Dynamic optimal estimation with atmospheric correction smoothing for sea surface skin temperature retrieval from infrared satellite imagery

Article

Accepted Version

Dynamic Optimal Estimation with Atmospheric Correction Smoothing for Sea Surface Skin Temperature Retrieval from Infrared Satellite Imagery

Mingkun Liu, *Member, IEEE*, Lei Guan, *Member, IEEE*, Fanli Liu, Zhicheng Sheng, Zhuomin Li, and Christopher J. Merchant

***Abstract*—This study offers an in-depth exploration into Sea Surface Skin Temperature (SST_{skin}) from the Haiyang-1D (HY-1D) Chinese Ocean Color and Temperature Scanner (COCTS). The main components include inter-calibration, cloud detection, and SST_{skin} retrieval. First, we conduct the inter-calibration of COCTS infrared channels utilizing Visible Infrared Imaging Radiometer Suite (VIIRS) as the reference instrument. A double-differencing methodology is employed to evaluate and correct the COCTS calibration. Next, we introduce a physically based deep learning algorithm for cloud detection, designed to interpret complex textures in satellite imagery. The algorithm demonstrates the superior performance across diverse conditions and geographical areas, especially reducing false flagging of ocean fronts. Lastly, we propose an Optimal Estimation (OE) methodology for COCTS SST_{skin} retrieval. One focus is on estimating appropriate covariance matrices within the OE algorithm, including an innovative method for dynamically setting the prior SST uncertainty appropriate to local spatial variability. A second focus is to employ atmospheric correction smoothing algorithm of OE. Both these measures combine to suppress noise and enhance sensitivity of SST_{skin}. We assign quality levels to the retrieved SST_{skin} data. The high-quality COCTS SST_{skin} is validated using iQuam in-situ data. Our results indicate the bias of -0.20 °C and the robust standard deviation of 0.27 °C between COCTS and in-situ SST, with an average sensitivity of 0.87. These findings affirm that the successful implementation of these methodologies significantly enhances the accuracy and reliability of SST_{skin} data from HY-1D COCTS. This advancement provides substantial benefits to expand the global high precision SST_{skin} dataset.**

***Index Terms*—infrared, inter-calibration, sea surface skin temperature, deep learning, cloud detection, optimal estimation retrieval**

I. INTRODUCTION

SEA Surface Temperature (SST) is a crucial parameter in understanding ocean dynamics and climate systems. It is essential across multiple disciplines, including climate studies, oceanography, ecological research, and weather prediction [1], [2]. Defined by Group for High Resolution Sea Surface Temperature (GHRSST), SST refers to the temperature of the ocean upper layer, generally less than 10 meters deep. For climate monitoring, satellite infrared radiometers are used to measure Sea Surface Skin Temperature (SST_{skin}), as recommended by the Global Climate Observing System (2011). Satellite SST_{skin} retrieval involves three main steps. First, satellite sensor data, measured as Digital Numbers (DN), are calibrated to reflect top-of-atmosphere radiance. Cloud detection is crucial since infrared waves cannot penetrate clouds, impacting the accuracy of SST_{skin} data. Misidentification of clouds leads to inaccuracies, while excessive cloud detection may lead to loss of valid data. Finally, atmospheric correction is applied to the radiance from cloud-free areas to obtain accurate SST_{skin} values.

Accurate and stable calibration is crucial for retrieving geophysical data. Despite SI-aligned prelaunch calibrations, post-launch methods remain non-standardized. Inter-calibration is key to resolving instrument discrepancies and measuring uncertainties [3]. The preferred method uses a well-calibrated reference, like a hyperspectral instrument, to compare with the target instrument [4], [5], [6]. When direct comparison is impractical, double-differencing with an intermediate standard, such as another instrument or a radiative transfer model, is employed [3], [7], [8], [9]. This method is particularly relevant for inter-calibrating infrared channels.

Cloud detection algorithms in satellite imagery are

This work was supported by the National Natural Science Foundation of China under grant 42006161, the Hainan Province Science and Technology Special Fund under grant ZDYF2024GXJS260/SOLZSKY2024006, and the Hainan Provincial Natural Science Foundation of China under Grant 122CXTD519. *(Corresponding author: Lei Guan).*

Mingkun Liu and Lei Guan are with College of Marine Technology, Faculty of Information, Science and Engineering, Ocean University of China / Laboratory for Regional Oceanography and Numerical Modeling, Qingdao Marine Science and Technology Center, Qingdao, China. They are also with the Key Laboratory of Ocean Observation and Information of Hainan Province, Sanya Oceanographic Institution, Ocean University of China / SANYA Oceanographic Laboratory, Sanya 572024, China.

Fanli Liu and Zhuomin Li are with College of Marine Technology, Faculty of Information, Science and Engineering / Key Laboratory of Ocean Observation and Information of Hainan Province, Sanya Oceanographic Institution, Ocean University of China, Qingdao 266100/ Sanya 572024, China.

Zhicheng Sheng is with School of Computer Science and Technology, Shandong University, Qingdao 266237, China.

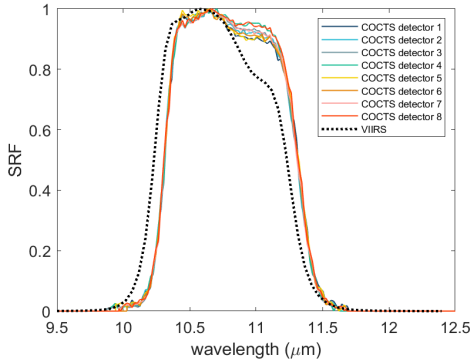
Christopher J. Merchant is with the Department of Meteorology and the National Centre for Earth Observation, University of Reading, Reading RG6 6ET, U.K.

categorized into threshold-based, Bayesian-based, and machine learning approaches. Threshold algorithms use specific threshold values based on spectral values and texture features for cloud identification, but struggle with variable conditions [10]. Bayesian algorithms improve accuracy by incorporating prior atmospheric knowledge, calculating clear-sky probabilities per pixel [11], [12], [13], [14], but require extensive data from weather forecasts and radiative transfer models. With AI advancements, deep learning is being explored for cloud detection, offering high accuracy and speed by classifying pixels directly and handling complex textures effectively [15], [16].

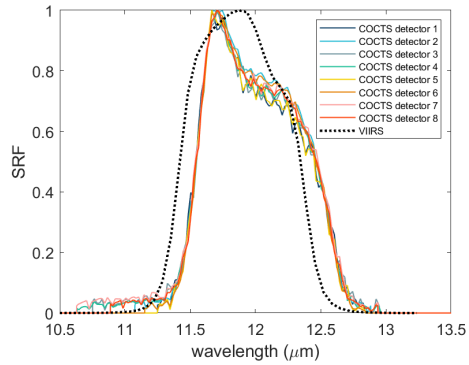
Infrared SST retrieval algorithms are mainly split into two types: Multi-Channel SST (MCSST) / Non-Linear SST (NLSST), which use split-window techniques [17], [18], and the Optimal Estimation (OE) method [19]. MCSST/NLSST algorithms, used in operational products from sensors like AVHRR, MODIS, and VIIRS, derive SST by correlating satellite BTs with buoy measurements through regression [17], [20], [21]. However, they reflect temperatures at buoy depths, not the true skin temperature. The ATSR series algorithms, overseen by ESA, rely on atmospheric radiative transfer models for brightness temperature (BT) simulation, that are fitted with SST_{skin} to determine the NLSST coefficients [22], [23]. Alternatively, the OE method integrates radiative transfer simulations with prior information for SST_{skin} retrieval, requiring accurate forward model errors, unbiased priors, and well-estimated error covariances [19], [24]. In recent years, scholars have made new technical explorations in the retrieval of infrared SST. Merchant et al., drawing on the concept of Kalman filtering, improved the calculation of the covariance matrix in the SST_{skin} retrieval by adopting a multi-iteration algorithm [25]. Luo et al. combined atmospheric radiative transfer models and multidimensional fitting techniques for nighttime SST retrieval in heavily aerosol-polluted sea areas [26]. Jia et al. improves the accuracy of MODIS SST retrievals in the Arctic by correcting for sea surface emissivity effects and adjusting algorithm weighting factors, resulting in significantly reduced biases [27].

China's first-generation oceanic satellites, HY-1C and HY-1D of the Haiyang-1 (HY-1) series, were launched in September 2018 and June 2020. The Chinese Ocean Color and Temperature Scanner (COCTS) onboard HY-1D has eight visible and near-infrared channels for ocean color and two thermal infrared channels for SST observation. Currently, scholars have conducted research related to the SST measured by the HY-1 COCTS. In terms of calibration evaluation, Liu et al. used the Infrared Atmospheric Sounding Interferometer (IASI), a hyperspectral sensor, to perform inter-calibration on the thermal infrared channels of HY-1B and HY-1C COCTS instruments [28], [29]. For cloud detection, Ye et al. and Li et al. applied threshold discrimination algorithms based on the reflectance from visible channels and the BTs from infrared channels [30], [31], [32]. However, because fixed thresholds may not be applicable to global ocean regions, this method can introduce errors in SST retrievals. Liu et al. conducted cloud detection research based on Bayesian principles, combining background field data with atmospheric radiative transfer models [12]. In terms of SST retrieval, Ye et al. employed the NLSST algorithm using a dataset that matches satellite-observed BTs with in-situ measured SSTs. They achieved a retrieval accuracy of approximately 0.8°C [30], [31]. However, this algorithm retrieves the bulk temperature of the seawater rather than the skin temperature. Li et al. used the NLSST algorithm for SST retrieval of HY-1D COCTS, based on a dataset matching simulated BTs from atmospheric radiative transfer with SST_{skin} [32]. The results indicated a retrieval accuracy of about 0.5°C [32]. Additionally, Liu et al. applied the OE retrieval algorithm to HY-1B COCTS data over the Northwest Pacific Ocean, achieving a regional retrieval accuracy of 0.51°C [12].

Our study focuses on retrieving accurate SST_{skin} from HY-1D COCTS data. We first address the inter-calibration of COCTS two thermal infrared channels with VIIRS (Section II). We then apply a U-Net based deep learning approach for cloud detection in satellite imagery, creating a dedicated cloud detection model for HY-1D COCTS (Section III). Section IV explores the use of the OE method for SST_{skin} retrieval by integrating atmospheric radiative transfer models with COCTS specifics, with a focus on the dynamic covariance matrix estimation and atmospheric correction smoothing. By integrating the cloud detection results with the OE, we develop a quality control process for SST_{skin} data from HY-1D COCTS and categorize its quality (Section V). The validation of the retrieved SST_{skin} using iQuam in-situ data is presented in Section VI, while Section VII concludes the study.



(a) 11 μm



(b) 12 μm

Fig.1. The spectral response functions of COCTS and VIIRS. (a) 11 μm channel, (b)12 μm channel.

II. Inter-calibration of COCTS infrared channels with VIIRS

In this study, we conduct an evaluation and correction of the calibration of the two thermal infrared channels from the HY-1D COCTS. We employ the double-differencing method for inter-calibration, selecting the VIIRS on-board the Joint Polar Satellite System-1 (JPSS-1) satellite as the inter-calibration reference instrument. Simulations from radiative transfer models are selected as the intermediate reference. Previous assessments of VIIRS calibration have indicated excellent stability, with its overall performance continuing to meet design requirements [33], [34].

A. COCTS and VIIRS Characteristics

In our study, we assess the calibration of COCTS infrared channels using VIIRS as a reference. Fig. 1 showcases the Spectral Response Functions (SRFs) for COCTS and VIIRS, with (a) and (b) corresponding to the 11 and 12 μm channels. The figure differentiates between the SRFs of COCTS eight detectors (solid lines) and VIIRS SRFs (dotted lines), revealing distinct spectral characteristics despite similar center wavelengths. Both instruments calibrate on-orbit using cold space and an onboard blackbody [29], [34].

