TECHNICAL REPORT **Report No.: ESSO-INCOIS-ODG-TPG-TR-04(2018)**

Marine Meteorological Atlas of Tropical Indian Ocean - Fluxes and Observational error analysis

by

Kameshwari N, TVS Uday Bhaskar, E Pattabhi Rama Rao

Indian National Centre for Ocean Information Services (INCOIS)

(Earth System Science Organization (ESSO),Ministry of Earth Sciences (MoES))

HYDERABAD, INDIA

www.incois.gov.in

October, 2018

DOCUMENT CONTROL SHEET

Earth System Science Organization (ESSO)

Ministry of Earth Sciences (MoES)

Indian National Centre for Ocean Information Services (INCOIS)

ESSO Document Number: ESSO-INCOIS-ODG-TPG-TR-04(2018)

Title of the report: Marine Meteorological Atlas of Tropical Indian Ocean - Fluxes and Observational error analysis

Author(s) [Last name, First name]:

Kameshwari Nunna, TVS Uday Bhaskar, E Pattabhi Rama Rao

Originating unit

Ocean Observations and Data Management Group (ODG) and Training and programme Planning and Management Group (TPG), INCOIS

Type of Document:

Technical Report (TR)

Number of pages and figures: 28, 13

Number of references: 23

Keywords:

Marine-MET observations, ICOADS, Objective Analysis, Semi Variogram technique, Net heat flux

Security classification:

Open

Distribution:

Open

Date of publication:

October, 2018

Abstract (170 words) :

This paper discusses the preparation of a new climatology Marine Meteorological Atlas for Tropical Indian Ocean (MaMetAtTIO) from ship observations that are used in preparation of International Comprehensive Ocean-Atmosphere Dataset (ICOADS) Release 3 and those obtained from the Indian Meteorological Department (IMD) and Naval Operations Data Processing and Analysis Centre (NODPAC). The enhancement in MaMetAtTIO is checked with reference to ICOADS R3.0 climatology. ICOADS R3.0 dataset has been found to be self robust, as there is no much significant improvement in the monthly and annual climatology of MaMetAtTIO even after adding new records. However the individual year-month summaries have shown some improvement and is discussed. The random error is calculated for SLP, DBT, DPT, SST are 1.85 ± 0.32 hPa, 0.78 ± 0.13 ^{o} C, 1.72 ± 0.28 ^{o} C, 1.15 ± 0.2 ^{o} C. For WS, systematic error correction is performed. The annual net heat flux for TIO calculated from error corrected variables is reduced by 14 W / m^2. For validation purpose, MAMetAtTIO is compared with insitu data from buoys and Tropflux dataset.

Contents

ABSTRACT

This paper discusses the preparation of a new climatology Marine Meteorological Atlas for Tropical Indian Ocean (MaMetAtTIO) from ship observations that are used in preparation of International Comprehensive Ocean-Atmosphere Dataset (ICOADS) Release 3 and those obtained from the Indian Meteorological Department (IMD) and Naval Operations Data Processing and Analysis Centre (NODPAC). The enhancement in MaMetAtTIO is checked with reference to ICOADS R3.0 climatology. ICOADS R3.0 dataset has been found to be self robust, as there is no much significant improvement in the monthly and annual climatology of MaMetAtTIO even after adding new records. However the individual year-month summaries have shown some improvement and is discussed. The random error is calculated for SLP, DBT, DPT, SST are 1.85 ± 0.32 hPa, 0.78 ± 0.13 ^{o} C, 1.72 ± 0.28 ^{o} C, 1.15 ± 0.2 ^{o} C. For WS, systematic error correction is performed. The annual net heat flux for TIO calculated from error corrected variables is reduced by 14 W / m^2. For validation purpose, MAMetAtTIO is compared with insitu data from buoys and Tropflux dataset.

1 INTRODUCTION

One of the oldest method for observing weather variables is the ship observations dating back to 1800s. In the nineteenth century many countries have taken up the initiative of observing the marine meteorological parameters on board the ship. Voluntary Observing Ship (VOS) scheme which is a fleet moving across various oceans and seas, observe the marine-met parameters. This collection is a very important contribution to the historical marine meteorological datasets. The observations made on board the ship are usually in machine readable form and are to be extensively organized before further processing. The data onboard the ship is transmitted to the corresponding data collection centers offshore. The data is then extracted to a normal format , quality controlled, etc. ICOADS is such an extensive project consisting of millions of ship reports observed since 1854 collected from various sources (Slutz *et al* 1985). There are several products derived from ICOADS. The NOCS flux dataset by Berry and Kent 2011 is a gridded dataset of air-sea fluxes calculated from bias corrected independent variables observed onboard the ships. The Atlas of Surface Marine data by DaSilva *et al* 1994 is also derived from ICOADS. He also discusses the bias corrections to the measured variables and fine tuning of fluxes so that a closure can be obtained between incoming and outgoing heat energy.

In the Indian ocean, study of air-sea interaction processes is specifically more important. The Indian subcontinent is highly dependent on the Indian summer monsoon (ISM) which is in turn dependent on air-sea interaction across TIO that has been studied by several authors (Rajeevan *et al.* 2002, Gadgil *et al.* 2003,Annamalai *et al* 2005, Chowdary *et al.* 2015, Goswami *et al.* 2016, etc,). The North Indian ocean (NIO) is always considered unique in its characteristics both in meteorology and oceanographic perspective. The monsoon reversal of winds and currents, the amount of fresh water influx from a number of rivers into the seas of NIO are some of the primary reasons behind the unique characteristics of TIO. But in contrast, Indian ocean is considered as a data sparse region. Observation systems like, moored buoys, drifting buoys, ships, etc, and programs like BOBMEX, ARMEX (Arabian Sea Monsoon EXperiment), OMM (Ocean Mixing and Monsoon), etc. are carried out to obtain observations for studying the air-sea interaction processes. Any amount of data available from any reliable source is always beneficial. Owing to this fact, an attempt is made in this work to combine the ship observations obtained from IMD, NODPAC to the already existing collection of ship observations of ICOADS Release 3.0 within Indian ocean region (20E - 120E , 30S - 30N), with an aim of possible enhancement in the existing climatology.

The description of all the datasets used in this work namely, IMDdata, ICOADS records, NODPAC records is briefly presented in section 2.1. The QC of this dataset i.e. the data obtained from IMD and NODPAC is discussed in Kameshwari *et al*. 2016. Meteorological parameters onboard ships are measured at different heights which are determined based on ship dimensions and other factors the details about the ship metadata and the height correction procedure can be referred from Kameshwari *et al*. 2016. Section 2.2 describes the derived variables and their formulation. Methodology used for gridding, number of records used in gridding are detailed in section 2.3. The random error analysis for correcting independent variables is discussed in section 2.4. Section 3 details about the validation of the new climatology by comparing with other insitu observations, flux datasets and also discusses the affect of removal of observational error. Section 4 briefly discusses the application of this climatology. Section 5 discusses the conclusion and limitations faced in this work and the scope for future tasks.

2 DATA and METHODOLOGY

This section describes the data used in the preparation of the climatology. Along with the independent Met parameters and SST certain derived quantities like sensible heat flux, latent heat flux are also calculated and gridded and their formulation is elaborated. Details about the objective analysis method is also discussed. Semivariogram analysis used for random error correction is also discussed.

2.1 Data used

The data used in this work are the marine-met observations made onboard ships across Indian Ocean and several buoy data used for validation. The observations include a set of records obtained from IMD (henceforth referred to as "IMDdata"), the individual ship records taken from ICOADS Release 3, the observations collected onboard Indian NAVY ships (NODPAC dataset).

ICOADS is the first climatology of marine-met parameters utilizing majorily the ship observations (Slutz *et al.* 1985). ICOADS releases consists of the climatologies of individual parameters, some derived variables like difference between DBT and SST, pseudo fluxes, etc. It is a collection of monthly statistical summaries across 1° x 1° (1960-2007) and 2° x 2° (1800-2007) globally analyzed fields and approximately 261 million individual observations from ships, buoys, sea stations, etc. The first release of ICOADS was in 1985 with 70 million individual observations, comprising observations recorded during the period 1854 to 1979. Gridded monthly statistical summaries were released in 1996. The latest version of ICOADS is Release 3.0 covering the time period from 1662-2017. Apart from the construction of climatologies, the individual records of ICOADS were used in several studies by daSilva *et al.* 1994, Woodruff *et al.* 2008, Fan *et al*. 2014, etc. In depth detailed description of ICOADS dataset can be referred from Freeman *et al.* 2017. In this work, individual ship records of ICOADS release 3.0 which are both quality controlled (QCed) and non-QCed are used. Non-QCed records of ICOADS are used to find out duplicate records between ICOADS and IMDdata. Both the non QCed ship records and QCed ship records are downloaded from CISL research data archive at " http://rda.ucar.edu/". The data downloaded was in IMMA (International Maritime Meteorological Archive) format. The spatial density of ICOADS individual records is complete, across all the ocean basins.

