Surface temperature and salinity in the northern Bay of Bengal: in situ measurements compared with satellite observations and model output

Md. Masud-Ul-Alam,a,b,* Md. Ashif Imam Khan,a Bradford S. Barrett,c,d and Sara Rivero-Calleb

aBangabandhu Sheikh Mujibur Rahman Maritime University, Department of Oceanography and Hydrography, Dhaka, Bangladesh
bUniversity of Georgia, Skidaway Institute of Oceanography, Department of Marine Sciences, Savannah, Georgia, United States
cU.S. Naval Academy, Oceanography Department, Annapolis, Maryland, United States
dAir Force Office of Scientific Research (AFOSR), Arlington, Virginia, United States

Abstract. The northern Bay of Bengal (BoB) has been traditionally understudied and undersampled. Satellite and modeling products could compensate for the scarcity of in situ measurements, but this requires evaluating the accuracy of satellite and modeling products first. We present a comparison of sea surface temperature (SST) and sea surface salinity (SSS) products (satellite and model output) with 46 in situ observations in the northern BoB. We used satellite and modeled SST (daily) and SSS (weekly and daily) in this comparison. The results are as follows. (1) Both model and satellite-derived SSTs agreed well with in situ observations and with each other, with small biases (<1°C) and large correlation coefficients (r > 0.77). (2) Neither model nor satellite SSSs agreed well with in situ observations (biases > 0.5 PSU, r < 0.54). (3) Calculations of the d-index support the argument that model and satellite SSTs agreed well with in situ observations (d-index values of 0.68 and 0.65, respectively), while the model and satellite SSSs did not agree well with observations (d-index values of 0.31 and 0.40, respectively). The results suggest that additional work is needed to improve both model prediction and satellite retrieval algorithms for SSS in the northern BoB. © The Authors. Published by SPIE under a Creative Commons Attribution 4.0 International License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: 10.1117/1.JRS.16.018502]

Keywords: Bay of Bengal; sea surface temperature; sea surface salinity; in situ measurements; satellite observations; model output.

Paper 210423 received Jul. 8, 2021; accepted for publication Dec. 16, 2021; published online Jan. 5, 2022.

1 Introduction

Accurate and timely remote measurements of ocean temperature and salinity are essential for numerical prediction models to forecast the ocean and atmosphere dynamics. Obtaining precise global sea surface temperature (SST) estimates has been the focus of the study of many groups for the last few decades. In a classical study, Brown et al. analyzed calibration methods of the U.S. National Oceanographic and Atmospheric Administration (NOAA) Advanced Very High Resolution Radiometer (AVHRR), by deriving accurate SST fields from satellite infrared (IR) observations, based on vacuum test datasets. Similarly, Kumar et al. explicitly examined the global Pathfinder algorithm’s performance in regional conditions, by comparing satellite data to buoy data. They concluded that a variation of ~5°C existed between these two sources. Recently, a saildrone instrument was used by Vazquez-Cuervo et al. to study SST retrievals. Among the six different Group for High Resolution Sea Surface Temperature (GHRSSST) Level-4 SST and Level-2 SSS products, they found good agreement with satellite-derived SST and less correlation to the SSS datasets.

*Address all correspondence to Md. Masud-Ul-Alam, masud.ocn@bsmrmu.edu.bd

Journal of Applied Remote Sensing 018502-1 Jan–Mar 2022 • Vol. 16(1)

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Applied-Remote-Sensing on 01 Aug 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use
Modern numerical ocean models capitalize on remote observations of temperature and salinity to help us understand both large- and small-scale processes in the Indian Ocean (IO). For example, Jensen et al.\(^5\) used the global Hybrid Coordinate Ocean Model and the Regional Ocean Modeling System (ROMS) to illustrate exchanges of salinity between the Bay of Bengal (BoB) and the IO. Moreover, Benshila et al.\(^6\) used the Nucleus for European Modeling of the Ocean (NEMO) system to provide one of the first a high-resolution structures of salinity in the BoB. Finally, important details on temperature inversions in the BoB were provided by Babu and Rao\(^7\) using the Princeton Ocean Model. These examples show how useful it is for researchers to use model output to evaluate physical, chemical, and biological parameters in the IO.

Recent improvements in numerical modeling\(^8\) and remote sensing (RS) systems include greater horizontal and vertical resolution and the introduction of new instruments, theories, and methodologies. Moreover, continuous monitoring of SST and sea surface salinity (SSS) is now possible over large spatial and temporal scales.\(^9\) Even on days when the satellite measurements contain missing values, other available products, including aerial and drone photography, buoy measurements, conductivity, temperature, and depth (CTD) profiles, and water sample analyses, can be used to fill the gaps.\(^10\) Indeed, high-resolution satellite-based observations of oceanographic fields,\(^11\) along with their assimilation into numerical ocean prediction models,\(^12\) enable advances in research and operational forecasting in marine sciences.\(^13\) Both active and passive satellite scanning strategies can be utilized to acquire a variety of oceanic parameters, including suspended particulate matter (SPM),\(^14\) sea surface roughness,\(^15\) and wave height.\(^16\) These satellite-derived products are very essential to the modeling study.

One way to evaluate numerical prediction model performance is to compare model output to satellite products.\(^17\) The AVHRR product of the NOAA is one such scanning system that has been used since 1981 to determine ocean SST in near real time.\(^18\) However, the use of satellite-derived SST has some associated caveats. Satellites measure the skin SST (at depth of ~10 to 20 \(\mu\)m), which is colder than the bulk SST (of the topmost few meters) by 0.1°C to 0.5°C.\(^19\) In addition, it depends on whether the satellite is measuring emitted radiation in the IR or in the visible as an approximate solution to the radiative transfer equation.\(^20\) Furthermore, there are some satellites (GHRSSST) those measure at night and others measure during the beginning of the day initial when the heat buildup from absorption of solar radiation surpasses the heat loss at the sea surface (also known as foundation SST). This type of satellite-derived estimates with a diurnal correction provide SST with negligible (<0.05°C) biases.\(^21\) Although there are different methods to estimate SST, scientists are consistently introducing new techniques to measure other parameters, such as SSS, from the space.