B. Double-differencing method

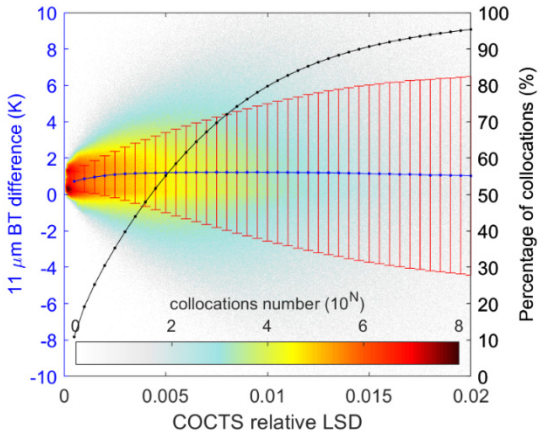
Because the spectral characteristics of COCTS and VIIRS infrared channels differ, the comparison of BTs from COCTS and VIIRS must consider the spectral difference. The double-differencing method can be represented using the following (1). The actual BT difference between COCTS and VIIRS is the observed BT difference minus the spectral difference. We apply the MODerate resolution atmospheric TRANsmission (MODTRAN) simulation to calculate the spectral difference between COCTS and VIIRS.

$$DD = (BT^{O}_{COCTS} - BT^{O}_{VIIRS}) - (BT^{M}_{COCTS} - BT^{M}_{VIIRS}) \qquad (1)$$

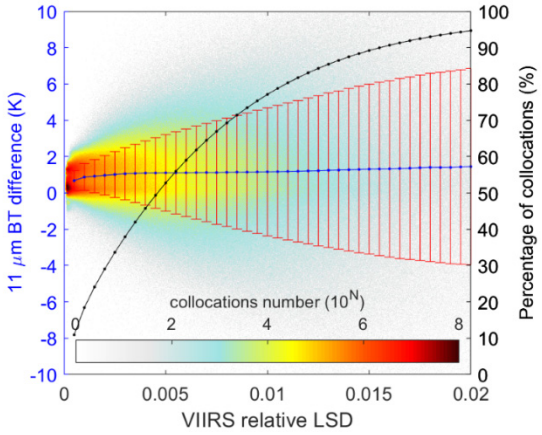
In (1), BT_{COCTS} and BT_{VIIRS} represent the COCTS and VIIRS BT, respectively, superscript O indicates observation and superscript M represents the modelled BT. DD represents the double difference between the COCTS and VIIRS BT.

To compare BTs between COCTS and VIIRS, we match their spatiotemporal data, using 0.01°×0.01° grid spacing and 10-minute time windows, and focusing on simultaneous nadir overpasses to minimize atmospheric path differences. Homogeneity of the collocated grid must be considered due to the different point spread functions between COCTS and VIIRS. We use a relative local standard deviation (LSD) in 9 × 9 boxes to assess homogeneity. Analysis of June 2021 data (Fig. 2) reveals that BT discrepancies increase with both COCTS and VIIRS relative LSDs. We determine a relative LSD threshold of 0.0005 to identify uniform regions, which represent about 10% of the data, thus reducing the BT uncertainties.

In this study, we utilize the MODTRAN, a well-established atmospheric "narrow band model," for simulating BTs of COCTS and VIIRS, given the atmospheric profiles, surface parameters, and SRFs of COCTS and VIIRS [35]. The requisite atmospheric profiles (including pressure, temperature, water vapor, ozone, carbon dioxide) and surface parameters for the MODTRAN simulation are derived from European Centre for Medium-Range Weather Forecasts (ECMWF) ERA5 data, which is projected onto a 0.25°× 0.25° grid map every hour [36]. The sea surface emissivity is acquired through a lookup of emissivity derived from the refractive index for a range of wavenumbers, view angles, temperature and wind speed [37]. Given that the spatial window for COCTS and VIIRS collocations is 0.01°×0.01°, multiple points fall within the same ERA5 grid. As a result, the input profiles and parameters for the MODTRAN simulations for these matchups are nearly identical, with only the satellite zenith angles (SZAs) differing. The simulated BT differences due to varying SZAs within each 0.25°×0.25° grid are minor. Consequently, to expedite the MODTRAN simulation, we retain only one matchup in each 0.25° × 0.25° grid for the subsequent double-differencing comparison. This method yields a total of 955352 collocations for inter-comparison.

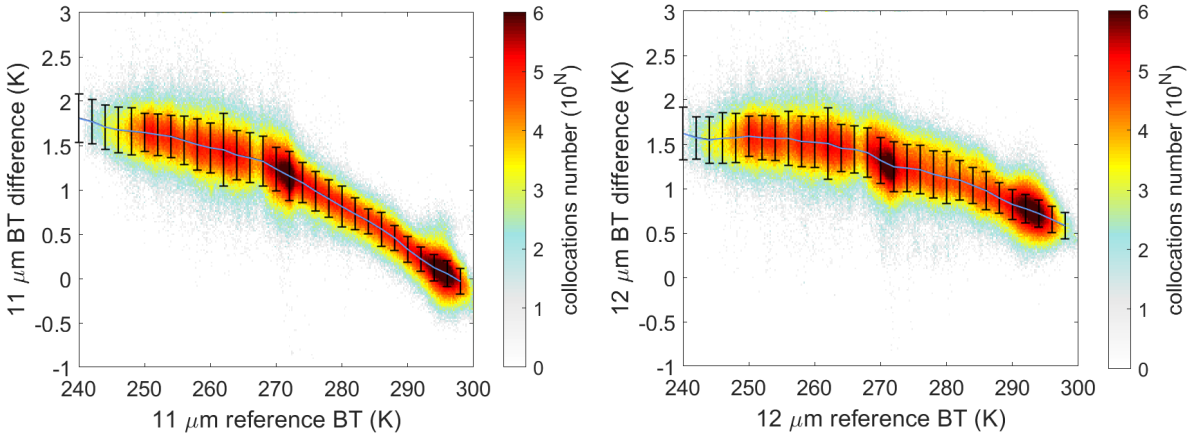


(a)



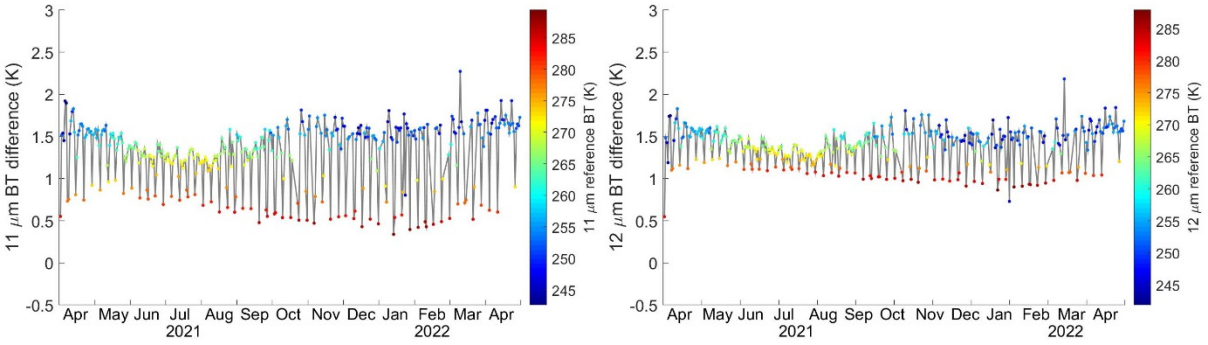
(b)

Fig.2. The variations of COCTS minus VIIRS 11 μm BT difference against the (a) COCTS relative LSD and (b) VIIRS relative LSD. The collocation number is depicted by the background color. The mean difference for each 0.0005 bin of COCTS relative LSD is shown by the blue line, while the standard deviation (SD) of this difference is illustrated by the red lines. The black curve indicates the proportion of matching points within each bin to the total number of matchups, corresponding to the y-axis on the right.



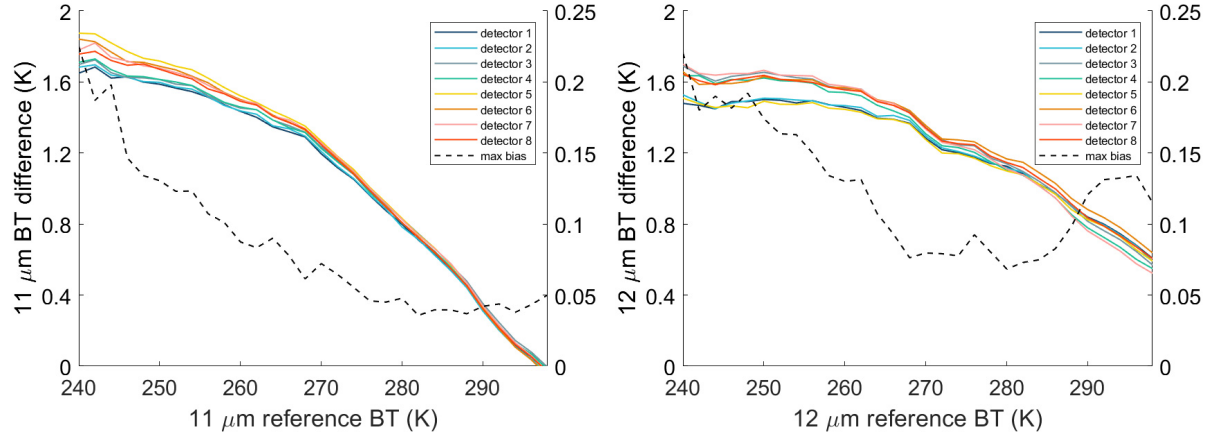
(a) 11 µm (b) 12 µm

Fig. 3. The variation of COCTS minus reference BT difference against the reference BT for (a) 11 and (b) 12 µm channel.



(a) 11 µm (b) 12 µm

Fig. 4. The time series plots of COCTS minus reference BT difference of the period from April 2021 to April 2022 for (a)11 and (b) 12 µm channel.



(a) 11 µm (b) 12 µm

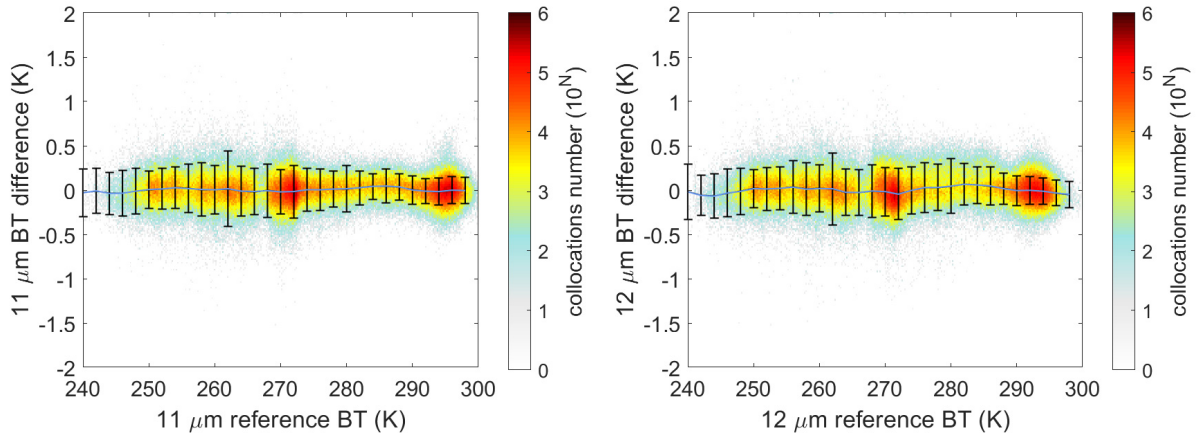
Fig. 5. The variants of BT double difference against the reference BT observing from eight detectors: (a) 11 and (b) 12 µm channel.

C. Evaluation of COCTS infrared BTs

According to the double-differencing comparison method, the reference BT for COCTS evaluation can be calculated based on the VIIRS observation by subtracting the spectral difference obtained from the MODTRAN simulation, shown as (2).

