

Physical retrieval of sea-surface temperature from INSAT-3D imager observations

By SATYA P. OJHA^{1*}, and RANDHIR SINGH², ¹MOSDAC Research and Training Division, MOSDAC Research Group, EPSA Space Applications Centre, Ahmedabad, India; ²Atmospheric Sciences Division, Atmospheric and Oceanic Sciences Group, EPSA, Space Applications Centre, Ahmedabad 380015, India

(Manuscript received 9 November 2018; in final form 14 August 2019)

ABSTRACT

This paper presents results of physical retrieval of sea-surface temperature (SST) from the INSAT-3D imager observations. Radiance measurements from two thermal infrared imager channels are used for retrieving SST. Prior to their use in the retrieval process, the INSAT-3D radiances are corrected for the biases using a newly developed algorithm. The physical retrieval of SST is performed using an optimal estimation method. The optimisation is performed under two scenarios. In the first scenario, the SST retrieval problem is assumed to be linear while in the second one, the assumption of linearity is relaxed. Further, retrievals are performed by considering the reduced, as well as, the full state vector. Overall, physical retrieval performs much better than the regression-based approach. Moreover, in the physical retrieval, the best performance is obtained in the second scenario with the use of the full state vector. Validation of these retrievals with the iQuam SST resulted in a standard deviation (bias) of ~ 0.65 (0.19) K while those from the regression-based method are found to have standard deviation (bias) of ~ 0.68 (−0.17) K.

Keywords: sea-surface temperature, INSAT-3D, physical retrieval, RTTOV

1. Introduction

Sea-surface temperature (SST) is of fundamental importance for climate and weather-related studies as it is responsible for the transfer of heat fluxes across the air-sea interface. It is one of the boundary layer input variables in the ocean, atmosphere and climate models (Donlon, et al., 2009). Accurate SST is required in these numerical models in order to make a good forecast/prediction (Emery et al., 2001). SST also regulates the exchange of gases such as carbon dioxide across the air-sea interface, providing a link to the carbon cycle thus controlling the evolution of ocean biology. It plays a vital role in marine navigation too. The present accuracy requirements of SST observations for various applications are given in Prigent et al. (2013). The current SST observing system consists of in-situ and satellite observations. In-situ observations are made from ships and buoys (both moored and drifting). Although, in-situ SST observations are of good quality, they are not frequently

available. On the other hand, satellite observations have wide coverage and are available at much higher resolution, in both space and time.

As satellite observations have global coverage and they provide frequent observations, satellite remote sensing of SST is used to address the key science issues and the demanding needs in various other applications. Satellite-based SST retrievals are generally performed using radiance observations in the thermal infrared (TIR) region. TIR-based SST is more accurate than microwave-based SST due to the smaller variation of surface emissivity in the TIR region and the stronger dependence of the radiation on the temperature (Singh et al., 2016a). In the microwave region, the variation in emissivity is much larger and the SST dependence on the radiance is linear, thus leading to larger SST uncertainties. Furthermore, the radiance observations in the microwave region have more noise, as compared to TIR, which also adds to uncertainty in SST retrievals. Another advantage of TIR-based SST retrievals is their high spatial resolution, which is very useful in the coastal waters and the regions with strong thermal gradients. Therefore, most of the SST

*Corresponding author. e-mail: satyaprakash.ojha@gmail.com

This article has been republished with minor changes. These changes do not impact the academic content of the article.

observations are retrieved from satellite sensors which operate in the thermal infrared region (Ricciardulli and Wentz, 2004).

The Advanced Very High Resolution Radiometer (AVHRR) on NOAA satellites (Cracknell, 1997), the Along Track Scanning Radiometer (ATSR) on the European Remote Sensing satellites (Mutlow et al., 1994), the Moderate Resolution Imaging Spectroradiometer (MODIS) on the NASA EOS platform (Esias et al., 1998), the Spinning Enhanced Visible and Infrared Imager (SEVIRI) on Meteosat Second Generation satellites (Le Borgne et al., 2006) and the Advanced Baseline Imager (ABI) on the Geostationary Operational Environmental Satellites-16 (GOES-16, Schmit et al., 2005) are successful examples of IR sensors which have been used for operational SST retrievals. Currently, Indian National Satellites (INSAT-3D and INSAT-3DR) provide the thermal infrared measurements over the Indian subcontinent from a geostationary platform.

Most of the operational SST retrievals from satellite observations are performed using regression-based methods (Koner and Harris, 2016). The coefficients in the retrieval equation are derived by the regression of in-situ SST measurements against the observed (e.g. Walton et al. 1998) or simulated radiances using the radiative transfer model (e.g. Zavody et al. 1995). Regression-based methods are easy to implement. However, these methods lead to significant systematic errors in SST retrievals that have complex spatial and temporal characteristics (Merchant et al., 2006). The biases in SST retrieved using regression-based approaches may arise from an error in specifying the retrieval coefficients from either forward modelling or due to instrumental biases (Merchant and Le Borgne, 2004; Merchant et al., 2008). The coefficients used in the regression-based approaches are fixed and the same set of coefficients are used for all regions throughout the year. As the atmospheric and oceanic states are highly dynamic in space and time, the use of a single set of regression coefficients globally throughout the year is not justified. In the regression-based SST retrieval methods, the characterisation of the retrieval error is also difficult (Koner and Harris, 2016). Because of these serious limitations associated with regression-based approaches, physical retrievals of SST are being explored and attempted by various researchers (Merchant et al., 2008; Koner et al., 2015; Nielsen-Englyst et al., 2018). Though the physical SST retrieval methods are computationally expensive, with the advent of high-speed computational resources, these methods are now feasible. Most of the physical retrieval methods have their foundations in the Bayesian theory and they make use of *a priori* estimates of the retrieved parameters along with their, as well as observation, error covariances and thus often referred as

stochastic methods. Optimal estimation (OEM) is among the most popular physical SST retrieval methods. Recently, Koner et al. (2015) proposed a deterministic physical SST retrieval algorithm which is built upon a total least squares (TLS) approach. This new algorithm is called the modified total least squares (MTLS). It does not require the error covariances of the parameters and observations. The MTLS is shown to be better than the optimal estimation. However, the MTLS works only for over-determined problems (Koner et al., 2015).

The present INSAT-3D operational algorithm, hereafter OPR, for SST retrieval is regression-based and is quadratic in nature. The angular dependency is not incorporated in the OPR algorithm directly, rather the regression coefficients are generated for four sets of satellite viewing angles. The INSAT-3D radiances are reported to have large biases (Singh et al. 2016a, 2016b). In this study, we have analysed the OPR SST in terms of its accuracy as compared with in-situ observations. In order to address the question of whether the radiance bias and the complex nature of the OPR algorithm is limiting its accuracy, we have developed a radiance bias correction method and applied it on INSAT-3D radiances. The bias corrected radiances are used in the non-linear SST (NLSST) algorithm proposed by Merchant et al. (2008). Further, we have explored the application of OEM and its variants on INSAT-3D radiances and assessed their performance on SST retrieval accuracy.

The structure of the paper is as follows. A brief description of the radiative transfer model, used in this study, is given in section 2. The data used in the present study are described in section 3. The various pre-processing steps of INSAT-3D observations are presented in section 4. The OPR, NLSST and OEM algorithms along with their variants, used in the present study, are discussed in section 5. The results of the SST retrievals obtained from different algorithms are presented in section 6.

2. Forward model: Radiative transfer model

In order to successfully retrieve atmospheric variables using an OEM approach, the underlying physics of the radiation measurement needs to be modelled properly by a forward model solving the radiative transfer (RT) equation. The derivative of the forward model with respect to the state (also called *Jacobian*) is also required in the OEM method. Considering the nonlinearities of the RT equation and the iterative nature of the retrieval scheme, the computational efficiency of the forward model and its derivative are essential. The general forward model equation mapping the atmospheric state into measurement space (satellite-measured radiance or brightness temperature spectrum) takes the following form (e.g. Rodgers, 2000).

Table 1. INSAT-3D imager channel characteristics (Singh et al., 2016a).

Channel number	Channel wavelength (μm)	Description	Purpose	Spatial resolution (km)
1	0.55–0.75	Visible	Clouds, surface features	1.0
2	1.55–1.70	Short-wave infrared	Snow, ice and water phase in clouds	1.0
3	3.80–4.00	Medium-wave infrared	Clouds, Fog, Fire	4.0
4	6.50–7.10	Water vapour	Upper troposphere water vapour	8.0
5	10.30–11.30	Long-wave infrared window	Cloud top and surface temperature	4.0
6	11.50–12.50	Long-wave infrared window	Lower-tropospheric moisture	4.0

$$\mathbf{y} = H(\mathbf{x}) + \epsilon \quad (1)$$

where $\mathbf{y}$ is the measurement vector, H is the forward model operator, $\mathbf{x}$ is the state vector, and ϵ is the measurement error. The Radiative Transfer for TIROS Operational Vertical Sounder (RTTOV) model version 12.2 is used to simulate the INSAT-3D imager radiances ($\mathbf{y}$) using the profiles of atmospheric temperature and specific humidity along with the essential surface variables (surface pressure, 2 m temperature, 2 m specific humidity, 10 m wind components, and SST). RTTOV is a fast radiative transfer model that calculates level-to-space transmittances (Saunders et al., 1999). The RTTOV simulated clear-sky IR brightness temperatures match the corresponding line-by-line model simulations with an accuracy of ± 0.1 K (Saunders et al. 2007).

Simulation of satellite radiances also requires the information of surface emissivity, particularly for surface sensitive channels (Singh et al., 2016a). In this study, the Infrared Surface Emissivity Model for RTTOV-6 (ISEM-6: Sherlock, 1999) is used for computing the IR surface emissivity.

3. Data used

3.1. Satellite observations

India successfully launched the INSAT-3D satellite on 26th July 2013. It is positioned at 82° E longitude. There are two meteorological instruments carried by INSAT-3D, the imager and the sounder. The imager has six channels, one in the visible range of the electromagnetic spectrum, and five within the IR region (Table 1, Singh et al., 2016a). In this study, the INSAT-3D imager brightness temperature observations, from two long-wave IR channels (channel no. 5 (10.3–11.3 μm) and 6 (11.5–12.5 μm), see Table 1), are used for the retrieval of SST. These two channels are the most commonly used channels for SST retrievals since they are in the atmospheric window region (McMillin and Crosby, 1984). High temporal frequency data are useful for tracking the SST diurnal cycle (Merchant et al., 2013) which requires the

consistent retrievals throughout the diurnal cycle. In order to maintain the consistency of the retrievals during day and night, the 3.9 μm channel is not used in the retrieval. The retrieval of SST is performed for eight months (October 2017–May 2018). Hereafter, channel no. 5 and 6 are referred to as TIR1 and TIR2 respectively.

3.2. First-guess

The first-guess temperature and specific humidity (along with essential surface variables), required by the forward model to compute the first-guess radiances, are taken from the Global Forecast System (GFS) model's six-hour forecast available at the spatial resolution of $0.5^\circ \times 0.5^\circ$ while the first-guess SST is taken from a daily climatology (Banzon et al., 2014). Since we intend to use one of the best algorithms from this study for near-real-time operational compatibility, therefore the forecast data is used instead of analysis/reanalysis. The GFS model forecast and the daily SST climatology are interpolated at the INSAT-3D imager observation locations using bilinear interpolation.