Remotely sensed observations of SSS are more recent and they are not as developed as SST. However, two platforms were recently launched: the Soil Moisture and Ocean Salinity (SMOS) satellite in 2009\(^22\) and the launch of the joint U.S./Argentinian Aquarius/Satélite de Aplicaciones Científicas (SAC)-D satellite in 2011. Although Aquarius came to halt in 2015, the SMOS platform is still engaged in routine monitoring of SSS.

Widespread usage of model output and satellite observations leads to questions regarding their accuracy. One way to measure that accuracy is to compare RS data and model output to \textit{in situ} observations, particularly for continental shelf regions where those observations are more common. However, \textit{in situ} observations themselves have uncertainty,\(^23\) and that possible error should be considered when making any comparisons. Nevertheless, a comparison of both model output and RS observations with \textit{in situ} observations on similar spatial and temporal scales can give an estimate of the accuracy of both model output and RS data. This is particularly important in enclosed regions such as the BoB that feature high to extreme sediment inputs and freshwater discharges that can complicate satellite retrieval algorithms.

Noteworthy differences between satellite SST (hereafter abbreviated as SST\(_{sat}\)) products and \textit{in situ} measurements of coastal water temperature (hereafter abbreviated as SST\(_{in}\)) have been noted in previous continental shelf studies in other parts of the world. For example, Castillo and Lima\(^24\) found negative biases between \textit{in situ} readings and Moderate Resolution Imaging Spectroradiometer (MODIS) SST\(_{sat}\) products in reef-dominated coastal waters in southern Belize. Additionally, large biases, up to 6°C, between \textit{in situ} and RS measurements were detected in 87 sites spanning the South African coastline. Although smaller in magnitude, Wu et al.\(^25\) found 1°C root mean square error (RMSE) for the Atlantic Zone Monitoring Program’s
ship-based SST$_{in}$ in the eastern Canadian shelf waters. That error was found when comparing the SST$_{in}$ with operational SST analyses from both the Canadian Meteorological Centre and the U.S. National Centers for Environmental Prediction. Stobart et al. found high correlations between annual SST$_{a}$ and SST$_{sat}$ for 32 stations in the southern Australia coastal region, but they noted significant seasonal and regional variability. Pramanik et al. found significant correlations ($r = 0.78$) between ROMS and MODIS satellite observations at the Sri Lankan dome region, near the southern BoB. Furthermore, Lin et al. compared the Aquarius SSS with ARGO in situ measurements and detected large negative biases in the northern BoB.

Differences between in situ (IS) observations, RS measurements, and numerical model output, particularly on regional and seasonal scales, make comparison studies crucial. It is generally accepted that remotely sensed measurements are well suited for large spatiotemporal scales (e.g., in the open ocean or on weekly to monthly time scales). However, RS retrievals often perform poorly in the coastal shelf waters and on sub-weekly time scales. In these coastal shelf areas, SST and SSS products with horizontal grid spacing $>1$ km are unable to accurately capture all of the features located near the coastlines. Coastal waters can be dynamic and unstable, particularly where there are large riverine discharges. That is especially the case in the northern BoB because the Ganges–Brahmaputra river transports $\sim 1 \times 10^9$ tons year$^{-1}$ of sediment to the coast, ranking first (along with the Amazon outflow) among the world’s estuaries and rivers in terms of sediment discharge. The continental shelf of the northern BoB, especially from 21.5°N to 22.5°N, contains very high concentrations of total SPM, peaking at 95 kg m$^{-3}$ in summer, which is more than four orders of magnitude higher than in the open-ocean parts of the IO (0.06 kg m$^{-3}$). The presence of SPM significantly reduces sea surface radiative emissivity of coastal water, thus disrupting the measurements of thermal radiometers that form the core of satellite-based RS.

The presence of estuaries and rivers causes the sediment composition to be extremely complex. Thus, the higher concentration of sediments in the northern BoB makes it more prone to these emission errors. The northeastern part of the continental shelf of the tropical IO is predominantly composed of low saline (10 to 15 PSU) surface water. These relatively low salinity values are the result of significant river water influxes as well as strong overturning from the seasonally dependent monsoon wind system. The surface salinity is coupled with a marked annual cycle in SSTs. When SSTs in the BoB are warm, there is more precipitation, and surface salinity is lower due to freshwater rain and river influx, again particularly in the continental shelf region. The northern BoB tends to feature small-scale SSS boundaries in all four seasons. This regional variability in SSS is due to advection-induced freshwater fluxes. The relatively small scale of the SSS features indicates that detailed high-resolution modeling and RS products are needed to capture the salinity structure. However, comparisons between RS products, model output, and in situ measurements remain scarce, partly because of the scarcity of in situ measurements. Moreover, the fixed buoy network has degraded since about 2015, and there is a lack of both research expeditions and routine cruises to collect in situ physicochemical parameters. Thus, studies such as this one are crucial to understand the surface salinity and temperature structure in the BoB.

Uncertainties of the regional application of the global model and satellite-derived SST and SSS are the key inspirations to do this study. The applications of such models are often problematic, especially on a regional scale. The concern further escalates if the study region (e.g., BoB), lacks cruise-based observations and is dependent on model or satellite measurements. Additionally, the presence of clouds, especially in the tropical region, limits some sensors’ capabilities, resulting in data gaps. These gaps suggest a continued need for in situ measurements. Finally, greater temporal and spatial coverage of in situ observations in the northern BoB, especially in the coastal shelf regions, would help to derive improved algorithms for satellite retrieval of SST and SSS. These improved retrieval algorithms could lead to improved model parameterizations in numerical models, allowing them to better capture important regional processes.

Accurate representation of both SST and SSS would thus improve both ocean and atmosphere numerical model skills. Comparative studies such as this one are imperative, as these would provide critical justification for subsequent statistical adjustments to satellite SST and SSS retrieval algorithms, resulting in improved representation of regional and small-scale disturbances in the RS measurements. This study addresses two critical questions: (1) How do the
model and remotely sensed data differ from in situ observations? (2) What are the spatial and temporal biases in model output and satellite observations? The remainder of this article is organized as follows: the data and analytical methods are presented in Sec. 2. Results are presented in Sec. 3. Discussion and conclusions are presented in Secs. 4 and 5, respectively.

