



**NMRF/TR/02/2020**



सत्यमेव जयते

**TECHNICAL REPORT**

**Implementation of NEMO based Global 3D-  
Var Ocean Data Assimilation System at  
NCMRWF: Technical Aspects**

Imranali M. Momin, Ashis K. Mitra and E. N. Rajagopal

**June 2020**

*National Centre for Medium Range Weather Forecasting  
Ministry of Earth Sciences, Government of India  
A-50, Sector-62, NOIDA, India 201 309*

# **Implementation of NEMO based Global 3D-Var Ocean Data Assimilation System at NCMRWF: Technical Aspects**

Imranali M. Momin, Ashis K. Mitra and E. N. Rajagopal

**June 2020**

**National Centre for Medium Range Weather Forecasting**

**Ministry of Earth Sciences, Government of India**

**A-50, Sector 62, NOIDA-201309, INDIA**

**[www.ncmrwf.gov.in](http://www.ncmrwf.gov.in)**

**Ministry of Earth Sciences**  
**National Centre for Medium Range Weather Forecasting**  
**Document Control Data Sheet**

|    |                                 |   |
|----|---------------------------------|---|
| 1  | Name of the Institute           | National Centre for Medium Range Weather Forecasting  |
| 2  | Document Number                 | <i>NMRF/TR/02/2020</i>  |
| 3  | Month of publication            | June 2020   |
| 4  | Title of the document           | Implementation of NEMO based Global 3D-Var Ocean Data Assimilation system at NCMRWF: Technical Aspects  |
| 5  | Type of Document                | Technical Report  |
| 6  | No of pages, Figures and Tables | 26 Pages, 9 Figures and 7 Tables  |
| 7  | Number of References            | 22  |
| 8  | Author (S)                      | Imranali M. Momin, Ashis K. Mitra, and E. N. Rajagopal  |
| 9  | Originating Unit                | NCMRWF  |
| 10 | Abstract                        | A high resolution global NEMO based Ocean Data Assimilation (ODA) system is implemented at National Centre for Medium Range Weather Forecasting (NCMRWF) high performance system (HPC) system named “ <i>Bhaskara, IBM iDataPlex</i> ” during February-2017 and later in “ <i>Mihir, CRAY XC40</i> ” during April-2020. This variational ODA system NEMOVar uses the Nucleus European Modelling of the Ocean (NEMO) ocean model and the Los Alamos sea ice model (CICE) as physical model at 1/4 degree horizontal resolution with 75 layers in vertical for dynamic and thermodynamics. This system assimilates the satellite and in-situ sea surface temperature (SST), in-situ temperature and salinity profiles in vertical, satellite sea level anomaly (SLA) observations and sea ice concentrations using a 24 hour data assimilation window. In this report, we have discussed the technical aspect of implementation of said ODA system which is based on ROSE/Cylc framework for managing and running the operational or research meteorological suites. It includes the preparation of the atmospheric surface boundary conditions (SBCs) from operational Global NCMRWF Unified Model (NCUM) model to force the ocean model, building and running the ocean data assimilation suite. The preparation of SBCs and output from global NEMOVar ODA are also discussed. Few examples of Indian Ocean features from the NEMOVar ODA system are also discussed. |
| 11 | Security classification         | Non-Secure  |
| 12 | Distribution                    | Unrestricted Distribution   |
| 13 | Key Words                       | NCUM, NEMO, Variation Ocean Assimilation, Ocean Initialisation  |

## Table of Contents

|    | <i>Topic</i>                           | <i>Page</i> |
|----|--|-------------|
|    | Abstract                               | 5           |
| 1. | Introduction                           | 6           |
| 2. | Ocean Observations                     | 7           |
| 3. | Atmospheric fluxes for NEMOVar         | 9           |
| 4. | NEMOVar suite structure                | 11          |
| 5. | Technical description of NEMOVar suite | 13          |
| 6. | Ocean Analysis Outputs                 | 16          |
|    | References                             | 25          |