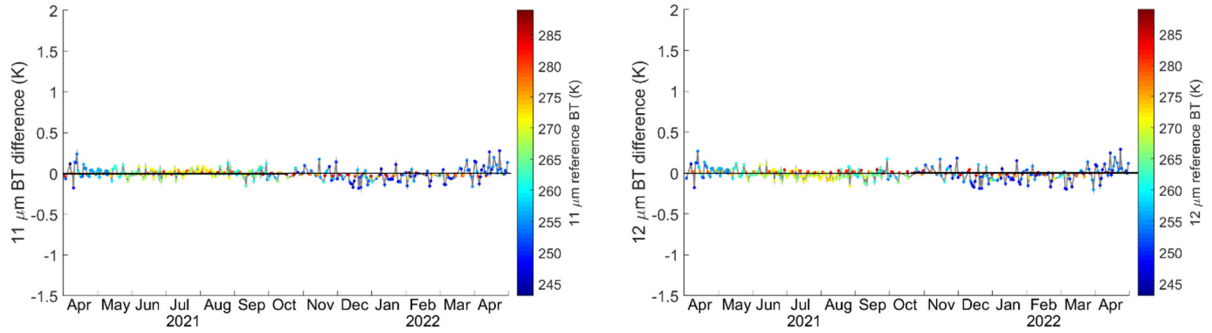
$$BT_{ref} = BT^{O}_{VIIRS} - (BT^{M}_{VIIRS} - BT^{M}_{COCTS}) \quad (2)$$

The calibration of COCTS 11 and 12 µm channels is assessed by comparing their BTs with reference BTs, as shown in Fig. 3. This figure illustrates the BT differences across 0.2 K reference BT bins, with color indicating collocation frequency. A blue line indicates the average BT discrepancy per 2 K reference BT bin, and black bars show the SDs. Significant differences between the COCTS and reference BT are present, with an average bias of 0.97 K and the Robust Standard Deviation (RSD) [38] of 0.65 K for the 11 µm channel, and an average bias of 1.21 K and RSD of 0.42 K for the 12 µm channel. These differences negatively correlate with reference BTs, especially



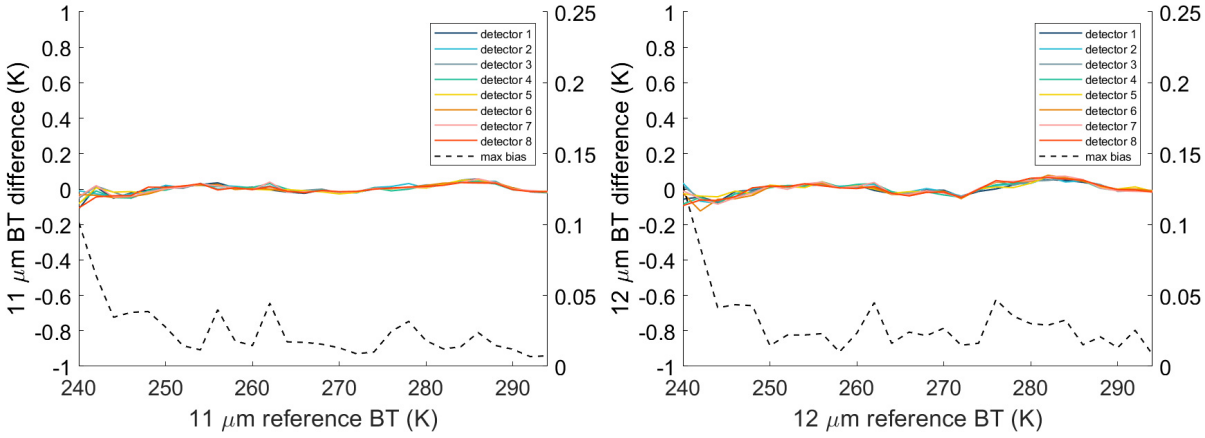
(a) 11 μm (b) 12 μm

Fig. 6. The variation of COCTS minus reference BT difference against the reference BT after correction for (a) 11 and (b) 12 μm channel.



(a) 11 μm (b) 12 μm

Fig. 7. The daily bias change of COCTS BT over the reference BT after correction for (a) 11 and (b) 12 μm channel.

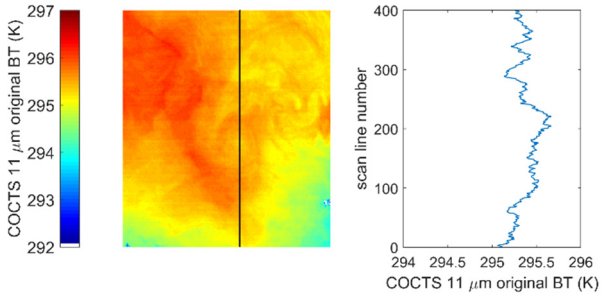


(a) 11 μm (b) 12 μm

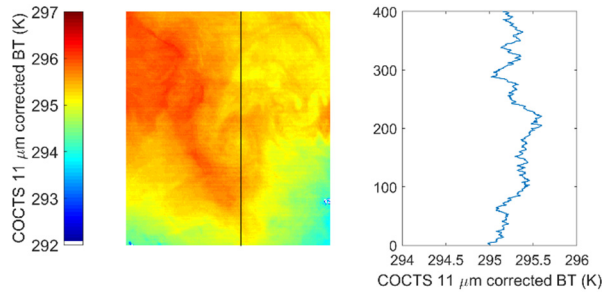
Fig.8. The variants of BT double difference over the reference BT after correction for (a) 11 and (b) 12 μm channel.

for the 11 μm channel, indicating BT-dependent nature of calibration errors. Such BT-dependent bias may result from inadequate nonlinear correction, a critical calibration step often assessed pre-launch. However, if these nonlinearities are not properly accounted for, variations in calibration bias against the scene temperature may become apparent [39]. Hence, uncertainties in the nonlinear correction for COCTS could be a probable cause for the observed BT-dependent bias.

The calibration stability of a radiometer is another vital metric. To monitor the infrared calibration stability of COCTS, a daily global BT comparison with VIIRS is conducted based on daily collocations. Fig. 4 shows variations in daily BT biases for the 11 and 12 μm channels in relation to the reference BTs. The gray lines in the graph, which correspond to the left y-axis, trace these bias variations. Meanwhile, the colors of the individual dots scattered across the graph represent the daily average values of the reference BT. The daily BT difference displays apparent fluctuation, and a strong negative correlation

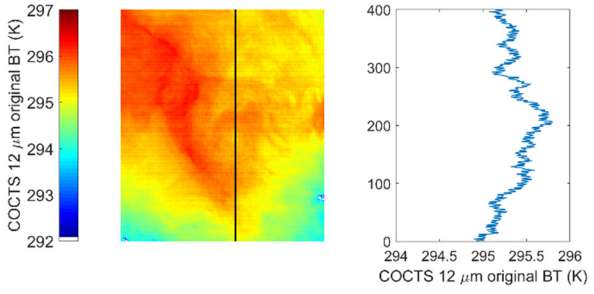


(a) original BT

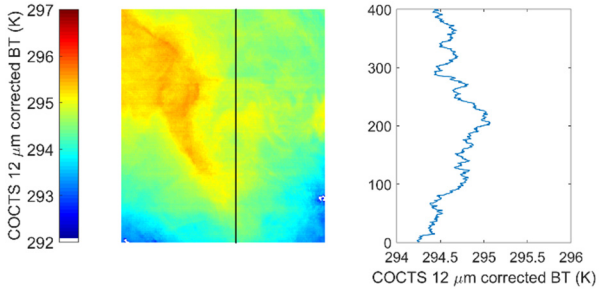


(b) corrected BT

Fig. 9. HY-1D COCTS 11μm striping demo: (a) original BT, (b) corrected BT.



(a) original BT



(b) corrected BT

Fig. 10. HY-1D COCTS 12 μm striping demo: (a) original BT, (b) corrected BT.

can be observed between the mean BT difference and the reference BT. The daily bias reaches a local minimum under high reference BT conditions. Thus, we infer that the daily bias is heavily influenced by the matched location and the corresponding reference BT.

HY-1D COCTS BT images exhibit noticeable striping noise along the track direction due to the different responses of its eight scanning detectors, a problem also reported in previous HY-1B and HY-1C COCTS [28], [29]. To assess the BT differences among these detectors, we compared their BTs with reference BTs. Fig. 5 shows BT difference variations for each detector, with solid lines with different colors indicating individual detectors and a dotted line marking the range of differences corresponding to right y-axis. The extent of overlap among the different lines illustrate the variance between detectors, with a greater divergence at BTs below 270 K for the 11 μm channel. The 12 μm channel shows more significant differences than the 11 μm, particularly below 260 K or above 290 K.

D. COCTS BT correction

Findings from the comparison of COCTS infrared BTs with the reference BTs suggest re-calibration of the COCTS infrared channels is needed. Given the nearly negative linear correlation observed between the BT difference and the reference BT, we employ robust linear regression to fit this relationship which is expressed in the following (3):

$$BT_{COCTS} - BT_{reference} = a \times BT_{reference} + b \quad (3)$$

where a and b represent the slope and offset of the linear relationship, respectively. BT_{COCTS} signifies the BTs observed by COCTS, and $BT_{reference}$ represents the reference BTs. Subsequently, we acquire corrected COCTS BTs, $BT_{COCTS'}$, by replacing $BT_{reference}$ with $BT_{COCTS'}$ in (3). This process is defined as follows:

$$BT_{COCTS'} = \frac{BT_{COCTS} - b}{a+1} \quad (4)$$

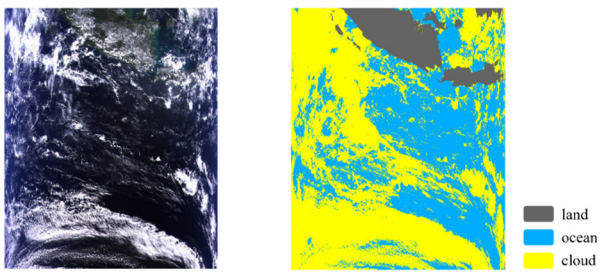
Our methodology involves randomly dividing the dataset, allocating two-thirds for regression analysis and one-third for model validation. To address variable BT difference trends, regression was applied distinctly for COCTS 11 μm BTs at ≥ 270 K and < 270 K. We also conducted separate regressions for each detector to reduce striping noise due to detector inconsistence, resulting in 32 sets of correction coefficients for COCTS BT. Table I provides the 32 sets of correction coefficients for COCTS BT. Subsequent sections display these corrected results.

Post-correction, Fig. 6 shows that the previous biases in COCTS 11 and 12 μm channels versus reference BTs are resolved, with mean differences near 0 K and RSDs reduced to 0.15 K and 0.16 K respectively. The BT dependencies that originally contributed to the error have been addressed, as evidenced by Fig. 6. To assess the post-correction stability of COCTS calibration, Fig. 7 shows the daily BT difference fluctuations for the corrected BTs in comparison to the reference, with a color scheme indicating reference BT values. The daily differences are now more consistent, with approximately zero bias (indicated by the black line), and the negative correlation with reference BTs has been eliminated, reflecting a reduction in overall uncertainty. Fig. 8 presents the double difference variance relative to the reference BT for each detector. Compared to Fig. 5, the improved overlapping extent among the lines for all eight detectors indicates successful mitigation of calibration differences, demonstrating the effectiveness of the separate regressions for each detector. To visually assess the de-striping improvements, BT images before

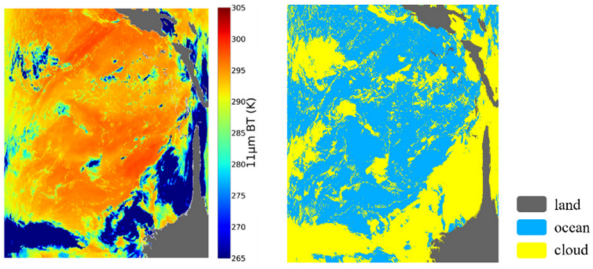
TABLE I
THE 32 SETS OF CORRECTION COEFFICIENTS FOR COCTS BT.

(a/b)	BT<270K		BT>=270K	
	11μm channel	12μm channel	11μm channel	12μm channel
Detector 1	-0.016/5.514	-0.005/2.699	-0.045/13.449	-0.023/7.439
Detector 2	-0.016/5.482	-0.004/2.614	-0.046/13.652	-0.023/7.634
Detector 3	-0.015/5.436	-0.008/3.742	-0.046/13.580	-0.027/8.593
Detector 4	-0.016/5.520	-0.009/3.774	-0.046/13.600	-0.027/8.677
Detector 5	-0.019/6.559	-0.004/2.484	-0.048/14.195	-0.023/7.522
Detector 6	-0.019/6.401	-0.006/3.218	-0.048/14.160	-0.024/7.954
Detector 7	-0.018/6.139	-0.009/3.961	-0.047/14.026	-0.029/9.223
Detector 8	-0.017/5.971	-0.007/3.384	-0.047/13.889	-0.026/8.284

and after correction are compared in Figs. 9 and 10. In these figures, (a) represents the original BT and (b) the corrected BT, with right panels showing BT fluctuation profiles along the satellite track as indicated by the vertical lines in the left panels. For the 11 μm channel, Fig. 9 reflects negligible striping noise, consistent with the statistical results. Conversely, Fig. 10 for the 12 μm channel illustrates clear striping noise at approximately 295 K, visible in the scene image and in the BT fluctuation profile (a). After correction, the BT images and corresponding fluctuation profiles (blue curves) in (b) indicate a significant reduction in striping noise, thereby validating the effectiveness of employing distinct correction coefficients for the eight detectors.



(a) day



(b) night

Fig. 11. Two samples of the cloud detection dataset: (a) day; (b) night. The left images include RGB image and 11μm BT image. The right images are labels.

III. CLOUD DETECTION USING A PHYSICALLY BASED DEEP LEARNING ALGORITHM

To detect clouds from the HY-1D COCTS observational data, we employ a deep learning approach due to its robust and accurate performance. However, deep learning methodologies require substantial amounts of labelled cloud data for training. In our previous study, we went to a lot of effort to produce a large cloud detection dataset on HY-1C COCTS. Considering the spectral characteristics of HY-1C and HY-1D COCTS are relatively similar, we directly use the model trained on HY-1C COCTS dataset to perform the cloud detection of HY-1D COCTS. It is also verified that the model has good generalization ability.

