Evaluation and Improvement of FY-4A/AGRI Sea Surface Temperature Data

Quanjun He[®], Xin Hu, and Yanwei Wu

Abstract—The advanced geosynchronous radiation imager (AGRI) aboard the Chinese Fengyun-4A (FY-4A) satellite can provide operational hourly sea surface temperature (SST) product. However, the temporal and spatial variation of the errors for this product is still unclear. In this article, FY-4A/AGRI SST is evaluated using the in situ SST from 2019-2021, and a cumulative distribution function matching method is adopted to reduce the errors. Statistical results show that the mean bias and root-mean-square error (RMSE) of FY-4A/AGRI SST are -0.37 °C and 0.98 °C, the median and robust standard deviation (RSD) are -0.30 °C and 0.90 °C. The variations in daily and monthly errors are large and there are no prominent seasonal variations during the period analyzed. There are negative biases exceeding -1.0 °C in low-mid latitude regions and larger positive biases in southern high latitude region. There are dependencies of satellite SST minus in situ SST on satellite zenith angle and on SST itself. After the bias correction, the bias and RMSE are reduced to -0.02 °C and 0.72 °C, and the median and RSD are reduced to 0.00 °C and 0.60 °C. On the time scale, the fluctuation ranges of bias and median are smaller. The difference of satellite SST minus in situ SST can reflect the diurnal variation of SST. The biases are generally within $\pm 0.2~^{\circ}\mathrm{C}$ in full disk. The error dependencies on satellite zenith angle and SST are also greatly reduced.

Index Terms—Advanced geosynchronous radiation imager (AGRI), bias correction, error evaluation, fengyun (FY)-4A, sea surface temperature (SST).

I. INTRODUCTION

SEA surface temperature (SST) is an essential variable for ocean and atmospheric prediction systems and climate change studies. After nearly half a century of development, satellite remote sensing has been the most important method to derive SST [1], [2]. The major advantage of satellite remote sensing SST is that it can obtain large coverage of ocean observation data in near real time. With the development of refined weather forecasting and regional ocean and climate change research, people have higher and higher requirements for the accuracy of SST. Climate applications require SST data with an accuracy

Manuscript received 14 July 2022; revised 25 October 2022 and 24 November 2022; accepted 25 November 2022. Date of publication 30 November 2022; date of current version 7 December 2022. This work was supported in part by the science and technology research project of Guangdong meteorological service under Grant GRMC2020M04, and in part by the Pan Pearl River Delta science and technology innovation open fund under Grant FZSJ202112. (Corresponding author: Quanjun He.)

The authors are with the Guangzhou Meteorological Satellite Ground Station, Guangzhou 510640, China, and also with the Guangdong Meteorological Satellite Remote Sensing Center, Guangzhou 510640, China (e-mail: hequanjunsx@163.com; huxin95@outlook.com; wuyw1132@163.com).

Digital Object Identifier 10.1109/JSTARS.2022.3225729

of 0.10 °C and a stability of 0.04 °C per decade [3]. Although no satellite retrieval product can achieve this goal at present, it is also a development goal of remote sensing retrieval SST. The global ocean data assimilation experiment believes that the accuracy of SST products is required to reach more than 0.40 °C for accurate ocean models [4]. A high-quality SST data for understanding and quantifying the variation of SST and its global impacts is therefore crucial and in high demand.

Estimates of the errors are imperative for satellite-derived SST assimilation into climate models. The precision has been validated when developing the SST retrieval algorithm. However, the quality of satellite-derived SST product is affected by several factors, such as anomalous atmospheric conditions, instrument calibration problems, cloud detection failures, in situ observation errors, sea surface emissivity, which may produce lots of biases and uncertainties in SST [5], [6], [7], [8]. In order to further understand and reveal the variation characteristics of the error of satellite-derived SST products, researchers have carried out a lot of research work and used a variety of evaluation indicators and methods to evaluate the error of SST products [9], [10], [11], [12].

After understanding the error characteristics of existing SST products, it is necessary to carry out quality control and deviation correction to further improve the quality and application ability of satellite SST products. Reynolds et al. [13], [14] successively demonstrated that a bias correction was necessary for satellitebased SST to remove errors associated with volcanic or other aerosols. Researchers attempt to correct the bias of the retrieved SST data using simultaneous matched in situ data. A web-based SST quality monitor [15] is employed by the national environmental satellite, data, and information service to continuously control the quality of operational SST products in near-real time. The in situ SST quality monitor (iQuam) [16] has been developed with the primary goal to support satellite calibration and validation at the national oceanic and atmospheric administration (NOAA). For the purpose of reconstructing SST analysis products, Reynolds and Smith [17] corrected the bias by solving Poisson's equation before optimum interpolation. Later, a new bias correction method was designed using empirical orthogonal teleconnection functions by Reynolds et al. [18]. Høyer et al. [19] developed a multi-sensor bias correction method by combining multisensor data within five days to generate reference data, and then subtracting the reference data from the given day's satellite data. A piecewise regression SST was designed by Petrenko et al. [20] to estimate the sensor-specific error statistic for SST in the advanced clear-sky processor for oceans. To reduce the bias in the coastal region, Kwon et al. [21] designed two spatial-temporal bias correction functions. The first bias correction consisted of an exponential function depending on distance from the land and a cosine function depending on time, the second function was defined as the difference between the two daily climatology datasets on day. Han et al. [22] developed two offline bias correction methods for SST forecasts based on the neural network and empirical orthogonal function. Some researchers tried to use the probability density function (PDF) [23] to adjust the satellite SST.

Fengyun-4 (FY-4) [24] is the second generation of Chinese geostationary meteorological satellite series, and FY-4A is the first scientific research and experimental satellite of this series launched on December 17, 2016. The satellite is located at 104.7 °E, 36500 km above the equator. The advanced geosynchronous radiation imager (AGRI) [24] aboard the FY-4A has 14 spectral bands that are quantized with 12 bits per pixel and sampled at 1 km at nadir in the visible, 2 km in the near-infrared, and 4 km in the remaining infrared spectral bands. The AGRI has the split windows channels and can provide the important observation data for hourly SST. SST with the high temporal resolution is very important for the air-sea interaction studies and coastal ocean modeling. The diurnal variation (DV) is one of the dominant variations in SST due to the solar radiation and the Earth's rotation. Nowadays, geostationary satellites are the only practical way to obtain SST data with sufficient frequency across the extensive oceans to resolve DV. The operational SST product of FY-4A/AGRI are developed using nonlinear SST (NLSST) algorithm. The coefficients of NLSST algorithm are obtained by regression between satellite data and in situ data. Different retrieval coefficients are used for daytime and nighttime data, but the same set of coefficients is used for full disk data. The product developer periodically evaluates the errors of SST product. When large errors are found, the developer will use the in situ data to conduct regression calculation again to obtain a new set of retrieval coefficients. Since it takes a period of accumulation to obtain the matchups between the satellite data and in situ data for regression, there will be a certain delay in updating the coefficients of SST retrieval algorithm.