## Abstract

A high resolution global NEMO based Ocean Data Assimilation (ODA) system is implemented at National Centre for Medium Range Weather Forecasting (NCMRWF) high performance system (HPC) systems named “*Bhaskara, IBM i-DataPlex*” during February-2017 and later in “*Mihir, CRAY XC40*” during April-2020. This variational ODA system NEMOVar uses the Nucleus European Modelling of the Ocean (NEMO) ocean model and the Los Alamos sea ice model (CICE) as physical model at 1/4 degree horizontal resolution with 75 layers in vertical for dynamic and thermodynamics. This system assimilates the satellite and in-situ sea surface temperature (SST), in-situ temperature and salinity profiles in vertical, satellite sea level anomaly (SLA) observations and sea ice concentrations using a 24 hour data assimilation window. In this report, we have discussed the technical aspect of implementation of said ODA system which is based on ROSE/Cylc framework for managing and running the operational or research meteorological suites. It includes the preparation of the atmospheric surface boundary conditions (SBCs) from operational Global NCMRWF Unified Model (NCUM) model to force the ocean model, building and running the ocean data assimilation suite. The preparation of SBCs and output from global NEMOVar ODA are also discussed. Few examples of Indian Ocean features from the NEMOVar ODA system are also discussed.

## 1. Introduction

Ocean data assimilation is a mathematically rigorous process, which combines the ocean observations and ocean models to take out the important information of the ocean circulation and associated thermocline fields. We know that the ocean observations are sparse and incomplete over the time. The main purposes of ocean assimilation are suitably combine model and observation to monitor the ocean circulation, and predict the ocean circulation at different spatial and temporal scales. The data assimilation approaches vary significantly in term of the assimilation method, observations assimilated, and also in term of forecast error covariance, model biases, observation errors and the quality control procedure for different types of observation. Various ocean data assimilation products were developed during the Global Ocean Data Assimilation Experiment (GODAE, Bell et al., 2009). In these data assimilation products, both model and observation are assumed as erroneous. The models have errors due to deficiencies in the model physics, grid resolution, lateral boundary conditions, and atmospheric forcing while the observations have error due to instrument or representative error. One important impact of the data assimilation is to counter the tendency of ocean models to drift away from reality. A large number of methods for combining model and observational data are described in the literature. These methods are classified in three classes: Variational methods such as 3DVar or 4DVar (Lorenz, 1986) based on the minimization of a cost function that measures the differences between the model and the observations, the various levels of approximation to the extended Kalman filter also called as sequential schemes (Daley, 1991), and ensemble-based schemes such as ensemble Kalman filter (Evensen, 1994). Each of these approaches has its own advantages and disadvantages with respect to the approximations made, complexity and computational cost.

The Nucleus European Modelling of the Ocean (NEMO, Madec, 2008) based variational data assimilation system called NEMOVar and its related research, development and operational implementation is initiated by collaboration between CERFACES, European Centre Medium Weather Forecast (ECMWF), INRIA and UK Met Office. This NEMOVar consortium aims to develop a multi-incremental variational assimilation system with NEMO. ECMWF was first to implement the NEMOVar system (Mogensen et al., 2012; Balmaseda et al., 2013) at ORCA1 configuration based on multivariate incremental three dimensional assimilation system (Weaver et al., 2005; Daget et al., 2009) which produced the ocean reanalysis on non-operational basis called NEMOVar-Combine (Balmaseda et al., 2010) and operational basis called as Ocean Re Analysis System 4 (ORAS4; Balmaseda et al., 2013) with basic difference of the parameter choices in data assimilation. Both reanalysis used 42 vertical levels out of which the 15 vertical levels are in upper 200m. Later, Met Office implemented the high resolution ( $1/4^\circ$ ) variational data assimilation system for operational global ocean analysis and forecast system which includes the flow-dependent vertical background error covariance of temperature, salinity and statistical background error covariance for Sea Surface Height (SSH) and sea ice concentration (Water et al., 2013; 2014). The flow dependent error co-variances are specified as combination of statistical error variance based on NMC method (Parrish and Derber, 1992) and vertical parameterization based on the mixed layer depth. The horizontal background error correlation for temperature, salinity and sea ice concentration are computed based on the first baroclinic Rossby radius while set at  $4^\circ$  for SSH. Initial comparisons of NEMOVar with its preceding Analysis Correction