3.3. Data for validation

The quality of the SST retrievals from INSAT-3D observations are compared with the in-situ SSTs obtained from the iQuam (in-situ SST Quality monitor). The iQuam is a system developed at NESDIS/STAR for performing the near real-time quality check of in-situ measurements and for monitoring its statistics (Xui and Ignatov, 2013). In-situ data measured by a large variety of sensors on different platforms, e.g. ships, drifting and mooring buoys, drifters, tropical and coastal moored buoys, ARGO and from individual platforms are included in the iQuam dataset. The iQuam SST is assigned a quality flag (QC) from 1 to 5 for bad to good data. For validation purpose, only the best quality (QC = 5) iQuam SSTs are used.

The SEVIRI SSTs obtained from the Ocean and Sea Ice Satellite Application Facility (OSI-SAF), available over the Indian Ocean since April 2017, are also used for

the validation of the SST retrievals. The SEVIRI SST is hereafter referred to as OSI-SAF-SST.

4. Observation pre-processing: Quality control and bias correction

Quality control is an integral part of any retrieval algorithm. Observation quality control consists of detecting bad observations which may result in poor retrieval. Any radiance observation contaminated with too much noise is considered as a bad observation. The noise may be due to the instrument error or due to the anomalous scans which may arise when the sensor reaches saturation level because of the intrusion of straylight. Cloud contaminated observations should also be screened out during the retrieval, as they may result in sub-optimal retrievals. Brightness temperature thresholds tests are usually used to remove cloudy observations, the same approach has been used in this study. However, it is difficult to remove observations near the cloud edges using these tests. Prior to applying any threshold test, a brightness temperature spatial gradient check was applied to remove the effects of the cloud edges. The observations at the locations where the gradient was found to be greater than 4 K/pixel were discarded as cloud contaminated observations. Next, an absolute quality check was applied; if $|T_{b_{obs}} - T_{b_{fg}}| > 12 \sigma_{b_{obs}}$, the observations were considered as cloud contaminated, thus discarded. Generally, due to the high spatial resolution of the visible (VIS) channel, the identification of clear and cloudy pixels is less challenging using the VIS channel as compared to the IR channel. However, unavailability of VIS channel during night-time necessitated the development of an IR channel based technique. Therefore, the IR based threshold for absolute quality check was finalised using a two steps procedure. In the first step, VIS channel images that contain only clear-sky pixels were generated, during different seasons. In the second step, infrared images obtained with the application of different threshold values were compared with the corresponding visible images to extract only the clear-sky pixels in IR images. The value which had the best skill in segregating clear and cloudy pixels, using VIS image as a reference, was selected as a final threshold value, which was found to be $12\sigma_{b_{obs}}$. Here, $T_{b_{obs}}$ and $T_{b_{fg}}$ stands for the observed brightness temperature and the first-guess brightness temperature (model simulated) respectively; $\sigma_{b_{obs}}$ is the error standard deviation in the observed brightness temperature, which is considered as the sum of the sensor noise ($\epsilon_{TIR1}, \epsilon_{TIR2}$) and the forward model error (ϵ_{RT}). The method which is given by Hillger and Haar (1988), hereafter HH88, was used to estimate the noise in the thermal infrared channels of INSAT-3D. The HH88 method is based on statistical structure analysis and provides the estimate of the absolute noise.

The mean (temporal) noise for the INSAT-3D thermal infrared channels obtained from HH88 method is ~ 0.3 K. However, in order to include the uncertainties in the forward model, a higher value (0.4 K) of $\sigma_{b_{obs}}$ was considered.

Using the criterion mentioned above, most of the cloudy observations were removed. To remove the remaining, if any, cloud contaminated observations, the following additional cloud detection checks were applied

- observations were considered as cloudy if the window channel (TIR1) brightness temperature was too cold, i.e. $T_{obsTIR1} < 270$ K.
- $$\frac{[(T_{b_{obsTIR1}} - T_{b_{obsTIR2}}) - (T_{b_{fgTIR1}} - T_{b_{fgTIR2}})] (T_{b_{obsTIR1}} - 260.0)}{(T_{b_{obsTIR2}} - 260.0)} > 0.7$$

For any given imager observation if any of the above conditions were met, then the whole set of selected channel observations was discarded. The last check was obtained from Zhuge and Zou (2016), however, the threshold value has been changed. With the use of these cloud-screening checks, it was found that these checks were able to filter all the cloudy pixels, along with a few non-cloudy pixels. Since the objective of the present study was the development of the SST algorithm, the above conservative cloud-detection checks were chosen to avoid any possibility of error due to cloud leakage.

Our recent studies (Singh et al., 2016a, 2016b) show that the INSAT-3D radiances have significant biases and they need to be corrected prior to their use. It was further verified by comparing the INSAT-3D observations (TIR1 and TIR2) with the corresponding SEVIRI (on-board Meteosat second generation-1 satellite, MSG-1) observations. SEVIRI radiances are reported to be of good quality (Hewison, 2013; <https://www.star.nesdis.noaa.gov/smcd/GCC/ProductCatalogImages.php>). More than 24000 SEVIRI and INSAT-3D imager collocated pairs were obtained from one month (Oct 2017) data. Prior to comparison, it was verified theoretically that the difference between the simulated brightness temperatures of SEVIRI and INSAT-3D imager sensors are within ± 0.1 K. The INSAT-3D imager observations, both TIR1 and TIR2, were found to be ~ 1.0 K colder than corresponding SEVIRI observations (Figure not shown). The bias in the INSAT-3D observations with respect to the numerical weather prediction (NWP) model first-guess is also found to be of similar order (~ 1.5 K). This confirms that the INSAT-3D observations are biased. Prior to their use in the retrieval algorithm, the INSAT-3D imager observations were bias-corrected using a scheme based on histogram specification. This method is widely used in image processing for matching the statistical properties of a pair of images. For completeness, the bias correction method used here is described briefly. Let $F_o(y_o)$ be the probability distribution function (PDF) of the

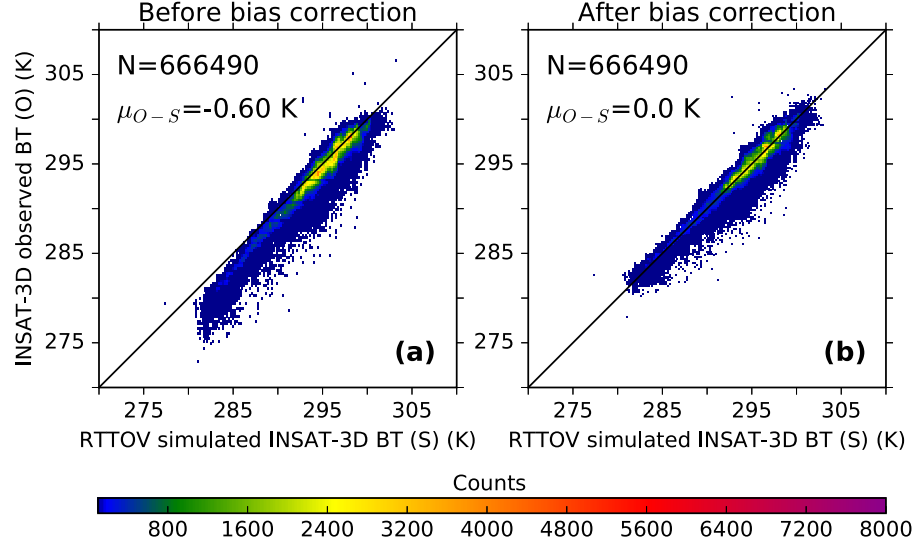


Fig. 1. Two-dimensional colour histogram plots of the RTTOV simulated and observed INSAT-3D brightness temperature of TIR1 channel, (a) before bias correction and (b) after bias correction.

observation and $F_b(y_b)$ be the PDF of the background. Let G_o and G_b be the cumulative distribution function (CDF) corresponding to F_o and F_b respectively. We made use of kernel density estimation for the calculation of F_o and F_b . Given G_o and G_b , the bias adjusted/corrected observations, y_{bc} , can be calculated as follows

$$y_{bc} = G_b^{-1}(G_o(y_o)) \quad (2)$$

It is very difficult to find the functional form of the CDFs, so the problem was solved using discrete lookup tables. For multiple spectral channel observations, the above mentioned procedure can be applied to each channel separately and it is possible to match the individual distributions of the channels (observation and background), but in doing so the correlations among the spectral channels get neglected. However, when the channels are correlated this method may lead to inconsistent results. In order to take the channel correlations into account, Inamdar et al. (2008) proposed a multidimensional probability density function matching for multispectral images. They proposed to project the observations and the background to a new space using a rotation matrix multiplication and applied the above mentioned CDF matching to individual channels in the projected space. The method for finding the suitable rotation matrix is described in detail by Inamdar et al. (2008). In this study too, the data (observation and background both) were projected to a new space before applying the CDF matching procedure. Post CDF matching, the observations were projected back to the original spectral feature space.

Unbiased observations are one pre-requisite for obtaining the optimal solution from the optimal estimation method. The bias correction procedure removes potential inconsistencies between the measurements and the forward model used in the retrieval. A sample plot of the observed radiances against the model simulated radiances, before and after bias correction, is shown in Figs. 1 and 2. It is evident from the plots (Figs. 1 and 2) that the above mentioned bias correction scheme is able to remove the bias in the INSAT-3D observations.

5. SST algorithms

5.1. INSAT-3D operational SST retrieval algorithm (OPR)

In the operational INSAT-3D algorithm (Fig. 3), the SST is derived from split thermal window channels (TIR1 and TIR2) over cloud free oceanic regions. The observation vector is augmented with the mid IR window channel (3.8-4.0 μm) during night-time. For day-time SST retrieval, the regression equation has the following form

$$SST = A_0 + A_1 T_{TIR1} + A_2 (T_{TIR1} - T_{TIR2}) + A_3 (T_{TIR1} - T_{TIR2})^2 \quad (3)$$

while during night-time the following regression equation is used

$$SST = A'_0 + A'_1 T_3 + A'_2 (T_{TIR1} - T_{TIR2}) + A'_3 (T_{TIR1} - T_{TIR2})^2 \quad (4)$$

where T_{TIR1} and T_{TIR2} are brightness temperatures for the split-window channels and T_3 is the brightness temperature corresponding to the 3.9 μm channel. The

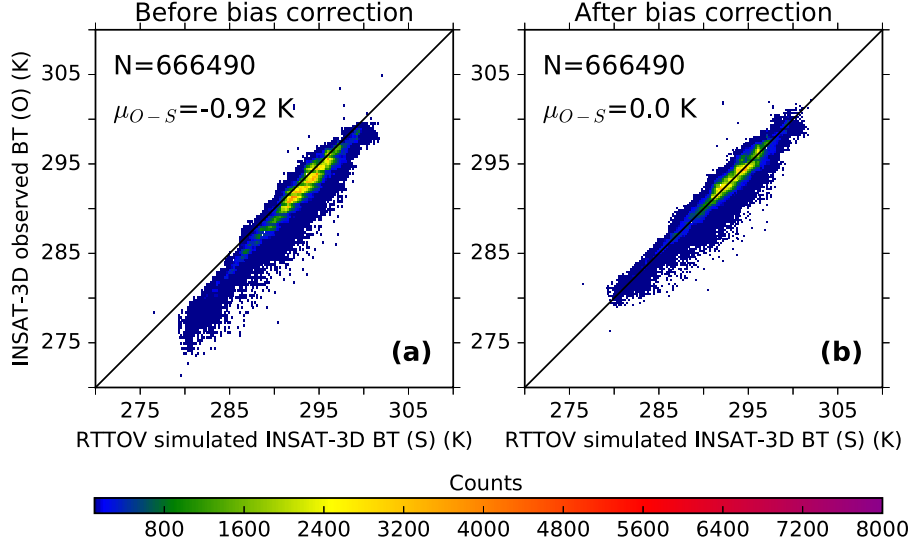


Fig. 2. Two-dimensional colour histogram plots of the RTTOV simulated and observed INSAT-3D brightness temperature of TIR2 channel, (a) before bias correction and (b) after bias correction.

