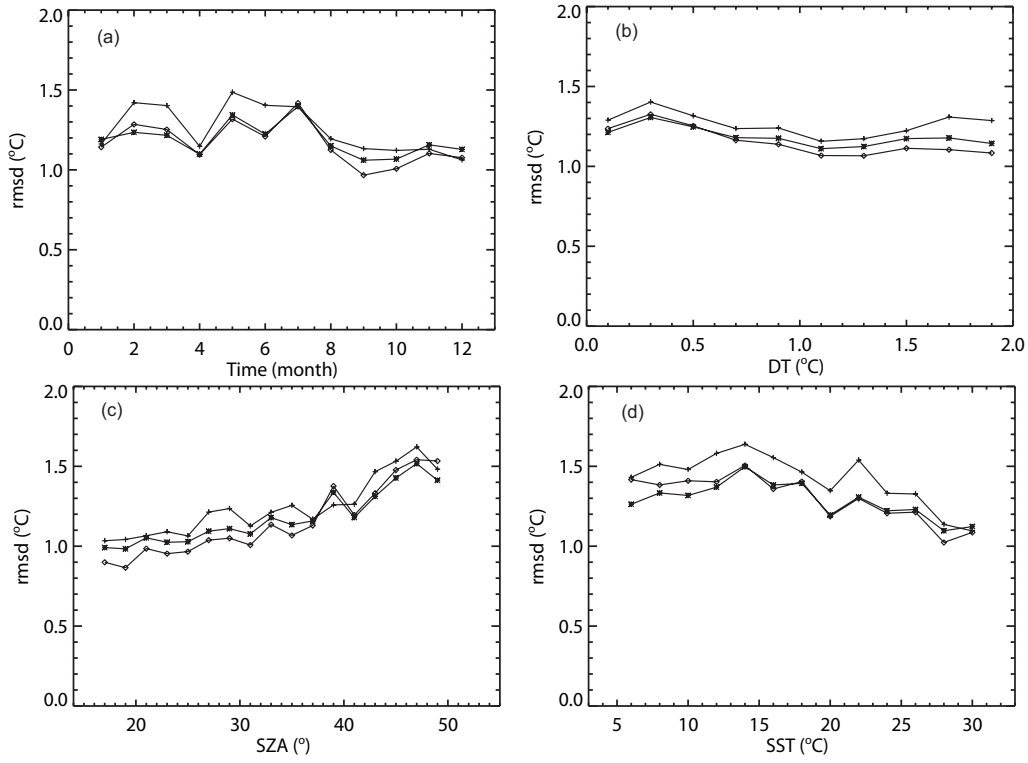

Fig. 15. Same as Fig. 14 but for rmsd.

eric moisture, viewing angle, and SST. It was believed that a regionally optimized algorithm would alleviate most of these problems.

To carry out this investigation, ocean buoy data are collocated with GMS-5 measurements from 1997 to 2001. A quality control procedure is developed for the buoy data to limit the daily and 5-day variations of the reported SST. This procedure removes about 19% of the erroneous buoy reports. A stringent cloud screening procedure is developed for the satellite data that includes the use of a global SST value. About 20% of the original collocated data survive all the quality control tests, and the percentage of reliable data is about equal between daytime and nighttime. This suggests that the nighttime procedures are as effective as the daytime procedures.

The data from 1997 to 1999 are used for the derivation of the regression coefficients while the remaining data are used for validation. All of the regionally optimized algorithms eliminate the overall bias and reduce the rmsd by 0.1°C. More remarkably, the regionally optimized algorithms change the characteristics of the residual error. For the global algorithm, for example, the bias tends to be positive when the atmosphere is humid, when the target is close to the nadir, when the sea is warm, or from August to October. All these dependencies are removed or significantly reduced by the regionally optimized algorithms, except that the seasonal dependence largely remains.

The seasonal dependence is mitigated by further optimization of the algorithm based on season, deriving one set of regression coefficients for August to October and one for the other months. The final algorithm, called RTMCSST for Regionally and Temporally optimized Multi-Channel SST algorithm, further reduces the rmsd by 0.1°C and basically removes the seasonal dependence of bias. A sensitivity study confirms that the sample size is sufficient and the derived regression coefficients are valid.

The current study is expected to be applicable for upcoming new satellites in the East Asia region such as Japan's MTSAT-1R and China's FY-2 series which measures radiances at the split window channels. For a better operational SST derivation, the combination of infrared and microwave data is highly required, especially when dealing with the East Asia region where monsoon fronts linger for usually more than a week at a certain latitude band. Although the spatial resolution of microwave radiometers such as TMI (TRMM Microwave Imager) and AMSR-E (Advanced Microwave Scanning Radiometer) is limited, these can provide SST for most of the cloudy area associated with the monsoon front. Thus the combination of infrared and microwave data is highly recommended.

Acknowledgments. The authors thank the Division of Satellite Meteorology of the Korea Meteorological Administration (KMA) for their provision of buoy and satellite data. This study was supported by the Basic Research Project of the Meteorological Research Institute (METRI) and the "Development of Meteorological Data Processing System for COMS" of KMA.