scheme (Storkey et al., 2010) show considerable improvements on the ocean fields in NEMOVar especially over high variability areas as well as Atlantic meridional overturning circulation (Water et al., 2014). Blockley et al. (2013) shows the detail description of new NEMO based global ocean assimilation and forecast system including the model physics, data assimilation, observation assimilated, operational implementation, and its forecast assessment. The fundamental equation of NEMOVar is an incremental cost function.

$$J(\delta x) = \frac{1}{2} \delta x^T \mathbf{B}^{-1} \delta x + \frac{1}{2} (d - \mathbf{H} \delta x)^T \mathbf{R}^{-1} (d - \mathbf{H} \delta x)$$

Where, the increment  $\delta x = x - x_b$  is the difference between the state vector  $x$  and its background estimate  $x_b$ ,  $d = y - H(x_i)$  is the innovation vector.  $y$  is the observation vector and  $x_i = M_{t_n \rightarrow t_i}(x_b)$ . the notation  $M_{t_n \rightarrow t_i}$  indicates the nonlinear propagation of model from the background state to the state at  $i^{\text{th}}$  time.  $\mathbf{B}$  and  $\mathbf{R}$  are the model background error covariance and observed error covariance respectively.  $\mathbf{R}$  is a diagonal matrix with the assumption of uncorrelated observation errors. The operator  $\mathbf{H}$  is the observation operator while matrix  $\mathbf{H}$  denotes the linearized observation operator.

In this report, we discuss the implementation and technical details of NEMOVar run on Bhaskara (IBM i-DataPlex) in February-2017 and later on Mihir (Cray XC40) in April-2020 HPCs. There are three stages for the NEMOVar, i) Preparation of atmospheric fluxes ii) Building and Running the NEMOVar suite iii) Preparation of ocean analysis. Section 2 describes the detail of preparation of atmospheric fluxes which includes the extraction of variables from the global atmospheric model and its interpolation to NEMO ocean model tripolar grid resolution. Section 3 and 4 show the structure of NEMOVar suite and its technical description. Section 5 shows the post processing of NEMOVar output which remaps the global NEMO output from tripolar grid to regular latitude-longitude grid resolution.

## 2. Ocean Observations

The NEMOVar system assimilates the in-situ sea surface temperature (SST), sub-surface temperature and salinity profiles, satellite derived SST, sea level anomaly, and sea ice concentration data. The satellite SST is obtained from the Global High Resolution Sea Surface Temperature (GHRSSST) project which is biased corrected by the reference or unbiased in-situ SST. The sea level anomaly observations are provided by the MyOcean project. The sea ice concentration data are derived from the Special Sensor Microwave Imager/Sounder (SSMIS) data provided by the EUMETSAT Ocean Sea Ice Satellite Application Facility (OSI-SAF). Figure 1(a)-(d) shows the various observations from in-situ, and satellite SST, sub-surface T/S profiles, and sea ice concentration for particular day respectively.