The raw ship records obtained from IMD were approximately 5.9 million. The records were of various byte sizes observed in the period 1961-2012. The variables present in the raw ship records from all byte sizes are given in table 1. The spatial coverage of the entire IMDdata is shown in figure 1(a). Three levels of duplicate check is performed before quality control procedure. The first two levels correspond to each individual dataset, i.e. intra dataset duplicate check (Level-1), intra dataset hard duplicate check (Level-2). The IMDdata records were compared with the non QCed individual ship records of ICOADS for duplicates, i.e. inter dataset duplicate check (Level-3). Figure 1(b) shows the data before and after Level - 2 duplicate check. Figure 1(c) shows the spatial coverage of unique records of IMDdata after Level - 3 check. Details about the comparison between IMDdata and ICOADS can be referred from Kameshwari *et al* 2016, Kameshwari *et al.* 2018. *The total number of unique records are 7,03,994 which is 11.8 % of total data obtained from IMD i.e. IMDdata.* Only this part of the data was further processed. In case of NODPAC data there are no duplicates. NODPAC data are basically observations made onboard Indian Navy ships. The NODPAC data is in standard GTS format and is concentrated in North Indian ocean. This data is spanning from 2002 to 2017. The year wise count of NODPAC records is shown in figure 2. The list of variables are same as in Table 1 except that metadata of ship and Quality control flags of the data are absent. QCed records of ICOADS, unique records in IMDdata and all of the NODPAC records that have passed the quality control check are used for the new TIO climatology MaMetAtTIO (20E - 120E , 30S - 30N). The ICOADS R3.0 records used in MaMetAtTIO are QCed, the QC procedure is applied only to IMDdata_unique and NODPAC data. Detailed description about the QC procedure can be referred from Kameshwari *et al* 2018. Table 2 gives the number of records before and after each level of QC upon IMDdata_unique and NODPAC data. As most of the data of IMDdata had already been utilized in ICOADS R3.0 climatology, the aim was to see the possible enhancement in MaMetAtTIO after adding the IMDunique records and NODPAC data to ICOADS R3.0.

S. No	**Variable name**	**S.No**	**Variable name**
1	Air Pressure (SLP)	7	Swell wave parameters
2	Dry Bulb Temperature (DBT)	8	Weather and cloud conditions
3	Dew point Temperature (DPT)	9	Details of ship position
4	Sea surface Temperature (SST)	10	Time of observation
5	Wet bulb Temperature (WBT)	11	Details of the ship

6	Wind speed and direction (WS, WDIR)	12	Quality Control flags of all the parameters

Table 1 Variables available in the IMD data

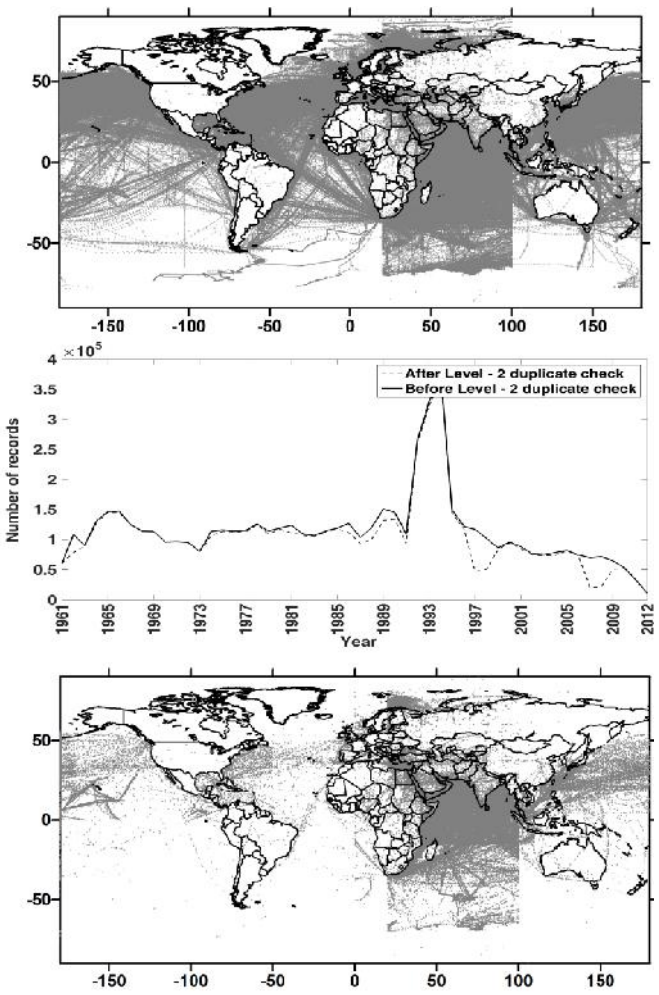


Figure 1 Spatial coverage of the entire dataset obtained from IMD (top - a), yearwise distribution of IMDdata before and after Level - 2 duplicate check(middle - b) and IMDunique(bottom - c) data

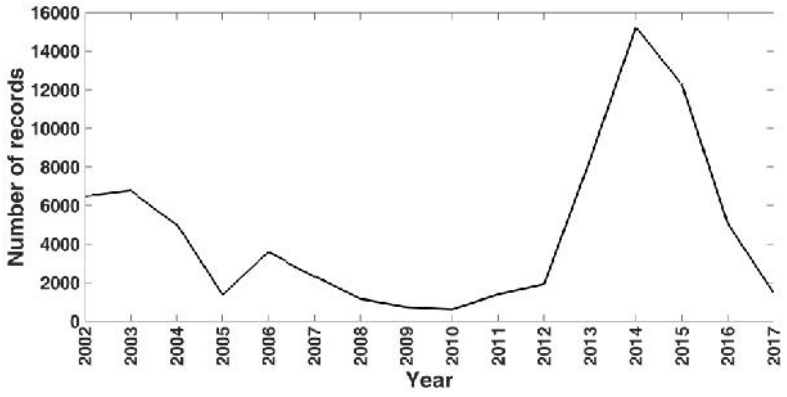


Figure 2 No.of records in NODPAC data yearwise

Data	No of records after MQC check	Removing records with missing MET value and land points	STDEV	% lost
IMDunique - SLP	703,994	6,85,905	6,71,576	4.6
IMDunique -DBT	703,994	5,61,158	5,31,762	25
IMDunique - DPT	703,994	4,58,081	4,26,414	40
IMDunique - SST	703,994	4,25,199	3,78,809	47
IMDunique - WS	703,994	6,22,630	5,66,204	12.5
NODPAC - SLP	37,447	34,971	33,102	11.6
NODPAC - DBT	37,447	32,163	29,833	20.3
NODPAC - DPT	37,447	32,163	29,833	20.3
NODPAC - SST	37,447	26,916	22,092	0.41
NODPAC - WS	37,447	35,096	32,358	13.5

Table 2 shows the number of records before and after each level of QC.

2.2 Parameters and their formulation

Momentum flux

The rate of transfer of horizontal momentum in vertical direction is called momentum flux. It is basically wind stress. The zonal and meridional component of momentum flux (N m^{-2}) are :

$$\tau_x = \rho * C_D * WS * u$$

$$\tau_y = \rho * C_D * WS * v$$

ρ = Density of air (kg m^{-3})

$C_D =$ Drag coefficient at 10 m

u = zonal component of WS at 10m = - WS*sin(WDIR*3.14159*(1/180)) (ms^{-1})

v = meridional component of WS at 10m = - WS*cos(WDIR*3.14159*(1/180)) (ms^{-1})

Sensible Heat Flux

The turbulent transfer of heat from Earth's surface to the air above it is called sensible heat flux. It is basically transfer of heat energy without involving any phase change. The magnitude ranges from 5 W m^{-2} to 60 W m^{-2}.

$$\mathrm{SHFX} = \rho * \mathrm{C_p} * \mathrm{C_t} * \mathrm{WS} * \Delta\Theta$$

$\mathrm{C_p}$ = Specific heat capacity of air (KJ Kg^{-1} K^{-1})

C_T = Transfer coefficient for temperature considering stability at 10 m

$\Delta\Theta$ = Difference of SST and temperature at 10 m

Latent Heat flux

The turbulent transfer of heat and water vapor from the sea surface to the air above it is called Latent heat flux. It involves phase change hence the term latent heat.

$$LHFX = \rho * \mathrm{L} * \mathrm{C_E} * \mathrm{WS} * \Delta\mathrm{q}$$

L = Latent heat of vaporization of saline water (KJ Kg^{-1})

C_E = Transfer coefficient for moisture considering stability at 10 m

$\Delta\Theta$ = Difference of specific humidity closer to the sea surface and 10 m

Net Shortwave Radiation

The incoming solar radiation at the sea surface after passing through the atmosphere is the net shortwave radiation. It includes both direct and diffuse components of shortwave radiation.