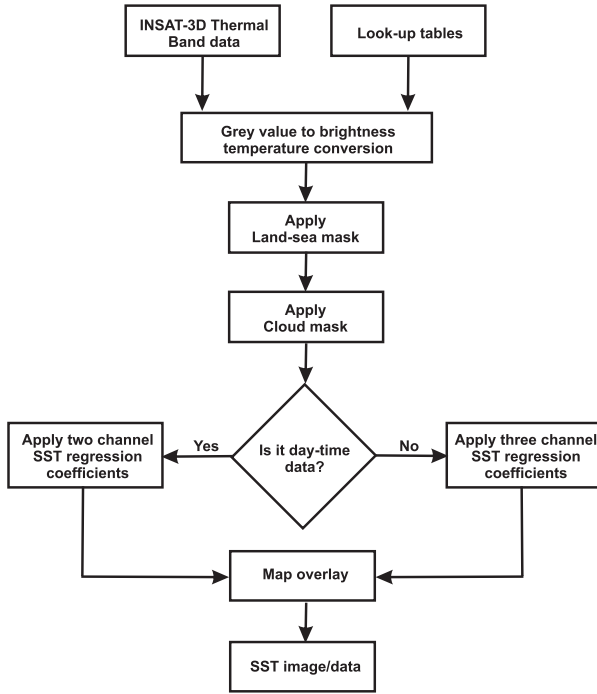


Fig. 3. Various steps of the operational INSAT-3D SST retrieval algorithm.

regression coefficients A_1 , A_2 , A_3 , A'_0 , A'_1 , A'_2 , and A'_3 are derived by regressing the model simulated INSAT-3D imager channel brightness temperatures (BTs) against the in-situ SSTs. To include the effect of satellite zenith angle (θ), the regression coefficients are generated for four values of $\sec(\theta)$, namely 1.0, 1.5, 1.75 and 2.0. INSAT-3D imager channel BTs were simulated using the

MODTRAN (MODerate resolution atmospheric TRANsmission) model (INSAT-3D ATBD, EPSA 2015). For further details on the operational INSAT-3D SST algorithm, the reader is requested to refer to INSAT-3D ATBD (EPSA 2015).

5.2 NLSST algorithm

The functional form of NLSST is given as

$$x = (a + bS)TIR1 + (c + dS + ex_c)(TIR1 - TIR2) + fS + g \quad (5)$$

where, x is the retrieved SST, $S = \sec(\theta) - 1$, θ is the satellite zenith angle and x_c is the climatological SST. For x_c , SST obtained from a daily climatology (Banzon et al., 2014) is used. More details about the NLSST can be obtained from Merchant et al. (2008). The coefficients a to g in Eq. (5) are determined by regressing the INSAT-3D simulated BTs against the in-situ SSTs using the twelve months of collocated data for the year 2016. This SST regression algorithm is referred to as REG in the subsequent text.

5.3 Optimal estimation

Let $\mathbf{y}$ be a given set of radiance observations (or equivalent black body brightness temperatures) and $\mathbf{x}_b$ a background state vector, representing the atmospheric and surface state at the observation location. The optimal estimation method seeks an optimal atmospheric state $\mathbf{x}$ that simultaneously minimises the distance between the state and the background and the state and the

observations. Under the assumption of normally distributed and uncorrelated background and observation errors, the maximum likelihood estimator of the state vector $\mathbf{x}$ can be obtained as the minimum of the following cost function $J(\mathbf{x})$:

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2}(\mathbf{y} - H(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y} - H(\mathbf{x})) \quad (6)$$

where H is the forward model simulating the observed brightness temperature $\mathbf{y}$ given the model state $\mathbf{x}$. $\mathbf{B}$ and $\mathbf{R}$ are the background and observation error covariance matrices, respectively. The optimal state $\mathbf{x}_a$ can be found by solving

$$\begin{aligned} \nabla J(\mathbf{x}) &= 0 \\ \text{or } \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) - \mathbf{H}^T \mathbf{R}^{-1}[(\mathbf{y} - H(\mathbf{x}_b)) - \mathbf{H}(\mathbf{x} - \mathbf{x}_b)] &= 0 \end{aligned} \quad (7)$$

where $\mathbf{H} = \frac{\partial H}{\partial \mathbf{x}}$.

The state vector $\mathbf{x}$ consists of vertical profiles of temperature, specific humidity, and essential surface variables. This is the general form of OEM. The solution of Eq. (7) retrieves the full state vector (at given levels). As the RT problem is non-linear, the problem is solved iteratively.

5.3.1 Optimal estimation of SST with a reduced state vector, solved in a single iteration (OEM_{rsvd}). Under the assumption that the spectral channels, used in SST retrievals, are mainly dominated by the SST and total column water vapour (TCWV), the size of the OEM can be reduced by retrieving a reduced state vector $\mathbf{z}(\mathbf{x}) = [\text{SST}, \text{TCWV}]^T$, instead of the full state vector $\mathbf{x}$. In this implementation of OEM, the background and the observation error covariance matrices, $\mathbf{B}$ and $\mathbf{R}$, were assumed to be diagonal. The diagonal elements of $\mathbf{B}$ contain the error variances of SST and TCWV. Further, assuming that the SST retrieval problem is linear, the solution can be obtained directly (in a single step), given as

$$\mathbf{z}_a = \mathbf{z}_b + \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}(\mathbf{y} - H(\mathbf{x}_b)) \quad (8)$$

The *Jacobian* matrix $\mathbf{H}$ is corresponding to the reduced state vector. For more details the reader is requested to refer to Merchant et al. (2008).

It should be noted here that although a reduced state vector is used in the retrieval problem a full state vector, $\mathbf{x}_b$, is used in the forward model. The background SST is taken from a daily climatology (SST_{clim}), while $\mathbf{T}$ and $\mathbf{q}$ profiles (for $\mathbf{x}_b$) are taken from a GFS model six-hour forecast. The first-guess TCWV is calculated by vertically integrating the $\mathbf{q}$ profiles. The observation vector $\mathbf{y}$ contains TIR1 and TIR2 observations available from the INSAT-3D imager. The background error for SST is

kept as 1.5 K while the domain (lon: 20°–130° E, lat: 50° S–50° N) averaged value is used for TCWV error. The observation error of both the IR channels is taken as 0.4 K.

5.3.2 Optimal estimation with the full state vector, solved in a single iteration (OEM_{fsvd}). It is assumed in OEM_{rsvd} that the spectral channels, used in the SST retrievals, are mainly dominated by the SST and TCWV. If this assumption is relaxed then the full state vector needs to be retrieved. Again, if the problem is still assumed linear the solution can be obtained in a single step, given as

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}(\mathbf{y} - H(\mathbf{x}_b)) \quad (9)$$

The background error covariance matrix $\mathbf{B}$ and the *Jacobian* matrix $\mathbf{H}$ are corresponding to the full state vector. Note that instead of a reduced state vector, a full state vector is retrieved in this case. The full state vector $\mathbf{x}$ consists of vertical profiles of temperature and specific humidity, at 43 vertical levels, and SST (i.e. $\mathbf{x}_b = (\mathbf{T}, \mathbf{q}, \text{SST})$). The background error covariance matrix, $\mathbf{B}$, is taken as a full covariance matrix and is averaged over the domain (oceanic region) of interest (lon: 20°–130° E, lat: 50° S–50° N). The background error covariance matrix, $\mathbf{B}$, is calculated using the NMC method (Parrish and Derber, 1992). Twelve and twenty-four hour GFS model forecasts, valid at the same time, are used for calculating the background error covariance matrix. The background SST and the observation errors are kept the same as used in OEM_{rsvd} .

5.3.3 Simple least squares solution (LSQ). If we put $\mathbf{B} = \mathbf{I}$ and $\mathbf{R} = \mathbf{0}$ in Eq. (9) then it reduces to

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}(\mathbf{y} - H(\mathbf{x}_b)) \quad (10)$$

which is the simple least squares (LSQ) solution (Rodgers, 2000) to the problem. In fact, OEM is the weighted least squares solution. Given the weights are assigned properly the optimal solution can be obtained. We have used the SST retrievals from LSQ as the baseline results for OEM_{fsvd} . Note that a full state vector is used (and retrieved) in LSQ.

5.3.4 Optimal estimation with the full state vector, solved iteratively (OEM_{fsvi}). The problem of SST retrieval is assumed linear in OEM_{fsvd} , so the solution is obtained in a single step. However, the SST retrieval problem is not linear. In a heavy water vapour loading scenario, due to the non-linear relation between the atmospheric moisture and SST, it may be difficult to find a reliable solution using only a single iteration (Koner et al. 2015). So, in this implementation of OEM, the assumption of linearity is relaxed and Eq. (7) is solved iteratively, using the

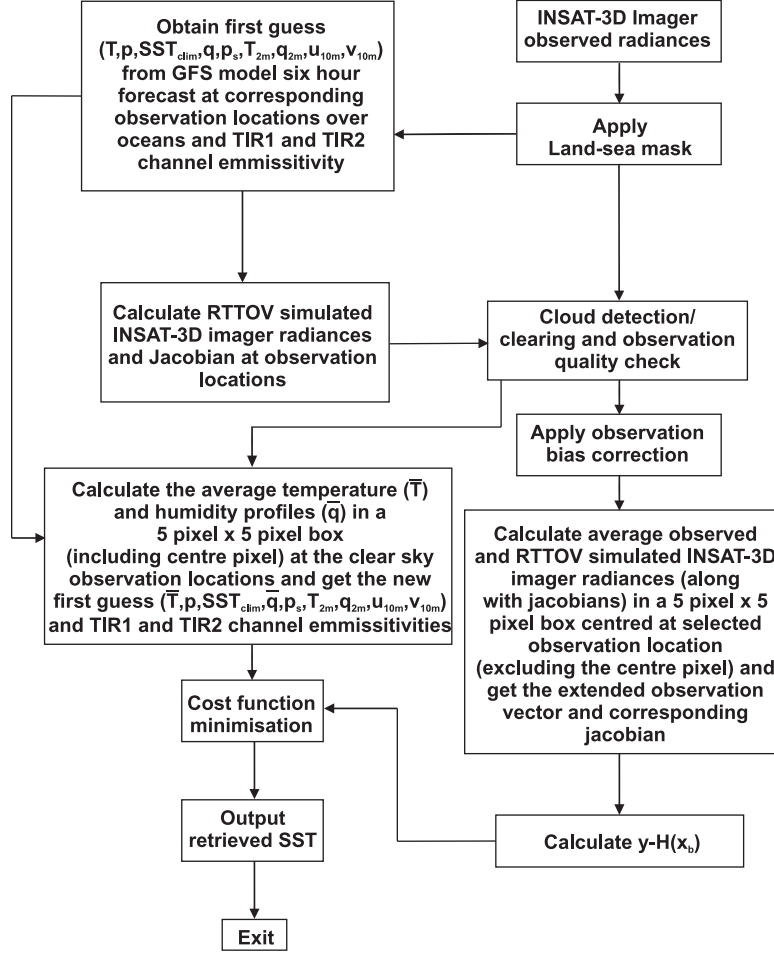


Fig. 4. Various steps of the OEM_{fsvi} SST retrieval algorithm used for INSAT-3D.

conjugate gradient algorithm. The same $\mathbf{B}$ and $\mathbf{R}$ matrices are used as in OEM_{fsvd} .