2 Materials and Methods

2.1 Model Description

Daily surface ocean analyses from the Copernicus Marine Environment Monitoring Service (CMEMS) NEMO ocean model (v3.1) at a horizontal resolution of 9 km at the equator and a tri-polar ORCA12 grid were compared with in situ surface measurements. The comparisons were made for measurements taken between December 2018 and March 2020. The NEMO modeling system uses two bathymetry products: it uses ETOPO1, which is a surface relief model on a 1-arcmin grid, for deeper water (depth > 300 m) and the interpolated General Bathymetric Chart of the Oceans (GEBCO08) for shallower water (depth < 200 m). A 7-day assimilation cycle of reduced-order Kalman filter was applied to the three-dimensional (3D) multivariate model, which itself was calculated from a singular extended evolutive Kalman (SEEK) filter. The model large-scale temperature and salinity biases were calibrated in a 3D-var scheme, and in situ and RS salinity and temperature profiles were used to set the initial conditions in the model. For more details on this assimilation cycle and 3D-var scheme, see Lellouche et al.

To compare ocean model output with in situ observations from January 2016 to February 2016, another daily 3D global ocean forecasting product from a coupled ocean-atmosphere system, also distributed by CMEMS, was used. This model was produced by the Met Office (UK) and coupled (ocean-atmosphere) hourly to the NEMO ocean model (v3.4) with a resolution of 0.25 deg on the tri-polar horizontal grid of ORCA025 (with a horizontal grid spacing of 28 km at the equator). The model dataset uses a daily updated 3D ocean analyses at an equirectangular projection in regular latitude and longitude, accompanied by 50 vertical levels down to 5500 m. Bathymetry from ETOPO1 and GEBCO was implemented in this coupled model as well. The model retained surface temperature and salinity through Haney retroaction, 3D Newtonian damping, and correction of the pressure gradient in the tropics (see Bell et al., for more details).

For this study, gridded daily mean values of model surface temperature (hereafter abbreviated as SSTmod) and surface salinity (SSSmod) in the BoB were extracted from two model versions (described below) at 0.083 deg horizontal resolution in the continental shelf region of the BoB. In those cases where the high-resolution model data were not available (from January 2016 to February 2016), model output at 0.25 deg resolution was the only one used. These values of SSTmod and SSSmod were then compared with in situ SST and SSS (see Sec. 2.2) and RS SST and SSS (see Sec. 2.3).

2.2 In Situ Measurements

In situ measurements of temperature, salinity, and density were acquired using a single-fire module with CTD probe (Sea and Sun Technology GmbH). This CTD probe is capable of measuring to 2000-m depth with a temperature accuracy of ±0.002°C and a conductivity accuracy of ±0.002 mS cm⁻¹. The most recent CTD profiles were collected in three field sessions. The first session took place in January 2020, during a field expedition by the Department of Oceanography and Hydrography of Bangabandhu Sheikh Mujibur Rahman Maritime University (BSMRMU) aboard a fishing boat. The second session took place in February 2020, aboard the Bangladesh Navy Ship, the “Sangu.” The final session took place in March 2020, where the profiles were obtained aboard fishing vessel “Salman-2.”

These transects encompassed a total of 15 sites in both the eastern and western portions of the continental shelf region of the northern BoB (see Fig. 1). The cruises occurred over 3 months, and they sampled vertical depths down to 50 m. The raw CTD data were processed into a more user-friendly NetCDF format for efficient storage, usage, and sharing. Vertical profiles of temperature and salinity were taken from this processed CTD data. In this study, only surface values...
of temperature and salinity were analyzed; a follow-up study is planned to compare subsurface temperature and salinity to model output. Another four CTD profiles were taken 1 year earlier, in December 2018, at the mouth of the Karnapahuli river estuary using the same single-fire module with CTD (Sea and Sun Technology GmbH). These four profiles were complemented by four more CTD profiles of temperature and salinity taken in December 2018 on the north, south, and east sides of St. Martin’s Island. Finally, 23 CTD profiles were obtained along 20.00°N to 21.20°N and 89.37°E to 92.20°E aboard the fishing vessel “Agro food-4” of Sea Resource Ltd. during a winter fishing period, and these profiles were taken between January 2016 to February 2016. Thus, a total of 46 measurements from CTD profiles from the northern BoB spanning the winter and spring months (December to March) from 2016 to 2020 were analyzed in this study.

2.3 Satellite Observations

Optimum Interpolation Sea Surface Temperature (OISST), also known as Reynolds SST, is the RS dataset used here for comparison. OISST (hereafter abbreviated as SST_sat) is a gridded product of 0.25 deg spatial resolution available daily. These OISST values are interpolated from direct observations from the AVHRR. In both the model and satellite datasets, the grid values closest to the stations were considered representative of the station conditions. Similarly, both the model and satellite SST datasets were available daily, thus the daily value of each was compared with the in situ observations made on that day. Finally, although OISST is a hybrid product, including both satellite and in situ measurements, for the continental shelf area of the BoB, we are not aware of any in situ surface observations available for assimilation into the near real-time OISST product from January 2016 to March 2020 (the duration of our study). Thus, the OISST product analyzed here is primarily based on satellite observations.

Satellite salinity measurements were obtained from the global SSS L4 dataset, which was optimally interpolated onto a regular grid of 0.25 deg and later distributed by CMEMS. This salinity dataset incorporates the near real-time European Space Agency’s (ESA) SMOS product. High levels of noise generated during retrieval of SSS from satellites, along with substantial data gaps from clouds, were filled using a multidimensional optimal interpolation method. This interpolation was later validated from an in situ dataset. Near real-time weekly SSS_sat values...
were extracted from this product. Like the method used for SST, the average weekly value of SSS from satellite was matched to the model and \emph{in situ} SSS values from the nearest grid-point value. All retrieved satellite and model temperature and salinity were gathered from open sources, and those are briefly summarized in Table 1.