$\mathrm{SW_{sea}}$ is calculated using the below formulation (da Silva *et al.* 1994).

$$\mathrm{SW_{sea}} = \mathrm{SW_{clear}} * (1 - 0.62 * \mathrm{c} + 0.0019 * \beta) * (1 - \alpha)$$

$$SW_{clear} = SW_{dir} + SW_{diff}$$

$$SW_{dir} = Q_o * \tau_a^{\sec z}$$

$$SW_{diff} = [(1 - A_a) * Q_o - SW_{dir}] / 2$$

$$T = A_i + B_i * sin\theta$$

Equation 1 where SW_{sea} is the solar radiation at sea surface, SW_{clear} is the incident radiation under clear skies, c is the cloud fraction, β is the solar noon altitude, SW_{dir} is the direct component of shortwave radiation, SW_{diff} is the diffuse component of shortwave radiation, Q_o is the incident radiation at the top of atmosphere, τ_a is the atmospheric transmission coefficient which is taken as 0.7, z is the solar zenith angle, A_a is the absorption due to water vapor and ozone, T is transmission factor, $\propto$ is the albedo of the sea surface, θ is the solar elevation angle , where A_i, B_i are coefficients based on cloud cover

We used the formulation similar to that followed by da Silva *et al.* 1994 to calculate solar elevation angle and incident shortwave radiation at the top of atmosphere. It is basically based on the equations from Smithsonian Meteorological tables (List 1958). In his report, da Silva has clearly explained the formulation for calculation of solar elevation angle. The Julian day number is calculated based on the local noon time. From which the Sun's declination and right ascension are calculated at local noon time value. The solar elevation angle is calculated for each observation. Albedo is obtained based on transmission factor and solar elevation angle(Payne 1972). The nearest value of the former mentioned parameters are searched and corresponding albedo value is taken. The transmission factor is obtained based on okta model(Dobson and Smith 1988).

Net Longwave Radiation (NLWR)

The wavelength of the radiation emitted by a body is inversely proportional to the temperature of the body. As the temperature of Earth surface and atmosphere is less, the wavelength of the radiation emitted is in Infrared region. The NLWR is the summation of the longwave radiation (LWR) form sea surface, LWR from atmosphere and clouds back to sea surface.

$$NLWR = \epsilon * \sigma * SST^4 * \left(0.39 - 0.05 * \sqrt{e}\right) * (1 - \chi * c^2) + 4 * \epsilon * \sigma * SST^3 * (SST - DBT10)$$

ϵ = Emissivity of ocean (0.97)

σ = Stefan - Boltzman's constant (W m^{-2} K^{-4})

e = Atmospheric vapor pressure (hPa)

χ = a numerical value dependent on latitude

c = fractional cloudiness

$DBT10$ = DBT converted to 10 m height

2.3 Objective Analysis

The concentration of IMDunique data is high in the Indian ocean. The data density of NODPAC records is also highly concentrated in the North Indian ocean. As the aim of the work is to see the enhancement in the existing ICOADS climatology after adding the records from IMDunique and NODPAC, the climatology is restricted to 20°E - 120°E and 30°S - 30°N. The records of ICOADS-QCed, IMDunique, NODPAC are all combined month wise and separately for each variable. The gridding method is the one used by Da Silva *et al.* 1994, i.e. Cressman iterative difference-correction scheme with Barne's weight function. Four iterations of the analysis method is performed upon the data with reducing radii of influence i.e. 1541 km, 1211km, 881km and 771km. There is no first guess taken. The scattered observation values are ' V_{ij} ' . The query grid location is' X_q 'i.e. the grid locations of 1° x 1° gridded field. The lat-lon locations of scattered observations is' X_v '. The correction' $Verr_{ij}$ 'is the difference between the observation and the interpolated values of grid locations at observation locations. w are the weights of the observations upon the grid point.

The procedure can be outlined as follows:

1. The first guess V_{q_old} is taken as zeros for the first iteration and the resultant gridded field of the previous iteration is taken for the subsequent iterations.
2. Calculate Verr

   for the first iteration

   Verr = V

   for the subsequent iterations

Verr = V - Interpolation (newly gridded field values $V_{q(i,j)}$ to observation locations, X_v)

3. Determine the observations which are within a distance less than radius of influence from X_q
4. (a) If the number of observations within the radius of influence are greater than or equal to 30, compute the weights W of the observations upon the queried grid point.

$$w = e^{\frac{-4*r^2}{R^{-2}}}$$

(w : weight of all observations within the radius of influence upon the grid point)

$$W_{ij} = w \,/\, sum(w)$$

$$V_{q(i,j)} = V_{q(i,j)}\_old + W_{ij} * Verr_{ij}$$

(b) If the number of observations within the radius of influence is less than 30, $V_{q(i,j)} = undef$

5. Repeat step 4 for all grid points, and from step 1 to step 4 at different radius of influence.

Two filters are applied to the final gridded field obtained from the iterative objective analysis function. First is the non-linear filter wherein the smoothed value at grid point is equal to the median of grid values in a 3^ox3^o box around it. Second is the Linear filter called Shapiro filter wherein the smoothed value at a grid point is given by

$$S_{ij} = V_{ij} + \frac{\alpha}{4}(V_{i-1,j} + V_{i+1,j} + V_{i,j-1} + V_{i,j+1} - 4 * V_{i,j})$$

Two iterations with α=0.5 and α=-0.5 are run. All the details regarding objective analysis and filter passes of the gridded data are obtained from Da Silva *et al* 1994.

A total of three kinds of summaries are prepared, namely, annual climatology, monthly climatology, individual year-month summaries. After the entire quality control check, SLP, SST, cloud, SW, LW, SHFX, LHFX are gridded by the above mentioned procedure to 1^o x 1^o gridded field over 20^oE - 120^oE and 30^oS - 30^oN. In case of DBT, SPHUM, WS, gridding is done to the 10 m recalculated values. For annual gridded fields the entire data is taken into consideration. From every grid point the distance from all the observations is calculated and those lying within the radius of influence are multiplied with Barnes weight and averaged to give a grid value. The observations lying closer to the grid value are given more weights when compared to those lying farther. For monthly gridded fields, the entire data is divided into twelve months data and each month is separately objectively analyzed by the above mentioned

procedure. For the individual year-month summaries, the entire data is divided into individual years and each year is divided into months. The radius of influence is taken as 551 km. Each monthly data of each year is separately objectively analyzed. For all the above climatologies and summaries, no initial guess was considered in objective analysis. In case of objective analysis of fluxes, sampling approach is followed instead of classical method thereby retaining the correlation between the meteorological variables. The Figures 3 show the number of records utilized in producing the gridded fields.