Although the FY-4A/AGRI SST product has been provided operationally by the national satellite meteorological center (NSMC) of the China meteorological administration (CMA), the accuracy of this product is not yet known. No detailed studies have been performed to evaluate the error and correct the bias for FY-4A/AGRI SST. The purpose of this article is to evaluate the quality of the operational FY-4A/AGRI SST product by comparing it with the in situ observations and reanalysis data, and to perform the bias correction for this product to obtain more accurate SST data. This article is based on the operational FY-4A/AGRI hourly SST data, in situ SST data and reanalysis SST data from the period January 2019–December 2021.

II. MATERIALS AND METHODS

A. Datasets

FY-4A/AGRI SST (hereinafter referred to as SATSST) data is the hourly level 2 (L2) normalized geostationary projection

data developed by the NCMC of CMA, and is stored in network common data form (NetCDF), the resolution is 4 km and the full disk size is 2478×2478 pixels. The L2 data contains the SST dataset and data quality flag dataset, and can be downloaded from the service website of Fengyun satellite data center (http://data.nsmc.org.cn/portalsite/default.aspx).

The in situ SST used to validate the SATSST is from the iQuam dataset [16]. The iQuam dataset was developed at NOAA center for satellite applications and research (STAR) and the current version is 2.1. One month's in situ data is stored as a single NetCDF file and can be available online (https://www.star.nesdis.noaa.gov/sod/sst/iquam/data.html). The iQuam dataset includes a variety of buoy and ship observation data. These observation data are processed by quality control and marked with quality level, and the quality level with the highest accuracy is 5. The in situ SST is the temperature at a certain depth, but the satellite-derived SST from FY-4A/AGRI is the skin temperature.

Another reference data uses the 1/4° daily optimum interpolation SST (OISST) data [18], [25]. Daily OISST is a blend of in situ ship and buoy SSTs with satellite SSTs derived from the AVHRR on a regular global grid. OISST are available at national centers for environmental information of NOAA (https://www.ncei.noaa.gov/data/sea-surface-temperature-optimum-interpolation/v2.1/).

B. Matchup Samples Extraction

The matchup data are collected by combining the SATSST data with the corresponding in situ SSTs and OISST on the grid resolution of 0.05°. First, the full disk of hourly satellite data is converted into equal longitude and latitude projection with 0.05° spatial resolution from normalized geostationary projection. Second, the best quality data with quality level equals 5 and within 30 minutes before and after the satellite observation time are chosen from the iQuam dataset to be resampled to the hourly grid resolution. If there are multiple observation data in the same grid point, the average value shall be taken. And, the daily OISST is also resampled to the grid resolution of 0.05°. Finally, sample pairs are extracted from satellite data and referenced data at the same time and at the same location.

To minimize errors from the outliers, different methods are used to remove the outliers from matchups by researchers [26], [27], [28]. In this article, the outliers are removed using three times the robust standard deviation (RSD) from the median of the satellite SST minus in situ SST [15]. Here, RSD is defined as: (75th percentile–25th percentile)/1.348. And, only the SATSST data with a quality index of zero (the best quality) are reserved. The final matchups are 1 155 536 for three years. Fig. 1 is the distribution and density of total matchups from 2019 to 2021. It can be seen that the in situ data is less distributed in tropical oceans near the equator and high latitude regions. In particular, the in situ data are extremely sparse in the high latitude ocean region of the south part of the full disk. The in situ data are relatively densely distributed in the mid-latitude ocean regions of the Pacific Ocean and the Indian Ocean.

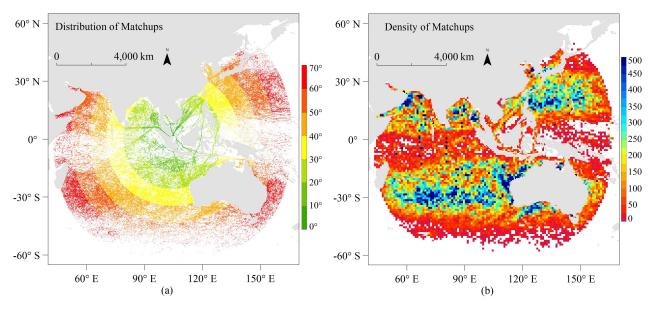


Fig. 1. Distributions (a) and density (b) of matchups between satellite SST and in situ SST from 2019 to 2021. In (a), the color represents the distribution of satellite zenith angle from 0° to 70° . In (b), the color represents the number of matchups in $1^{\circ} \times 1^{\circ}$ latitude and longitude boxes.

C. Error Evaluation

The conventional statistic of mean bias and root-meansquare error (RMSE) are employed to evaluate the difference between SATSST and in situ SST. The bias and RMSE can be calculated using the following formulas:

Bias =
$$\frac{1}{N} \sum_{i=1}^{N} (X_i - Y_i)$$
 (1)

RMSE =
$$\sqrt{\frac{1}{N} \sum_{i=1}^{N} (X_i - Y_i)^2}$$
 (2)

where i is the number of collocated data points, N is the total number of collocated data points, X is the satellite-derived SST, and Y is the in situ SST.

However, the conventional statistics do not characterize fairly the center and spread of the distribution of differences between the retrieved SST and in situ SST because they are strongly influenced by a small percentage of outliers which do not fit the approximately Gaussian distribution of the majority of the data [9], [15]. So, the robust statistics of median and RSD [9], [15] of SATSST minus in situ SST are employed to circumvent this problem. Also, the correlation coefficient (*R*) is used to evaluate the linear correlation between satellite-derived SST and in situ SST. *R* can be calculated using the following formula:

$$R = \frac{\sum_{i=1}^{N} (X_i - \bar{X}) (Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{N} (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^{N} (Y_i - \bar{Y})^2}}$$
(3)

where \bar{X} is the mean value of satellite-derived SST, and \bar{Y} is the mean value of in situ SST.