REFERENCES

Ahn, M. H., B. J. Hwang, E. H. Son, M. J. Kim, and A. S. Suh, 2001: Derivation of sea surface temperature from GMS-5 for the high resolution numerical weather prediction models. *Korean Journal of Atmospheric Sciences*, **4**, 41–56.

Ahn, M. H., E. H. Son, and B. J. Hwang, 2003a: A new algorithm for sea fog/stratus detection using GMS-5 IR data. *Adv. Atmos. Sci.*, **20**, 899–913.

Ahn, M. H., J. M. Koo, C. Y. Chung, and J. C. Nam, 2003b: Effects of the tropospheric dust on the sea surface temperature derivation from the GMS-5 IR data. *Journal of Korean Meteorological Society*, **39**, 653–666.

Barton, I. J., 1995: Satellite-derived sea surface temperature current status. *J. Geophys. Res.*, **106**, 8777–8790.

Bates, J. J., and W. L. Smith, 1985: Sea surface temperature: Observations from geostationary satellites. *J. Geophys. Res.*, **90**, 11609–11618.

Emery, W. J., S. Castro, G. A. Wick, P. Schluessel, and C. Donlon, 2001: Estimating sea surface temperature from infrared satellite and in situ temperature data. *Bull. Amer. Meteor. Soc.*, **82**, 2773–2786.

Gill. A. E., 1982: *Atmosphere-Ocean Dynamics*. Academic Press. 643pp.

Hansen, D. V., and P. M. Poulain, 1996: Quality control and interpolations of WOCE/TOGA drifter data. *J. Atmos. Oceanic Technol.*, **13**, 900–910.

Kilpatrick, K. A, G. P. Podesta, and R. Evans, 2001: Overview of the NOAA/NASA Advanced Very High Resolution Radiometer Pathfinder algorithm for sea surface temperature and associated matchup database. *J. Geophys. Res.*, **106**, 9179–9197.

Legeckis, R., and T. Zhu, 1997: Sea surface temperatures from the GOES-8 geostationary satellite. *Bull. Amer. Meteor. Soc.,* **78**, 1971–1983.

McClain, E. P, W. G. Pichel, and C. C. Walton, 1985: Comparative performance of AVHRR-based multichannel sea surface temperature. *J. Geophys. Res.*, **90**, 11587–11601.

Minnett, P. J., 1990: The regional optimization of infrared measurements of sea surface temperature from space. *J. Geophys. Res.*, **95**, 13497–13510.

MSC, 1997: The GMS User's Guide. Meteorological Satellite Center, Japan Meteorological Agency, 190pp.

Park, K. A., 1996: Spatial and temporal variability of sea surface temperature and sea level anomaly in the East Sea using satellite data (NOAA/AVHRR, TOPEX), Ph. D dissertation, Seoul National University, 294pp.

Pearce, A. F., A. J. Prata, and C. R. Manning, 1989: Comparison of NOAA/AVHRR-2 sea surface temperatures with surface measurements in coastal waters. *Int. J. Remote Sens.*, **10**, 37–52.

Rodon, G. I., 1975: On the North Pacific temperature, salinity, sound velocity and density fronts and their relation to the wind and energy flux field. *J. Phys. Oceanogr.*, **4**, 168–182.

Saunders, P. M., 1967: Aerial measurement of sea surface temperature in the infrared. *J. Geophys. Res.*, **72**, 4109–4117.

Shirakawa, Y., 1996: Sea surface temperature. Meteordagical Satettite Center Technical Note, Special Issue, 95–101.

Shenoi, S. C., 1999: On the suitability of global algorithms for the retrieval of SST from the north Indian Ocean using NOAA/AVHRR data. *Int. J. Remote Sens.,* **20**, 11–29.

Stramma, L., P. Cornillon, R. A. Weller, J. F. Price, and M. G. Briscoe, 1986: Large diurnal sea surface temperature variability: Satellite and in situ measurements. *J. Phys. Oceanogr.*, **16**(5), 827–837.

Walton, C. C., 1988: Nonlinear multichannel algorithms for estimating sea surface temperature with AVHRR satellite data. *J. Appl. Meteor.*, **27**, 115–124.

Walton, C. C., W. G. Pichel, and J. F. Sapper, 1998: The development and operational application of nonlinear algorithms for the measurement of sea surface temperatures with the NOAA polar-orbiting environmental satellites. *J. Geophys. Res.*, **103**(C12), 27999–28102.

Webster, P. J., C. A. Clayson, and J. A. Curry, 1996: Cloud, radiation, and the diurnal cycle of sea surface temperature in the tropical western Pacific. *J. Climate*, **9**, 1712-1730.

Wu, X., P. Menzel, and G. S. Wade, 1999: Estimation of sea surface temperature using GOES-8/9 radiance measurements. *Bull. Amer. Meteor. Soc.*, **80**, 1127–1138.

Yasuda, H., and Y. Shirakawa, 1999: Improvement of the derivation method of sea surface temperature from GMS-5 data. *Meteorological Satellite Center Technical Note*, **37**, 19–33.