A. Creation of a Large-Scale COCTS Cloud Detection Dataset

We have constructed a comprehensive HY-1C COCTS cloud detection dataset covering the global ocean region from 2020 to 2021, leveraging both the Bayesian algorithm and manual labeling for cloud/clear masks using 11 and 12μm BT data. Cloud detection within the Bayesian theory is considered a probability issue, and pixels are classified as clear sky pixels when its cloud probability is larger than 0.9; any misclassifications are manually corrected. The dataset is consisted of 447,854 patches (256 pixels × 256 pixels each) from 7238 scenes, optimizing the balance between computational load and spatial feature coverage. Fig. 11 displays the example labels from day and nighttime datasets for cloud detection, including an RGB image (left in subfigure a), and 11 μm BT (left in subfigure b). The corresponding cloud, ocean, and land pixels are marked in yellow, blue, and grey, respectively.

B. Designing a Physically based U-Net Model

We employ a U-Net deep learning model for cloud detection in HY-1D COCTS data, leveraging encoders for feature compression and decoders for result generation [16]. The model processes COCTS Level 1B radiation data using 14 features reflecting cloud physical properties. We incorporate reflectance from the 443, 490, 565, 670, and 865 nm bands, and the 11 and

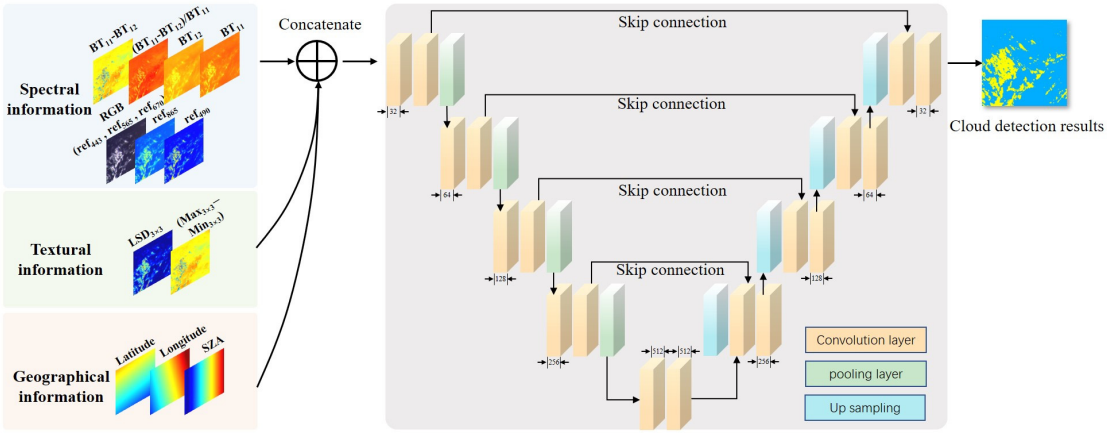
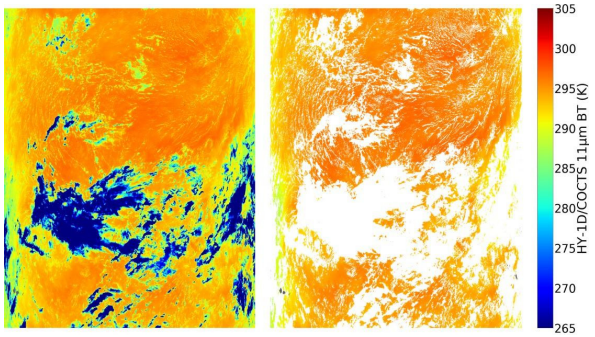
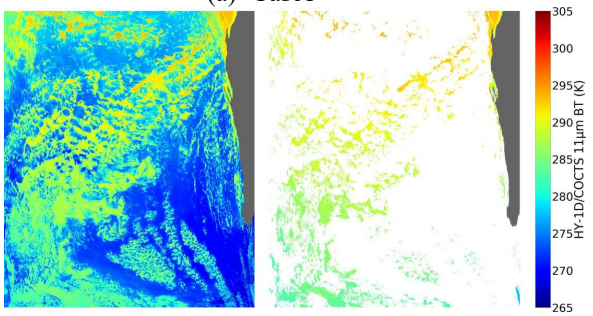


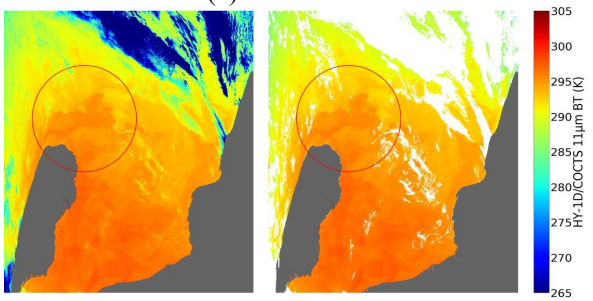
Fig. 12. The physically based U-Net architecture for cloud detection.



(a) Case1



(b) Case2



(c) Case3

Fig. 13. The demo region of COCTS cloud detection results, (a) – (c) are three cases. The left and right panels are COCTS L1B 11 μm BT before and after cloud detection, respectively, where grey region represents the land.

12 μm channel BTs—key for identifying thick clouds due to their high reflectance and low temperature. As the BT difference between 11 and 12 μm channels is strongly correlated with water vapor, we also incorporated (11 μm BT − 12 μm BT) and (11 μm BT − 12 μm BT)/ 11 μm BT as model inputs. Textural features like LSD and the BT variation within 3 × 3 boxes assist in detecting cloud edges, while geolocation and SZA data are also supplemented. Visible bands are excluded at night, necessitating separate day and night U-Net models. The model generates a probability score for cloud presence, with scores above 0.5 indicating clear conditions. These scores are further applied to define quality levels discussed in Section V.

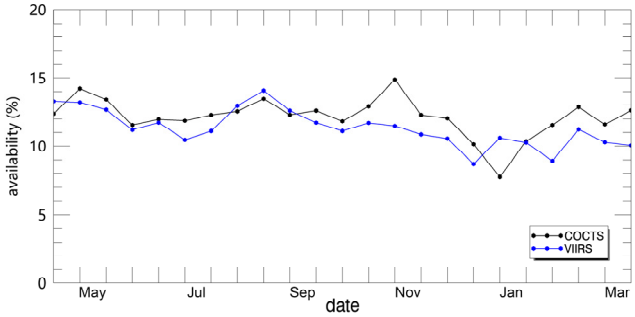
Our U-Net model, designed for cloud detection, utilizes an encoder-decoder structure with a focus on feature extraction and image reconstruction. The encoder, essential for extracting features, consists of four convolutional blocks—each with two 3×3 convolutions, BatchNorm, ReLU activation, and 2×2 down-sampling. The decoder mirrors the encoder and incorporates skip connections for image reconstruction. This architecture is visualized in Fig. 12, showing data downsampled by the encoder and upsampled by the decoder. Softmax is used as the output activation function, determining the clear or cloudy status per pixel. To counteract the imbalanced dataset issue that could bias the classifier, we apply Focal loss to focus training on challenging samples, defined by the following (5) [40]:

$$L_{Focal} = -(1 - P_t)^{\gamma} log P_t \quad (5)$$

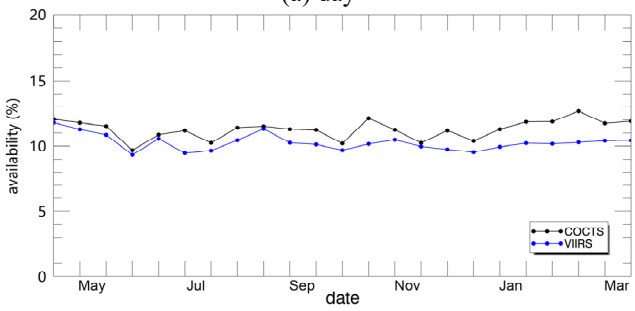
where P_t denotes is the predicted probability, and γ adjusts the rate of down-weighting easy samples. The Adam optimizer is adopted and the learning rate was 0.0001.

C. Cloud Detection Results from COCTS

Our cloud detection model, trained on the HY-1C COCTS dataset, achieved 96% accuracy against ground truth. We applied the model to HY-1D COCTS to evaluate transferability, given that the SRFs of HY-1C COCTS and HY-1D COCTS bear a resemblance. Fig. 13 shows the successful application of



(a) day



(b) night

Fig. 14. The SST availability comparison conducted between COCTS and VIIRS: (a) day; (b) night.

the model across various cloud types without noticeable instances of false negative or positive errors. Notably, the model accurately detects clouds in oceanic frontal areas. The red-circled area in Fig. 13(c) indicates cloud detection over oceanic frontal areas. Traditional cloud detection algorithms often overestimate clouds at these fronts due to the significant gradient variation that resembles the boundary fluctuations between clouds and clear ocean. However, it can be observed that the oceanic gradient was adequately preserved post our cloud detection, indicating its superior performance in challenging frontal regions.

The accuracy of retrieved SST after cloud detection can also reflect the situation of cloud missing detection to a certain extent, because the SST value obtained in the missed cloud region must be low. Section VI will validate SST results to assess missed detections. Additionally, SST availability reflects the extent of cloud over-detection. By comparing the clear-sky pixel proportion after cloud detection between COCTS and VIIRS, as shown in Fig. 14, we find similar usability rates, with COCTS slightly outperforming VIIRS at night, indicating no significant cloud over-detection with our model. The results above indicate that the cloud detection model has good transferability and performs excellently on HY-1D COCTS.

IV. Optimal Estimation for COCTS SST_{SKIN} Retrieval

A. Introduction to Optimal Estimation (OE)

Our study applies atmospheric radiative transfer simulations and the Optimal Estimation (OE) algorithm to retrieve SST_{skin} from HY-1D COCTS data, an approach initially explored by [19]. The fundamental concept of the OE retrieval algorithm is to ascertain the deviation between a prior background state

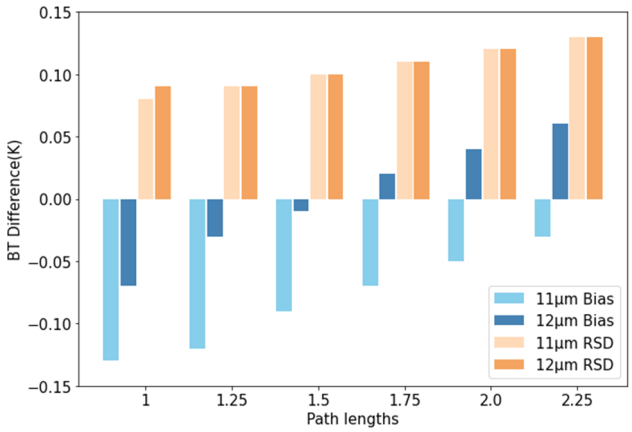


Fig. 15. The comparison results of MODTRAN minus LBLRTM simulated BTs against different path lengths.

values and actual state values. This is achieved through the combination of prior information on atmospheric and oceanic states, satellite-observed BTs, BTs simulated from atmospheric radiative transfer models, and their corresponding uncertainties. This method refines the estimation of true state values. The OE retrieval algorithm, as described by [19], uses $\hat{\boldsymbol{x}}$ for the retrieved SST_{skin} and Total Column Water Vapor (TCWV), $\boldsymbol{x_a}$ for the background state, $\boldsymbol{y}^o$ for observed BTs, $\boldsymbol{F}(\boldsymbol{x}_a)$ for simulated BTs via MODTRAN, $\boldsymbol{K}$ for the tangent linear matrix, $\boldsymbol{S}_a$ for background state uncertainty, and $\boldsymbol{S}_\varepsilon$ for the combined observational and simulation uncertainties. The OE equation is provided as (6):

$$\hat{\boldsymbol{x}} = \boldsymbol{x_a} + (\boldsymbol{K}^T\boldsymbol{S}_\varepsilon^{-1}\boldsymbol{K} + \boldsymbol{S}_a^{-1})^{-1}\boldsymbol{K}^T\boldsymbol{S}_\varepsilon^{-1}\left(\boldsymbol{y}^o - \boldsymbol{F}(\boldsymbol{x}_a)\right) \quad (6)$$

B. Determining Uncertainty for OE SST_{skin} Retrieval

To achieve a genuinely "optimal" estimation, precisely quantifying the uncertainties of both observations and priors is crucial. The uncertainty of the observed BT is determined based on the results of the inter-calibration between COCTS and VIIRS, with uncertainties for the 11 and 12 μm channels at 0.15 K and 0.16 K, respectively. To accurately ascertain the uncertainty of the simulated BT, we compare the BT differences between MODTRAN and LBLRTM simulation. LBLRTM, a line-by-line atmospheric radiative transfer model, is renowned for its high simulation accuracy and therefore serves as a reference to assess the uncertainty of MODTRAN simulations [41]. We utilize 5000 representative atmospheric profiles provided by ECMWF as input for the MODTRAN and LBLRTM simulations with six different path lengths (that equal to sec(SZA)) including 1, 1.25, 1.5, 1.75, 2.0, and 2.25. Fig. 15 shows the comparison results of MODTRAN minus LBLRTM simulated BTs including bias and RSD for 11 and 12 μm, showing minimal SZA impact. For both channels, the overall mean biases are close to zero, with the same RSD of 0.12 K. Thus, the uncertainties of the 11 and 12 μm simulated BT in the COCTS OE SST_{skin} retrieval are set to 0.12 K.