Taylor diagram [29] describes the *R*, standard deviation (SD) and centered pattern root mean square difference in a single diagram, and can easily identify error differences between satellite-derived SST and in situ SST. The centered pattern root

mean square difference [29], also named as unbiased RMSE (ubRMSE) [10], [30], has been widely used to evaluate the error between observations and models. The RMSE and ubRMSE differ in that the ubRMSE excludes errors from biases and simply considers the difference in amplitude between variations. The SD and ubRMSE can be calculated using the following formulas:

$$SD = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(X_i - \bar{X} \right)^2} \tag{4}$$

ubRMSE =
$$\sqrt{\frac{1}{N} \sum_{i=1}^{N} [(X_i - \bar{X}) - (Y_i - \bar{Y})]^2}$$
. (5)

Finally, the three-way error analysis method [11] is used to further estimate the error between the SATSST, in situ SST and OISST. The error variances σ_i^2 for SST type i (where i=1,2 or 3) can be given by

$$\begin{cases}
\sigma_1^2 = (V_{12} + V_{31} + V_{23})/2 \\
\sigma_2^2 = (V_{23} + V_{12} + V_{31})/2 \\
\sigma_2^2 = (V_{21} + V_{22} + V_{12})/2
\end{cases}$$
(6)

$$V_{ij} = \frac{1}{N} \sum_{i=1}^{N} \left(SST_{ij} - \overline{SST_{ij}} \right)^2 \tag{7}$$

$$SST_{ij} = SST_i - SST_j \tag{8}$$

where SST_{ij} is the difference between SST types i and j, $\overline{SST_{ij}}$ is the mean value of SST_{ij} , V_{ij} is the variance of SST_{ij} , and subscripts 1, 2, and 3 refer to the SATSST, in situ SST and OISST, respectively.

D. Bias Correction

To reduce the errors in the operational FY-4A/AGRI hourly SST, the cumulative distribution function (CDF) matching method is used to adjust satellite-derived SST data against the in situ SST data in this article. The CDF method was first proposed

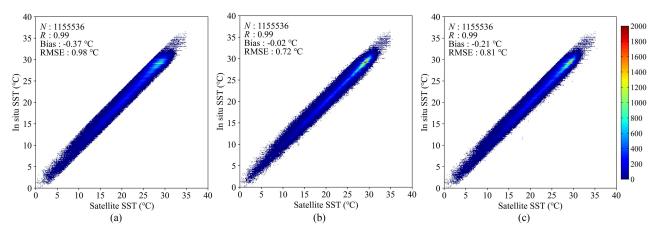


Fig. 2. Density scatter plots between satellite SST and in situ SST. (a) SATSST and in situ SST. (b) SATSST $_{\mathrm{CDF}}$ and in situ SST. (c) SATSST $_{\mathrm{LSR}}$ and in situ SST.

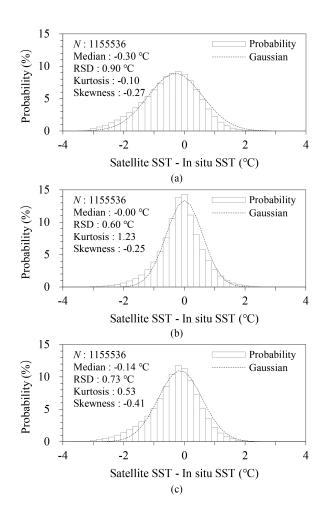


Fig. 3. PDF of satellite SST minus in situ SST. (a) SATSST minus in situ SST. (b) SATSST $_{\rm CDF}$ minus in situ SST. (c) SATSST $_{\rm LSR}$ minus in situ SST. N is the number of matchups used in PDF.

by Reichle and Koster [31] to reduce the bias of satellite soil moisture. It has been widely used to remove systematic biases between observation data and reference data [32], [33], [34]. The CDF of the satellite-derived SST should be the same as that of

the in situ SST, so the bias of the satellite observations could be defined as the difference between the satellite-retrieved SST and the in situ SST at the same percentiles. The basic principle of CDF matching technique is to find a transfer function that allow matching the CDFs of in situ SST data and satellite-derived SST data. The linear transfer function can be simply descripted as

$$X_{c} = a + b \times X_{m} \tag{9}$$

where X_c is the corrected SST, X_m is the uncorrected SST, and a and b are the linear fitting coefficients found when comparing the uncorrected SST with the observed SST using the least square regression (LSR) method. In this application, the uncorrected SST is the operational FY-4A/AGRI hourly SST, the observed SST is the in situ SST.

The CDF curve is divided into 12 segments with the values of 0, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 95, and 100 percentiles [32], and then the linear regression is performed for each segment to rescale data falling into different segments. Finally, the piecewise linear CDF matching approach is applied for each grid cell in the full disk of hourly SATSST. Here, the size of grid cell is $1^{\circ} \times 1^{\circ}$ in this article for reducing the computational time. In order to get the enough matchups for each grid cell, the local matchup dataset is selected from the total matchup dataset during a period prior to the target day and the initial local window length will be increased by 1° for each time if the number of local matchups is less than 300. The initial local window length is 1°. The pixels within the target grid cell and its nearby regions are used so that at least 300 pairs of data are included to define the CDF functions, and a set of bias correction coefficients are calculated with the piecewise linear CDF matching method. Then, the correction coefficients of all grid cell within the whole coverage of SATSST data can be achieved through cyclic calculation and stored in the coefficient lookup table. Finally, the bias correction is performed for the pixels within $1^{\circ} \times 1^{\circ}$ grid cell using a same set of bias correction coefficients.

In the previous study, different periods were selected to generate reference data or train data for the similar method [23], [34], [35]. Here, simple tests are performed to decide the best

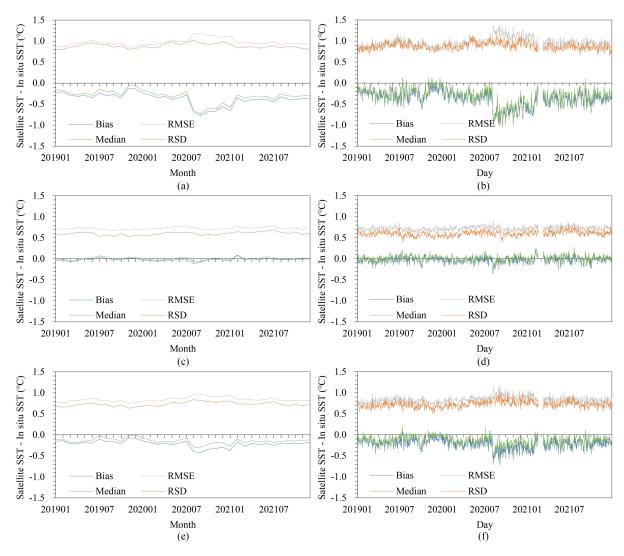


Fig. 4. Monthly (left) and daily (right) errors between satellite SST and in situ SST. (a) and (b) SATSST minus in situ SST. (c) and (d) SATSST $_{CDF}$ minus in situ SST. (e) and (f) SATSST $_{LSR}$ minus in situ SST.

TABLE I ERROR STATISTICS BETWEEN SATELLITE SST AND IN SITU SST BEFORE AND AFTER BIAS CORRECTION

	Bias (°C)	RMSE (°C)	Median (°C)	RSD (°C)	R
SATSST	-0.37	0.98	-0.30	0.90	0.99
$SATSST_{CDF}$	-0.02	0.72	0.00	0.60	0.99
$SATSST_{LSR}$	-0.21	0.84	-0.14	0.73	0.99

period for calculating regression coefficients using the data in January and July in 2019, respectively. A specific time window is selected and the observation and reference data from 510,15 until 60-day periods are collocated. Then, the collocated data are corrected with CDF method and the errors are calculated. As the selected period is increased, the bias of satellite SST minus in situ SST appears to be increased. However, the SD, RMSE and RSD show a change from decline to rise, with inflection point appearing around 15-day. Also, based on the previous study, the 15-day period is selected in this article.