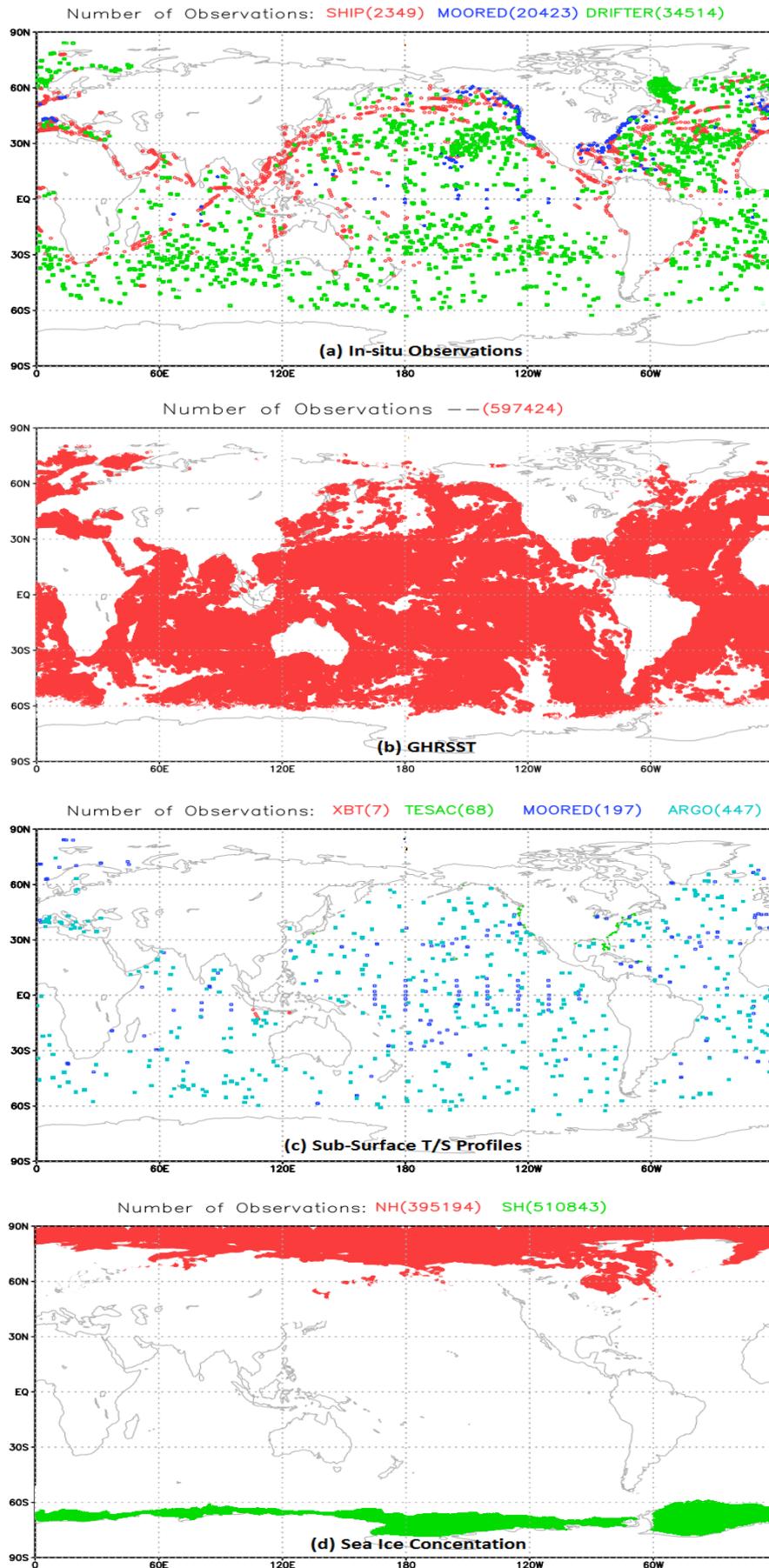


Figure 1(a)-(d): Ocean observations from in-situ, and satellite SST, T/S profiles and sea ice concentration.

### 3. Atmospheric Fluxes for NEMOVar

NEMOVar ODA system requires the surface boundary conditions (SBCs) from the global atmospheric model. Table-1 shows the list of SBCs, its units and frequencies. At National Centre for Medium Range Weather Forecasting (NCMRWF), the Unified Model based global atmospheric model called NCUM was implemented for the seamless numerical modelling system by Rajagopal et al., (2012). Later, the NCUM model upgraded in term of physics (global atmosphere version 6.1) and dynamic (ENDGame dynamic) at 17 km (Rakhi et al., 2016) and at 12 km with Hybrid 4D-Var (Sumit et al., 2018). To extract the SBCs from the NCUM and convert into NetCDF format, “subset.tcl” script is used. All the SBCs are extracted at 3 hourly time interval except wind forcings at 1 hourly time interval. After this, the interpolate suite is used to interpolate the NCUM SBCs to NEMO tripolar grid resolution using the Spherical Coordinate Remapping and Interpolation Package (SCRIP) remapping tool. Figure 2(a)-(h) shows for a typical day the zonal, and meridional winds(m/s), downward shortwave, and downward longwave radiation( $w/m^2$ ), 2 m air temperature (K) and specific humidity (kg/kg), precipitation and snowfall rate ( $kg/m^2/s$ ) from 12 km NCUM global model.