<table>
<thead>
<tr>
<th>Product type</th>
<th>Product name</th>
<th>Production unit</th>
<th>Grid (deg)</th>
<th>Processing level</th>
<th>Temporal extent</th>
<th>Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>SST</td>
<td>Model</td>
<td>Mercator Ocean</td>
<td>0.083</td>
<td>L4</td>
<td>Daily</td>
<td>July 2018 to Present</td>
</tr>
<tr>
<td></td>
<td>Model</td>
<td>Met Office</td>
<td>0.25</td>
<td>L4</td>
<td>Daily</td>
<td>December 2015 to Present</td>
</tr>
<tr>
<td></td>
<td>Satellite OISST</td>
<td>NOAA (AVHRR)</td>
<td>0.25</td>
<td>L4</td>
<td>Daily</td>
<td>September 1981 to Present</td>
</tr>
<tr>
<td>SSS</td>
<td>Model</td>
<td>Mercator Ocean</td>
<td>0.083</td>
<td>L4</td>
<td>Daily</td>
<td>July 2018 to Present</td>
</tr>
<tr>
<td></td>
<td>Model</td>
<td>Met Office</td>
<td>0.25</td>
<td>L4</td>
<td>Daily</td>
<td>December 2015 to Present</td>
</tr>
<tr>
<td></td>
<td>Satellite Global Observed Ocean Physics</td>
<td>Collected Localisation Satellites (CLS)</td>
<td>0.25</td>
<td>L4</td>
<td>Weekly</td>
<td>July 2018 to Present</td>
</tr>
</tbody>
</table>

2.4 \textbf{Quantifying Agreement between Datasets}

To investigate and quantify the agreement between the \emph{in situ}, RS, and model values of SST and SSS, we calculated the Pearson correlation coefficient\textsuperscript{69} using the following equation:

\begin{equation}
    r = \frac{n \left( \sum xy \right) - \left( \sum x \right) \left( \sum y \right)}{\sqrt{n \sum x^2 - \left( \sum x \right)^2} \sqrt{n \sum y^2 - \left( \sum y \right)^2}},
\end{equation}

where \(x\), in our case, is \emph{in situ} measurements and \(y\) is either model output or satellite observations. This method was used to calculate the correlation of (1) \(\text{SST}_{\text{in}}\) and \(\text{SSS}_{\text{in}}\) with (2) \(\text{SST}_{\text{mod}}\) and \(\text{SSS}_{\text{mod}}\) and (3) \(\text{SST}_{\text{sat}}\) and \(\text{SSS}_{\text{sat}}\). Additionally, scatter plots of these parameters were fitted to a linear model with a 95% significance level using the ggpmisc package in R.\textsuperscript{70} Another commonly used metric, the RMSE, was calculated to quantify differences between model and satellite parameters \((P_i)\) and the \emph{in situ} values \((O_i)\) using the mltools package\textsuperscript{71} in R:

\begin{equation}
    \text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (P_i - O_i)^2}{n}}.
\end{equation}

In addition to RMSE, biases were calculated for the model and satellite parameters. The biases were calculated by following the method of Thakur et al.,\textsuperscript{72} where the overlying grid values of both model and satellite datasets were subtracted from \emph{in situ} observations in the corresponding grid. We calculated the standard deviation (SD) of the biases to depict overall variations for different locations and sources.
The index of agreement, or \( d \)-index, is a standardized measurement of error in model prediction that ranges between 0 and 1. A \( d \)-index of 1 implies a perfect match with the observations.\(^{73}\) This index is capable of distinguishing proportional and additive differences of mean and variances of the simulated and observed values.\(^{74}\) Here, we calculated \( d \)-indices between \textit{in situ} (\( O_i \)) and satellite as well as model values (\( P_i \)). The following equation was used to calculate the \( d \)-index (\( md \)), using the \textit{hydroGOF} package\(^{75}\) in R:

\[
md = 1 - \frac{\sum_{i=1}^{N} |O_i - S_i|^d}{\sum_{i=1}^{N} |S_i - \overline{O}| + |O_i - \overline{O}|^d},
\]

where \( O \) and \( S \) represent the observed and simulated values, respectively, and \( \overline{O} \) is the mean of the observed values. Additionally, \( i \) represents the initial states of both parameters and \( j \) is the exponent applied in calculation of the \( d \)-index.

The final quantity used to assess the agreement between the \textit{in situ} and model, and \textit{in situ} and satellite, was the concordance correlation coefficient (CCC),\(^{76}\) calculated using the \textit{DescTools} package in R.\(^{77}\) The CCC (\( \rho_c \)) represents precision, bias, and agreement with respect to a true value or magnitude of the observation coinciding with the concordance line.\(^{78}\) The \( \rho_c \) quantifies biases and the fit with the concordance line

\[
\rho_c = rC_b,
\]

where the bias correction factor is given as \( C_b \). The CCC thus depicts the deviation of the best-fit line, while \( r \) is the correlation coefficient between \( x \) and \( y \). \( C_b \) is given as

\[
C_b = \left[ \frac{(\nu + 1/\nu + u^2)}{2} \right]^{-1},
\]

where \( \nu \) and \( u \) represent the scale bias (slope shift) and location bias (height shift), respectively. These two terms can be expressed as

\[
\nu = \frac{\sigma_1}{\sigma_2},
\]

\[
uu = \frac{\mu_1 - \mu_2}{\sqrt{\sigma_1 \times \sigma_2}},
\]

where \( \mu_1 \) and \( \mu_2 \) depict the means for the measurement and true values, respectively, and \( \sigma_1 \) and \( \sigma_2 \) represent their SDs. The CCC thus quantifies the magnitude of deviation of a dataset from a perfect agreement. CCC can be interpreted the same way as the Pearson’s correlation coefficient, where values closer to 1 imply a stronger agreement with the observed value. CCC tends to be closer to zero than Pearson’s correlation coefficient \( r \).\(^{79}\)

### 3 Results

#### 3.1 Statistical Comparison

Scatter plots between \textit{in situ} (\( \text{SST}_{\text{in}} \)) and satellite (\( \text{SST}_{\text{sat}} \)) [Fig. 2(a)], using the Student’s \( t \)-test, show that model and \textit{in situ} SST are positively correlated (\( r = 0.82, p < 0.001 \)). We also used the Student’s \( t \)-test to compare \( \text{SST}_{\text{in}} \) and \( \text{SST}_{\text{mod}} \) estimates [Fig. 2(b)], and these remain significantly correlated (\( r = 0.77, p < 0.001 \)). In both cases, \( r \) values were statistically significant.