Precise estimation of background SST_{skin} uncertainty is essential for the OE retrieval process. Underestimation can lead to excessive reliance on the background field, obscuring detailed SST variations, while overestimation may introduce errors and amplify sensor noise effects. The background SST_{skin}

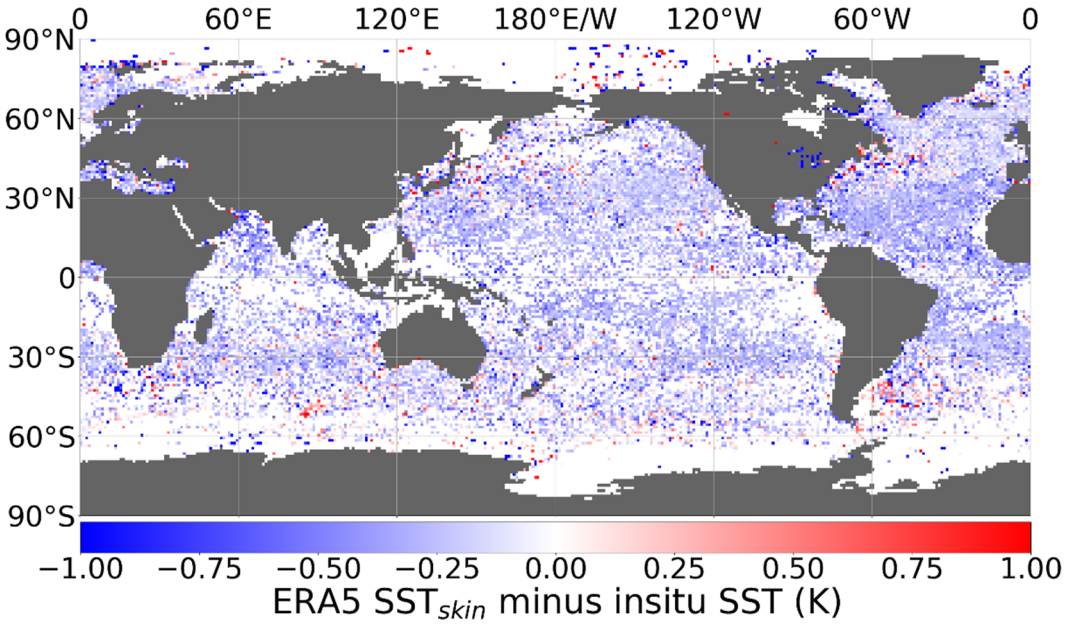
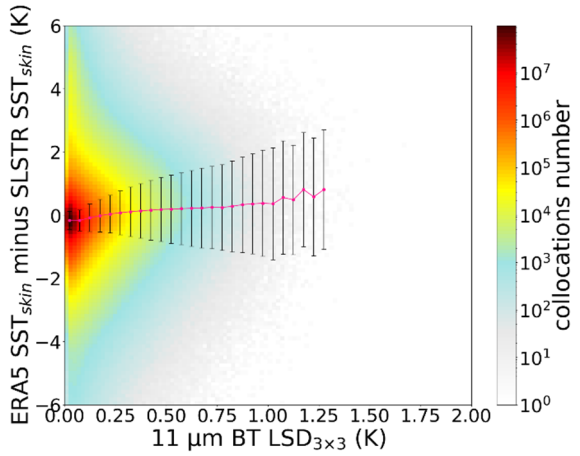


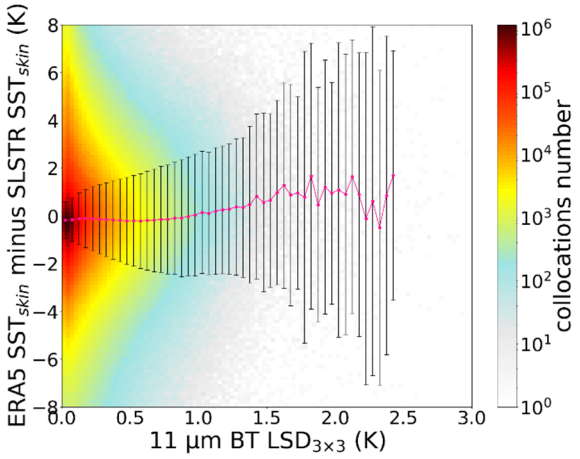
Fig. 16. The geophysical distribution of ERA5 SST_{skin} minus iQuam in-situ SST difference

from the ECMWF ERA5 dataset, with a 25 km resolution, is coarser than the COCTS 1 km observational resolution. Consequently, it is essential to evaluate the accuracy of interpolating the background SST_{skin} field to the COCTS spatial resolution under varying SST spatial patterns. This assessment is vital for dynamically setting the uncertainty associated with the background SST_{skin} field. The ERA5 background uncertainty comprises systematic uncertainty, e^s_{SSTa}, assessed by comparing original ERA5 SST_{skin} (~25 km) with in-situ SST, and spatial resolution uncertainty, e^d_{SSTa}, determined by comparing interpolated ERA5 SST_{skin} with a 1 km SST_{skin} reference. Systematic uncertainty, e^s_{SSTa}, is obtained by comparing ERA5 with iQuam in-situ SST. Fig. 16 illustrates the geophysical distribution of SST_{skin} differences between ERA5 and iQuam in-situ measurements, with larger discrepancies observed near land. Accordingly, the data is analyzed in two spatial contexts: within 25 km of land and beyond. Near land, the ERA5 SST_{skin} shows an average bias of -0.58°C with an RSD of 1.02°C, while further from land, the bias is -0.22°C with an RSD of 0.26°C. Consequently, the systematic uncertainty for ERA5 SST_{skin} is set at 1.02°C for coastal areas and 0.26°C for open ocean regions, based on these RSD values.

We use SLSTR SST_{skin} to validate ERA5 SST_{skin} interpolated to 1 km, employing bilinear interpolation based on the SLSTR orbit position and scan time to align ERA5 with satellite spatial resolution. The LSD within a 3×3 box ($LSD_{3\times3}$) of the SLSTR 11 μm BT serves as an indicator of the spatial variation scale. We excluded points near cloud edges to prevent SLSTR cloud contamination from affecting results. The comparison, split into near-land and off-land scenarios, is illustrated in Fig. 17 (a) and (b), showing SLSTR and ERA5 SST_{skin} differences against SLSTR 11 μm BT $LSD_{3\times3}$. As the LSD increases, the bias (shown as the purple curve) remains relatively stable while the RSD (shown as the black vertical line) gradually increases. Following this comparison, we established a lookup table for the background field SST_{skin} uncertainty, thereby determining

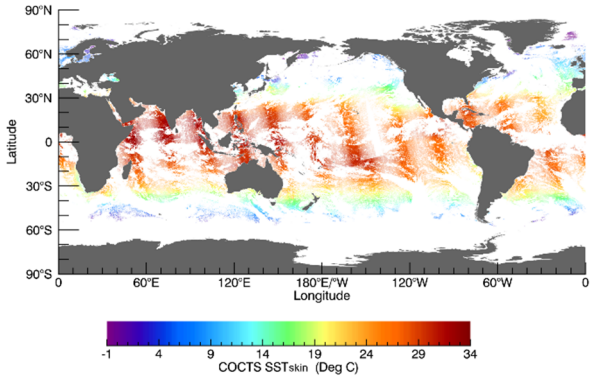


(a) off-land

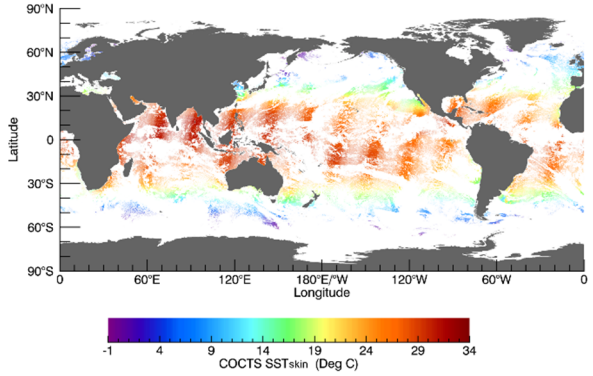


(b) near-land

Fig. 17. The relationship between the difference of SLSTR SST_{skin} minus ERA5 SST and the LSD of the SLSTR 11 μm BT.



(a) day



(b) night

Fig.18. The global COCTS OE SST_{skin} on 1, May, 2021: (a) day, (b) night.

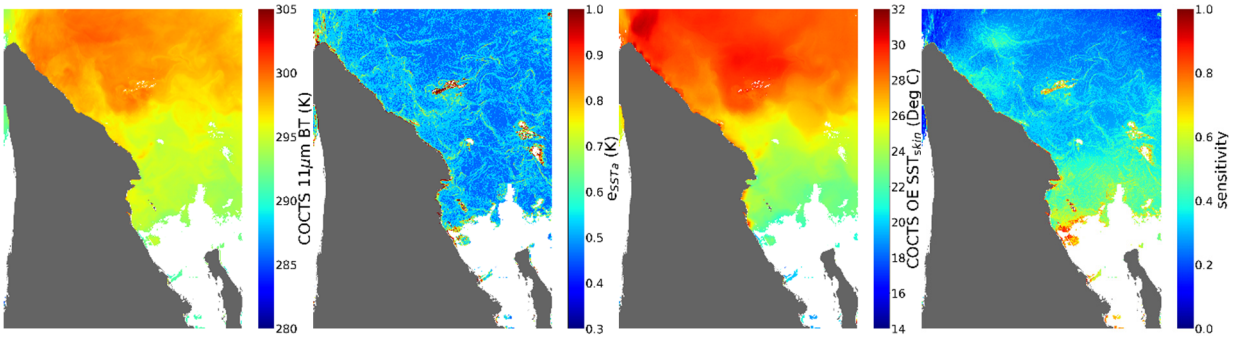
the dynamically changing uncertainty of the background SST_{skin} based on different LSD values of the 11 μm BT. Given that the SLSTR SST_{skin} itself carries errors, we calculated spatial uncertainty, e_{SSTa}^{d}, using (7), where $e_{ERA5-SLSTR}$ is the RSD across BT LSD intervals. SLSTR uncertainty, e_{SLSTR} is set at 0.2K based on the validation of SLSTR SST_{skin} [42]. By summing e_{SSTa}^{s} and e_{SSTa}^{d}, we derived total uncertainty, e_{SSTa}.

$$e_{SSTa}^{d} = \sqrt{e_{ERA5-SLSTR}^{2} - e_{SLSTR}^{2}} \quad (7)$$

C. Results from the COCTS OE SST_{skin} Retrieval

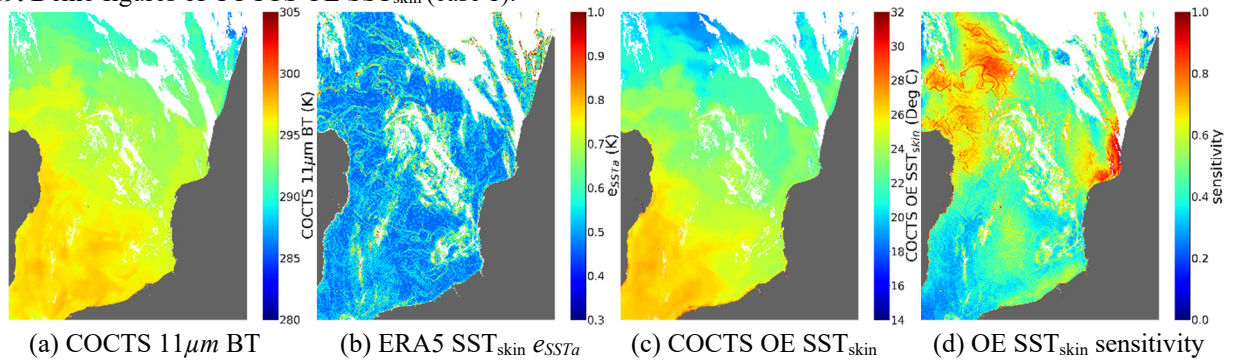
Fig. 18 shows the global COCTS SST on May 1, 2021, with day (a) and night (b) maps. Two frontal ocean areas are exemplified in Figs. 19 and 20, with (a) and (b) showing the 11 μm BT and background SST_{skin} uncertainty, and (c) and (d) representing the COCTS OE SST_{skin} and its sensitivity. From these two representative frontal regions, it can be observed that within the frontal zone, where the SST manifests significant spatial variations, the ERA5 SST—due to its lower spatial resolution—is assigned a larger uncertainty. In these frontal regions, the OE SST retrieval relies more heavily on satellite observations, resulting in pronounced spatial variations in the retrieved SST.