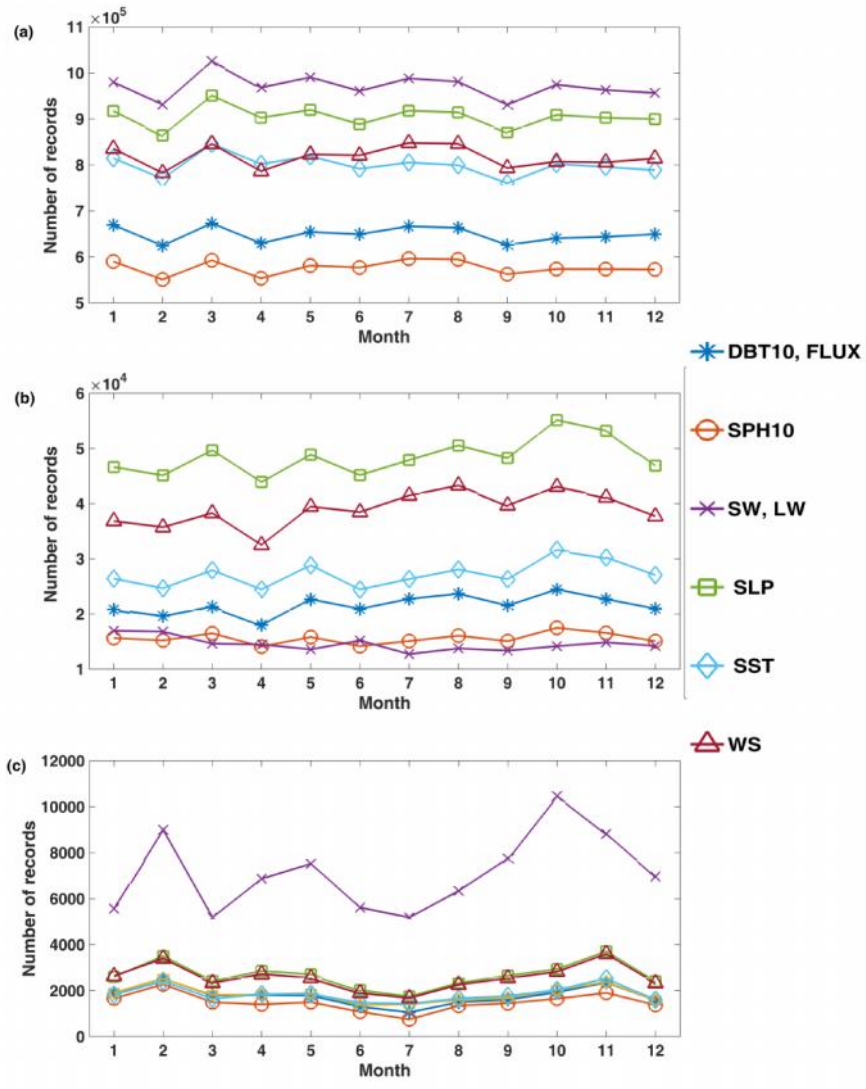


Figure 3 : No . of records that are used in gridding for all the variables

2.4 Correction for observational error

An attempt is made to correct the MET parameters and SST for random and systematic observational error. Calculating the error estimates has various applications. Data devoid of random error would give more accurate model outputs and hence more reliable. Also the fluxes calculated would be more accurate resulting into accurate heat budget analysis arriving to a closure. In this study, the random error estimates of DBT, SLP, DPT and SST are calculated exclusively for TIO following the semivariogram statistical technique described in Kent *et al*. 1998 owing to the unique regional characteristics of TIO. Briefly, this technique can be defined as the estimation of random error in observations by pairing simultaneous observations (i.e. pairing observations measured in the same hour and are at some fixed distances apart), calculating variance between these simultaneous pairs of observations and extrapolating this variance fit to a point where the distance between the pairs is zero. The elaborate description is given in the above mentioned reference. The random error for SLP is 1.85 ± 0.32 hPa, for DBT is 0.78 ± 0.13 ^{o} C, for DPT is 1.72 ± 0.28 ^{o} C, for SST is 1.15 ± 0.2 ^{o} C. In case of WS which had very high variability, systematic bias correction following Kameshwari *et al.* 2018 is applied instead of random error. The affect of removing the above mentioned observational errors is discussed in heat budget analysis. Though there is lack of clarity about the correlation between the terms used in flux calculation, an attempt is made to check the affect of removal of observational error. The results are discussed in section 3.2.

3 RESULTS AND DISCUSSIONS

As the major aim is to check the enhancement of MaMetAtTIO with respect to ICOADS, in this section, we evaluate the new climatology by performing a comparison between both the climatologies and a relative comparison between the both with respect to other datasets. Gridded fields of both independent variables and derived quantities of MaMetAtTIO are produced before and after removing the observational error and then compared. The affect of removal of observational error is also discussed in terms of heat budget analysis.

3.1 Characteristics of objectively analyzed fields

This section describes the flux fields generated by Barnes's interpolation described in **section** 2.4 . The chief characteristics of the flux fields are the same when compared to the existing flux climatologies, however, our aim is to add new data to ICOADS R3.0 and look for any enhancement.

The annual climatology of net heat flux is shown in figure 4. The net heat flux is obtained by subtracting LHFX, SHFX, LWR from SWR. In this study, positive(negative) values of Qnet signify ocean heat gain(heat loss). The three sub figures are of ICOADS R3.0, MaMetAtTIO and MaMetAtTIO corrected for observational error. All the three datasets show a net heat gain upto 100 W/m^2 across TIO except near the NE boundary region of Australia.

For LHFX in TIO shown in figure 5, positive(negative) values denote heat loss(heat gain). The heat loss across TIO is ranging from 80 W/m^2 to 130 W/m^2 during boreal winter and ranging from 90 W/m^2 to 160 W/m^2 during austral winter. The strongest heat loss is in arabian sea and southern subtropics of TIO in both boreal winter and austral winter where in austral winter, it is the highest and can be owed to the increased WS during SW monsoon.

In case of SHFX in TIO shown in figure 6, positive(negative) values denote heat loss(heat gain) There is almost all uniform heat loss in the month of Jan. In the month of July, i.e. boreal summer, there is a gradual variation from heat loss in STIO to heat gain in NTIO. There is a clear evidence of heat gain in Western Arabian sea which is a region of strong upwelling under the influence of strong winds of Somali jet. This strong upwelling causes colder water to rise up thereby reversing the direction of heat as SST would be less than air temperature, resulting into a heat gain.

The overall range of LW radiation is from 30 W/m^2 to 65 W/m^2 where positive(negative) values denote heat loss(heat gain). It can be observed in figure 7 that the LW peaks in boreal winter and austral winter in the respective hemispheres due to relatively clear sky conditions. The downward SW radiation at sea surface in figure 8 is positive for ocean heat gain. The range of SW radiation varies from 100 W/m^2 to 260 W/m^2. SW radiation peaks in boreal summer and austral summer in their respective hemispheres.

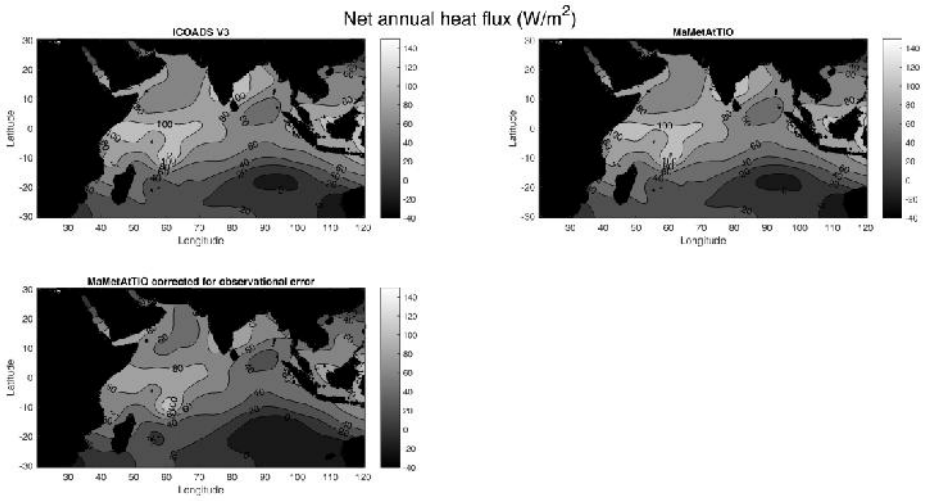


Figure 4 Net annual heat flux W/m^2 . Left top - ICOADS R3.0, right top - MaMetAtTIO, left bottom - MaMetAtTIO corrected for observational error

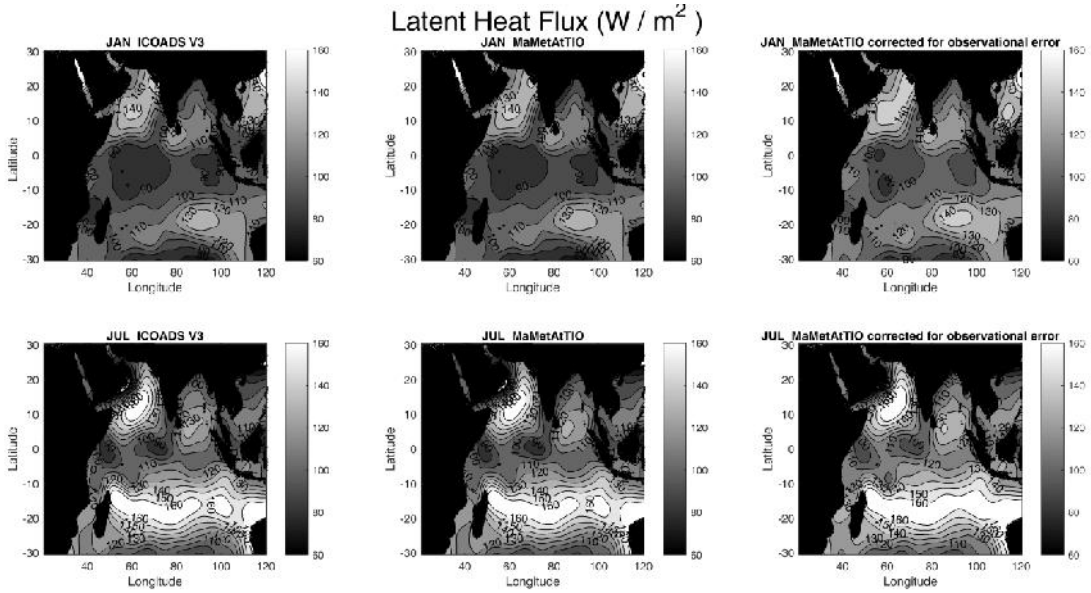


Figure 5 Latent heat flux W/m^2 . Top - January climatology, below - July climatology. Left - ICOADS R3.0, middle - MaMetAtTIO, right - MaMetAtTIO corrected for observational error