The mean bias and SD of bias for \( \text{SST}_{\text{mod}} \) were 0.43°C and 0.96°C, respectively (Table 2). For \( \text{SST}_{\text{sat}} \), the SD of bias was larger (1.06°C) than it was for \( \text{SST}_{\text{mod}} \), even though the mean bias decreased to 0.09°C. \( \text{SST}_{\text{mod}} \) and \( \text{SST}_{\text{sat}} \) had similar ranges of RMSE, 1.05°C and 1.07°C, respectively. Moreover, \( \text{SST}_{\text{mod}} \) and \( \text{SST}_{\text{sat}} \) had similar indices of agreement (\( d \)-indices) of 0.66 and 0.65. \( \text{SST}_{\text{mod}} \) and \( \text{SST}_{\text{sat}} \) agree well with the \textit{in situ} measurements, with a model CCC of 0.92 and a satellite CCC of 0.95 (see Table 2).
Unlike SST, which featured good agreement between in situ and both model and satellite, satellite (SSS\textsubscript{sat}) and model (SSS\textsubscript{mod}) SSS differed from in situ observations by an average of 1.84 PSU. SSS\textsubscript{mod} was only weakly positively correlated with in situ measurements ($r = 0.54$), and SSS\textsubscript{sat} was uncorrelated ($r = 0.26$) with in situ measurements [Figs. 3(a) and 3(b)].

**Table 2** Summary table of all statistical analyses with the main results in this study.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Correlation coefficient ($r$)</th>
<th>P-value</th>
<th>SD of bias</th>
<th>Mean of bias</th>
<th>RMSE</th>
<th>Index of agreement</th>
<th>CCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>SST\textsubscript{mod}</td>
<td>24.3°C</td>
<td>0.82</td>
<td>&lt;0.001</td>
<td>0.96</td>
<td>0.43°C</td>
<td>1.05°C</td>
<td>0.68</td>
<td>0.92</td>
</tr>
<tr>
<td>SST\textsubscript{sat}</td>
<td>24.6°C</td>
<td>0.77</td>
<td>&lt;0.001</td>
<td>1.06</td>
<td>0.09°C</td>
<td>1.07°C</td>
<td>0.65</td>
<td>0.95</td>
</tr>
<tr>
<td>SSS\textsubscript{mod}</td>
<td>27.9 PSU</td>
<td>0.53</td>
<td>&lt;0.001</td>
<td>1.99</td>
<td>2.74 PSU</td>
<td>3.38 PSU</td>
<td>0.31</td>
<td>0.45</td>
</tr>
<tr>
<td>SSS\textsubscript{sat}</td>
<td>29.7 PSU</td>
<td>0.26</td>
<td>0.075</td>
<td>1.42</td>
<td>0.90 PSU</td>
<td>1.68 PSU</td>
<td>0.40</td>
<td>0.52</td>
</tr>
</tbody>
</table>

Fig. 2 Scatter plots of (a) in situ temperature versus model temperature, and (b) in situ temperature versus satellite temperature, showing the Pearson correlation coefficient ($r$) between SST\textsubscript{mod} and SST\textsubscript{sat} with SST\textsubscript{in}. A regression line (red), based on a linear model, is drawn to show the best fit of the dataset, along with a semitransparent error bar (light green) showing a 95% confidence interval. P-value of the correlation coefficient is given.

Fig. 3 As in Fig. 2, but for (a) in situ salinity versus model salinity, and (b) in situ salinity versus satellite salinity.

Unlike SST, which featured good agreement between in situ and both model and satellite, satellite (SSS\textsubscript{sat}) and model (SSS\textsubscript{mod}) SSS differed from in situ observations by an average of 1.84 PSU. SSS\textsubscript{mod} was only weakly positively correlated with in situ measurements ($r = 0.54$), and SSS\textsubscript{sat} was uncorrelated ($r = 0.26$) with in situ measurements [Figs. 3(a) and 3(b)].
The SSS\textsubscript{mod} had a mean bias of +2.74 PSU when compared with in situ measurements. There was also large variance in SSS\textsubscript{mod}, with a SD of 1.99 PSU (Table 2). The mean bias of SSS\textsubscript{sat} was +0.90 PSU. The largest SD of the biases was 1.42 PSU, which was observed for SSS\textsubscript{sat} (Table 2). Similarly, SSS\textsubscript{mod} had a maximum RMSE of 3.38 PSU, whereas the RSME of the satellite was 1.68 PSU. This suggests that neither the model nor the satellite captured SSS well, and moreover, the model significantly underperformed the satellite. SSS\textsubscript{mod} and SSS\textsubscript{sat} had d-index values of only 0.31 and 0.40, respectively, indicating low agreement between both model and satellite and in situ SSS measurements. Additionally, both SSS\textsubscript{mod} (CCC of 0.45) and SSS\textsubscript{sat} (CCC of 0.52) had weaker concordance correlations than they did with SST. We discuss several possible reasons for the poor model performance in the next section.

### 3.2 Spatial Variability in SST and SSS Biases

Model SST products seem to have a positive bias. Most model SST grid points were overestimated by between +0.5°C and +2.7°C. This indicates that model SSTs were generally warmer than the observations [Fig. 4(a)]. The warm bias was strongest near the southeastern coastlines of the northern BoB, over the continental shelf region. While most of the model biases were warm, negative (cold) biases of relatively smaller magnitude (−1.0°C to −1.5°C) were also noticed in the study area. Fourteen station locations had small model SST biases (between −0.5°C and +0.5°C), and we consider those locations to be similar to the in situ measurements. While model SST biases were mostly positive, satellite SSTs [Fig. 4(b)] had predominately negative biases, ranging from −0.5°C to −1.5°C, with only six points depicting strong positive bias (+1.0°C to +2.7°C). Both positive and negative SST\textsubscript{sat} biases did not have any distinct spatial pattern, and both were somewhat evenly dispersed on two fringes of the continental shelf of the bay. Overall, eight station locations had satellite SST small biases (between −0.5°C and +0.5°C).