III. RESULTS

A. General Error Statistics

The evaluations are conducted between satellite-derived SST and in situ SST using the total matchup data in 2019–2021. In order to evaluate the performance of the CDF method, a simple LSR method is used for comparison. The coefficients of LSR method are calculated using the same training samples as CDF method. The bias-corrected SATSST with CDF method is named as SATSST_{CDF}, the bias-corrected SATSST with LSR method is named as SATSST_{LSR}. The error statistics are given in Table I. The statistics show that SATSST has a stronger negative bias of -0.37 °C. The RMSE is 0.98 °C, indicating that the accuracy of FY-4A/AGRI operational SST product is low. Even the robust statistics of median and RSD have reduced the impact of outliers, they still show that the SATSST has large negative median and higher RSD compared with in situ SST. After bias correction with CDF method, the errors decreased significantly. The bias and RMSE of SATSST_{CDF} are reduced to -0.02 °C and 0.72 °C, and the median and RSD are reduced to 0.00 °C and 0.60 °C,

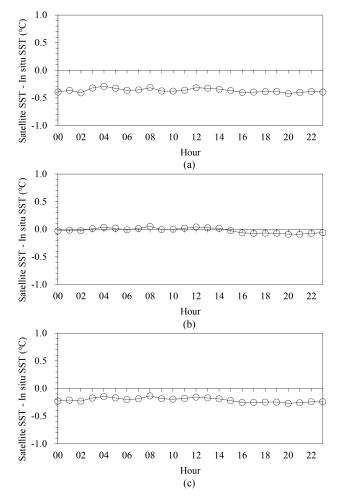


Fig. 5. Hourly difference between satellite SST and in situ SST. (a) SATSST minus in situ SST. (b) SATSST $_{\rm CDF}$ minus in situ SST. (c) SATSST $_{\rm LSR}$ minus in situ SST.

respectively. The errors of SATSST $_{\rm LSR}$ are also reduced to a certain extent, but these errors are still higher than those of SATSST $_{\rm CDF}$. The correlation coefficients before and after bias correction have no obvious change and their values reach 0.99, indicating that there has always been a high correlation between the satellite-derived SST and in situ SST.

Fig. 2 is the density scatter plots between satellite-derived SST and in situ SST. Before and after bias correction, there is little difference in the shapes of scatter distribution between the satellite SST and in situ SST. However, after bias correction, the scatter distribution density of satellite SST and in situ SST is more concentrated, especially at the high temperature section.

Fig. 3 shows the PDFs of difference between satellite SST and in situ SST. The PDFs are shown using 0.2 °C bins. The dotted lines show the Gaussian distributions defined using the median and RSD of satellite SST minus in situ SST. The kurtosis and skewness of PDFs are also noted in the figures. Although the distribution of SATSST minus in situ SST conforms to the Gaussian distribution in shape, but it is easy to note that there is a considerable negative deviation between the SATSST and in situ SST. After bias correction with CDF method, the median of satellite SST minus in situ SST close to zero. The kurtosis

of PDF increases from -0.10 to 1.23, which is higher than the result of LSR method. According to the statistics of SATSST, SATSST_{CDF} and SATSST_{LSR}, the percentages of samples with an error between ± 1.0 °C in the total number of matchup samples are 73.56%, 87.33% and 81.40%, respectively. These indicate that the distributions of satellite SST minus in situ SST are more centered after bias correction, and the CDF method is superior to LSR method. The skewness of SATSST_{CDF} minus in situ SST is the smallest in three PDFs, indicating that its normal distribution is better.

B. Spatial and Temporal Error Statistics

The density scatter plots and PDFs of the total matchups show the overall accuracy of the satellite-derived SST, more details need to be investigated from spatial and temporal variations of the errors. Fig. 4 shows the monthly and daily mean error variation between satellite SST and in situ SST from January 2019 to December 2021. The details of temporal variation of errors are displayed intuitively. The performance of original satellite SST product is not stable over different periods. The nonnegligible negative biases always exist and vary greatly. The monthly errors and daily errors are consistent in time and intensity change trend, but the variation ranges of daily errors are greater than that of monthly error. The monthly biases vary between -0.13 °C and -0.77 °C, but the daily biases vary between 0.06 °C and -1.03 °C. There is no prominent seasonal variation during the period analyzed. It is worth noting that there are significant bias jumps in the time series. In particular, the bias and median increased abnormally in August 2020 and lasted for about six months. The curves of RMSE and RSD also show this change process of sudden increase of error. This phenomenon is related to the operational calibration update of FY-4A/AGRI in this period. According to [36], the operational calibration update was carried out in August 2020. The calibration update can change the reflectance and brightness temperature of satellite data, which can lead to anomalies in the remote sensing products based on the threshold and empirical formula methods. For example, when an error occurs in the cloud detection result, the retrieval SST may be affected by the cloud and reduce the accuracy. In addition, since the operational SST products of FY-4A/AGRI are developed using the NLSST algorithm, after the calibration update, the original retrieval coefficients are no longer suitable for the updated data, which will also cause the accuracy of the retrieval SST product to change. In addition, in order to obtain enough matchups for calculating new retrieval coefficients, a period of data accumulation is required. This is also the reason why the abnormal errors have lasted for nearly 6 months. About the reason why the errors of FY-4A/AGRI operational SST products suddenly increased during this period, the author had communicated with the SST algorithm developer, and the author's view was confirmed.

After bias correction with CDF method, the monthly and daily errors are generally much smaller than those before bias correction, and the variation ranges of errors are also smaller. The monthly bias and median of SATSST $_{\rm CDF}$ are close to zero. The bias fluctuates between 0.08 °C and -0.11 °C, and the

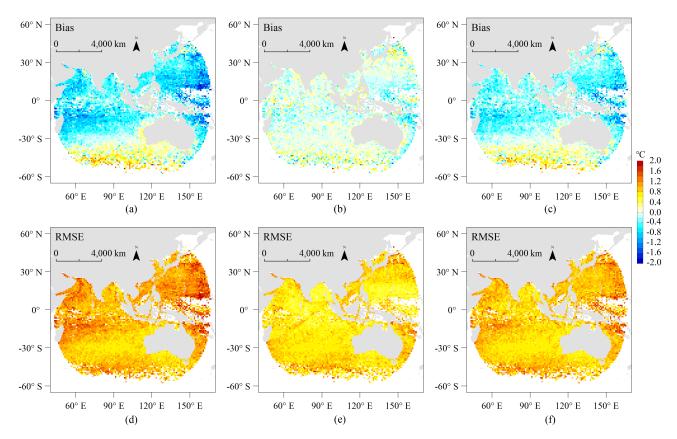


Fig. 6. Map of biases and RMSEs between satellite SST and in situ SST within $1^{\circ} \times 1^{\circ}$ latitude and longitude boxes. (a) and (d) SATSST minus in situ SST. (b) and (e) SATSST_{CDF} minus in situ SST. (c) and (f) SATSST_{LSR} minus in situ SST.