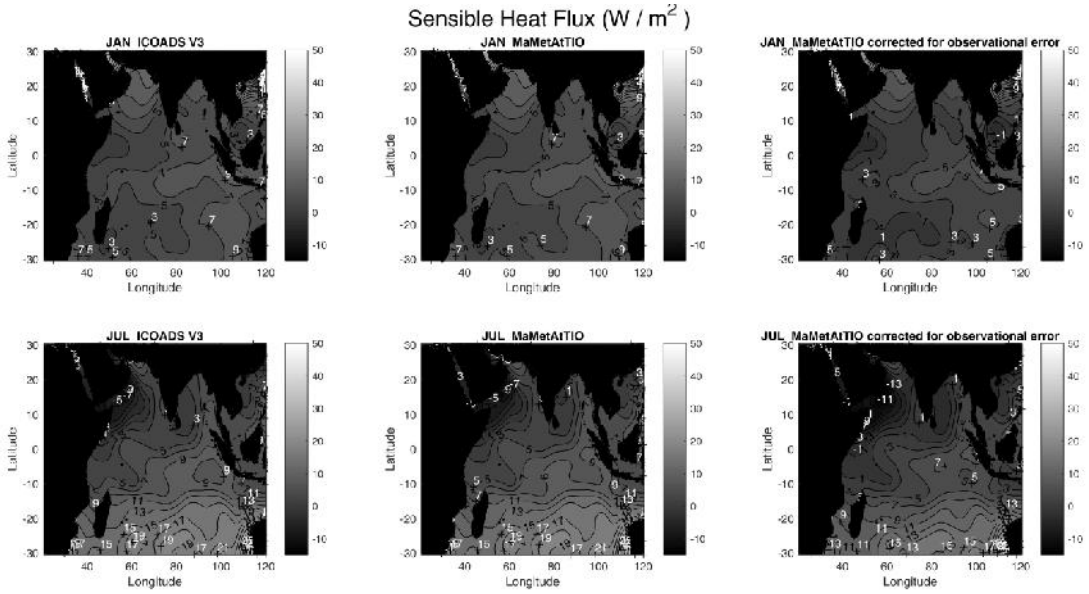


Figure 6 Sensible heat flux W/m^2 . Top - January climatology, below - July climatology. Left - ICOADS R3.0, middle - MaMetAtTIO, right - MaMetAtTIO corrected for observational error

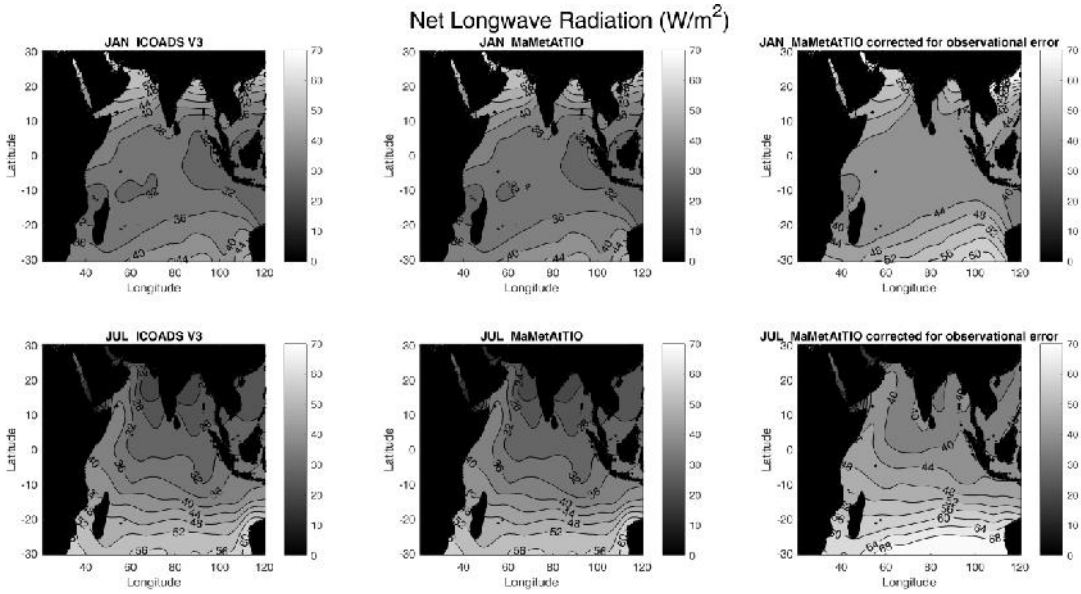


Figure 7 Longwave radiation W/m^2 . Top - January climatology, below - July climatology. Left - ICOADS R3.0, middle - MaMetAtTIO, right - MaMetAtTIO corrected for observational error

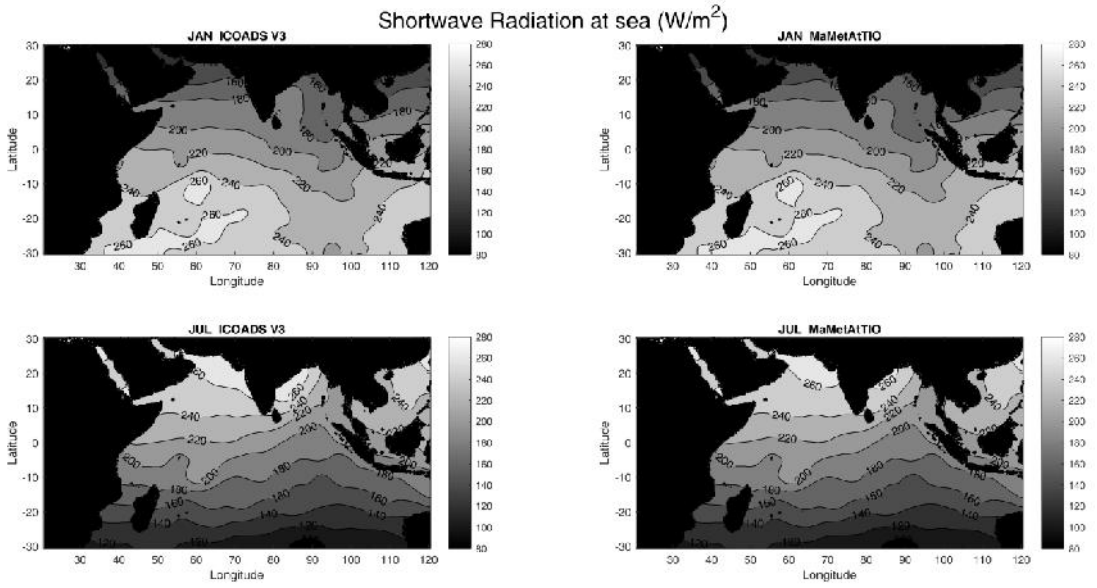


Figure 8 Shortwave radiation W/m^2 . Top - January climatology, below - July climatology. Left - ICOADS R3.0, right - MaMetAtTIO

3.2 Comparison With Icoads Climatology

The three gridded fields of MaMetAtTIO namely, annual and monthly climatology and individual year month summaries are compared with ICOADS R3.0 gridded fields. The following sections discuss the same in detail.

Monthly climatologies:

Bias between the monthly climatologies of ICOADS R3.0 and MaMetAtTIO has turn out to be of very less magnitude. For example this can be understood from figure 5 to figure 8 showcasing the heat fluxes and radiation fluxes calculated from ICOADS R3.0 and MaMetAtTIO**.** Also the figure 9 depicts the bias between DBT 10m of ICOADS R3.0 and MaMetAtTIO. The bias is too less of the order of magnitude less that 0.5 ^{o}C. *Hence the ICOADS R3.0 and MAMetAtTIO monthly climatology are not having significant differences even after adding the new records to ICOADS R3.0.* Some statistical analysis has been done to investigate further about the same. Frequency distribution of the ICOADS data with and without adding IMDunique and NODPAC data are explored to ascertain the cause of lack of improvement. ICOADS is a collection of enormous amount of data with spatial data density that is complete across all the ocean basins. As there are no gaps, any amount of data added would only blend into the existing data unless until when the new data is carrying a new signal. In such cases there

would be a new peak arising in a data distribution plot which is not witnessed in case of any variable of any month. This establishes the self-robustness of the ICOADS climatology. The frequency distribution plots of various variables for which the climatology has been prepared by adding the IMD and NODPAC data to ICOADS data, are shown in figures 10 . It can be observed that there is basically no much difference between the distributions after adding the new data. In the distribution plots of the combined data , there are slight increments in the existing bars of ICOADS records because of the addition of new records.