5.3.5 OEM_{fsvi} with atmospheric correction smoothing (OEM_{fsvis}). The component of the SST error variance which arises directly from the propagated radiometric sensor noise can be reduced by means of atmospheric correction smoothing (Harris and Saunders, 1996). The method of atmospheric correction smoothing in the context of optimal estimation is well described by Merchant et al. (2013), which is adopted here. This implementation of OEM is the same as OEM_{fsvi} but with the application of atmospheric correction smoothing. The $\mathbf{B}$ matrix is kept the same as in OEM_{fsvi} , while $\mathbf{R}$ is rectified as per Merchant et al. (2013), accordingly, the $\mathbf{H}$ matrix is also modified. A flowchart showing various steps of OEM_{fsvis} is shown in Fig. 4. A list of all the algorithms mentioned above, along with their key features, is provided in Table 2. The same bias corrected observations are used in REG, OEM_{rsvd} , OEM_{fsvd} , LSQ, OEM_{fsvi} and OEM_{fsvis} .

Once the optimal solution of Eq. (7) is obtained (in a single step or iteratively; for reduced as well as full state vector retrieval), the retrieval error covariance matrix ($\mathbf{S}_a$) and the averaging kernel matrix ($\mathbf{A}$) are calculated as

$$\mathbf{S}_a = (\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} + \mathbf{B}^{-1})^{-1} \quad (11)$$

$$\mathbf{A} = \mathbf{S}_a \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \quad (12)$$

The averaging kernel matrix provides the sensitivity of the retrieved variables to the true state (Rodgers, 2000), i.e.

$$\mathbf{A} = \frac{\partial \mathbf{x}_a}{\partial \mathbf{x}} \quad (13)$$

It provides an idea of the partitioning of the information determined from the observation and the background. The diagonal elements of $\mathbf{A}$ equal to unity indicates that the parameter corresponding to the diagonal element is well characterised by the observations. The value less than unity implies that *a priori* information must have been used in determining the retrieval (Line

Table 2. List of all the SST retrieval algorithms used in this study.

Algorithms	State vector used, reduced (R) or full (F)	Solution method, direct (D) or iterative (I)	Smoothing applied (Y/N)
REG	—	—	N
OEM _{rsvd}	R	D	N
LSQ	F	D	N
OEM _{fsvd}	F	D	N
OEM _{fsvi}	F	I	N
OEM _{fsvis}	F	I	Y

The algorithms are arranged in order of increasing complexity.

et al., 2012). It can be shown that for an ideal observing system, $\mathbf{A}$ is an identity matrix. But since the observations have some noise contribution also, the matrix $\mathbf{A}$ is never identity.

The averaging kernel matrix, $\mathbf{A}$, is used as an important diagnostic for determining the quality of the retrieval. The diagonal elements of $\mathbf{A}$ can be thought of as the degrees of freedom of signal (DFS) per parameter. The total DFS can be determined by calculating the trace of $\mathbf{A}$ (Rodgers, 2000). In an ideal case, the total DFS is equal to the total number of observations. The SST sensitivity for OEM_{rsvd}, OEM_{fsvd}, OEM_{fsvi} and OEM_{fsvis} were obtained from the calculation of $\mathbf{A}$ while for REG, it was calculated by differentiating the Eq. (5) with respect to SST.

It is possible to quantify the quality of the retrieval using an OEM method by evaluating the following expression (Merchant et al. 2008):

$$\chi^2 = (\mathbf{H}\mathbf{x}' - \mathbf{y}')^T [\mathbf{R}(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} \mathbf{R}]^{-1} (\mathbf{H}\mathbf{x}' - \mathbf{y}') \quad (14)$$

where $\mathbf{y}' = \mathbf{y} - \mathbf{H}(\mathbf{x}_b)$ and $\mathbf{x}' = \mathbf{x}_a - \mathbf{x}_b$. In the present study, the χ^2 value is used as an indicator of solution convergence (Rodgers, 2000).

6. Results and discussion

The skills of the algorithms namely, OPR, REG, OEM_{rsvd}, OEM_{fsvd}, LSQ, OEM_{fsvi} and OEM_{fsvis} are evaluated in terms of *bias*, root mean square error (*RMSE*) and standard deviation (*SD*) with respect to the reference dataset (iQuam SST). The *bias*, *RMSE*, and *SD* are suitable for the errors following the Gaussian distribution, but this is not always the case. Very often the error distribution is skewed and may have many outliers. In such cases, the conventional statistics may be misleading and it is recommended to report robust statistics which are not affected by the outliers (Merchant, et al. 2009). The median and median absolute deviation (*MAD*) are two such measures of robust statistics. We have used scaled median absolute deviation, *MAD_s* ($=1.4826MAD$) instead

of *MAD* so as to compare it with *SD*. In the case of Gaussian error distribution, scaled *MAD_s* equals *SD*.

The skill of the SST retrieval methods should also be compared in terms of the SST sensitivity (Merchant et al. 2009, 2013). The SST sensitivity ($\frac{\partial SST_{retrieved}}{\partial SST_{true}}$) is defined as the amount of change in the retrieved SST (*SST_{retrieved}*), for a given change in the true SST (*SST_{true}*) for all other factors being same. The ideal SST retrieval should be unbiased, have low *SD* and have a sensitivity to true SST variability close to 1 (Merchant et al. 2013). In the case of optimal estimation techniques, the sensitivity can be easily calculated from the matrices available from the retrieval process. The diagonal elements of the averaging kernel matrix, $\mathbf{A}$ (Rodgers, 2000) contains the sensitivities of the retrieved parameters while the off-diagonal elements provides the cross-sensitivities between parameters. The SST sensitivities along with the DFS are also considered while comparing the skills of the retrieval algorithms against the iQuam SST.

6.1 Comparison of SST retrievals with iQuam SST

The SSTs retrieved using OPR, REG, OEM_{rsvd}, OEM_{fsvd}, LSQ, OEM_{fsvi} and OEM_{fsvis} algorithms are validated against the iQuam SST for the eight months (October 2017–May 2018). The overall validation statistics of all the algorithms are presented in Table 3. The error statistics of the background SST (BAK) used in the SST retrievals in all the algorithms are also presented in the same table. To isolate the effect of solar illumination in the SST algorithms, the validation of the INSAT-3D SST, from the OPR, REG, OEM_{rsvd}, OEM_{fsvd}, LSQ, OEM_{fsvi} and OEM_{fsvis} algorithm, was performed during day and night separately. A pixel was defined as a day-time pixel if its solar zenith angle is less than or equal to 85°, otherwise, considered as a night-time pixel. The validation statistics separated for day and night hours is presented in Table 4. The variations of *RMSE* and *bias*, of the OPR, REG, OEM_{rsvd}, OEM_{fsvd}, LSQ, OEM_{fsvi} and OEM_{fsvis} algorithms, with respect to the iQuam SST, latitude, satellite zenith angle and TCWV are shown in Fig.

Table 3. Overall validation statistics of the INSAT-3D imager SST retrievals obtained using OPR, REG, LSQ, OEM_{rsvd} , OEM_{fsvd} , OEM_{fsvi} and OEM_{fsvis} against iQuam SST.

	Bias (K)	RMSE (K)	SD (K)	Median (K)	MAD_s (K)	$\frac{\partial SST_{ret}}{\partial SST_{true}}$	$\overline{DFS}$
OPR	-0.15	0.92	0.91	-0.21	0.89	—	—
REG	-0.17	0.70	0.68	-0.19	0.65	0.81	—
LSQ	-0.15	0.98	0.97	-0.15	0.93	1.00	2.00
OEM_{rsvd}	0.14	0.82	0.81	0.14	0.70	0.69	1.23
OEM_{fsvd}	0.23	0.71	0.67	0.21	0.63	0.66	1.13
OEM_{fsvi}	0.19	0.68	0.65	0.16	0.61	0.66	1.13
OEM_{fsvis}	0.06	0.63	0.63	0.03	0.59	0.89	1.69
BAK	0.48	0.86	0.71	0.47	0.66	—	—

The statistics of the background SST (BAK) used in these algorithms is also presented. This statistics is obtained from eight months data (October 2017–May 2018). A total of 60314 collocated pairs were used in calculating this statistics.

Table 4. Validation statistics of the SST retrievals obtained using OPR, REG, LSQ, OEM_{rsvd} , OEM_{fsvd} , OEM_{fsvi} and OEM_{fsvis} against iQuam SST, for day and night separately.

n	Day 36796			Night 23518		
	Bias (Median) (K)	RMSE (K)	SD (MAD_s) (K)	Bias (Median) (K)	RMSE (K)	SD (MAD_s) (K)
OPR	-0.26 (-0.32)	0.91	0.87 (0.84)	0.02 (-0.01)	0.95	0.95 (0.92)
REG	-0.19 (-0.20)	0.70	0.67 (0.64)	-0.15 (-0.16)	0.71	0.69 (0.67)
LSQ	-0.21 (-0.20)	1.01	0.99 (0.95)	-0.05 (-0.07)	0.93	0.93 (0.90)
OEM_{rsvd}	0.12 (0.10)	0.81	0.80 (0.70)	0.17 (0.17)	0.84	0.82 (0.70)
OEM_{fsvd}	0.20 (0.17)	0.71	0.68 (0.64)	0.29 (0.25)	0.71	0.65 (0.61)
OEM_{fsvi}	0.16 (0.14)	0.68	0.66 (0.61)	0.24 (0.21)	0.68	0.63 (0.60)
OEM_{fsvis}	0.04 (0.0)	0.63	0.63 (0.58)	0.01 (0.07)	0.64	0.63 (0.61)
BAK	0.51 (0.50)	0.91	0.76 (0.72)	0.42 (0.43)	0.77	0.64 (0.58)

The statistics of the background SST (BAK) used in these algorithms is also presented. This statistics is obtained from eight months data (October 2017–May 2018).