median fluctuates between $-0.07\,^{\circ}\mathrm{C}$ and $0.09\,^{\circ}\mathrm{C}$. The curves of RMSE and RSD are smoother than those without bias correction. The daily bias also fluctuates around the zero value, with the maximum and minimum values of $0.25\,^{\circ}\mathrm{C}$ and $-0.35\,^{\circ}\mathrm{C}$. Compared with the original SATSST, the SATSST_{CDF} eliminates the significant negative bias. Even for the data with a sudden increase in errors after August 2020, the bias correction has achieved excellent results. However, there are still distinguishable biases in August 2020 and February 2021. The former may be due to inadequate correction, and the latter may be due to overcorrection. After the LSR method is used for bias correction, the errors between satellite SST and in situ SST have also been reduced to some extent, but these errors are still higher than those of the CDF method.

FY-4A/AGRI provides hourly SST products, which is helpful to study the DV of SST. Fig. 5 demonstrates the DV characteristics of satellite SST minus in situ SST in local time (LT). Before the bias correction, the satellite SST shows a stronger negative bias during all of 24 hours. After the bias correction with CDF method, the difference between the satellite SST and in situ SST fluctuates around zero, with the maximum value of 0.05 °C and the minimum value of -0.09 °C. It can be found that the difference of SATSST_{CDF} minus in situ SST is subject to daytime warming and nighttime cooling. The cooling effect is significant after 15:00 LT. These features are similar to the study of Tu and Hao [27]. The negative value of SATSST_{CDF} minus in situ SST indicates that the satellite-derived SST is lower

than in situ SST. According to Donlon et al. [4], the temperature measured by an infrared radiometer is the skin temperature at a depth of $\sim\!10\text{--}20~\mu\text{m}$, however, the temperature measured using drifting buoys, vertical profiling floats, or deep thermistor chains is the depth temperature at depths ranging from $10^{\text{-}2}$ to 10^3 m. At night, the skin temperature is lower than the depth temperature, but in daytime, the skin temperature is higher than the depth temperature. The difference of SATSST_{CDF} minus in situ SST also reflects this fact. The variation of SATSST_LSR minus in situ SST can also reflect the process of daytime warming and nighttime cooling, but the temperature difference is always negative.

Fig. 6 shows the geographical distributions of biases and RMSEs for the full disk of satellite SST and in situ SST produced from the matchup dataset by aggregating matchups within 1° × 1° latitude and longitude boxes. The original satellite SST is obviously lower than in situ data in low and middle latitudes and higher in high latitudes. Larger negative biases exceeding -1.0 °C can been found at the low latitudes and at the north edge of disk view. There are also positive biases in the south part of the full disk that cannot be ignored. This bias distribution may be related to the use of the same set of coefficients for the full disk in FY-4A/AGRI SST retrieval algorithm. As shown in Fig. 1, although the matchups almost cover the entire FY-4A observation domain, the spatial density distribution of matchups is not uniform. This affects the distribution of error statistics to a certain extent. It can be seen from Fig. 1 that there are many in

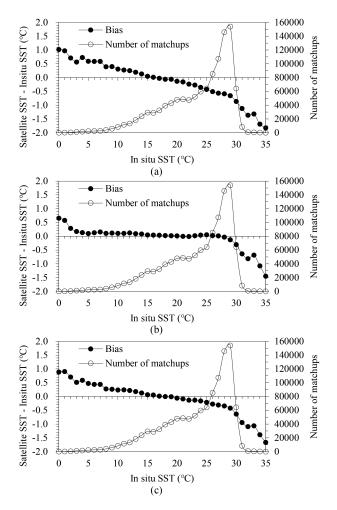


Fig. 7. Dependency of satellite SST minus in situ SST on SST distribution. (a) SATSST minus in situ SST. (b) SATSST $_{\rm CDF}$ minus in situ SST. (c) SATSST $_{\rm LSR}$ minus in situ SST.

situ data in mid-latitude regions, but few in low latitude regions and high latitude regions of the south part of the full disk, and it can be seen from Fig. 2 that the in situ data are mainly below 30 °C. If the in situ data with such distribution characteristics are used to calculate the retrieval coefficients for the full disk data, the values in the high temperature region will be underestimated and the values in the low temperature region will be overestimated. In the eastern tropical ocean, number of SATSST are masked due to the contamination of clouds or precipitation, resulting in fewer matchups in this region. Therefore, the bias in this region is significantly larger than that in the surrounding areas. The bias and RMSE in the area near the latitude of 30 $^{\circ}$ in the south part of the full disk are relatively small, which may be related to the open ocean area in this area and more in situ data. After bias correction with CDF method, the error between satellite SST and in situ SST is significantly reduced, and the spatial distribution of low error is more extensive and uniform. The biases are generally within ± 0.2 °C in full disk. From the spatial distribution map, the bias correction result of LSR method is worse than that of CDF method.

Fig. 7 is the dependency of difference between satellite SST and in situ SST on SST distribution. The number of in situ SST

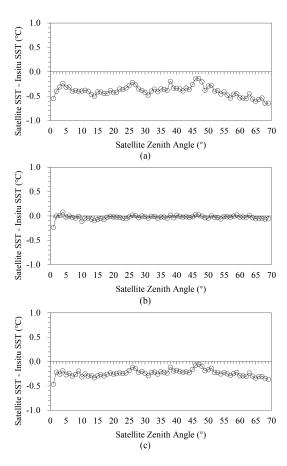


Fig. 8. Dependency of satellite SST minus in situ SST on satellite zenith angel. (a) SATSST minus in situ SST. (b) SATSST $_{\rm CDF}$ minus in situ SST. (c) SATSST $_{\rm LSR}$ minus in situ SST.

for each temperature grade is also shown on the figure. It can better visualize the satellite SST anomalies over the whole range of observed SST values. There is obvious dependency of residual between SATSST and in situ SST on SST distribution. In the low temperature section, SATSST is higher in situ SST. But in the high temperature section, the SATSST is lower than in situ SST. This can also explain the reason to some extent why the retrieval SST from FY-4A/AGRI is lower at low latitude regions and higher in high latitude regions of the southern hemisphere. After bias correction with CDF method, the dependency of satellitederived SST within 5 °C and 28 °C on SST distribution decreases significantly, and the difference of SATSSTCDF minus in situ SST is within ± 0.1 °C. However, the errors of SATSST_{CDF} located outside the range of 5 °C to 28 °C are still large. This may be related to the uneven distribution of in situ data in various temperature zones. It can also be found the LSR method has insufficient ability to correct the error caused by the uneven temperature distribution.