Apart from commonly used metrics such as bias and RSD, we also evaluate the retrieved SST using an additional metric—SST sensitivity, which represents the rate of change of the retrieved SST_{skin} relative to the true SST [19]. Ideally, this value should be one, but it typically falls short in real-world applications. Higher sensitivity indicates that the SST_{skin} closely reflects real-time diurnal and frontal changes. Conversely, low sensitivity points to a stronger dependence on the background field. Analysis of subfigure (d) from the presented cases shows that COCTS OE SST_{skin} has greater sensitivity in frontal zones and less in uniform



(a) COCTS 11μm BT (b) ERA5 SST_{skin} e_{SSTa} (c) COCTS OE SST_{skin} (d) OE SST_{skin} sensitivity

Fig.19. Demo figures of COCTS OE SST_{skin} (case 1).



(a) COCTS 11μm BT (b) ERA5 SST_{skin} e_{SSTa} (c) COCTS OE SST_{skin} (d) OE SST_{skin} sensitivity

Fig.20. Demo figures of COCTS OE SST_{skin} (case 2).

areas, that is related to the dynamic setting of the background field SST_{skin} uncertainty. This variation is expected, as we expect the retrieval algorithm to reflect the rich variations in SST when they occur, such as in frontal regions. The next sub-section will focus on the atmospheric correction smoothing algorithm, which significantly improves SST sensitivity.

D. Atmospheric Correction Smoothing

The key notion, initially proposed in [19], is to conceive of atmospheric variations as comparatively smoother in comparison to SST which can exhibit more rapid spatial variations. The main premise of the atmospheric correction smoothing algorithm is that the water vapor component can be smoothed during atmospheric correction due to the slower atmospheric variations.

Considering that the spatial scale of atmospheric variations typically ranges between 15 and 50 km, we adopt a 5×5 grid of COCTS for averaging in the smoothing process. The matrices employed in the OE SST_{skin} retrieval algorithm need to be reconfigured accordingly. The adjusted matrices are illustrated in the subsequent equations. Consequently, $\boldsymbol{x}_a$ becomes the background field SST and the averaged TCWV, shown as the (8). The observation matrix is also expanded to include the observed BT of the central pixel and the averaged BT of the surrounding clear-sky pixels (the number is n), shown as (9). In this context, the water vapor term involved in the average is common across all the surrounding clear points in the 5×5 area excluding the center pixel, and the BT term involved in the average is all clear sky points in the 5×5 area excluding the center pixel. The uncertainty in the averaged BT therefore equals that of the observed BT divided by the root of n, shown as (10)–(12). When there are a greater number of clear-sky pixels among the surrounding pixels, the uncertainty of the averaged BT term will be significantly reduced, implying that during the retrieval, the averaged BT of the surrounding pixels will be more heavily relied upon relative to the BT of the central pixel. The prior state covariance $\boldsymbol{S}_a$ is not changed shown as (13). The **K** matrix must be constructed to mirror this relationship: all observations are sensitive to water vapor, but only the observations corresponding to a given pixel are sensitive to the SST_{skin} of that pixel shown as (14).

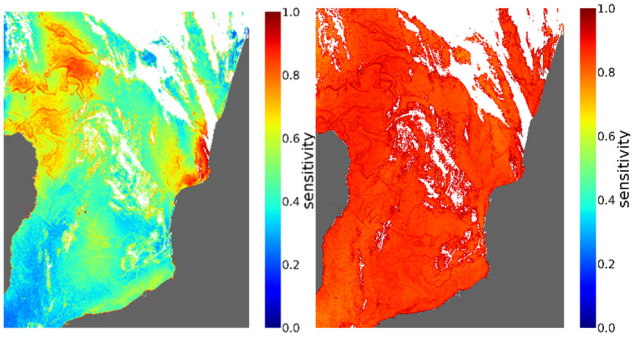
$$\boldsymbol{x}_a = \left[\frac{SST}{TCWV}\right] \tag{8}$$

$$\boldsymbol{y} = \begin{bmatrix} y_{11} \\ y_{12} \\ \overline{y_{11}} \\ \overline{y_{12}} \end{bmatrix} \tag{9}$$

$$\mathbf{S}_\varepsilon = \boldsymbol{R}_m + \boldsymbol{R}_e \tag{10}$$

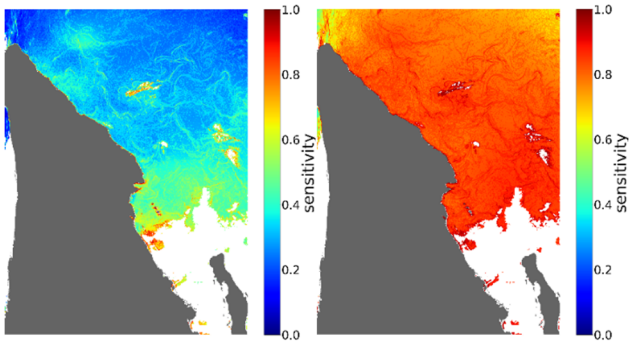
$$\boldsymbol{R}_m = \begin{bmatrix} e_{RTM11}^2 & 0 & 0 & 0 \\ 0 & e_{RTM12}^2 & 0 & 0 \\ 0 & 0 & e_{RTM11}^2 & 0 \\ 0 & 0 & 0 & e_{RTM12}^2 \end{bmatrix} \tag{11}$$

$$\boldsymbol{R}_e = \begin{bmatrix} e_{ob11}^2 & 0 & 0 & 0 \\ 0 & e_{ob12}^2 & 0 & 0 \\ 0 & 0 & \frac{e_{ob11}^2}{n} & 0 \\ 0 & 0 & 0 & \frac{e_{ob12}^2}{n} \end{bmatrix} \tag{12}$$



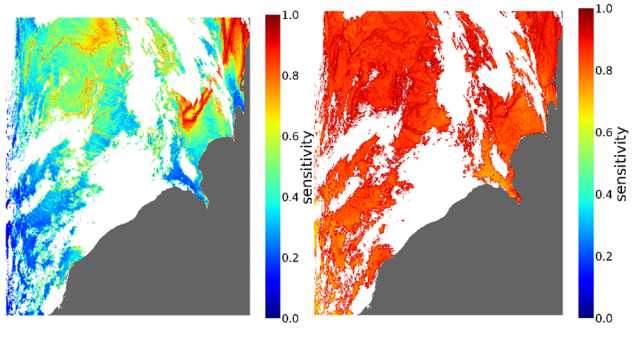
(a) (b)

Fig. 21. Case 1: The comparison demos of sensitivity for OE SST_{skin}: (a) without the application of atmospheric correction smoothing; (b) with the atmospheric correction smoothing applied.



(a) (b)

Fig. 22. Case 2: The comparison demos of sensitivity for OE SST_{skin}: (a) without the application of atmospheric correction smoothing; (b) with the atmospheric correction smoothing applied.



(a) (b)

Fig. 23. Case 3: The comparison demos of sensitivity for OE SST_{skin}: (a) without the application of atmospheric correction smoothing; (b) with the atmospheric correction smoothing applied.

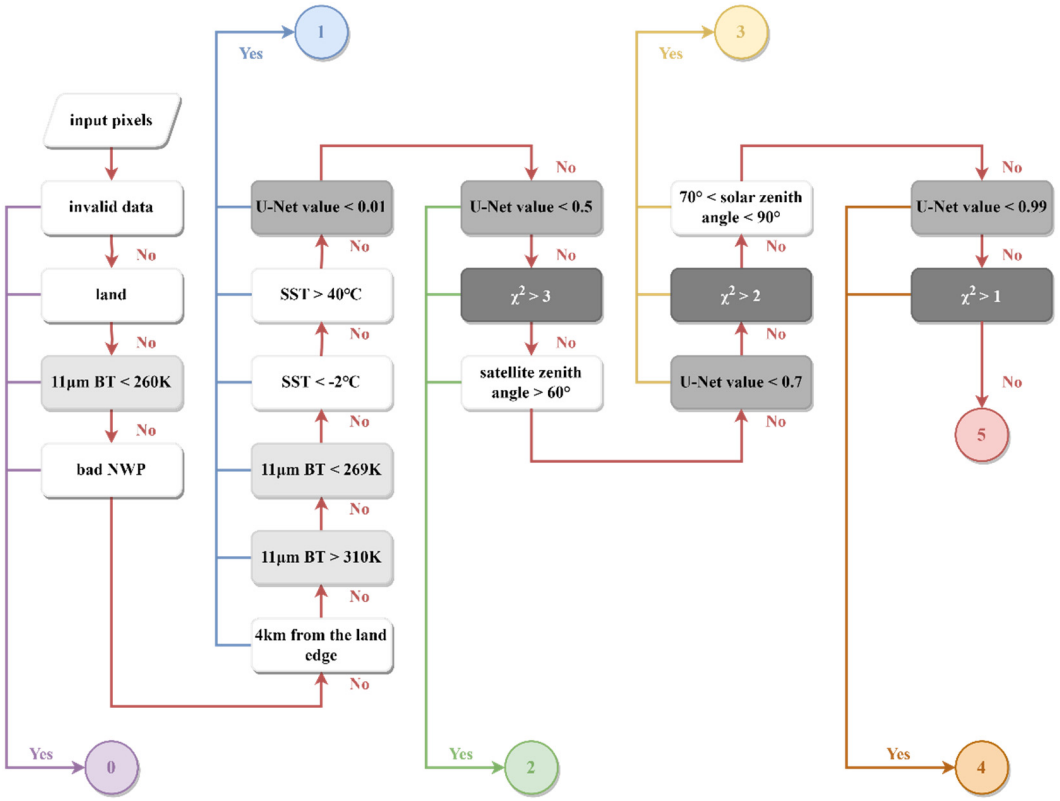


Fig. 24. The processing flow and criteria for assigning pixel quality level flags to COCTS SST_{skin}.

$$\boldsymbol{S}_a = \begin{bmatrix} e_{SSTa}^2 & 0 \\ 0 & (0.5 \times \left(0.1 + \frac{75 - TCWV}{150}\right))^2 \end{bmatrix} \tag{13}$$

$$\boldsymbol{K} = \frac{\partial y}{\partial x_a} = \begin{bmatrix} \frac{\partial y_{11}}{\partial SST} & \frac{\partial y_{11}}{\partial \overline{TCWV}} \\ \frac{\partial y_{12}}{\partial SST} & \frac{\partial y_{12}}{\partial \overline{TCWV}} \\ 0 & \frac{\partial \overline{y_{11}}}{\partial \overline{TCWV}} \\ 0 & \frac{\partial \overline{y_{12}}}{\partial \overline{TCWV}} \end{bmatrix} \tag{14}$$

The atmospheric correction smoothing method significantly enhances the sensitivity of COCTS OE SST_{skin}. Fig. 21–23 illustrate this with three comparative demonstrations. Sensitivity improvements are evident, with the first case rising from 0.59 to 0.85, the second from 0.42 to 0.78, and the third from 0.56 to 0.86. These results confirm the efficacy of this method, particularly in areas of previously low sensitivity. Section VI will further examine this by comparing sensitivity and accuracy against in-situ SST measurements.

V. Quality Indication

SST_{skin} data quality is assigned via a quality flag ranging from zero (lowest quality) to five (highest quality), as outlined in Fig. 24, with comprehensive quality control in two key areas. Firstly, "clear-sky" confidence stems from the deep learning cloud detection network. Outputs range from zero to one, with values above 0.5 typically indicating clear skies. However, the errors of algorithm primarily occur in cloud edge detection. Therefore, we statistically analyze the cloud edges and classify regions with values greater than 0.5 but less than 0.7 as clear-sky edges, indicating lower confidence in these areas. Pixels with values greater than 0.7 but less than 0.99 are categorized as probable clear-sky points, while those with values exceeding 0.99 are classified as confidently clear-sky pixels. Secondly, SST_{skin} confidence levels are determined using the chi-square value from the OE retrieval algorithm. In OE algorithm, the consistency between observed and simulated BTs provides a reliable metric for assessing the confidence index of the OE SST_{skin}. Rodgers et al. [43] proposed a method for calculating the confidence index of the OE SST_{skin} with the formula provided below:

$$\hat{\chi}^2 = \left(\boldsymbol{K}(\hat{\boldsymbol{x}} - \boldsymbol{x}_a) - \left(\boldsymbol{y}^o - \boldsymbol{F}(\boldsymbol{x}_a)\right)\right)^T (\boldsymbol{S}_\varepsilon(\boldsymbol{K}\boldsymbol{S}_a\boldsymbol{K}^T + \boldsymbol{S}_\varepsilon)^{-1}\boldsymbol{S}_\varepsilon)^{-1}\left(\boldsymbol{K}(\hat{\boldsymbol{x}} - \boldsymbol{x}_a) - \left(\boldsymbol{y}^o - \boldsymbol{F}(\boldsymbol{x}_a)\right)\right) \tag{15}$$

For all clear-sky pixels, we set $0 < \hat{\chi}^2 < 1$ as the quality level of 5. For $1 < \hat{\chi}^2 < 2$ cases, we assign a quality level of 4. For $2 < \hat{\chi}^2 < 3$ situations, a quality level of 3 is assigned.