5. The relative performance of all the algorithms is discussed in the following sections.

6.1.1 Analysis of OPR vs REG. The results (in Table 3) show that, overall, the values of the bias (median), RMSE and SD (MAD_s) for the OPR are -0.15 (-0.21), 0.92 and 0.91 (0.89) K respectively. The RMSE values are higher during night time (0.95 K) as compared to day-time (0.91 K). However, the bias during night-time is quite low (0.02 K) as compared day-time (-0.26). Large improvement in error statistics is observed with the use of bias-corrected radiances and the NLSST algorithm (REG). The REG retrievals have RMSE of 0.70 K and the bias almost similar to OPR. Although there is a small difference in the REG bias during the day (-0.19 K) and night (-0.15 K) but almost no change in the RMSE values is reflected in the results. Overall, REG is found to be better than the OPR but the plots in Fig. 5 suggest that the error statistics of both the algorithms are highly

dependent on the temperature (SST), latitude, satellite zenith angle, and TCWV. Too much variation in OPR error statistics (RMSE: 0.8 to 1.25 K and bias: -0.8 to 0.5 K) is seen with the change in satellite viewing angle from 10° to 50° . Similarly, the effect of TCWV variability on the OPR and REG retrievals is also clearly evident from the plots in Fig. 5. The OPR and REG are both regression-based algorithms but have different forms. The radiance observations used in REG are bias-corrected while those in OPR are not. The contrast difference in the error statistics of both the retrievals suggests that the poor performance of the OPR may be due to the biased observations or due to the unnecessarily complex nature of the algorithm. It should be noted here that $3.9 \mu\text{m}$ channel observations are also used in OPR during night-time.

6.1.2 Analysis of REG and OEM_{rsvd} . When compared with REG, the OEM_{rsvd} SST retrievals are not very good,

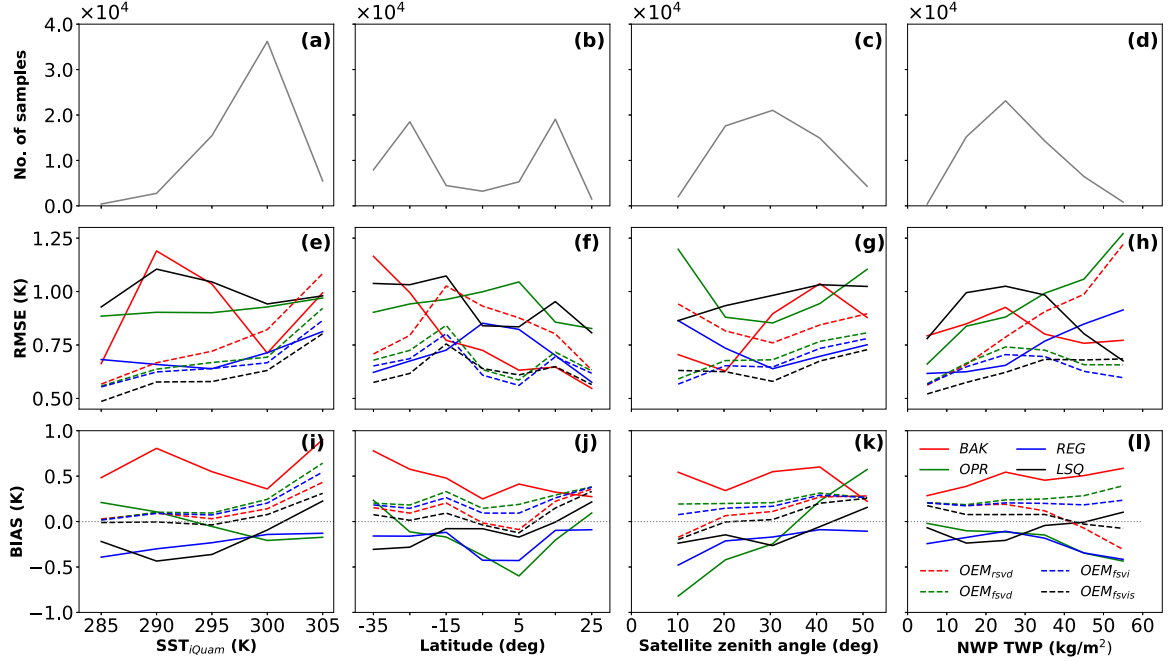


Fig. 5. Variation of the SST bias and RMSE with respect to (a) iQuam SST, (b) latitude, (c) satellite zenith angle and (d) TCWV.

though there is an improvement over the BAK. The OEM_{rsvd} retrievals have overall RMSE (bias) of about 0.82 (0.17) K. The effect of solar illumination is also reflected in the retrieval statistics of OEM_{rsvd} . The average sensitivity of OEM_{rsvd} retrievals is about 0.69 while that of REG is 0.81. The bias in OEM_{rsvd} retrievals show less variation with SST, latitude, satellite zenith angle and TCWV while RMSE seems to have much dependency on these variables (Fig. 5). The χ^2 value is used as a quality indicator of the OEM_{rsvd} retrievals. To investigate further, we have calculated the error statistics of the OEM_{rsvd} retrievals with varying χ^2 values. The REG algorithm has no quality indicator similar to χ^2 . The $\chi^2 \leq 1$ indicates optimal retrievals. The plots in Fig. 6 reveals that the OEM_{rsvd} retrievals, with $\chi^2 \leq 1$, have RMSE of about 0.68 K and bias of about 0.21 K. The percentage of good retrievals ($\chi^2 \leq 1$) is about 80% (Table 5).

6.1.3 Analysis of OEM_{rsvd} vs OEM_{fsvd} and LSQ. All the three methods, OEM_{rsvd} , OEM_{fsvd} and LSQ, have some similarities among them. The difference in OEM_{rsvd} and OEM_{fsvd} is the size of the state vector, the rest of the things are the same. In both these methods, the background and the observations are assigned weights according to their assumed errors. In LSQ, the observations are assumed to be correct thus $\mathbf{R}$ is set to zero while $\mathbf{B}$ is taken as an identity matrix. As expected (considering their formulations) the LSQ retrievals have the worst overall quality (RMSE = 0.98 and bias = -0.15) amongst

the three. The impact of using the full state vector in OEM_{fsvd} is reflected in its overall error statistics. The RMSE got reduced to 0.71 K as compared to 0.81 K for OEM_{rsvd} . The overall bias in OEM_{fsvd} retrievals is larger (0.23 K) relative to OEM_{rsvd} bias (0.14 K). The effect of the solar zenith angle is seen in the bias of OEM_{fsvd} retrievals. However, RMSE values are the same during day and night. In terms of sensitivity, OEM_{rsvd} retrievals are better with about 5% higher average sensitivity than OEM_{fsvd} retrievals. The average DFS for OEM_{rsvd} is also higher (1.23) than that of OEM_{fsvd} (1.13). The χ^2 plots (Fig. 6) show that the OEM_{fsvd} has better performance for all values of χ^2 , in terms of bias and RMSE, when compared with OEM_{rsvd} . More than 82% OEM_{fsvd} retrievals have achieved good convergence ($\chi^2 \leq 1$) with RMSE and bias of 0.67 and 0.24 K respectively (Table 5).

6.1.4 Analysis of OEM_{fsvd} vs OEM_{fsvi} . The OEM_{fsvd} and OEM_{fsvi} are similar in their formulation but differ in their solution method. The solution in OEM_{fsvd} is obtained in a single step while the problem is solved iteratively in OEM_{fsvi} . The idea of comparing OEM_{fsvd} and OEM_{fsvi} is to address the question if an iterative solution to the problem of SST retrieval (using optimal estimation) can achieve any gain, in retrieval accuracy, over the direct method (single step)? The answer to this question is clear from the results presented in Table 3. The iterative solution results in a reduction of the overall retrieval bias

(0.19 K from 0.23 K for OEM_{fsvd}) and RMSE (0.68 K from 0.71 K for OEM_{fsvd}). The OEM_{fsvi} performs equally well during day and night in terms of RMSE (0.68 K). However, the OEM_{fsvi} retrievals bias is found to be larger during night (0.24 K) as compared to day (0.16 K). The

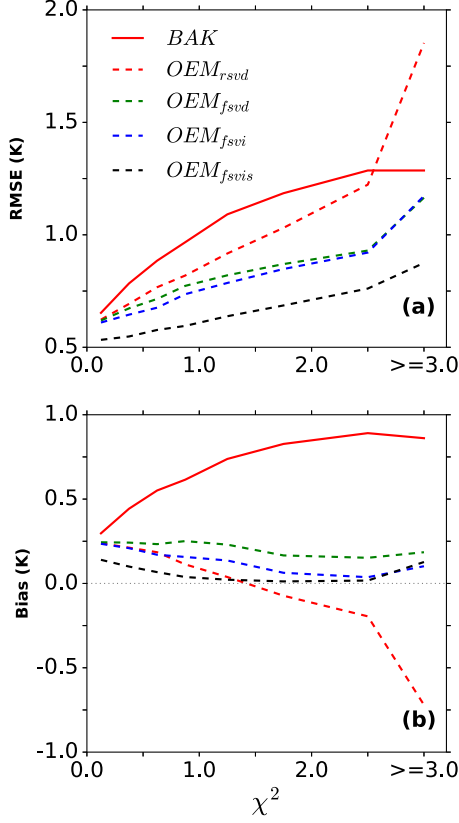


Fig. 6. RMSE and bias plots of SST retrievals, from different methods, as a function of corresponding χ^2 values. Values of the χ^2 for the corresponding methods are used for generating these plots. For BAK, χ^2 values of OEM_{rsvd} are used.

retrieval errors are less for OEM_{fsvi} as compared to OEM_{fsvd} for all almost all values of χ^2 (Fig. 6). About 84% OEM_{fsvi} retrievals have $\chi^2 \leq 1$ (Table 5) with low RMSE (0.64) and bias (0.21 K). The better performance of OEM_{fsvi} is also supported by the plots in Fig. 5. The bias and RMSE for OEM_{fsvi} is lower than that of OEM_{fsvd} for all values SST, latitude, satellite zenith angle and TCWV. These results suggest that solving the SST optimal estimation problem iteratively can improve its retrieval quality.

6.1.5 Analysis of OEM_{fsvi} vs OEM_{fsvis} . The results discussed above clearly demonstrate that a significant improvement in SST retrieval quality can be achieved with an optimal estimation if a full state vector is used and the problem is solved iteratively instead of using the reduced state vector and solving the problem in a single step. It is already shown by Merchant et al. (2013) that the SST retrieval accuracy can be improved with the use of atmospheric correction smoothing (ACS). However, they applied it on OEM_{rsvd} . We have applied ACS in OEM_{fsvis} and compared the results of the retrievals with those from OEM_{fsvi} . The retrieval error statistics in Table 3 shows that the application of ACS leads to a significant improvement in the retrieval bias (0.06 K) and RMSE (0.63 K). It is clear from the error statistics that the noise in the retrieved SST, from OEM_{fsvis} , has been reduced to a large extent as a result of atmospheric correction smoothing. The OEM_{fsvis} retrievals have almost similar bias and RMSE during day and night. The errors of OEM_{fsvis} are less than those of OEM_{fsvi} for all values of χ^2 (Fig. 6). More than 84% OEM_{fsvis} have RMSE (bias) of about 0.59 K (0.03 K) (Table 5). There is a remarkable improvement in the average SST sensitivity for OEM_{fsvis} (89% compared to 66% for OEM_{fsvi}). It is also supported by the spatial plots of the monthly

Table 5. Validation statistics of OEM_{rsvd} , OEM_{fsvd} , OEM_{fsvi} and OEM_{fsvis} calculated for different values of χ^2 .