Satellite zenith angle is a key factor affecting the accuracy of remote sensing retrieval SST. Although the influence of zenith angle is considered in the NLSST algorithm adopted by FY-4A/AGRI, the dependency of the difference between satellite SST and in situ SST relative to zenith angle still needs to be considered. Fig. 8 is the dependency of satellite SST minus in situ SST on satellite zenith angel distribution. Before bias correction,

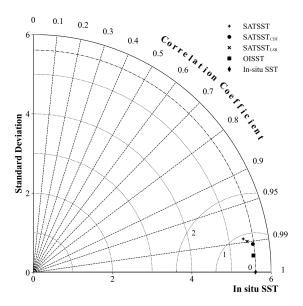


Fig. 9. Taylor diagram comparing the various SST.

the difference of SATSST minus in situ SST is negative in all zenith angle observation ranges. When the satellite zenith angle is greater than 50° and becomes larger, SATSST is more and more low than in situ SST. From 50° zenith angle to 70° zenith angle, a bias of about -0.3 °C is increased. Larger negative bias at high zenith angle may be related to the increase of surface radiation attenuation caused by the lengthening of atmospheric path radiation path. In addition, as the zenith angle increases, the resolution of satellite will decrease. After bias correction using CDF method, the bias of SATSST_{CDF} minus in situ SST is less dependent on the variation of satellite zenith angle. However, the negative bias near the 0° zenith angle is still obvious. According to Fig. 1, the position of sub-satellite point of FY-4A/AGRI is close to land and islands, and the in situ data in this area is relative scarce. At the same time, the detection accuracy of satellite instrument is insufficient in the sea land boundary zone. Therefore, there is a large bias in the low satellite zenith angle region near the sub-satellite point. Compared with CDF method, LSR method still has poor bias correction ability.

C. Evaluation With Taylor Diagram

The in situ SST is usually considered as truth value to evaluate the satellite-derived SST and reanalysis SST. Taylor diagram is used here to determine which SST performs best with respect to in situ data. Fig. 9 shows the overall comparison between the SATSST, SATSST_{CDF}, SATSST_{LSR}, OISST, and in situ SST. Compared with SATSST, the SDs of SATSST_{CDF}, SATSST_{LSR} and OISST are closer to that of in situ SST. Compared with in situ SST, all ubRMSEs of SATSST, SATSST_{CDF}, SATSST_{LSR}, and OISST are within the range of 1.0 °C. The position of OISST is the closest to the in situ SST, followed by SATSST_{CDF} and SATSST_{LSR}, and SATSST is the farthest. The SD of SATSST_{CDF} is at the same level as OISST, but the *R* is slightly lower than that of OISST.

TABLE II ERROR STATISTICS BETWEEN SATSST, IN SITU SST, AND OISST

	Satellite SST (°C)	in situ SST (°C)	OISST (°C)
SATSST	0.83	0.35	0.25
$SATSST_{CDF}$	0.64	0.33	0.27
SATSSTLSR	0.74	0.35	0.25

D. Evaluation With Three-Way Error Analysis

The precision of satellite-derived SST is estimated with the three-way error analysis method, which allows the simultaneous estimation of the precision of each of three observation types [37]. The three-way error analysis is carried out for the total matchups from 2019–2021. Table II is the SD of errors between satellite SST, in situ SST and OISST. The error of SATSST is 0.83 °C. After bias correction through CDF and LSR methods, the errors reduce to 0.64 °C and 0.74 °C, respectively. It shows that the precision of SATSST has been improved after bias correction. The correction result of CDF method is better than that of LSR method. The precision of in situ SST is better than satellite SST, but worse than OISST. This may be related to the fact that the in situ SST comes from various observation platforms such as ships, drifters, Argo floats, and tropical and coastal moorings. This result is also consistent with the analysis results of Xu and Ignatov [16], [38] and Sukresno et al. [39]. The precision estimates for OISST have the smallest error, which is close to the error of global regions given by Huang et al. [25]. This shows that the OISST product is very reliable as reference

IV. DISCUSSION AND CONCLUSION

SST is an essential data for the air-sea interaction studies and ocean modeling. In this article, the errors of operational FY-4A/AGRI hourly SST products are first evaluated in temporal and spatial distribution from January 2019 to December 2021, and the CDF matching method is adopted to reduce the error of FY-4A/AGRI hourly SST based on the in situ SST for a period.

The operational FY-4A/AGRI SST has stronger negative biases and large errors. Comparing against the in situ SST, the FY-4A/AGRI SST has the bias of -0.37 °C with higher RMSE of 0.98 °C for the data of 2019 to 2021. The robust statistics of median and RSD are -0.30 °C and 0.90 °C. These are much larger than the accuracy requirements of 0.5 °C to 0.8 °C for infrared radiometers onboard the geostationary satellite [4]. This result is worse than the evaluation of SST retrieval from similar geostationary satellites such as Himawari-8 [27] and geostationary operational environmental satellites 16 [40].

The quality of SST from FY-4A/AGRI is not stable in time and space. The temporal variations of the errors are displayed by computing the difference between FY-4A/AGRI SST and in situ SST at monthly, daily and hourly timescale. There is always a strong negative bias in the time series. And there is no prominent seasonal variation during the period analyzed. Satellite operational calibration update will cause jumps in the time series of SST bias, which is related to the failure of SST retrieval algorithm to update in time. The spatial distributions

of errors show that there are larger negative biases in low and middle latitude region and larger positive biases in southern high latitude region. In local area, larger biases even exceed $\pm 1.0\,^{\circ}\text{C}$. There are dependencies of the residuals between satellite SST and in situ SST on SST distribution and satellite zenith angle. The satellite SST is higher than in situ SST in the low temperature section, but lower than in situ SST in the high temperature section. As the satellite zenith angle is greater than 50° and becomes larger, the negative bias of satellite SST minus in situ SST becomes larger. At the maximum zenith angle of 70°, the negative bias almost increases by $-0.30\,^{\circ}\text{C}$. These show that the accuracy of operational FY-4A/AGRI SST products is quite poor.

The precision of FY-4A/AGRI SST has been greatly improved after bias correction. Compared with LSR method, the CDF methos is more effective to correct the biases of FY-4A/AGRI SST. Although the CDF and LSR methods use the same local matchup dataset, the CDF considers the probability distribution characteristics of the matchups and performs piecewise regression for the data within the local region. After bias correction with CDF method, the biases and RMSEs of FY-4A/AGRI SST are reduced to -0.02 °C and 0.72 °C, respectively. The median and RSD are 0.00 °C and 0.60 °C. Although the corrected SST accuracy has not yet met the demand target of climate research, these errors meet the absolute accuracy for SST retrieval from geostationary satellites [4]. And the evaluation result is even better than that of some satellite-derived SST products [41], [42]. On the time scale, the fluctuation ranges of bias and median are smaller. The difference of satellite SST minus in situ SST can reflect the DV of SST. The spatial distribution of error shows the bias within ± 0.2 °C in full disk. The error dependencies on satellite zenith angle and SST distribution are also greatly reduced. Taylor diagram and three-way error analysis also show that the accuracy of FY-4A/AGRI SST has been significantly improved after bias correction.