The salinity analysis revealed both positive (i.e., salty) and negative (i.e., fresh) biases over the region. SSS\textsubscript{mod} [Fig. 5(a)] had a mostly positive bias, with only three points having a negative bias. Most of the overestimations (24 stations) were between +0.5 and +4.0 PSU. The saltiest model overestimations (+4.0 to +7.5 PSU) were found in the eastern region of the study area. Model biases between −0.5 and +0.5 PSU were considered small, but unlike SST\textsubscript{mod} where 14 stations had small biases, only two stations had small SSS\textsubscript{mod} biases. SSS\textsubscript{mod} featured a distinct longitudinal variation in bias, where the salinity overestimations occurred in the eastern BoB and underestimations occurred in the western BoB. In contrast to model salinity, the satellite salinity had an opposite pattern: a positive or saltier bias (+0.5 to +2 PSU) in the west and fresher or negative bias (−0.5 to −2.1 PSU) in the eastern areas of the study area [Fig. 5(b)]. Moreover, positive satellite SSS\textsubscript{sat} salinity overestimations were generally much smaller in magnitude (<2 PSU) than the model biases, particularly in the eastern part of the BoB. Finally, SSS\textsubscript{sat} had 12 stations with small over or underestimations (between −0.5 and +0.5 PSU).

### 4 Discussion

The primary goal of this study was to evaluate model and satellite SST and SSS products in the northern BoB. We accomplished that goal by comparing both model output and satellite measurements with in situ surface observations obtained during three cruises: January 2016 to February 2016, December 2018, and January 2020 to March 2020. To the best of our knowledge, this is the first comparative study using these recent datasets in this part of the BoB. Our results suggest that among all the products considered, the model and satellite SST products best matched with the in situ measurements over the northern BoB. However, model output and satellite measurements of salinity did not agree at all with the in situ observations over the northern BoB. Our results agree with other comparative studies performed in coastal areas, who also found low biases and high correlations between in situ SST and model and satellite SST\textsuperscript{24,29} along with high biases and low correlations between in situ SSS and model and satellite SSS.\textsuperscript{72}

Recent work suggests the introduction of new parameterizations in the NEMO model, such as additional dynamic height statistics and assimilations from several sources, can explain the agreement with in situ measurements.\textsuperscript{80} Despite the reports of larger biases in SST\textsubscript{sat} in different
shelf regions, including off the coasts of China, South Africa, the United States, and Western Australia, the continental shelf areas of the BoB features significantly smaller bias in \( \text{SST}_{\text{sat}} \). Calculated model (0.68) and satellite (0.65) \( d \)-index values for SST in this study suggest both model and satellite SSTs agreed well with the \textit{in situ} measurements made during the cruise. We found RMSE for \( \text{SST}_{\text{mod}} \) and \( \text{SST}_{\text{sat}} \) to be 1.05°C and 1.07°C, respectively, which are larger than the global average RMSE of 0.4°C calculated by Brasnett, but smaller than the values (1.5°C) reported by Chen and Hu. The larger RMSE temperature values are possibly explained by small-scale horizontal temperature gradients not captured well by the model or resolved well by the satellite. Steep gradients typically exist between the northern and southern part of the BoB, and those are sometimes poorly represented in either satellite or model output. Indeed, the RSME was larger in regions where the climatological SST gradient is larger and smaller in regions where the gradient tends to be smaller. Nevertheless, the CCC calculation

![Fig. 4](https://example.com/figure4.png)

**Fig. 4** The bias of (a) model and (b) satellite were compared with \textit{in situ} measurements of SST over the continental shelf of the northern BoB. Points were categorically plotted where positive and negative biases are illustrated with red and blue points, accordingly. The deeper the corresponding color is, the higher the biases of corresponding color is.
affirmed that both $\text{SST}_{\text{mod}}$ and $\text{SST}_{\text{sat}}$ agreed with in situ SST: $\text{SST}_{\text{mod}}$ ($\text{CCC} = 0.92$) and $\text{SST}_{\text{sat}}$ ($\text{CCC} = 0.92$). This good fit could possibly be attributed to one of the inherent characteristics of the AVHRR satellite’s IR sensor, as IR sensors can retrieve SST within 1 km of the coastline. Most of the stations in this study were more than 1 km from the coastline.

The diurnal cycle of SST over the ocean is relatively small because of seawater’s large heat capacity.\textsuperscript{90} Consideration of diurnal skin SST remains crucial for the tropics as it yields enhanced forecast.\textsuperscript{91} Surface fluxes are subjectively affected by diurnal to intraseasonal scales due to the diurnal variability of skin SST.\textsuperscript{92} All but two of the in situ measurements were made during the day, and those observations were then compared with the daily averaged model and satellite SST. Therefore, diurnal variability could have contributed to some of the differences between $\text{SST}_{\text{in}}$ and $\text{SST}_{\text{sat}}$. Such diurnal variations have been reported in several regional seas, including the Mediterranean Sea,\textsuperscript{93} Sargasso Sea,\textsuperscript{94} equatorial tropical Pacific,\textsuperscript{95} western North Pacific,\textsuperscript{96} and western North Atlantic.\textsuperscript{97} But, there has been very limited study of the diurnal variability of SST

![Figure 5](https://www.spiedigitallibrary.org/journals/Journal-of-Applied-Remote-Sensing)
in the BoB, and one such study suggests that the diurnal cycle is strong only during warming phases of the intraseasonal oscillations of SST.\textsuperscript{98} Nevertheless, including the effects of the diurnal variation in the regional model algorithm would enhance the model’s performance and coupling, as it would represent the air–sea interactions more accurately.\textsuperscript{99} even on a longer time scale.\textsuperscript{100}

Turning to SSS, $d$-index values for the model (0.31) and satellite (0.40) SSS suggested those datasets agreed only weakly with \textit{in situ} cruise data. The RMSE increased to +3.4 PSU for $SSS_{\text{mod}}$ and +1.7 PSU for $SSS_{\text{sat}}$. CCC values for $SSS_{\text{mod}}$ and $SSS_{\text{sat}}$ were low (0.45 and 0.52, respectively), confirming that both model and satellite SSS did not agree with the \textit{in situ} measurements. The study area is subject to heavy freshwater influx from precipitation and river discharge, along with vertical stratification and horizontal advection. As such, high RMSE values for SSS products could be attributed to these physical processes, as they strongly affect general patterns of SSS and occur at scales that could be difficult to capture for both models and satellites. Furthermore, the radio frequency interference errors (ocean reflected radio frequency errors in passive microwave measurements) for RS products might be another possible cause for the $SSS_{\text{sat}}$ disagreements with the \textit{in situ} SSS.\textsuperscript{101}