Algorithm	χ^2	Bias (K)	RMSE (K)	SD (K)	Median (K)	MAD_s (K)	N
OEM_{rsvd}	$\chi^2 \leq 1$	0.21	0.68	0.65	0.19	0.62	48024 (~80%)
	$1 < \chi^2 \leq 2$	0.0	0.96	0.96	-0.16	0.87	7551 (~12%)
	$2 < \chi^2$	-0.43	1.54	1.48	-0.71	1.13	4739 (~8%)
OEM_{fsvd}	$\chi^2 \leq 1$	0.24	0.67	0.62	0.23	0.59	49731 (~82%)
	$1 < \chi^2 \leq 2$	0.21	0.84	0.81	0.11	0.77	7836 (~13%)
	$2 < \chi^2$	0.17	0.99	0.98	-0.01	0.81	2747 (~5%)
OEM_{fsvi}	$\chi^2 \leq 1$	0.21	0.64	0.60	0.19	0.58	50925 (~84%)
	$1 < \chi^2 \leq 2$	0.11	0.80	0.80	-0.02	0.72	7005 (~12%)
	$2 < \chi^2$	0.05	0.99	0.99	-0.16	0.75	2384 (~4%)
OEM_{fsvis}	$\chi^2 \leq 1$	0.08	0.56	0.56	0.06	0.53	34951 (~58%)
	$1 < \chi^2 \leq 2$	0.02	0.66	0.66	-0.04	0.62	15856 (~26%)
	$2 < \chi^2$	0.06	0.81	0.81	-0.07	0.80	9507 (~16%)

Values of the χ^2 for the corresponding methods are used while calculating these statistics.

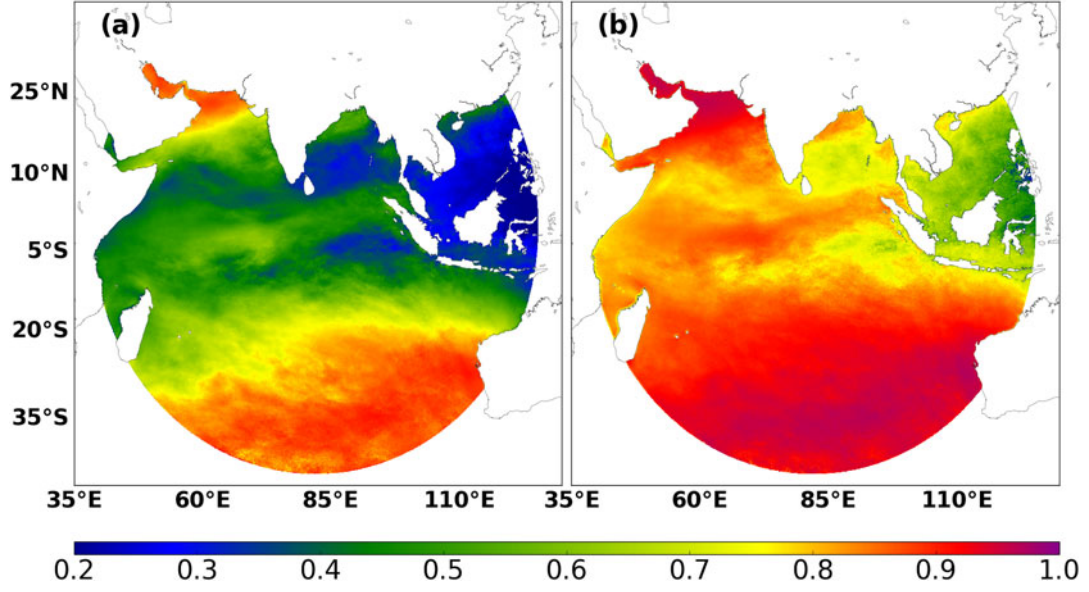


Fig. 7 Spatial maps of the monthly average SST sensitivity, for SST retrieved (a) from OEM_{fsvi} and (b) OEM_{fsvis} , for the month of October 2017.

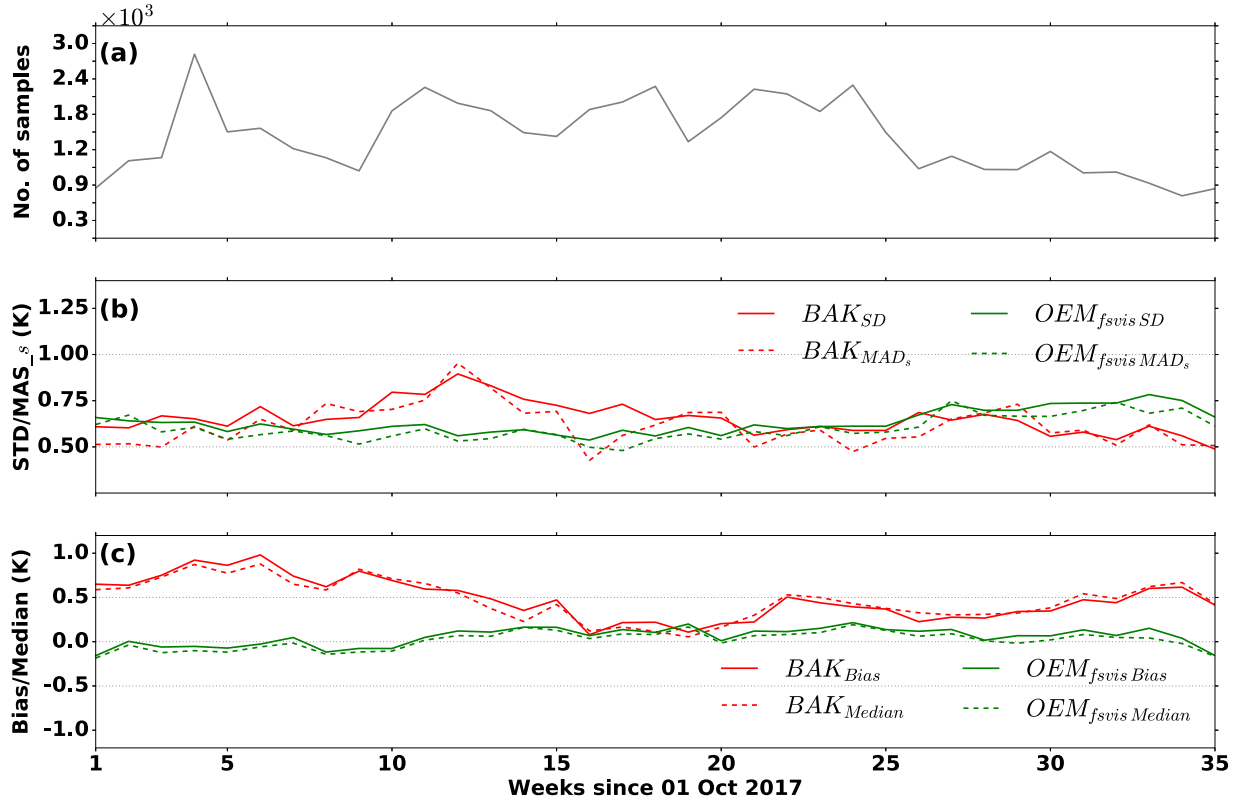


Fig. 8. Weekly time series of the error statistics, day and night-time combined, of BAK and OEM_{fsvis} SST.

average SST sensitivity for OEM_{fsvi} and OEM_{fsvis} shown in Fig. 7. Without ACS (OEM_{fsvi}), the average DFS is found to be 1.13 which has increased to about 1.5 fold

with ACS (OEM_{fsvis}). Retrieval errors of OEM_{fsvis} are less for almost all ranges of SST, latitude, satellite, and TCWV (except a few exceptions) as compared to errors

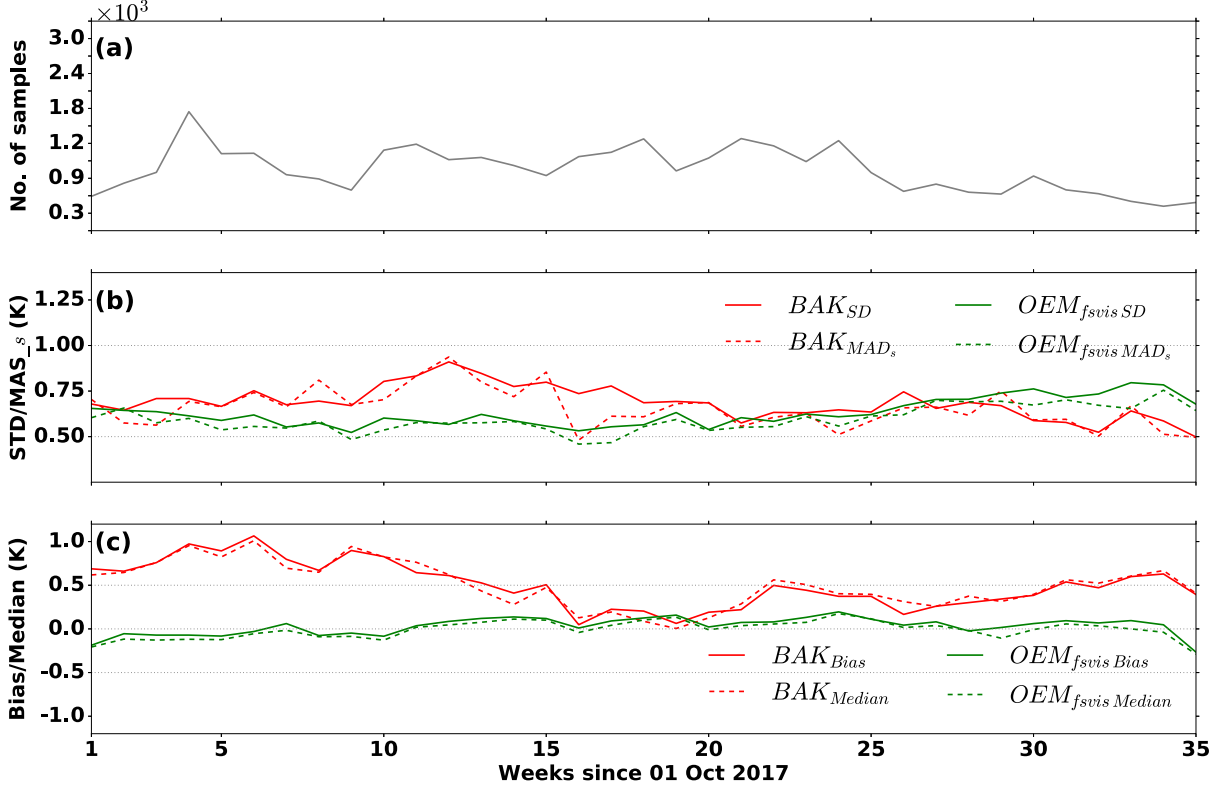


Fig. 9. Weekly time series of the error statistics, during day-time, of BAK and OEM_{fsvs} SST.

of all other methods. These results clearly suggest that the OEM_{fsvs} outperforms OPR, REG, OEM_{rsvd}, OEM_{fsvd} and OEM_{fsvi}.

To observe the skill of OEM_{fsvs} with varying time, the weekly time series of the error statistics, as obtained against the iQuam SST, for eight months (Oct 2017–May 2018) is plotted and shown in Fig. 8. It is evident from the time series plots of bias that the *prior* (BAK) SST has a large bias and is highly variable in time while the OEM_{fsvs} bias is much less and is almost constant in time. Overall, SD values of the OEM_{fsvs} are also consistently smaller as compared to BAK SST. The weekly time series of the error statistics are also plotted separately for day (Fig. 9) and night (Fig. 10) time. The performance of OEM_{fsvs} seems to be degraded, in terms of SD, after mid April 2018 (corresponds to 28th week in Fig. 9). During night-time, the skill of the OEM_{fsvs} retrievals is not good in terms of SD (Fig. 10). However, the OEM_{fsvs} night-time retrievals have smaller bias than the BAK SST.