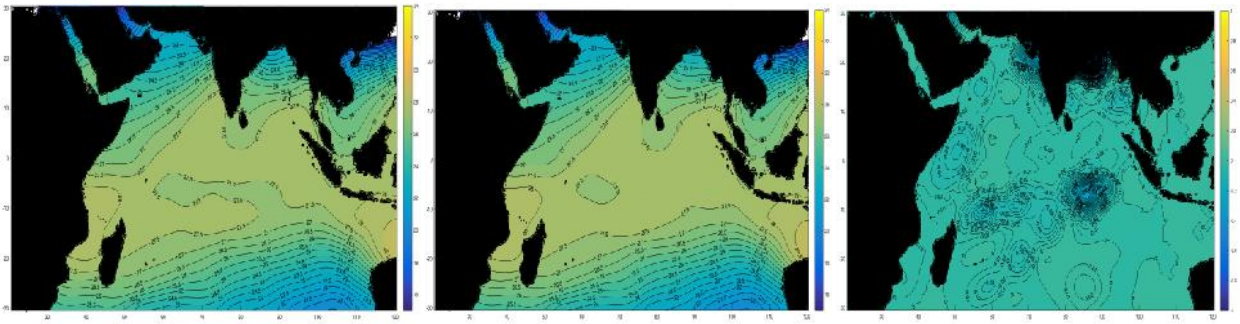
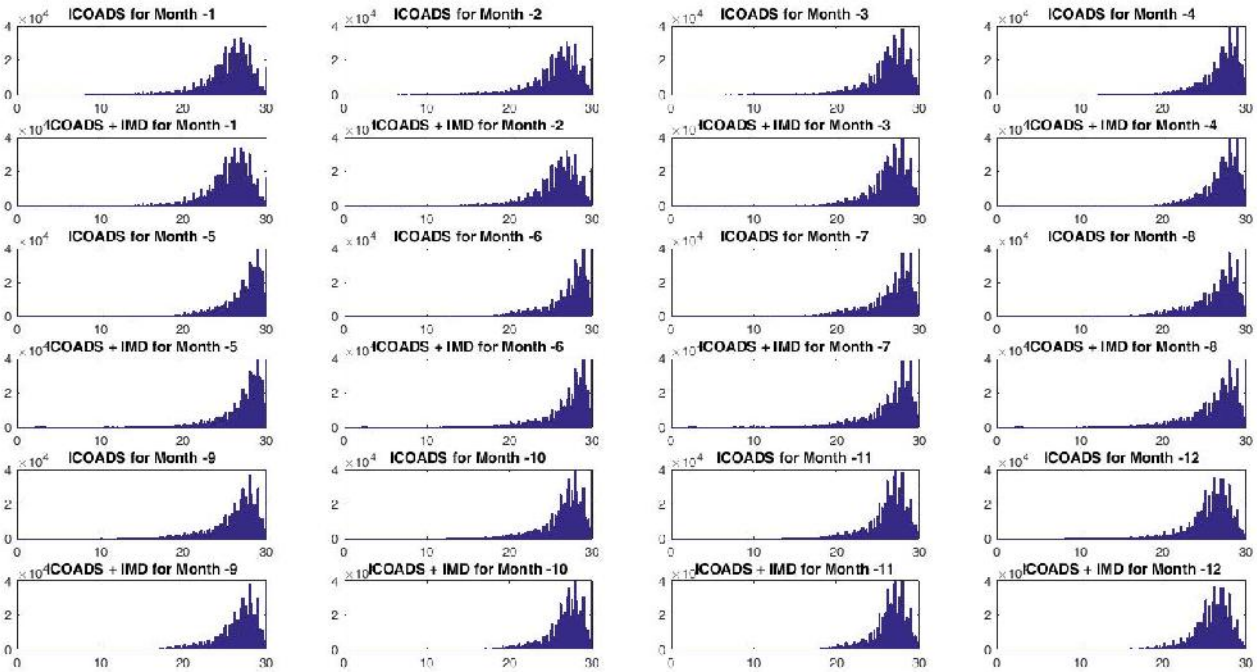


Figure 9 : Bias in DBT-10m of January month (^{o}C). MaMetAtTIO on the left, ICOADS in the middle and MaMetAtTIO - ICOADS on the right.



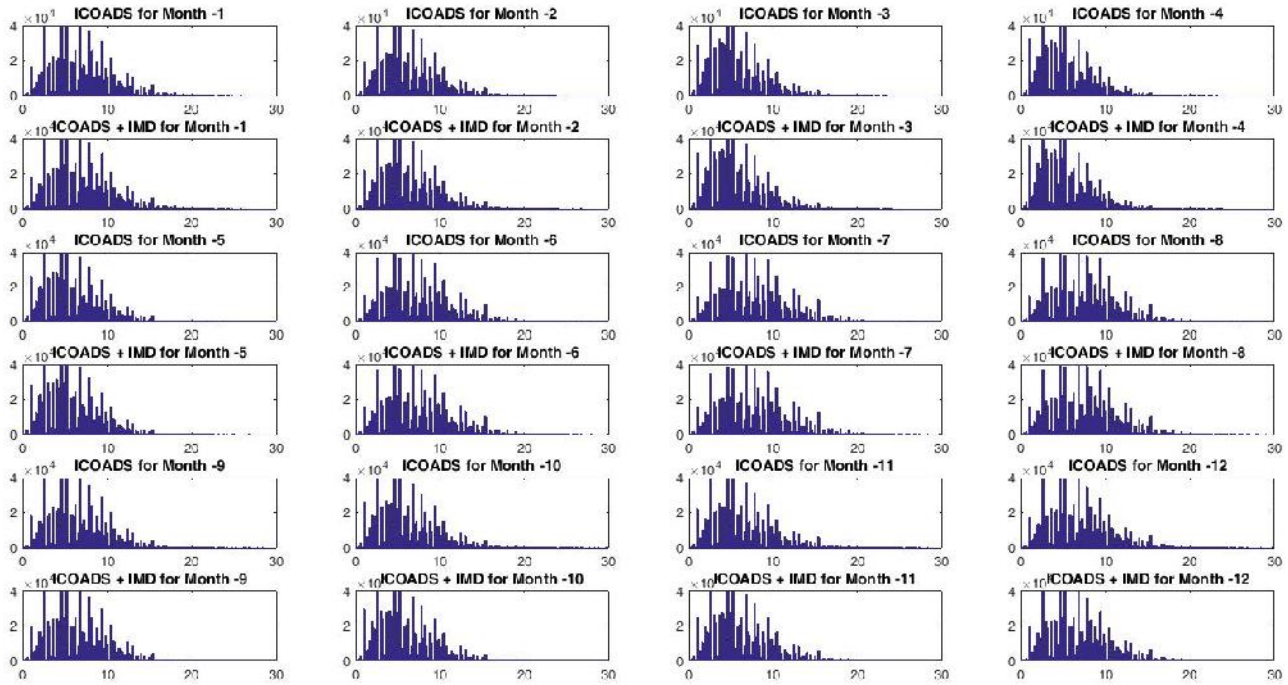


Figure 10 : Comparison between frequency distribution of ICOADS data with and without adding IMDunique and NODPAC data for DBT 10 (top) and WS10 (bottom)

The characteristics and patterns of the spatial gridded fields of monthly climatologies of ICOADS R3.0 and MaMetAtTIO are almost the same. However there would be a difference in the statistical reliability upon the dataset. Though ICOADS is a robust dataset, adding any number of observations will only improve the statistical reliability of the data. Though the bias observed in figure 9 is not very significant nonetheless the importance of adding new data into the existing data is demonstrated through statistical analysis by comparing individual records of MaMetAtTIO with insitu and flux data and is described below.

Individual year-month summaries:

The significance of the improvement in individual year-month summaries is examined by comparing the RMSE and stdev of the observations used in gridding at each gridpoint. Obviously the number of observations available at certain grid points will be more in case of MaMetAtTIO when compared to ICOADS. The standard deviation among the observations within the radius of influence and RMSE between these observations and the grid value is compared among both the datasets. It is observed that at most number of grid points, the STDEV among the observations within the radius of influence is reduced in case of MaMetAtTIO. Also the deviation from the final gridvalue and each observation

within the radius of influence is also less in case if MaMetAtTIO i.e. the RMSE is also reduced. *For example consider 1975 Jan summary of DBT(figure 11 shown below), it can be observed that standard deviation is reducing across most part of TIO and the same with RMSE. This is the region where we have significant number of observations added. It can be seen that there are areas where the RMSE or STDEV has been increased in case of MaMetAtTIO, but those are the grid points where the number of observations within the radius of influence are less than 30 hence not statistically significant. Another example is that of WS of the year 1991 and month May. It can be observed from bottom left subfigure that there is an addition of observations near the west coast of India. Also there is a reduction of STDEV and RMSE of the order of 0.15 ms^{-1} in the same region.*

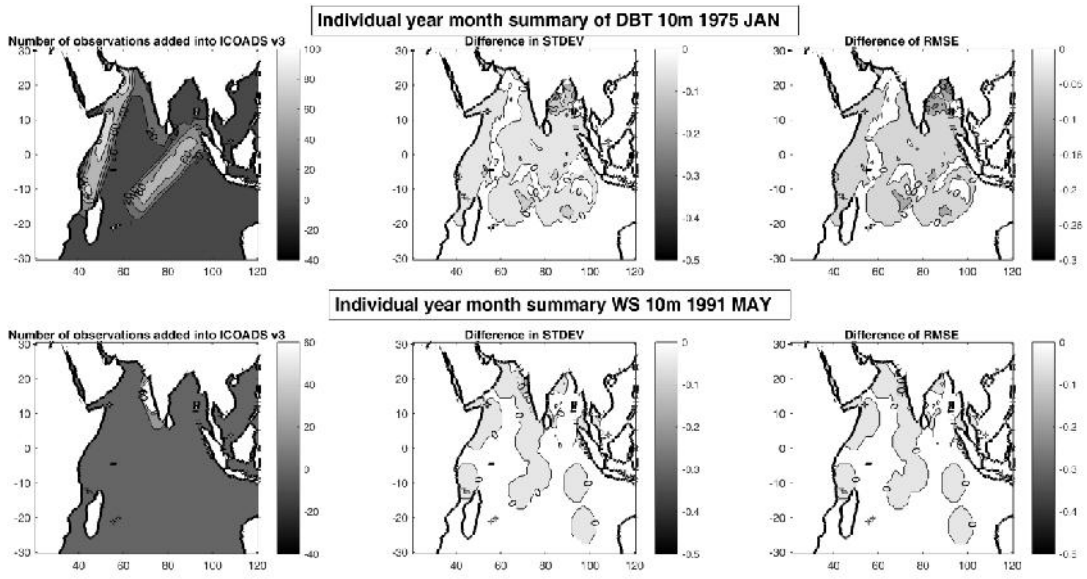


Figure 11 : Left - Number of observations added into ICOADS R3.0, Middle - Difference in STDEV, Right - Difference of RMSE (Difference : MaMetAtTIO - ICOADS R3.0)

3.3 Comparison with other datasets

Inorder to highlight the statistical significance of the addition of new data, the new climatology MaMetAtTIO and ICOADS R3.0 are compared with insitu buoy data and tropflux gridded dataset. Thus a relative comparison is made to figure out the level of agreement between the climatologies and other datasets.