The coarse spatial resolution of the model and satellite products could also explain the disparities. It is known that ESA’s SMOS data are often not suitable for coastal and estuarine studies as they can be too coarse at times.\textsuperscript{30} This could be one of the many reasons for the differences between $SSS_{\text{sat}}$ and $SSS_{\text{in}}$. Similar to the SST differences, it is also possible that the differences in SSS between \textit{in situ} and model and satellite observations could be explained by temporal offsets between model (daily) and satellite (weekly values) and the instantaneous diurnal cycle sampled at the moment of the \textit{in situ} SSS measurements. However, that would suggest that satellite values might be worse than model output, when the opposite was found here. The annual influx of 1300 km$^3$ of freshwater from the Ganga–Meghna–Brahmaputra river system\textsuperscript{102} causes very large fluctuations in the surface salinity in the continental shelf seas of the northern BoB, and those fluctuations could possibly occur on sub-weekly time scales. However, the freshwater fluctuations are not directly captured by the salinity model, thus forcing the model to rely upon pseudo-observations (climatological river-runoff) with a higher signal to noise ratio (see the processing manual for the SSS data by CMEMS for more details). Another possible reason for the deviation could be that model salinity parameterizations and satellite salinity retrieval algorithms were developed for open-ocean applications, rather than coastal sites. Correction for only the large-scale biases was performed in the numerical model, which possibly limited the applicability here. It must be mentioned that we are aware of no comprehensive studies on the variability of SSS in the continental shelf of the BoB. We suggest that further study is necessary, and particularly additional field campaigns are needed over the shelf region of the BoB where high freshwater influx and stratification are occurred.

To summarize, continental shelf areas in the BoB are data-poor regions, as no continuous or long-term physical parameter datasets are available for the region. The northern BoB remains a highly dynamic region, especially along the coast and shelf region. Coastal shelf basins are prone to high short-term sedimentation,\textsuperscript{103} and moreover, surface hydrographic forcing, with daily to even hourly variations, regularly alters the oceanographic conditions of coastal zones.\textsuperscript{104,105} These ever-shifting surface processes are, at present, better captured with \textit{in situ} measurements than numerical models or by satellite retrieval algorithms. Nevertheless, \textit{in situ} measurements are not always openly accessible and are expensive to arrange. Satellite and models can overcome these limitations and provide continuous physiochemical information on the continental shelf areas of the BoB, but they are only useful when they compare favorably with \textit{in situ} measurements.

5 Conclusions

This study presents evidence that in the data-poor region of the northern BoB, satellite observations and model analyses compare well with \textit{in situ} measurements of SST but poorly with \textit{in situ} measurements of SSS. We suggest that differences in the SSS in both satellite products and model analyses can be reduced by increasing the number of \textit{in situ} observations and increasing the resolution of satellite data to better constrain the model in this region. Field campaigns are needed to continue to gather \textit{in situ} data in both the coastal and the open-ocean regimes of the
BoB. More *in situ* measurements would assist in improving models or satellite algorithms that would more accurately capture the physicochemical conditions in the BoB.

6 Appendix A

The *in situ* measurements used in this study was done through the CTD deployments on several different occasions in 2016, 2018, and 2020. The stations’ information was given in the following table (Table 3).