**Table-1: List of SBCs and its units and frequencies**

| <i>Index</i> | <i>SBC</i>                   | <i>Units</i> | <i>Frequency</i> |
|--------------|------------------------------|--------------|------------------|
| 1            | Zonal wind                   | $m/s^2$      | 1 h              |
| 2            | Meridional wind              | $m/s^2$      | 1 h              |
| 3            | Downward Shortwave Radiation | $w/m^2$      | 3 h              |
| 4            | Downward Longwave Radiation  | $w/m^2$      | 3 h              |
| 5            | Specific humidity            | kg/kg        | 3 h              |
| 6            | Air temperature              | K            | 3 h              |
| 7            | Precipitation Rate           | $kg/m^2/s$   | 3 h              |
| 8            | Snowfall                     | $kg/m^2/s$   | 3 h              |

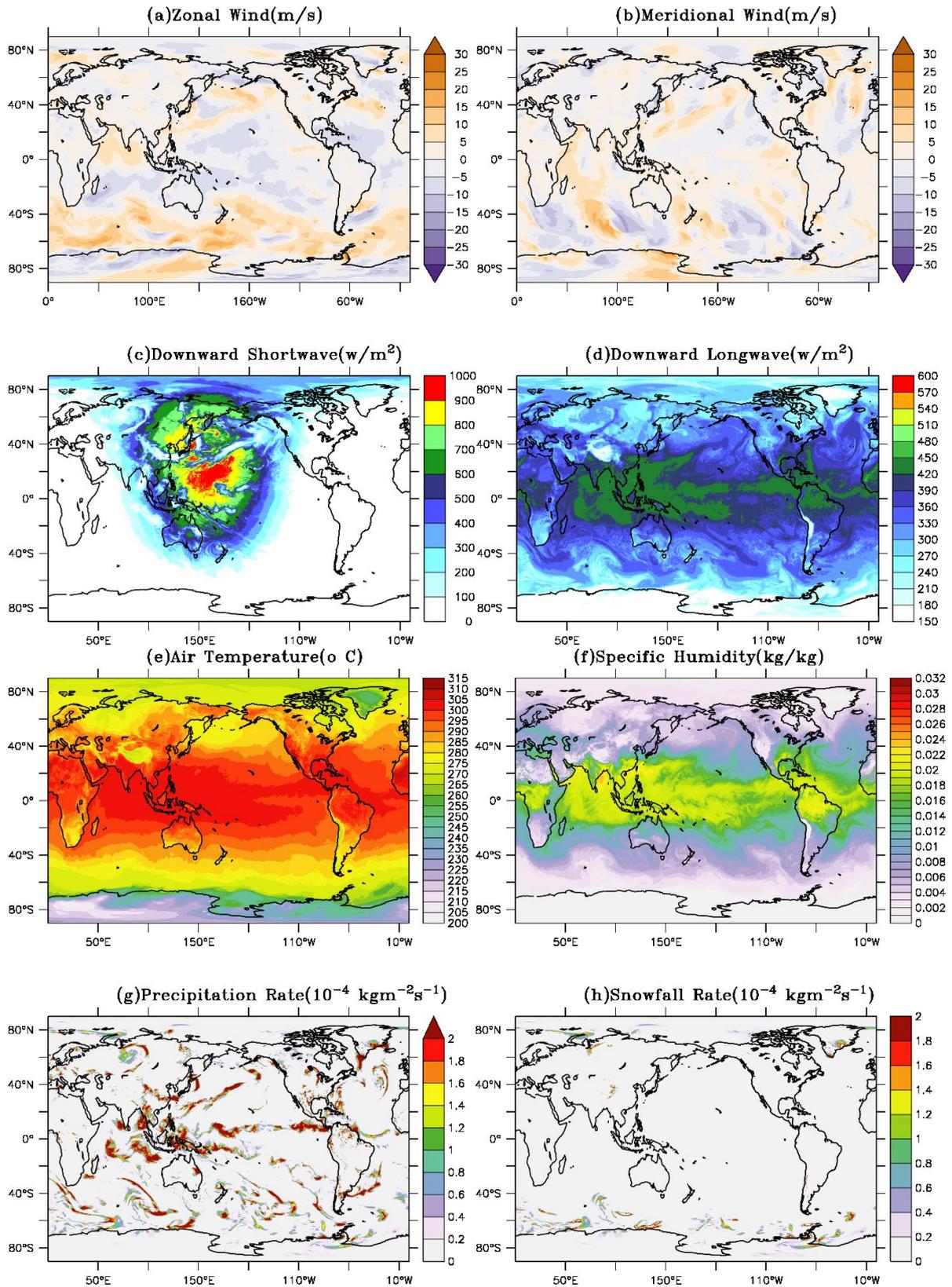


Figure 2(a)-(h): Surface Zonal, and Meridional wind (m/s), Downward shortwave and Downward longwave radiation ( $w/m^2$ ), 2 m air temperature (K) and Specific humidity (kg/kg), Precipitation and Snowfall rate ( $kg/m^2/s$ ) on typical day from 12 km NCUM global NWP model respectively.

#### 4. NEMOVar Suite Structure

We have discussed the operational NEMOVar suite based on the ROSE/Cylc environment. ROSE is a toolkit for writing, editing and running application configurations. ROSE uses the Cylc workflow engine for managing and running suites of inter-dependent application tasks (Oliver et al., 2018). Cylc's suite are very simple, human readable text format, configure various task scheduling using the dependency graph, no central server process for different suites (each suite has its own scheduler daemon), efficient scheduling for the cyclic process, easily identify the fail task, and removal of tasks at run time. The structure of the NEMOVar suite is defined here.

- Suite.rc : Suite structure and operations
- rose-suite.info: basic suite information
- rose-suite.conf: Configuration files controls the assimilation period, assimilation window, compute host, assimilation of various observations.