VI. Validation of COCTS OE SST_{skin}

The COCTS SST_{skin} data are validated against iQuam in-situ measurements, focusing on data from the 1st and 15th of each month between April 2021 and March 2022 due to the high computational time of global MODTRAN simulations. For accuracy analysis, only SST_{skin} data with quality levels four and five are considered. The matching process between COCTS SST_{skin} and in-situ data uses a spatial window of 0.01° by 0.01°

TABLE II
THE STATISTICS RESULTS OF COCTS OE SST_{SKIN} MINUS IN-SITU SST (BLUE FONT: WITHOUT ATMOSPHERIC CORRECTION SMOOTHING; RED FONT: WITH ATMOSPHERIC CORRECTION SMOOTHING APPLIED).

	Bias (°C)	SD (°C)	Median (°C)	RSD (°C)	Sensitivity	Number
Day	-0.15/-0.18	0.35/0.33	-0.15/-0.19	0.29/0.26	0.56/0.87	5270
Night	-0.24/-0.22	0.35/0.33	-0.21/-0.21	0.30/0.28	0.56/0.87	5416
All	-0.20/-0.20	0.36/0.33	-0.18/-0.20	0.29/0.27	0.56/0.87	10686

and a temporal window of one hour, resulting in 10,686 matchups. A histogram (Fig. 25) shows the COCTS SST_{skin} and in-situ SST differences, with an average bias of -0.20 °C, a SD of 0.33 °C and a RSD of 0.27°C. Table II shows the results of COCTS compared with in-situ SST, distinguishing results from algorithms with and without atmospheric correction smoothing; the former in red and the latter in blue. With smoothing, the daytime bias is -0.18 °C with the SD of 0.33 °C and RSD of 0.26 °C, and the nighttime bias is -0.22 °C (SD 0.33 °C, RSD 0.28 °C). The comparison results indicate that after adopting the atmospheric correction smoothing algorithm, the accuracy of the retrieved SST_{skin} has slightly improved, and the sensitivity has been significantly enhanced, increasing from 0.56 to 0.87.

The primary reason for the average negative bias between COCTS and in-situ SST is the distinction in the measured quantity, while the buoy measures the subsurface water temperature, COCTS observes the skin temperature, which is generally cooler. According to [44], under typical conditions with a wind speed of about 6 m/s, the skin temperature of the ocean surface is estimated to be around 0.17°C lower than the bulk temperature that is measured by buoys. Therefore, the difference in observing depths between COCTS and buoys is a major contributor to the negative bias. This validation result also suggests minimal cloud contamination in COCTS cloud detection. Fig. 26 is the spatial distribution of difference between COCTS and in-situ SST. The left and right panels correspond to daytime and nighttime, respectively. The central graph depicts the latitude-dependent deviation trend, with the line representing the average deviation and short bars indicating SDs within latitude intervals (orange: daytime, blue: nighttime), corresponding to the bottom axis. The gray rectangles represent the collocation numbers, corresponding to the top axis. Overall, deviations exhibit no pronounced latitude dependence. Most matching points fall within a deviation range of ±0.5°C. A few points in high-latitude regions exhibit deviations greater than 1°C. Fig. 27 presents the daily comparison results between COCTS OE SST_{skin} and in-situ SST. From a time series perspective, the deviation between the two remains stable, signifying the reliability of COCTS SST observations. The above validation results all indicate that, after inter-calibration, cloud detection, and OE SST_{skin} retrieval, the accuracy of the derived COCTS SST_{skin} is relatively high and has reached the

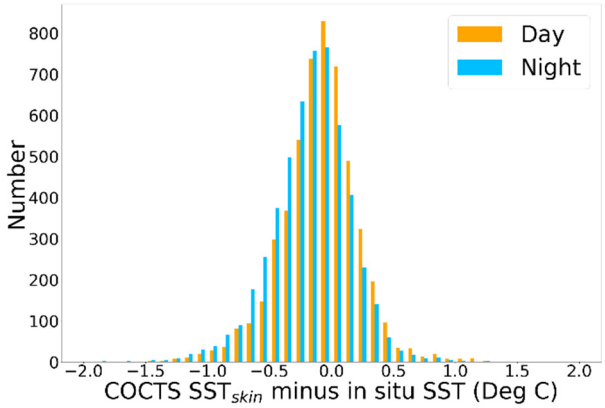


Fig. 25. The histogram of the deviation between the COCTS OE SST_{skin} and in-situ SST.

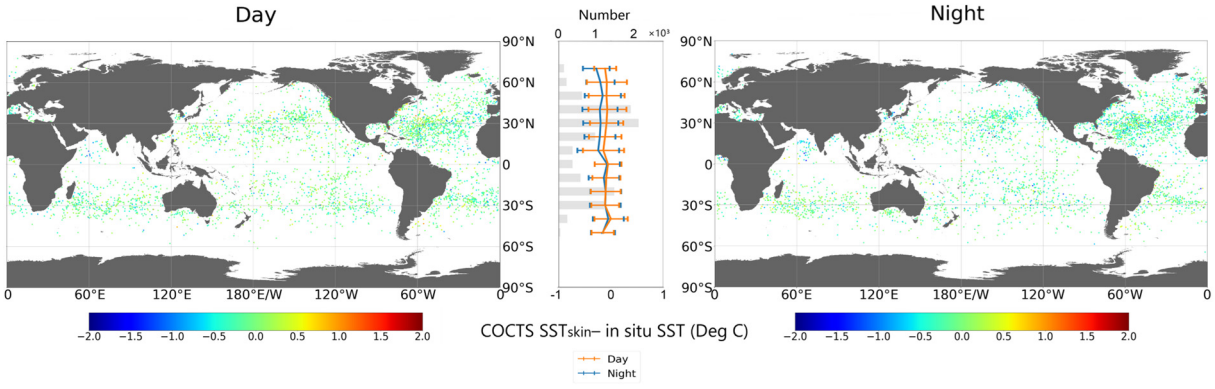


Fig. 26. The spatial distribution of difference between COCTS and in-situ SST.

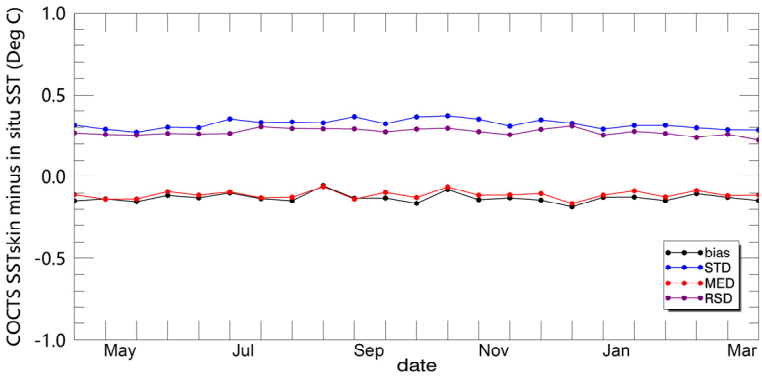


Fig. 27. The daily comparison results between COCTS OE SST_{skin} and in-situ SST.

level required for climatological research and similar applications.

VII. Conclusion

In summary, our research represents significant progress of SST_{skin} retrieval algorithm using HY-1D COCTS data. This is achieved by inter-calibrating the COCTS infrared channels with VIIRS via the double-differencing method, incorporating a physically based deep learning for cloud detection, and formulating a retrieval method for low noise and high sensitivity, yielding robust performance across diverse scenarios.

Our investigation underscores the substantial potential of OE for SST_{skin} retrieval from COCTS data. The new approach we introduce to dynamically estimate the uncertainty of background SST_{skin} has shown promise in improving data accuracy and reliability. The implementation of atmospheric correction smoothing, which treats atmospheric variations as being slower relative to SST, enable retrievals with better SST_{skin} sensitivity. The validation of the COCTS OE SST_{skin} against in-situ data show that the uncertainty of higher-quality HY-1D COCTS SST_{skin} is as low as 0.27 K (RSD), with higher sensitivity of 0.87. These findings affirm that the successful implementation of these methodologies significantly enhances the accuracy and reliability of SST_{skin} data from HY-1D COCTS. It is also of great significance for the subsequent establishment of climate data record from HY-1 series satellites. This advancement provides substantial benefits to expand the global high precision SST_{skin} dataset.

Future studies may build upon our findings by exploring alternative calibration techniques and further integrating machine learning algorithms. By persistently refining our methodologies, we can continue to strive for more accurate and reliable SST measurements from satellite data. Such advancements will provide invaluable support to critical research areas including climate science and oceanographic studies.

Acknowledgment

Haiyang-1B (HY-1B) Chinese Ocean Color and Temperature Scanner (COCTS) data were provided by the National Satellite Ocean Application Service (NSOAS). The ERA5 data were derived from European Centre for Medium-Range Weather Forecasts (ECMWF). The in-situ sea surface temperature (SST) data were provided by the iQuam System developed by the National Oceanic and Atmospheric Administration (NOAA)/National Environmental Satellite, Data, and Information Service (NESDIS) Satellite Application and Research (STAR).

References

[1] S. Martin, *An introduction to ocean remote sensing*. Cambridge University Press, 2014.

[2] P. J. Minnett *et al.*, "Half a century of satellite remote sensing of sea-surface temperature", *Remote Sensing of Environment*, vol. 233, p. 111366, Nov. 2019, doi: 10.1016/j.rse.2019.111366.

[3] G. Chander, T. J. Hewison, N. Fox, X. Wu, X. Xiong, and W. J. Blackwell, "Overview of Intercalibration of Satellite Instruments", *IEEE Trans. Geosci. Remote Sensing*, vol. 51, no. 3, pp. 1056–1080, Mar. 2013, doi: 10.1109/TGRS.2012.2228654.

[4] T. J. Hewison et al., "GSICS Inter-Calibration of Infrared Channels of Geostationary Imagers Using Metop/IASI," *IEEE Trans. Geosci. Remote Sensing*, vol. 51, no. 3, pp. 1160–1170, Mar. 2013, doi: 10.1109/TGRS.2013.2238544.

[5] Z. Gu, L. Chen, H. Dai, L. Tian, X. Hu, and P. Zhang, "The Uncertainty of SNO Cross-Calibration for Satellite Infrared Channels," *Remote Sensing*, vol. 15, no. 13, p. 3313, Jun. 2023, doi: 10.3390/rs15133313.

[6] D. Di, Y. Liu, J. Li, R. Zhou, Z. Li, and X. Gong, "Inter-Calibration of Geostationary Imager Infrared Bands Using a Hyperspectral Sounder on the Same Platform, " *Geophys. Res. Lett.*, vol. 50, no. 3, p. e2022GL101628, Feb. 2023, doi: 10.1029/2022GL101628.

[7] X. Liang and A. Ignatov, "Monitoring of IR Clear-Sky Radiances over Oceans for SST (MICROS)", *Journal of Atmospheric and Oceanic Technology*, vol. 28, no. 10, pp. 1228–1242, Oct. 2011, doi: 10.1175/JTECH-D-10-05023.1.

[8] R. W. Saunders, T. A. Blackmore, B. Candy, P. N. Francis, and T. J. Hewison, "Monitoring Satellite Radiance Biases Using NWP Models", *IEEE Trans. Geosci. Remote Sensing*, vol. 51, no. 3, pp. 1124–1138, Mar. 2013, doi: 10.1109/TGRS.2012.2229283.

[9] R. Chen, L. Guan, M. Liu, and L. Qu, "Cross-Calibration of HY-1D/COCTS Thermal Emissive Bands in the South China Sea," *Remote Sensing*, vol. 16, no. 2, p. 292, Jan. 2024, doi: 10.3390/rs16020292.

[10] K. A. Kilpatrick, G. Podestá, E. Williams, S. Walsh, and P. J. Minnett, "Alternating Decision Trees for Cloud Masking in MODIS and VIIRS NASA Sea Surface Temperature Products", *Journal of Atmospheric and Oceanic Technology*, vol. 36, no. 3, pp. 387–407, Mar. 2019, doi: 10.1175/JTECH-D-18-0103.1.

[11] Y. Kurihara, H. Murakami, and M. Kachi, "Sea surface temperature from the new Japanese geostationary meteorological Himawari-8 satellite", *Geophys. Res. Lett.*, vol. 43, no. 3, pp. 1234–1240, Feb. 2016, doi: 10.1002/2015GL067159.