The CDF matching method is a simple and easy-to-use dynamic piecewise linear regression algorithm, which can complete the bias correction only with a reference data. The period of data used in the regression process covers 15 days prior to the target day. This time period is obtained through simple tests on matching samples of FY-4A/AGRI SST and in situ SST. It may not be the optimal time period for other data, so it is necessary to use the corresponding matching sample to retrain for obtaining suitable time period. According to the statistical results of three years' data, the performance of CDF is stable. However, the researchers believe that the SST retrieval errors should account for the dependence on observational conditions [6], [7]. Some approaches [20], [35] of SST retrieval and error correction had considered the influences of satellite zenith angle, water vapor, wind speed, and climatological SST. Therefore, the influence of more physical factors needs to be considered in the future work of error evaluation and bias correction for satellite-derived SST products.

The in situ SSTs are usually used for accuracy verification and bias correction of satellite retrieval SST. However, these in situ SSTs are very heterogeneous in depth for coming from a variety of observation platforms. The research shows that there

are larger differences between the validation results from various types of in situ SST [27], [38]. The difference of observation data from different platforms is not considered during the error evaluation and bias correction of FY-4A/AGRI SST. In future work, it is necessary to consider the influence of observation platforms.

ACKNOWLEDGMENT

The authors would like to thank NSMC/CMA for providing FY-4A/AGRI data, also thank STAR/NESDIS/NOAA for providing iQuam data. The authors are very grateful to Peng Cui, a senior engineer of NSMC/CMA, for his explanation of FY-4A/AGRI SST retrieval algorithm.

REFERENCES

- [1] P. J. Minnett et al., "Half a century of satellite remote sensing of sea-surface temperature," *Remote Sens. Environ.*, vol. 233, 2019, Art. no. 111366.
- [2] C. J. Merchant et al., "Satellite-based time-series of sea-surface temperature since 1981 for climate applications," *Sci. Data*, vol. 6, no. 1, 2019, Art. no. 223.
- [3] G. Ohring, B. Wielicki, R. Spencer, B. Emery, and R. Datla, "Satellite instrument calibration for measuring global climate change: Report of a workshop," *Bull. Amer. Meteorol. Soc.*, vol. 86, no. 9, pp. 1303–1313, 2005.
- [4] C. Donlon et al., "The global ocean data assimilation experiment highresolution sea surface temperature pilot project," *Bull. Amer. Meteorol. Soc.*, vol. 88, no. 8, pp. 1197–1213, 2007.
- [5] J. J. Kennedy, "A review of uncertainty in in situ measurements and data sets of sea surface temperature," *Rev. Geophys.*, vol. 52, no. 1, pp. 1–32, 2014.
- [6] S. L. Castro, G. A. Wick, D. L. Jackson, and W. J. Emery, "Error characterization of infrared and microwave satellite sea surface temperature products for merging and analysis," *J. Geophys. Res.*, *Oceans*, vol. 113, 2008, Art. no. C03010.
- [7] F. Xu, A. Ignatov, and X. Liang, "Towards continuous error characterization of sea surface temperature in the advanced clear-sky processor for oceans," in *Proc. 16th Conf. Satell. Meteorol. Oceanographic*, 2009, [Online]. Available: https://ams.confex.com/ams/89annual/webprogram/ Paper143882.html
- [8] J. Cheng, X. Cheng, X. Meng, and G. Zhou, "A Monte Carlo emissivity model for wind-roughened sea surface," *Sensors-Basel*, vol. 19, no. 9, 2019, Art. no. 2166.
- [9] C. J. Merchant and A. R. Harris, "Toward the elimination of bias in satellite retrievals of sea surface temperature: 2. Comparison with in situ measurements," *J. Geophys. Res., Oceans*, vol. 104, no. C10, pp. 23579–23590, 1999
- [10] Z. Liao, Q. Dong, and C. Xue, "Evaluation of sea surface temperature from FY-3C data," Int. J. Remote. Sens., vol. 38, no. 17, pp. 4954–4973, 2017.
- [11] A. G. O'Carroll, J. R. Eyre, and R. W. Saunders, "Three-way error analysis between AATSR, AMSR-E, and in situ sea surface temperature observations," *J. Atmos. Ocean. Technol.*, vol. 25, no. 7, pp. 1197–1207, 2008.
- [12] C. Kumar, G. Podestá, K. Kilpatrick, and P. Minnett, "A machine learning approach to estimating the error in satellite sea surface temperature retrievals," *Remote Sens. Environ.*, vol. 255, 2021, Art. no. 112227.
- [13] R. W. Reynolds, C. K. Folland, and D. E. Parker, "Biases in satellite-derived sea-surface-temperature data," *Nature*, vol. 341, no. 6244, pp. 728–731, 1989.
- [14] R. W. Reynolds, "Impact of Mount Pinatubo aerosols on satellite-derived sea surface temperatures," J. Climate, vol. 6, no. 4, pp. 768–774, 1993.
- [15] P. Dash, A. Ignatov, Y. Kihai, and J. Sapper, "The SST quality monitor (SQUAM)," J. Atmos. Ocean. Technol., vol. 27, no. 11, pp. 1899–1917, 2010.
- [16] F. Xu and A. Ignatov, "In situ SST quality monitor (iQuam)," J. Atmos. Ocean. Technol., vol. 31, no. 1, pp. 164–180, 2014.
- [17] R. W. Reynolds and T. M. Smith, "Improved global sea surface temperature analyses using optimum interpolation," *J. Climate*, vol. 7, no. 6, pp. 929–948, 1994.