Some of variables used in rose-suite.conf are defined as

```
START POINT = YYYYMMDD      #Start point of run
FINAL POINT = YYYYMMDD      #Last Point of run
CYCLE LENGTH = 'P1D'        # (24hr of Assimilation cycle)
COMPUTER HOST = 'ncmlogin3'  #Host used for the compilation and model run
ASM_DATA = True/False       # Assimilation of Observations
ASM_ENACT = True/False      # Assimilation of Profile Observations
ASM_SLA = True/False        # Assimilation of Altimeter Sea Level Anomaly
ASM_SST = True/False        # Assimilation of Sea Surface Temperature
ASM_SEAICE = True/False     # Assimilation of Sea Ice
FORECAST_ENABLE = True/False # NEMO Forecast Run
FORECAST_LENGTH = 'P9D'     # NEMO Forecast Run length
```

Figure 4 shows an example of NEMOVar suite based on ROSE/Cylc environment for a typical day with 24 hours assimilation cycle length. The dependency and scheduling of all these tasks are managed through the Cylc. The rose suite contains several applications called as apps to perform various tasks. The apps and its description are as followed.

- 1) **app/fcm make nemo**: This app copies the NEMO, NEMOVar and configuration files from the repository to the HPC using the flexible configuration management (FCM) and builds the NEMO and NEMOVar code in HPC. In Bhaskara and Mihir HPCs, we have used the fcm\_make2.cfg configuration file to build the NEMO and NEMOVar code.
- 2) **app/daily fluxes**: This app copies the daily flux files from the INPUT\_DIR to OUTDIR\_FLUXES and unzips it. The bin/prepare flux script is the main script for the above process.
- 3) **app/daily observations**: This app copies the observations files for assimilation into the particular cycle. Everything in the opt directory controls a different set of observations. The bin/NemoQcProg\_ExtractAndProcess script copies the observations file in the place and run the NEMO QC to file convert into 'feedback' format. After, 09-Nov-2016, the

bin/NemoQcProg\_ExtractAndProcess script is modified which copies the observations and renamed it because the observations are already quality controlled and in feedback format.