[12] M. Liu, C. J. Merchant, O. Embury, J. Liu, Q. Song, and L. Guan, "Retrieval of Sea Surface Temperature From HY-1B COCTS", *IEEE Trans. Geosci. Remote Sensing*, vol. 60, pp. 1–13, Jul. 2022, doi: 10.1109/TGRS.2022.3190444.

[13] C. J. Merchant, A. R. Harris, E. Maturi, and S. Maccallum, "Probabilistic physically based cloud screening of satellite infrared imagery for operational sea surface temperature retrieval", *Q. J. R. Meteorol. Soc.*, vol. 131, no. 611, pp. 2735–2755, Oct. 2005, doi: 10.1256/qj.05.15.

[14] C. J. Merchant *et al.*, "A 20 year independent record of sea surface temperature for climate from Along-Track Scanning Radiometers: SST FOR CLIMATE FROM ATSRS", *J. Geophys. Res.*, vol. 117, no. C12, p. 117, Dec. 2012, doi: 10.1029/2012JC008400.

[15] X. Li *et al.*, "Deep-learning-based information mining from ocean remote-sensing imagery", *National Science Review*, vol. 7, no. 10, pp. 1584–1605, Oct. 2020, doi: 10.1093/nsr/nwaa047.

[16] J. H. Jeppesen, R. H. Jacobsen, F. Inceoglu, and T. S. Toftegaard, "A cloud detection algorithm for satellite imagery based on deep learning", *Remote Sensing of Environment*, vol. 229, pp. 247–259, Aug. 2019, doi: 10.1016/j.rse.2019.03.039.

[17] C. C. Walton, W. G. Pichel, J. F. Sapper, and D. A. May, "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.*, vol. 103, no. C12, pp. 27999–28012, Nov. 1998, doi: 10.1029/98JC02370.

[18] R. Li et al., "Land Surface Temperature Retrieval From Sentinel-3A SLSTR Data: Comparison Among Split-Window, Dual-Window, Three-Channel, and Dual-Angle Algorithms," *IEEE Trans. Geosci. Remote Sensing*, vol. 61, pp. 1–14, 2023, doi: 10.1109/TGRS.2023.3288584.

[19] C. J. Merchant, P. Le Borgne, H. Roquet, and G. Legendre, "Extended optimal estimation techniques for sea surface temperature from the Spinning Enhanced Visible and Infra-Red Imager (SEVIRI)", *Remote Sensing of Environment*, vol. 131, pp. 287–297, Apr. 2013, doi: 10.1016/j.rse.2012.12.019.

[20] K. A. Kilpatrick et al., "A decade of sea surface temperature from MODIS," *Remote Sensing of Environment*, vol. 165, pp. 27–41, Aug. 2015, doi: 10.1016/j.rse.2015.04.023.

[21] B. Petrenko, A. Ignatov, Y. Kihai, J. Stroup, and P. Dash, "Evaluation and selection of SST regression algorithms for JPSS VIIRS", *JGR Atmospheres*, vol. 119, no. 8, pp. 4580–4599, Apr. 2014, doi: 10.1002/2013JD020637.

[22] O. Embury and C. J. Merchant, "A reprocessing for climate of sea surface temperature from the along-track scanning radiometers: A new retrieval scheme", *Remote Sensing of Environment*, vol. 116, pp. 47–61, Jan. 2012, doi: 10.1016/j.rse.2010.11.020.

[23] C. J. Merchant, A. R. Harris, M. J. Murray, and A. M. Závody, "Toward the elimination of bias in satellite retrievals of sea surface temperature: 1. Theory, modeling and interalgorithm comparison", *J. Geophys. Res.*, vol. 104, no. C10, pp. 23565–23578, Oct. 1999, doi: 10.1029/1999JC900105.

[24] O. Embury et al., "Satellite-based time-series of sea-surface temperature since 1980 for climate applications," *Sci Data*, vol. 11, no. 1, p. 326, Mar. 2024, doi: 10.1038/s41597-024-03147-w.

[25] C. J. Merchant, S. Saux-Picart, and J. Waller, "Bias correction and covariance parameters for optimal estimation by exploiting matched in-situ references," *Remote Sensing of Environment*, vol. 237, p. 111590, Feb. 2020, doi: 10.1016/j.rse.2019.111590.

[26] B. Luo, P. J. Minnett, C. Gentemann, and G. Szczodrak, "Improving satellite retrieved night-time infrared sea surface temperatures in aerosol contaminated regions," *Remote Sensing of Environment*, vol. 223, pp. 8–20, Mar. 2019, doi: 10.1016/j.rse.2019.01.009.

[27] C. Jia and P. J. Minnett, "High latitude sea surface temperatures derived from MODIS infrared measurements," *Remote Sensing of Environment*, vol. 251, p. 112094, Dec. 2020, doi: 10.1016/j.rse.2020.112094.

[28] M. Liu, C. Merchant, L. Guan, and J. Mittaz, "Inter-Calibration of HY-1B/COCTS Thermal Infrared Channels with MetOp-A/IASI", *Remote Sensing*, vol. 10, no. 8, p. 1173, Jul. 2018, doi: 10.3390/rs10081173.

[29] M. Liu, L. Guan, J. Liu, Q. Song, C. Ma, and N. Li, "First Assessment of HY-1C COCTS Thermal Infrared Calibration Using MetOp-B IASI", *Remote Sensing*, vol. 13, no. 4, p. 635, Feb. 2021, doi: 10.3390/rs13040635.

[30] X. Ye, J. Liu, M. Lin, J. Ding, B. Zou, and Q. Song, "Sea Surface Temperatures Derived From COCTS Onboard the HY-1C Satellite", *IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing*, vol. 14, pp. 1038–1047, 2021, doi: 10.1109/JSTARS.2020.3033317.

[31] X. Ye *et al.*, "Evaluation of Sea Surface Temperatures Derived From the HY-1D Satellite", *IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing*, vol. 15, pp. 654–665, 2022, doi: 10.1109/JSTARS.2021.3137230.

[32] Z. Li, L. Guan, and R. Chen, "Sea surface skin temperature retrieval from HY-1D COCTS observations," *Front. Mar. Sci*., vol. 10, p. 1205776, Oct. 2023, doi: 10.3389/fmars.2023.1205776.

[33] C. Cao, F. J. De Luccia, X. Xiong, R. Wolfe, and F. Weng, "Early On-Orbit Performance of the Visible Infrared Imaging Radiometer Suite Onboard the Suomi National Polar-Orbiting Partnership (S-NPP) Satellite", *IEEE Trans. Geosci. Remote Sensing*, vol. 52, no. 2, pp. 1142–1156, Feb. 2014, doi: 10.1109/TGRS.2013.2247768.

[34] Y. Li, A. Wu, and X. Xiong, "Inter-Comparison of S-NPP VIIRS and Aqua MODIS Thermal Emissive Bands Using Hyperspectral Infrared Sounder Measurements as a Transfer Reference", *Remote Sensing*, vol. 8, no. 1, p. 72, Jan. 2016, doi: 10.3390/rs8010072.

[35] A. Berk, G. P. Anderson, and P. K. Acharya, "MODTRAN®5.3.2 USERS' MANUAL", 2016.

[36] P. Berrisford, "The ERA-Interim Archive, version 2.0", 2016.

[37] O. Embury, C. Merchant, and M. Filipiak, "Refractive indices (500-3500 cm-) and emissivity (600-3350 cm-1) of pure water and seawater." University of Edinburgh, Sep. 22, 2008. doi: 10.7488/DS/162.

[38] NIST/SEMATECH e-Handbook of Statistical Methods, 1.3.5.17. Robust Measures of Scale. (2012). Accessed: Oct. 18, 2024. [Online]. Available: https://www.itl.nist.gov/div898/handbook/eda/section3/eda35h.htm.

[39] L. Wang and C.38 Cao, "On-Orbit Calibration Assessment of AVHRR Longwave Channels on MetOp-A Using IASI", *IEEE Trans. Geosci. Remote Sensing*, vol. 46, no. 12, pp. 4005–4013, Dec. 2008, doi: 10.1109/TGRS.2008.2001062.

[40] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, "Focal Loss for Dense Object Detection", in *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice: IEEE, Oct. 2017, pp. 2999–3007. doi: 10.1109/ICCV.2017.324.

[41] S. A. Clough, M. J. Iacono, and J. Moncet, "Line-by-line calculations of atmospheric fluxes and cooling rates: Application to water vapor," *J. Geophys. Res.*, vol. 97, no. D14, pp. 15761–15785, Oct. 1992, doi: 10.1029/92JD01419.

[42] EUMETSAT, "Copernicus SLSTR Sea Surface Temperature Validation Report". Accessed: Oct. 13, 2023. [Online]. Available: https://www-cdn.eumetsat.int/files/2022-03/SLSTR_SST_Validation_Report.pdf.

[43] C. D. Rodgers, *Inverse Methods for Atmospheric Sounding: Theory and Practice*, vol. 2. in Series on Atmospheric, Oceanic and Planetary Physics, vol. 2. WORLD SCIENTIFIC, 2000. doi: 10.1142/3171.

[44] C. J. Donlon *et al.*, "Toward Improved Validation of Satellite Sea Surface Skin Temperature Measurements for Climate Research", *J. Climate*, vol. 15, no. 4, pp. 353–369, Feb. 2002, doi: 10.1175/1520-0442(2002)015<0353:TIVOSS>2.0.CO;2.

Mingkun Liu (Member, IEEE) received the B.S. degree in Marine Technology from Ocean University of China, Qingdao, China, in 2014, and Ph.D. degree in Marine Detection Technology from Ocean University of China, Qingdao, China in 2019. She visited University of Reading in UK and conducted research on infrared sea surface temperature from September 2017 to September 2018.

She has been working on research and teaching at Ocean University of China since graduation. She is an Associate Professor with the College of Marine Technology, Faculty of Information Science and Engineering. Her research interests include the validation and retrieval of sea surface temperature from satellite infrared and microwave radiometers. She teaches undergraduate courses on College Physics and Introduction to Remote Sensing.

Dr. Liu has been a member of the Group for High Resolution Sea Surface Temperature (GHRSST) Science Teams since 2022.

Lei Guan (Member, IEEE) received the B.S. degree in Electronic Engineering from Ocean University of Qingdao, Qingdao, China, in 1992 and Ph.D. degree in Marine Physics from Ocean University of Qingdao in 1998.

She has been working on research and teaching at the Ocean University of China (formerly the Ocean University of Qingdao) since graduation. She is a Professor with the College of Marine Technology, Faculty of Information Science and Engineering. She visited Tohoku University, Sendai, Japan, where she conducted research from September 2000 to September 2001. She has been a principal investigator for the national projects related to sea surface temperature (SST) studies and involved in international projects, such as the EC FP6 DRAGONESS project, the China/ESA Dragon project, and the IOC/WESTPAC New Generation SST project. She advises graduate students and teaches undergraduate and graduate courses on digital image processing, introduction to atmospheric science, atmospheric radiative transfer, and interactive data language (IDL) programming for satellite data processing. Her research interests include remote sensing of the oceanic parameters, primarily in the retrieval, validation, and merging of multisensor sea surface temperature.

Dr. Guan has been a member of the Group for High Resolution Sea Surface Temperature (GHRSST) Science Teams since 2011.

Fanli Liu received the B.S. degree in Computer Science and Technology from Ocean University of China, Qingdao, China, in 2020, and M.S. degree in Electronic Information, Ocean University of China, Qingdao, China, in 2023. She is currently pursuing a Ph.D. degree in Electronic Information under the supervision of Professor Lei Guan.

Her research interests include satellite ocean remote sensing and deep learning.

Zhicheng Sheng received the B.S. and M.S. degrees from the Ocean University of China and Zhejiang University in 2015 and 2018, respectively. After M.S, he worked as a research engineer in Huawei Technologies Co., Ltd. He is currently pursuing a Ph.D. degree in the School of Computer Science and Technology, Shandong University.

His current research interests include multimedia computing and deep learning image processing.

Zhuomin Li received the B.S. degree from the Taiyuan University of Technology, Taiyuan, China, in 2018, and the M.S. degree in photogrammetry and remote sensing from the Ocean University of China, Qingdao, China, in 2021. She is currently pursuing the Ph.D. degree in marine technology, Ocean University of China.

Her research interests include retrieving and validating sea surface temperature from satellite infrared observations.

Christopher J. Merchant received the B.A. degree in physics from the University of Oxford, Oxford, U.K., in 1989, and the Ph.D. degree in space and climate physics from University College London, London, U.K., in 1999.

He joined the Department of Meteorology, The University of Edinburgh, Edinburgh, U.K., as a Lecturer. Since 2013, he has been a Professor of ocean and Earth observation with the University of Reading, Reading, U.K. He has authored or coauthored more than 120 journal articles. His research interests focus on thermal remote sensing for climate applications.