<table>
<thead>
<tr>
<th>St. No</th>
<th>Latitude (°E)</th>
<th>Longitude (°N)</th>
<th>Date (month date, year)</th>
<th>Time (GMT + 6)</th>
<th>Depth (m)</th>
<th>Tide</th>
<th>Weather condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20.64</td>
<td>92.32</td>
<td>February 26, 2018</td>
<td>11:15 a.m.</td>
<td>1.82</td>
<td>High</td>
<td>Sunny</td>
</tr>
<tr>
<td>2</td>
<td>20.61</td>
<td>92.32</td>
<td>February 27, 2018</td>
<td>11:37 a.m.</td>
<td>2.32</td>
<td>High</td>
<td>Sunny</td>
</tr>
<tr>
<td>3</td>
<td>20.59</td>
<td>92.33</td>
<td>February 27, 2018</td>
<td>12:04 p.m.</td>
<td>2.07</td>
<td>High</td>
<td>Sunny</td>
</tr>
<tr>
<td>4</td>
<td>20.61</td>
<td>92.33</td>
<td>March 1, 2018</td>
<td>12:50 p.m.</td>
<td>0.16</td>
<td>Low</td>
<td>Sunny</td>
</tr>
<tr>
<td>5</td>
<td>21.25</td>
<td>91.87</td>
<td>December 22, 2018</td>
<td>09:26 a.m.</td>
<td>20.63</td>
<td>Low</td>
<td>Sunny</td>
</tr>
<tr>
<td>6</td>
<td>21.20</td>
<td>91.87</td>
<td>December 22, 2018</td>
<td>09:52 a.m.</td>
<td>23.17</td>
<td>Low</td>
<td>Sunny</td>
</tr>
<tr>
<td>7</td>
<td>21.20</td>
<td>91.82</td>
<td>December 22, 2018</td>
<td>01:52 p.m.</td>
<td>20.61</td>
<td>High</td>
<td>Sunny</td>
</tr>
<tr>
<td>8</td>
<td>21.25</td>
<td>91.82</td>
<td>December 22, 2018</td>
<td>02:31 p.m.</td>
<td>18.08</td>
<td>High</td>
<td>Partially cloudy</td>
</tr>
<tr>
<td>9</td>
<td>20.58</td>
<td>92.36</td>
<td>January 2, 2020</td>
<td>05:17 p.m.</td>
<td>7.18</td>
<td>High</td>
<td>Sunny</td>
</tr>
<tr>
<td>10</td>
<td>20.64</td>
<td>92.32</td>
<td>January 3, 2020</td>
<td>10:04 a.m.</td>
<td>11.60</td>
<td>Low</td>
<td>Sunny</td>
</tr>
<tr>
<td>11</td>
<td>20.51</td>
<td>90.67</td>
<td>February 7, 2020</td>
<td>2:29 p.m.</td>
<td>83.06</td>
<td>High</td>
<td>Partially cloudy</td>
</tr>
<tr>
<td>12</td>
<td>20.60</td>
<td>92.10</td>
<td>February 8, 2020</td>
<td>09:38 a.m.</td>
<td>28.90</td>
<td>Low</td>
<td>Sunny</td>
</tr>
<tr>
<td>13</td>
<td>21.31</td>
<td>90.47</td>
<td>February 9, 2020</td>
<td>11:06 a.m.</td>
<td>14.38</td>
<td>Low</td>
<td>Sunny</td>
</tr>
<tr>
<td>14</td>
<td>21.38</td>
<td>90.28</td>
<td>February 9, 2020</td>
<td>08:54 a.m.</td>
<td>4.79</td>
<td>High</td>
<td>Sunny</td>
</tr>
<tr>
<td>15</td>
<td>21.32</td>
<td>90.16</td>
<td>February 9, 2020</td>
<td>09:56 a.m.</td>
<td>26.57</td>
<td>Low</td>
<td>Sunny</td>
</tr>
<tr>
<td>16</td>
<td>21.29</td>
<td>89.81</td>
<td>February 9, 2020</td>
<td>11:57 a.m.</td>
<td>19.57</td>
<td>High</td>
<td>Sunny</td>
</tr>
<tr>
<td>17</td>
<td>21.29</td>
<td>89.70</td>
<td>February 9, 2020</td>
<td>04:42 p.m.</td>
<td>27.71</td>
<td>High</td>
<td>Sunny</td>
</tr>
<tr>
<td>18</td>
<td>21.61</td>
<td>89.80</td>
<td>February 9, 2020</td>
<td>03:50 p.m.</td>
<td>5.47</td>
<td>High</td>
<td>Sunny</td>
</tr>
<tr>
<td>19</td>
<td>20.85</td>
<td>90.97</td>
<td>February 9, 2020</td>
<td>11:18 a.m.</td>
<td>82.74</td>
<td>Low</td>
<td>Sunny</td>
</tr>
<tr>
<td>20</td>
<td>20.51</td>
<td>90.58</td>
<td>March 18, 2020</td>
<td>12:31 p.m.</td>
<td>90.41</td>
<td>Low</td>
<td>Sunny</td>
</tr>
<tr>
<td>21</td>
<td>20.85</td>
<td>92.14</td>
<td>March 19, 2020</td>
<td>11:53 a.m.</td>
<td>11.95</td>
<td>High</td>
<td>Sunny</td>
</tr>
<tr>
<td>22</td>
<td>20.98</td>
<td>91.83</td>
<td>March 19, 2020</td>
<td>06:07 p.m.</td>
<td>54.19</td>
<td>Low</td>
<td>Partially cloudy</td>
</tr>
<tr>
<td>23</td>
<td>21.13</td>
<td>91.09</td>
<td>March 21, 2020</td>
<td>10:45 a.m.</td>
<td>10.49</td>
<td>High</td>
<td>Sunny</td>
</tr>
<tr>
<td>24</td>
<td>20.96</td>
<td>91.61</td>
<td>January 18, 2016</td>
<td>12:10 p.m.</td>
<td>85.85</td>
<td>Low</td>
<td>Sunny</td>
</tr>
<tr>
<td>25</td>
<td>20.95</td>
<td>91.63</td>
<td>January 20, 2016</td>
<td>02:20 p.m.</td>
<td>79.01</td>
<td>High</td>
<td>Partially cloudy</td>
</tr>
</tbody>
</table>
Acknowledgments

This research was partially performed under a project funded by the University Grants Commission (UGC) of Bangladesh and Bangabandhu Sheikh Mujibur Rahman Maritime University, Bangladesh, Reference no. BSMRMU/PG Research-255 (part-04)/19/563. Additionally, the authors appreciate the Vice-Chancellor of BSMRMU, Dean of the Faculty of Earth and Ocean Science, and Bangladesh Navy for arranging the ship used in the field campaign. Special thanks to the University of Georgia Skidaway Institute of Oceanography for supporting Rivero-Calle and Md. Masud-Ul-Alam and funding this work as an open-access article. We also acknowledge the anonymous reviewers for their valuable comments. No potential conflict of interest was reported by the authors.

Data and Code Availability Statement

The data and code that support the findings of this study are available from the corresponding author, Md. Masud-Ul-Alam, upon request.
References

70. P. J. Aphalo, “ggpmisc: miscellaneous extensions to ‘ggplot2’,” 0.3.5, Comprehensive R Archive Network (CRAN) (2020).
75. M. Zambrano-Bigiarini, “hydroGOF: Goodness-of-fit functions for comparison of simulated and observed hydrological time series,” 0.4-0 (2020).

Md. Masud-Ul-Alam is a graduate student in the Department of Marine Science, University of Georgia, Athens. Additionally, he is a lecturer in oceanography at Bangabandhu Sheikh Mujibur Rahman Maritime University, Bangladesh. He started his career at the Physical and Space Oceanography Division, BORI, Bangladesh. He received his MS degree in oceanography from the University of Dhaka. He worked as a research fellow at AWI, Germany. His research interests include satellite oceanography, air–sea interactions, and ocean modeling.

Md. Ashif Imam Khan is a final year undergraduate student who majors in oceanography. He is currently performing a thesis upon paleo-climatic conditions during the Last Interglacial and Last Glacial Maximum in the BoB. His research interest lies in performing spatial and temporal analysis using remotely sensed and model datasets that incorporate physical, chemical, biological, and meteorological aspects of the ocean and atmosphere.

Bradford S. Barrett received his PhD in meteorology and has studied the intersection of climate, extreme events, and the ocean. His research spans multiple temporal and spatial scales, with focus on variability of the atmosphere, ocean, and cryosphere on the subseasonal scale.

Sara Rivero-Calle is an assistant professor at the University of Georgia. She received her BSc degree in biology from Complutense University, her MSc degree in biological oceanography from the University of Puerto Rico, and her PhD in oceanography from Johns Hopkins University. At UGA, she manages the Bio-Optical and Satellite Oceanography lab and is a science lead of SeaHawk CubeSat. She is interested in CubeSat Technology, subpixel variability, and the intersection of RS and numerical models.