- 4) **app/fixd ancillary**: This app copies the ancillary files like bathymetry, x/y positions, rivers, and XML configuration of NEMO into the output directory. The bin/operational\_ancillary is the scripts to create softlinks for all the ancillary files.
- 5) **app/fixd restart**: This app copies the initial restarts which are used as the background state on the first assimilation. In further iterations of the cycle, restarts are taken from the analysis of the previous iteration. The bin/copy\_restarts is the main script to create a link from RESTART\_DIR to OUTDIR\_RESTARTS.
- 6) **app/sst bias**: This app is used to correct the satellite estimates SST observations. The reference set of the observations assumed as unbiased are used to correct the bias satellite observations. At beginning, the matchup points are extracted from the bias observations and reference dataset within the distance of less than 25km using matchup.exe. Once the matchups are found, they are used in the calculation of observation differences where the reference data subtracted from the bias observations. These differences are then treated as “observations of bias” and assimilated using the NEMOVar into a pre-existing background field of bias.
- 7) **app/NEMOVar**: This app is called NEMOVar. The bin/init\_NEMOVar.sh and bin/run\_NEMOVar.sh are the two main scripts for the NEMOVar. The bin/init\_NEMOVar.sh creates the directory named DIR\_INPUT and copies the innovations, the assimilation background and the altimeter bias into them. The bin/run\_NEMOVar.sh linked the coordinates, bathymetry, and assim\_background files into DIR\_WORK. The function interp\_sd is used to interpolate between covariances in different seasons. The covariances used are for the profiles, SST, altimeter, and sea ice. The current and previous seasons are determined and the weights are produced based on how far through the seasons. After the interpolation has been performed, more links for Rossby radii, x/y positions are produced. The namelist NEMOVar file is also updated and finally, NEMOVAR itself is run.
- 8) **app/nemo cice**: This app configures NEMO-CICE. There are three stages: obsoper, Increment Analysis Update (IAU) and forecast. The governing script for the obsoper, IAU and forecast is run\_nemo\_cice. Both NEMO and CICE are configured before running them.

**Configuring NEMO and CICE:** The flux files are linked into the CYLC\_TASK\_WORK\_DIR/fluxes directory. The ancillary files, background restarts, altimeter bias and the observations are also linked into CYLC\_TASK\_WORK\_DIR. For the first cycle, the initial restart files are used; in subsequent cycles, the files produced by the IAU stage are used. The function handle\_observations is used to update the list of input files in namelist. The CICE namelist is also configured and called ice\_in. After the NEMO and CICE have been configured, they are run using the MPI. When the job finishes the innovations are linked to the directory OUTDIR\_INNOVATIONS.