Importance of newly added individual records:

The National Institute of Ocean Technology under the National Data Buoy Program has deployed a number of buoys in NIO and are operational since 1997. All the marine meteorological parameters and oceanic parameters are being retrieved from these buoys and a number of scientific studies utilizing the data are present. Here in this study comparison with buoy data is made between the individual observation level in order to highlight the importance of adding new observations. The comparison between short term mean summaries of daily time series of our data and buoy data is presented. Data for the mean summaries of daily time series are chosen based on the availability of continuous data from buoy platform at a particular location over a period of few years. Such availability was observed in case of few buoys and statistical analysis is presented in the following table 3.

Buoy	Location	Time Period	Parame-ter	RMSE between ICOADS and buoy data	RMSE between MaMetTIO and buoy data
DS01	15 ^{o}N, 69 ^{o}E	1998 - 2008	SLP	0.96	0.94
DS03	13 ^{o}N, 87 ^{o}E	1997 - 2007	SLP	1.59	1.57
SW06	13 ^{o}N, 80 ^{o}E	1997 - 2007	SLP	2.35	2.33
DS01	15 ^{o}N, 69 ^{o}E	1998 - 2008	WS	1.57	1.55
DS03	13 ^{o}N, 87 ^{o}E	1997 - 2007	WS	1.83	1.79
DS04	18 ^{o}N, 88 ^{o}E	1997 - 2007	WS	2.84	2.72
SW06	13 ^{o}N, 80 ^{o}E	1997 - 2007	WS	2.13	2.05

Table 3 Statistics obtained on comparing the ICOADS R3.0 and MAMetAtTIO with NIOT buoy data

It can be observed that the RMSE has decreased after the addition of the new data. Comparison has been made only between SLP and WS in case of NIOT buoys because of lack of sufficient data in case of other parameters. In case of few buoys, the student t-test conducted (not shown here) also gave a reduced t-value for MaMetAtTIO when compared to ICOADS R3.0 whereas there is no significant difference in t-value in case of other buoys.

All the above validation exercises and other (not shown here) show only a slight enhancement in the existing ICOADS R3.0 dataset thereby exhibiting the robustness of ICOADS. However there is atleast a

mark of slight enhancement in the statistics due to the addition of new data. This shows the statistical importance of addition of new data though less in number.

Correlation between ICOADS R3.0, MaMetAtTIO and Tropflux

Tropflux dataset is a combination of ERA - I reanalysis dataset and International Satellite Cloud Climatology Project (ISCCP) dataset. Turbulent and heat fluxes are calculated with the bulk variables obtained from ERA-I using COARE R3.0 algorithm. SW radiation is calculated from surface radiation data obtained from ISCCP. In detail description about the product can be obtained from Praveen Kumar *et al*. 2012 . The spatial plot of correlation coefficient of net heat flux between the climatologies and tropflux is given in figure 12. The spatial average value of correlation coefficient between Tropflux and MaMetAtTIO without observational error is 0.753 and with that of MaMetAtTIO it is 0.752 and with that of ICOADS R3.0 it is 0.74. Hence there is only slight improvement that is observed. The spatial distribution of correlation coefficient between Tropflux and MaMetAtTIO for the components of net heat flux is shown in figure 13. The correlation is reasonably good for both heat and radiation fluxes except that there is negative correlation value in case of longwave radiation in the equatorial region. The cause of this negative correlation has to be worked out.

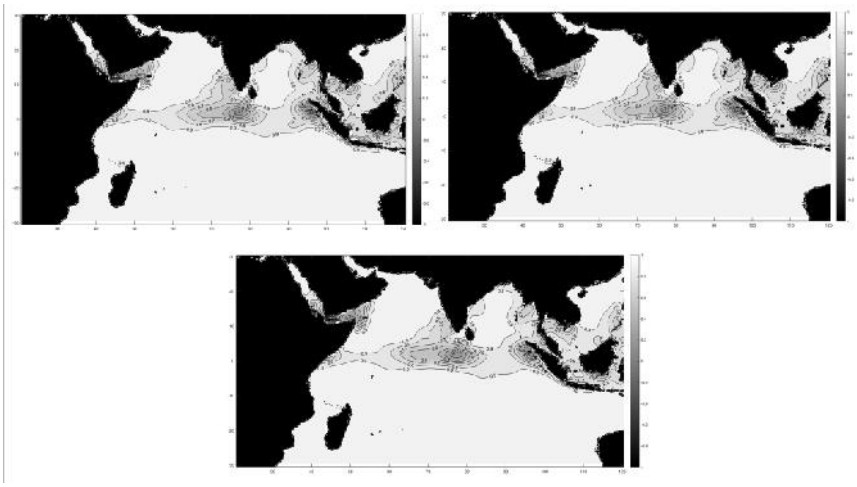


Figure 12 Spatial plot of correlation coefficient. Left top : Tropflux-ICOADS R3.0, right top : Tropflux-MaMetAtTIO, bottom : Tropflux-MaMetAtTIO corrected for observational error.

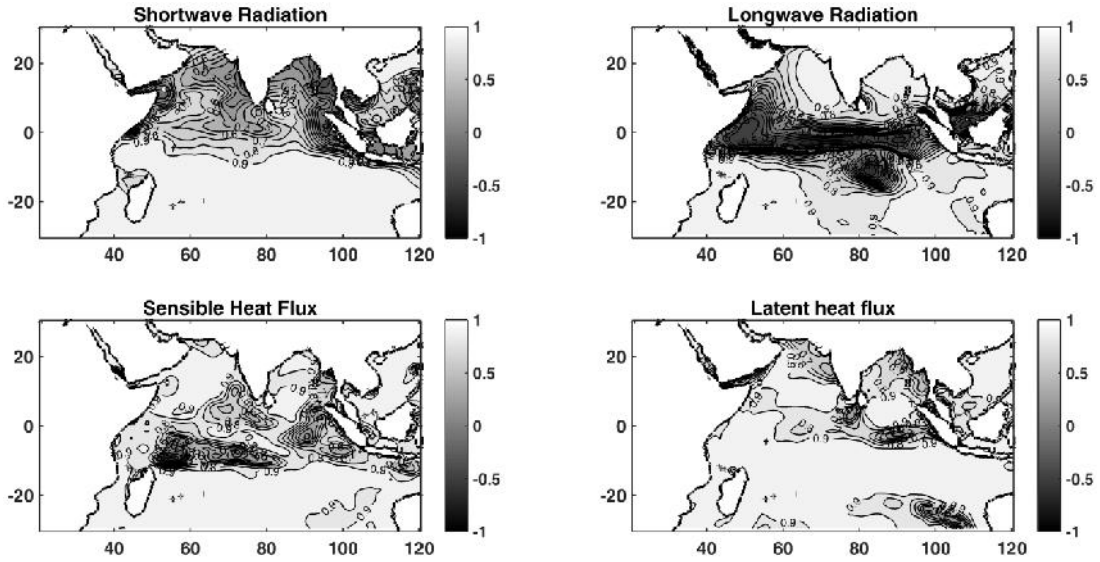


Figure 13 Spatial plot of correlation coefficient between Tropflux-MaMetAtTIO for the components of net heat flux

3.4 Affect of removing observational error on heat budget

One of the aim in heat budget analysis when dealt with climatological datasets, is to bring the net heat flux close to a very small values of 2-3 W/m^2. However it is still an unsolved problem. By applying the bias corrections, Josey *et al.* 1999 could obtain a global net heat gain of 30 W/m^2. As the present study is limited to TIO, the net heat budget would be greater than the value for the global mean. From the figure 6, the positive values of the net heat flux denotes ocean heat gain. The spatial average of annual net heat flux for TIO calculated for ICOADS R3.0 dataset alone gave a value of 57.39 W / m^2, whereas for MaMetAtTIO, it is 57.02 W / m^2 . On addition of new records from IMDdata and NODPAC, there is a reduction of 0.37 W/m^2 in the net heat flux. In case of MaMetAtTIO corrected for observational error, the net heat flux is 43.16 W / m^2. As a reduction of 14 W / m^2 is quiet significant, the correction made for observational error is considered to be a worthful attempt.