- [18] R. W. Reynolds, T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax, "Daily high-resolution-blended analyses for sea surface temperature," *J. Climate*, vol. 20, no. 22, pp. 5473–5496, 2007.
- [19] J. L. Høyer, P. L. Borgne, and S. Eastwood, "A bias correction method for Arctic satellite sea surface temperature observations," *Remote Sens. Environ.*, vol. 146, no. 5, pp. 201–213, 2014.
- [20] B. Petrenko, A. Ignatov, Y. Kihai, and P. Dash, "Sensor-specific error statistics for SST in the advanced clear-sky processor for oceans," *J. Atmos. Ocean. Technol.*, vol. 33, no. 2, pp. 345–359, 2016.
- [21] K. Kwon, B.-J. Choi, S.-D. Kim, S.-H. Lee, and K.-A. Park, "Assessment and improvement of global gridded sea surface temperature datasets in the Yellow Sea using in situ ocean buoy and research vessel observations," *Remote Sens.*, vol. 12, no. 5, 2020, Art. no. 759.
- [22] G. Han et al., "Bias correction of sea surface temperature retrospective forecasts in the South China Sea," *Acta Oceanologica Sinica*, vol. 41, no. 2, pp. 41–50, 2022.
- [23] W. Wang and P. Xie, "A multiplatform-merged (MPM) SST analysis," J. Climate, vol. 20, no. 9, pp. 1662–1679, 2007.
- [24] J. Yang, Z. Zhang, C. Wei, F. Lu, and Q. Guo, "Introducing the new generation of Chinese geostationary weather satellites, Fengyun-4," *Bull. Amer. Meteorol. Soc.*, vol. 98, no. 8, pp. 1637–1658, 2017.
- [25] B. Huang et al., "Improvements of the daily optimum interpolation sea surface temperature (DOISST) version 2.1," *J. Climate*, vol. 34, no. 8, pp. 2923–2939, 2021.
- [26] Q. He, Y. Zhang, and J. Wang, "Development and evaluation of regional SST regression algorithms for FY-3C/VIRR data in the western north pacific," *Remote Sens. Lett.*, vol. 11, no. 12, pp. 1090–1099, 2020.
- [27] Q. Tu and Z. Hao, "Validation of sea surface temperature derived from Himawari-8 by JAXA," *IEEE J. Sel. Topics Appl. Earth Ob*serv. Remote Sens., vol. 13, pp. 448–459, Jan. 2020, doi: 10.1109/JS-TARS.2019.2963773.
- [28] S.-B. Duan et al., "Validation of collection 6 MODIS land surface temperature product using in situ measurements," *Remote Sens. Environ.*, vol. 225, pp. 16–29, 2019.
- [29] K. E. Taylor, "Summarizing multiple aspects of model performance in a single diagram," *J. Geophys. Res. Atmos.*, vol. 106, no. D7, pp. 7183–7192, 2001.
- [30] D. Entekhabi, R. H. Reichle, R. D. Koster, and W. T. Crow, "Performance metrics for soil moisture retrievals and application requirements," *J. Hy-drometeorol.*, vol. 11, no. 3, pp. 832–840, 2010.
- [31] R. H. Reichle and R. D. Koster, "Bias reduction in short records of satellite soil moisture," *Geophys. Res. Lett.*, vol. 31, no. 19, Oct. 2004, Art. no. L19501.
- [32] Y. Y. Liu et al., "Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals," *Hydrol. Earth Syst. Sci.*, vol. 15, no. 2, pp. 425–436, 2011.
- [33] D. Macias, E. Garcia-Gorriz, A. Dosio, A. Stips, and K. Keuler, "Obtaining the correct sea surface temperature: Bias correction of regional climate model data for the Mediterranean Sea," *Climate Dyn.*, vol. 51, no. 3, pp. 1095–1117, 2018.
- [34] J.-H. Heo, G.-H. Ryu, E.-H. Sohn, and J.-D. Jang, "Application of cumulative distribution functions to compositing precipitable water with low Earth orbit satellite data," *Asia-Pac. J. Atmos. Sci.*, vol. 56, no. 1, pp. 57–75, 2020.
- [35] Z. Liao, Q. Dong, and C. Xue, "A bias correction method for FY-3C VIRR SST data," *Remote Sens. Lett.*, vol. 8, no. 5, pp. 429–437, 2017.
- [36] X. He, N. Xu, X. Feng, X. Hu, H. Xu, and P. Yi, "Assessing radiometric calibration of FY-4A/AGRI thermal infrared channels using CrIS and IASI," *IEEE Tran. Geosci. Remote Sens.*, vol. 60, Sep. 2021, Art. no. 5514512, doi: 10.1109/TGRS.2021.3111975.
- [37] C. J. Merchant et al., "A 20 year independent record of sea surface temperature for climate from along-track scanning radiometers," *J. Geophys. Res. Oceans*, vol. 117, no. C12, 2012, Art. no. C12013.
- [38] F. Xu and A. Ignatov, "Error characterization in iQuam SSTs using triple collocations with satellite measurements," *Geophys. Res. Lett.*, vol. 43, no. 20, pp. 10826–10834, 2016.
- [39] B. Sukresno, R. Hanintyo, D. W. Kusuma, D. Jatisworo, and A. Murdimanto, "Three-way error analysis of sea surface temperature (SST) between Himawari-8, buoy, and MUR SST in Savu Sea," *Int. J. Remote Sens. Earth Sci.*, vol. 15, no. 1, pp. 25–36, 2018.

- [40] M. H. Azevedo, N. Rudorff, and J. A. Aravéquia, "Evaluation of the ABI/GOES-16 SST product in the tropical and Southwestern Atlantic Ocean," *Remote Sens.*, vol. 13, no. 2, 2021, Art. no. 192.
- [41] K.-A. Park, E.-Y. Lee, X. Li, S.-R. Chung, E.-H. Sohn, and S. Hong, "NOAA/AVHRR sea surface temperature accuracy in the East/Japan Sea," Int. J. Digit. Earth, vol. 8, no. 10, pp. 784–804, 2015.
- [42] Y. Hao, T. Cui, V. P. Singh, J. Zhang, R. Yu, and Z. Zhang, "Validation of MODIS sea surface temperature product in the coastal waters of the Yellow Sea," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 5, pp. 1667–1680, May 2017, doi: 10.1109/JSTARS.2017.2651951.



Quanjun He received the M.S. degree in geophysical prospecting and information technology from Jilin University, Changchun, China, in 2004, and the Ph.D. degree in physical oceanography from the South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou, China, in 2014.

He is currently a Senior Engineer with the Guangzhou Meteorological Satellite Ground Station (Guangdong Meteorological Satellite Remote Sensing Center), Guangdong Meteorological Service, Guangzhou, China. His research interests include re-

trieval algorithms and validation for sea surface temperature, data reconstruction and multi-source data fusion, and application of remote sensing.



Xin Hu received the B.S. degree in remote sensing science and technology from Chengdu University of Information Technology, Chengdu, China, in 2018, and the M.S. degree in atmospheric physics and atmospheric environment from the Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, China, in 2021.

He is currently with the Guangzhou Meteorological Satellite Ground Station (Guangdong Meteorological Satellite Remote Sensing Center), Guangdong Meteorological Service, Guangzhou, China. His re-

search interests include identification and monitoring of sea surface wind fields by remote sensing and application of remote sensing in land surface ecological environment.



Yanwei Wu received B.S. degree in computer science and technology from South China University of Technology, Guangzhou, China, in 2004.

He is currently an Engineer with the Guangzhou Meteorological Satellite Ground Station (Guangdong Meteorological Satellite Remote Sensing Center), Guangdong Meteorological Service, Guangzhou, China. His research interests include meteorological satellite operation, meteorological satellite application and service.