6.2. SST monthly maps

In order to observe the large scale features in the retrieved SST, the monthly average SST maps from REG, OEM_{fsvs} and OSI-SAF-SST, for the month of

October 2017, are generated. The monthly maps are generated by calculating the average of all the samples obtained (every hourly sample, since OSI-SAF-SST is available hourly) throughout the month. A plot of the monthly maps is shown in Fig. 11. The large scale features in the REG and OEM_{fsvs} SST match well with those observed in the OSI-SAF-SST monthly maps. A close look at the monthly maps suggest that the REG algorithm overestimates SST when compared with OSI-SAF-SST. In order to assess the improvement in the SST due to OEM_{fsvs} algorithm (over REG), we define the improvement parameter

$$\alpha = \text{abs}(SEVIRI_{\text{avg}} - REG_{\text{avg}}) - \text{abs}(SEVIRI_{\text{avg}} - OEM_{\text{avg}}) \quad (15)$$

Here, SEVIRI refers to OSI-SAF-SST, and OEM refers to OEM_{fsvs}. The positive (negative) value of α indicates the improvement (degradation) in OEM (OEM_{fsvs}) as compared to REG. The spatial map of α is shown in Fig. 12. Except for a few isolated locations the value of α is positive which indicates that overall the OEM_{fsvs} SST is better than the REG when a comparison is made with the SEVIRI SST. The improvement, as large as 1.5 K, is seen particularly over the equatorial Indian Ocean.

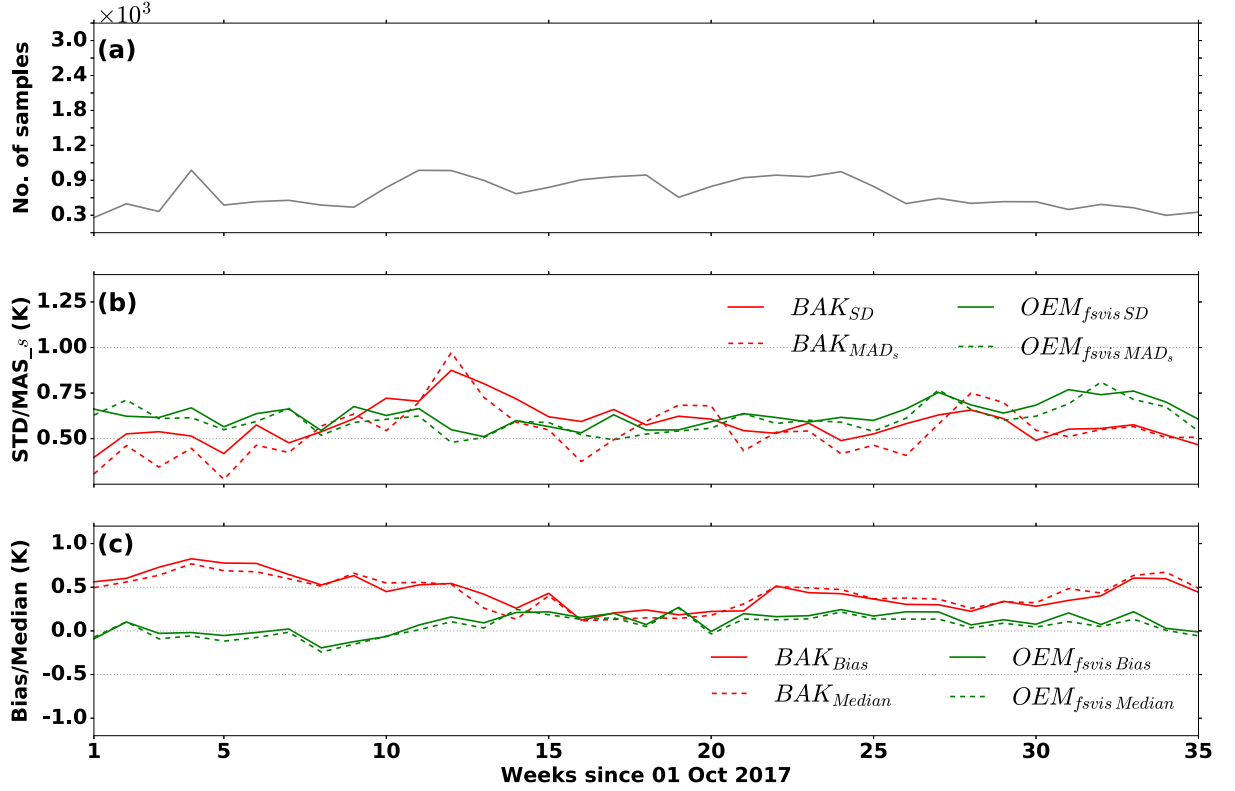


Fig. 10. Weekly time series of the error statistics, during night-time, of BAK and OEM_{fsvs} SST.

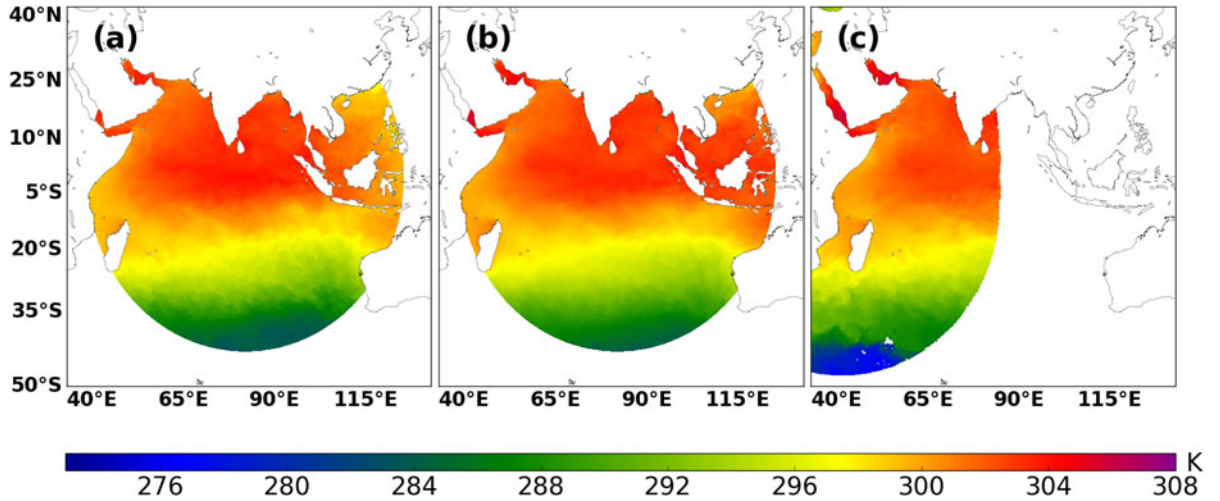


Fig. 11. Monthly average SST for the month of October 2017. (a) SST retrieved from INSAT-3D observations using REG, (b) SST retrieved from INSAT-3D observations using OEM_{fsvs} and (c) OSI-SAF-SST.

7. Conclusions

The accurate SST observations over the Indian Ocean have immense importance for various atmospheric and oceanic studies. Presently, only INSAT-3D/3R satellites are the dedicated geostationary satellites providing high

temporal resolution (~ 30 min) observations over the Indian Ocean. The operational SST from INSAT-3D imager is regression-based and has large biases. The large SST biases are due to the radiance bias and the inherent deficiency of the regression algorithm. With the intention

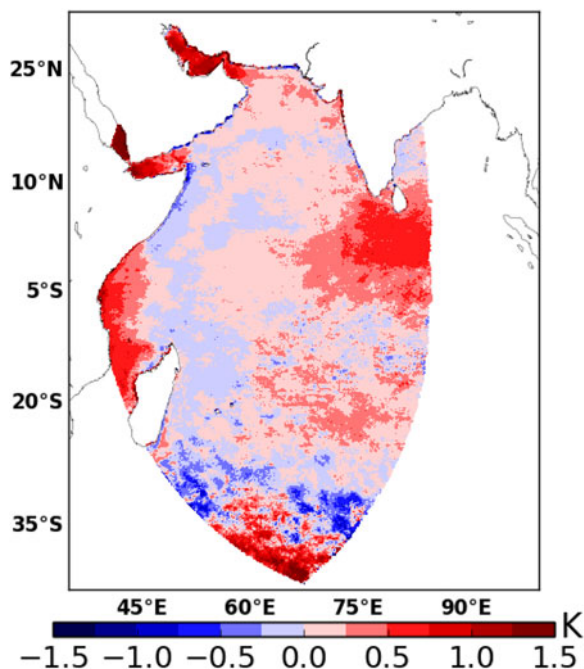


Fig. 12. Spatial distribution of the improvement parameter, α .

of providing good quality SST observations over the Indian Ocean, a radiance bias correction method is developed. A new regression-based SST algorithm similar to NLSST is developed and applied on the bias-corrected radiances. It is shown that with the use of bias-corrected INSAT-3D observations, in the newly developed regression algorithm, reasonable estimates of SST can be achieved. It is demonstrated that the optimal estimation of SST using the full state vector and iteratively solving the problem can further improve the SST retrieval accuracy over the regression-based algorithm. The atmospheric correction smoothing is applied to the full state vector and the iterative version of optimal estimation algorithm (OEM_{fsvs}) to quantify its impact on SST retrieval error. Results reveal that the atmospheric correction smoothing leads to overall improved SST retrievals in terms of all the skill parameters. The overall SD (MAD_s), with respect to the iQuam SST, of the OEM_{fsvs} SST retrieval was obtained as 0.63 (0.58). The bias in the SST retrieved using the OEM_{fsvs} algorithm is considerably smaller (0.06 K) as compared to the bias in the regression algorithm (−0.19 K). The weekly time series of the validation statistics shows a consistent performance of OEM_{fsvs} algorithm with time. The comparative statistics of the regression-based algorithm and OEM_{fsvs} clearly suggests that the OEM_{fsvs} is more robust. Improved high spatial and temporal resolution SST from INSAT-3D/3DR imager can enhance the SST observing system over the Indian Ocean. There is no doubt that the radiative

transfer model based retrieval algorithms (like optimal estimation) are computationally expensive. Further, solving the problem iteratively with the full state vector will make it more resource demanding. Now-a-days almost every computer is equipped with multi-core processors and the state-of-the-art radiative transfer models are also developed with the capability to utilise them. We would like to mention here that the iterative retrieval using full state vector is what is applied when satellite radiances are used directly in variational data assimilation (3D-var and 4D-var) for NWP. It was only when this step was introduced that one could prove a significant positive impact of satellite radiances in NWP. So, the computational resources should not be considered as a limiting factor in the application of advanced algorithms for SST retrieval.