4 APPLICATION

The dataset described in this paper has been into application. The first one being the correction of WMO 1100 (**reference)** scale exclusively for the TIO. The new Beaufort equivalent scale has been derived using ICOADS+IMDunique data observed between 1985-2005. The scale derived is from July

month data of that period. The scale significantly reduces the bias between anemometer measured WS and Beaufort estimated WS. Further details can be referred from Kameshwari *et al*. 2018.

The next application is in a GUI tool presented to NODPAC by INCOIS. This tool is being used by Indian Navy onboard the ships for navigation purpose. This tool produces plots of various MET parameters and SST across TIO instantaneously on mouse-click basis. The background data used for plotting the MET parameters and SST is MaMetAtTIO. Detailed description can be found in Pavan *et al.* 2018.

5 CONCLUSION AND FUTURE SCOPE

Ship observations are one of the oldest and reliable observational component for collecting marine-met data. The collection in ICOADS dataset is huge and will always be the best choice for studying meteorological parameters across oceans and also few of the ocean parameters. In simple words, these ship records are a boon to study air-sea interaction processes. Adding any number of records to this dataset makes it more robust.

We have made an attempt to make use of the new ship records obtained from IMD data and NODPAC data. Even though there is no significant improvement in case of monthly climatologies, the addition of new data into existing climatologies is always beneficial and it is hoped that this attempt will essentially be useful to the user community. Nevertheless, addition of new records (IMDunique and NODPAC) into ICOADS individual year-monthly summaries has shown a considerable improvement. The climatology datasets of various variables available are given in table 4.

S.No	Monthly Climatology	Annual Climatology	Individual year-month summaries
1	DBT 10	DBT 10	DBT 10
2	SPHUM 10	SPHUM 10	SPHUM 10
3	SST	SST	SST
4	WS 10	WS 10	WS 10
5	SLP	SLP	SLP
6	Cloud amount	Cloud amount	Cloud amount
7	U-Momentum Flux	U-Momentum Flux	
8	V-Momentum flux	V-Momentum flux	
9	Latent Heat flux	Latent Heat flux	
10	Sensible Heat Flux	Sensible Heat Flux	

11	Longwave radiation	Longwave radiation	
12	Shortwave radiation	Shortwave radiation	
13	Net heat flux	Net heat flux	

Table 4 List of climatologies available

Few of the limitations faced are described here. While comparing ICOADS R3.0 and IMD data records, in the hard duplicate check, all of the hard duplicates are eliminated without referring to which record could actually be a duplicate one. Also, wherever the height of measurement of MET parameters are absent and couldn't be fetched from other sources as discussed in section 2.3, default values of 18m , 20m (according to Josey *et al.* 1999) were taken as wind speed and DBT measurement height without reference to ship or platform type. This is because, there is no metadata information available and in addition to that there are no callsigns for many records . This would definitely introduce certain amount of error.

Systematic bias corrections appropriate for DBT, DPT, SST are not applied for a number of reasons. One reason is that, the absence of some required parameters in the record. For example, inorder to correct the bias in DBT, RWS (Relative wind speed) is needed. For calculating RWS, ship speed and direction are required. But in many records, these values are absent. Also, there are many records with missing callsigns. Without a callsign, the ship metadata cannot be extracted from WMO PUB 47 document. Preparing a decadal climatology of the latest years with bias corrected variables is another task that would be taken up in future.

6 REFERENCES

1. Annamalai, Hanna, Ping Liu, and Shang-Ping Xie. "Southwest Indian Ocean SST variability: Its local effect and remote influence on Asian monsoons." Journal of Climate 18.20 (2005): 4150-4167.
2. Berry, D. I. and Kent, E. C., 2011: Air-Sea Fluxes from ICOADS: The Construction of a New Gridded Dataset with Uncertainty Estimates. International Journal of Climatology, 31(7), 987-1001 (DOI: 10.1002/joc.2059).
3. da Silva, Arlindo M., Christine C. Young, and Sydney Levitus. Atlas of Surface Marine Data 1994 Vol. 1: Algorithms and Procedures.

4. Fan, T., Deser, C., & Schneider, D. P. (2014). Recent Antarctic sea ice trends in the context of Southern Ocean surface climate variations since 1950. Geophysical Research Letters, 41(7), 2419-2426.
5. Freeman, E., Woodruff, S. D., Worley, S. J., Lubker, S. J., Kent, E. C., Angel, W. E., Berry, D. I., Brohan, P., Eastman, R., Gates, L., Gloeden, W., Ji, Z., Lawrimore, J., Rayner, N. A., Rosenhagen, G. and Smith, S. R., 2017: ICOADS Release 3.0: a major update to the historical marine climate record. Int. J. Climatol., 37, 2211-2232 (DOI: 10.1002/joc.4775).
6. Gadgil, S., Vinayachandran, P., & Francis, P. (2003). Droughts of the Indian summer monsoon: Role of clouds over the Indian Ocean. Current Science, 85(12), 1713-1719. Retrieved from http://www.jstor.org/stable/24109976
7. Goswami, B.N., S.A. Rao, D. Sengupta, and S. Chakravorty. 2016. Monsoons to mixing in the Bay of Bengal: Multiscale air-sea interactions and monsoon predictability. Oceanography29(2):18–27.
8. Helber, R. W., Kara, A. B., Barron, C. N., & Boyer, T. P. (2009). Mixed layer depth in the Aegean, Marmara, Black and Azov Seas: Part II: Relation to the sonic layer depth. Journal of Marine Systems, 78, S181-S190.
9. http://sot.jcommops.org/vos/ (About VOS scheme)
10. https://www.wmo.int/pages/prog/amp/mmop/JCOMM/OPA/SOT/vos.html (About VOS scheme)
11. ICOADS data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at http://www.esrl.noaa.gov/psd/
12. Josey, S. A., Kent, E. C., & Taylor, P. K. (1999). New insights into the ocean heat budget closure problem from analysis of the SOC air-sea flux climatology. Journal of Climate, 12(9), 2856-2880.
13. Kameshwari N, Uday Bhaskar TV, Suprit K, Rama Rao EP. 2016. Marine Meteorological Atlas of Tropical Indian Ocean. (Technical Report : ESSO/INCOIS/DMG/TR/09(2016) , INCOIS)
14. Kent, E. C., & Taylor, P. K. (1997). Choice of a Beaufort equivalent scale. Journal of Atmospheric and Oceanic Technology, 14(2), 228-242.
15. Kent, E. C., Taylor, P. K., Truscott, B. S., & Hopkins, J. S. (1993). The accuracy of voluntary observing ships' meteorological observations-Results of the VSOP-NA. Journal of Atmospheric and Oceanic Technology, 10(4), 591-608.

16. Kent, E. C., Tiddy, R. J., & Taylor, P. K. (1993). Correction of marine air temperature observations for solar radiation effects. Journal of Atmospheric and Oceanic Technology, 10(6), 900-906.
17. Kent, Elizabeth C., Peter G. Challenor, and Peter K. Taylor. "A statistical determination of the random observational errors present in voluntary observing ships meteorological reports." Journal of Atmospheric and Oceanic Technology 16.7 (1999): 905-914.
18. National Climatic Data Center/NESDIS/NOAA/U.S. Department of Commerce, Data Support Section/Computational and Information Systems Laboratory/National Center for Atmospheric Research/University Corporation for Atmospheric Research, Earth System Research Laboratory/NOAA/U.S. Department of Commerce, and Cooperative Institute for Research in Environmental Sciences/University of Colorado. 1984, updated monthly. International Comprehensive Ocean-Atmosphere Data Set (ICOADS) Release 2.5, Individual Observations. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. http://dx.doi.org/10.5065/D6H70CSV. Accessed on 17/11/2015.
19. Nunna, K., Udaya Bhaskar, T. S., Pattabhi Rama Rao, E., & Raju, J. S, 2018. Correction to Beaufort estimated wind speed over the Tropical Indian Ocean. Meteorological Applications.
20. Pavan Kumar, J., Kameshwari, N., Udaya Bhaskar, T. V. S., & Rama Rao, E. P. (2017). MaMeAT-A tool for visualizing marine meteorological data for Naval applications.
21. Praveen Kumar, B., J. Vialard, M. Lengaigne, V. S. N. Murty, M. J. McPhaden, M. F. Cronin and K. Gopala Reddy, 2012: *Evaluation of Air-sea heat and momentum fluxes for the tropical oceans and introduction of TropFlux*, CLIVAR exchanges, Issue 58, February 2012, tropflux_clivar_final_acc.pdf
22. Rajeevan, M., Pai, D. S., & Thapliyal, V. (2002). Predictive relationships between Indian Ocean sea surface temperatures and Indian summer monsoon rainfall. Mausam, 53(3), 337-348.
23. Tetens formula from
" http://faculty.eas.ualberta.ca/jdwilson/EAS372_13/Vomel_CIRES_satvpformulae.html"