Acknowledgments

The authors would like to thank Director SAC, for the constant encouragement and guidance to undertake this work. The support from Deputy Director, EPSA/SAC, Group Head, MRG/EPSC and Head, MRTD/MRG/EPSC is gratefully acknowledged. The fortran codes used in this study were made parallel with the help of Mr. Nitesh Kumar Kaushik, SIPG/ODPD/SAC. Authors would like to thank Dr. Neeraj Agarwal, AOSG/EPSC/SAC for the valuable discussions during the course of this work. The anonymous reviewers critical comments and valuable suggestions led to the significant improvements in the quality of the manuscript. The authors acknowledge the NWP-SAF for making RTTOV freely available. The GFS model forecasts used in this study were obtained from ncdc.noaa.gov. The INSAT-3D imager radiances and the operational INSAT-3D SST product were obtained from MOSDAC (<http://www.mosdac.gov.in>). The OSI-SAF-SST product was obtained from <ftp://eftp.ifremer.fr> and iQuam data was obtained from <https://www.star.nesdis.noaa.gov/sod/sst/iquam>.

References

- Banzon, V. F., Reynolds, R. W., Stokes, D. and Xue, Y. 2014. A 1/4°-spatial-resolution daily sea surface temperature climatology based on a blended satellite and in situ analysis. *J. Climate* **27**, 8221–8228. doi:10.1175/JCLI-D-14-00293.1
- Cracknell, A. P. 1997. *The Advanced Very High Resolution Radiometer (OVAR)*. Taylor & Francis, London.
- Donlon, C. J., Casey, K. S., Robinson, I. S., Gentemann, C. L., Reynolds, R. W., and co-authors. 2009. The GODAE high-resolution sea surface temperature pilot project. *Oceanography* **22**, 34–45. doi:10.5670/oceanog.2009.64
- Emery, W. J., Castro, S., Wick, G. A., Schluessel, P. and Donlon, C. 2001. Estimating sea surface temperature from

- infrared satellite and in situ temperature data. *Bull. Amer. Meteor. Soc.* **82**, 2773–2786. doi:10.1175/1520-0477(2001)082<2773:ESSTFI>2.3.CO;2
- Esias, W. E., Abbott, M. R., Barton, I., Brown, O. B., Campbell, J. W., Carder, K. L., and co-authors. 1998. An overview of MODIS capabilities for ocean science observations. *IEEE Trans. Geosci. Remote Sens.* **36**, 1250–1265. doi:10.1109/36.701076
- EPSA. 2015. INSAT-3D Algorithm Theoretical Basis Document. Space Applications Centre, Government of India. Online at: http://mosdac.gov.in/doc/INSAT_3D_ATBD_MAY_2015.pdf.
- Xu, F. and Ignatov, A. 2014. In situ SST Quality Monitor (iQuam). *J. Atmos. Oceanic Technol.* **31**, 164–180. doi:10.1175/JTECH-D-13-00121.1
- Gentemann, C. L. 2014. Three way validation of MODIS and AMSR-E sea surface temperatures. *J. Geophys. Res. Oceans* **119**, 2583–2598. doi:10.1002/2013JC009716
- Harris, A. R. and Saunders, M. A. 1996. Global validation of the along-track scanning radiometer against drifting buoys. *J. Geophys. Res.* **101**, 12127–12140. doi:10.1029/96JC00317
- Hewison, T. J. 2013. an evaluation of the uncertainty of the GSICS SEVIRI-IASI intercalibration products. *IEEE-TGRS* **51**, 3,
- Hillger, D. W. and Vonder Haar, T. H. 1988. Estimating noise levels of remotely sensed measurements from satellites using spatial structure analysis. *J. Atmos. Oceanic Technol.* **5**, 206–214. doi:10.1175/1520-0426(1988)005<0206:ENLORS>2.0.CO;2
- Inamdar, S., Bovolo, F., Bruzzone, L. and Chaudhuri, S. 2008. Multidimensional probability density function matching for preprocessing of multitemporal remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **46**, 1243–1252. doi:10.1109/TGRS.2007.912445
- Kilpatrick, K. A., Podesta, G., Walsh, S., Williams, E., Halliwell, V. and co-authors. 2015. A decade of sea surface temperature from MODIS. *Remote Sens. Environ.* **165**, 27–41. doi:10.1016/j.rse.2015.04.023
- Koner, P. K., Harris, A. R. and Maturi, E. 2015. A physical deterministic inverse method for operational satellite remote sensing: An application for sea surface temperature retrievals. *IEEE Trans. Geosci. Remote Sens.* **53**, 5872–5888. doi:10.1109/TGRS.2015.2424219
- Koner, P. K. and Harris, A. 2016. Improved quality of MODIS sea surface temperature retrieval and data coverage using physical deterministic methods. *Remote Sens.* **8**, 454. doi:10.3390/rs8060454
- Le Borgne, P., Legendre, G. and Marsouin, A. 2006. SST retrieval from MSG/SEVIRI data paper presented at The Meteorological Satellite Conference Eur Organ for the Exploit of Meteorol Satell Helsinki.
- Line, M. R., Zhang, X., Vasisht, G., Natraj, V., Pin, C., and co-authors. 2012. Information content of exoplanetary transit spectra: An initial look. *Astrophys. J.* **749**, 93.
- McMillin, L. M. and Crosby, D. S. 1984. Theory and validation of the multiple window sea surface temperature technique. *J. Geophys. Res.* **89**, 3655–3661. doi:10.1029/JC089iC03p03655
- Merchant, C. J., Le Borgne, P., Roquet, H. and Legendre, G. 2013. Extended optimal estimation techniques for sea surface temperature from the Spinning Enhanced Visible and Infra-Red Imager (SEVIRI). *Remote Sens. Environ.* **131**, 287–297. doi:10.1016/j.rse.2012.12.019
- Merchant, C. J., Harris, A. R., Roquet, H. and Le Borgne, P. 2009. Retrieval characteristics of non-linear sea surface temperature from Advanced Very High Resolution Radiometer. *Geophys. Res. Lett.* **36**, L17604. doi:10.1029/2009GL039843
- Merchant, C. J., Horrocks, L. A., Eyre, J. and O’Carroll, A. G. 2006. Retrievals of sea surface temperature from infra-red imagery: Origin and form of systematic errors. *Quart. J. Royal Met. Soc.* **132**, 1205–1223.
- Merchant, C. J. and Le Borgne, P. 2004. Retrieval of sea surface temperature from space based on modeling of infrared radiative transfer: Capabilities and limitations. *J. Atmos. Oceanic Technol.* **22**, 1734–1746. <http://dx.doi.org/10.1175/JTECH1667.1>.
- Merchant, C. J., Le Borgne, P., Marsouin, A. and Roquet, H. 2008. Optimal estimation of sea surface temperature from split-window observations. *Remote Sens. Environ.* **112**, 2469–2484. doi:10.1016/j.rse.2007.11.011
- Minnett, P. J. (2010). The validation of sea surface temperature retrievals from spaceborne infrared radiometers. In: *Oceanography from Space, Revisited* (ed. V. Barale). Springer, Heidelberg, Germany, pp. 273–295.
- Mutlow, C. T., Zavody, A. M., Barton, I. J. and Llewellyn-Jones, D. T. 1994. Sea surface temperature measurements by the along-track scanning radiometer on the ERS-1 satellite: Early results. *J. Geophys. Res.* **99**, 22575–22588. doi:10.1029/94JC01758
- Nielsen-Englyst, P., Høyer, J. L., Toudal Pedersen, L., Gentemann, C. L., Alerskans, E. and co-authors. 2018. Optimal estimation of sea surface temperature from AMSR-E. *Remote Sens.* **10**, 229.
- Parrish, D. F. and Derber, J. C. 1992. The national meteorological center’s spectral statistical-interpolation analysis system. *Mon. Wea. Rev.* **120**, 1747–1763. doi:10.1175/1520-0493(1992)120<1747:TNMCS>2.0.CO;2
- Prigent, C., Aires, F., Bernardo, F., Orhac, J. C., Goutoule, J. M. and co-authors. 2013. Analysis of the potential and limitations of microwave radiometry for the retrieval of sea surface temperature: Definition of MICROWAT, a new mission concept. *J. Geophys. Res. Oceans* **118**, 3074–3086. doi:10.1002/jgrc.20222
- Ricciardulli, L. and Wentz, F. J. 2004. Uncertainties in sea surface temperature retrievals from space: Comparison of microwave and infrared observations from TRMM. *J. Geophys. Res.* **109**, C12013.
- Rodgers, C. D. 2000. *Inverse Methods for Atmospheric Sounding, Theory and Practice* World Scientific, Singapore.
- Saunders, R. W., Matricardi, M. and Brunel, P. 1999. An improved fast radiative transfer model for assimilation of satellite radiance observations. *Q.J.R. Meteorol. Soc.* **125**, 1407–1425. doi:10.1002/qj.1999.49712555615
- Saunders, R., Rayner, P., Brunel, P., von Engeln, A., Bormann, N. and co-authors. 2007. A comparison of radiative transfer

- models for simulating Atmospheric Infrared Sounder (AIRS) radiances. *J. Geophys. Res.* **112**, D01–S90.
- Schmit, T. J., Gunshor, M. M., Menzel, W. P., Gurka, J. J., Li, J. and co-authors. 2005. Introducing the next-generation Advanced Baseline Imager on GOES-R. *Bull. Am. Meteor. Soc.* **86**, 1079–1096. doi:[10.1175/BAMS-86-8-1079](https://doi.org/10.1175/BAMS-86-8-1079)
- Sherlock, V. J. 1999. ISEM-6: Infrared Surface Emissivity Model for RTTOV-6. *NWP Tech. Rep.* **287**.
- Singh, R., Ojha, S. P., Kishtawal, C. M., Pal, P. K. and Kiran Kumar, A. S. 2016a. Impact of the assimilation of INSAT-3D radiances on short-range weather forecasts. *QJR Meteorol. Soc.* **142**, 120–131. doi:[10.1002/qj.2636](https://doi.org/10.1002/qj.2636)
- Singh, R., Singh, C., Ojha, S. P., Kumar, A. S., Kishtawal, C. M. and co-authors. 2016b. Land surface temperature from INSAT-3D imager data: Retrieval and assimilation in NWP model. *J. Geophys. Res. Atmos.* **121**, 6909–6926. doi:[10.1002/2016JD024752](https://doi.org/10.1002/2016JD024752)
- Walton, C. C., Pichel, W. G., Sapper, J. F. and May, D. A. 1998. The development and operational application of nonlinear algorithms for the measurement of sea surface temperatures with the NOAA polarorbiting environmental satellites. *J. Geophys. Res.* **103**, 27999–28012. doi:[10.1029/98JC02370](https://doi.org/10.1029/98JC02370)
- Zavody, A. M., Mutlow, C. and Llewellynjones, D. T. 1995. A radiative-transfer model for sea-surface temperature retrieval for the along-track scanning radiometer. *J. Geophys. Res.* **100**.
- Zhuge, X. and Zou, X. 2016. Test of a modified infrared-only ABI cloud mask algorithm for ahi radiance observations. *J. Appl. Meteor. Climatol.* **55**, 2529–2546. doi:[10.1175/JAMC-D-16-0254.1](https://doi.org/10.1175/JAMC-D-16-0254.1)
- EUMETSAT/OSI SAF. 2017. GHRSSST Level 3C Indian-Ocean (IO) sub-skin Sea Surface Temperature from the Spinning Enhanced Visible and Infrared Imager (SEVIRI) on MSG1 in GDS2 format produced by OSISAF.