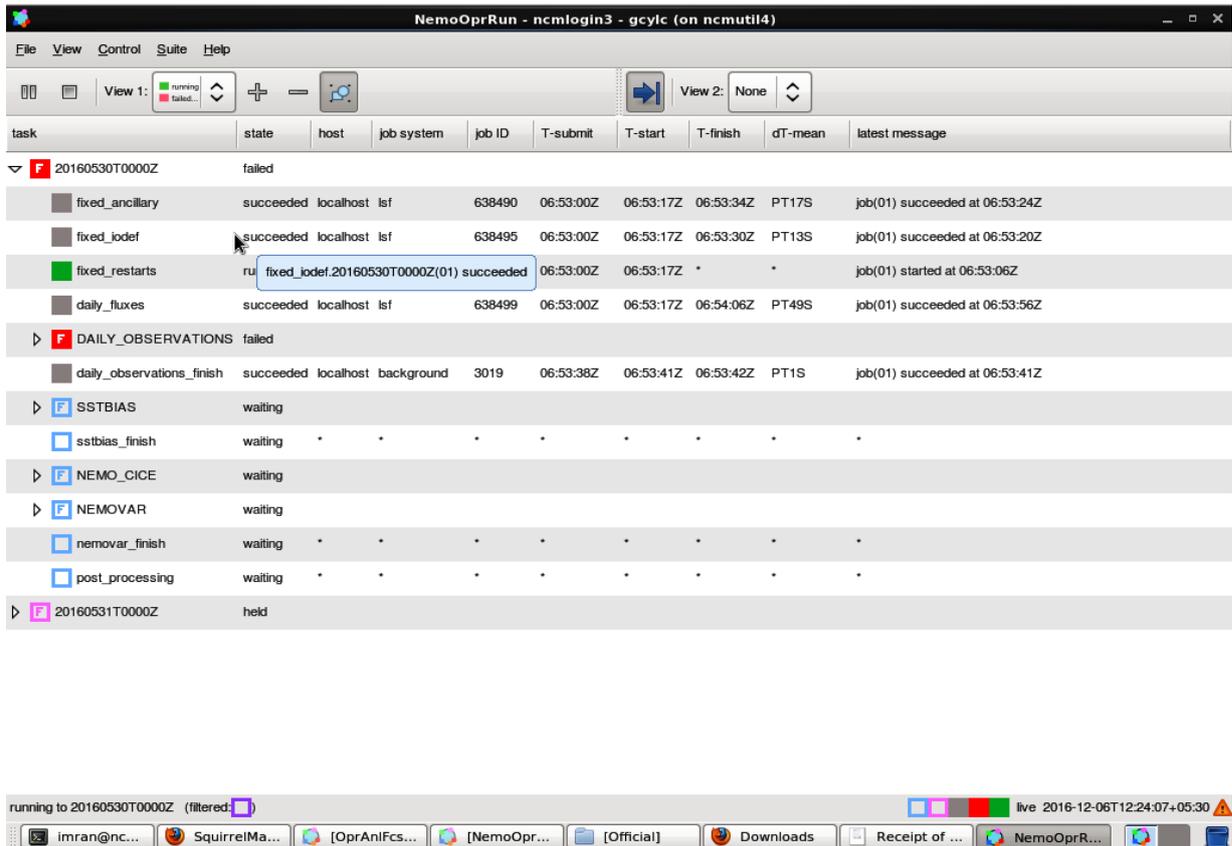


Figure 3: Various tasks of NEMOVar suite for the typical day

## 5. Technical description of NEMOVar Suite

The description of physical model used in NEMOVar system is available in literature (Blockley et al., 2013). Figure 4 shows the flow chart of NEMOVar suite run at NCMRWF. A detailed description of the daily NEMOVar suite is as follows:

- 1) The Quality controlled ocean observations are used here daily. All observations are then copied and extracted at for  $[T - 24 \text{ h}, T + 00 \text{ h}]$  time periods. The satellite SST bias correction is then performed using the reference data sets (at present only in situ SST) to correct for biases in the satellite SST data.
- 2) SBCs are processed from NCUM Global NWP system output, using analysis fields from  $T - 24 \text{ h}$  up to  $T + 00 \text{ h}$ . The resulting SBCs are then translated onto the NEMO grids using bilinear interpolation. The preparation of NCUM SBCs is described in section 2.
- 3) A 24 hour NEMO model forecast is then run for the period  $T - 24 \text{ h}$  to  $T + 0 \text{ h}$  and the model forecast fields are mapped into observation space using the NEMO observation operator to create model counterparts using bilinear interpolation in the horizontal and cubic splines in the vertical directions to create First Guess at Appropriate time (FGAT) model–observation differences (innovations) valid at the observation locations/times.

- 4) The FGAT innovations output by the observation operator are then used by the NEMOVar assimilation scheme to generate fields of daily increments as detailed in Waters et al. (2013, 2014).
- 5) The model is then rerun for the period  $T - 24$  h to  $T + 0$  h, and these increments are applied evenly over the 24 h period using an incremental analysis update (IAU) method (Bloom et al., 1996). It also runs on forecast mode from  $T + 0$ h period.

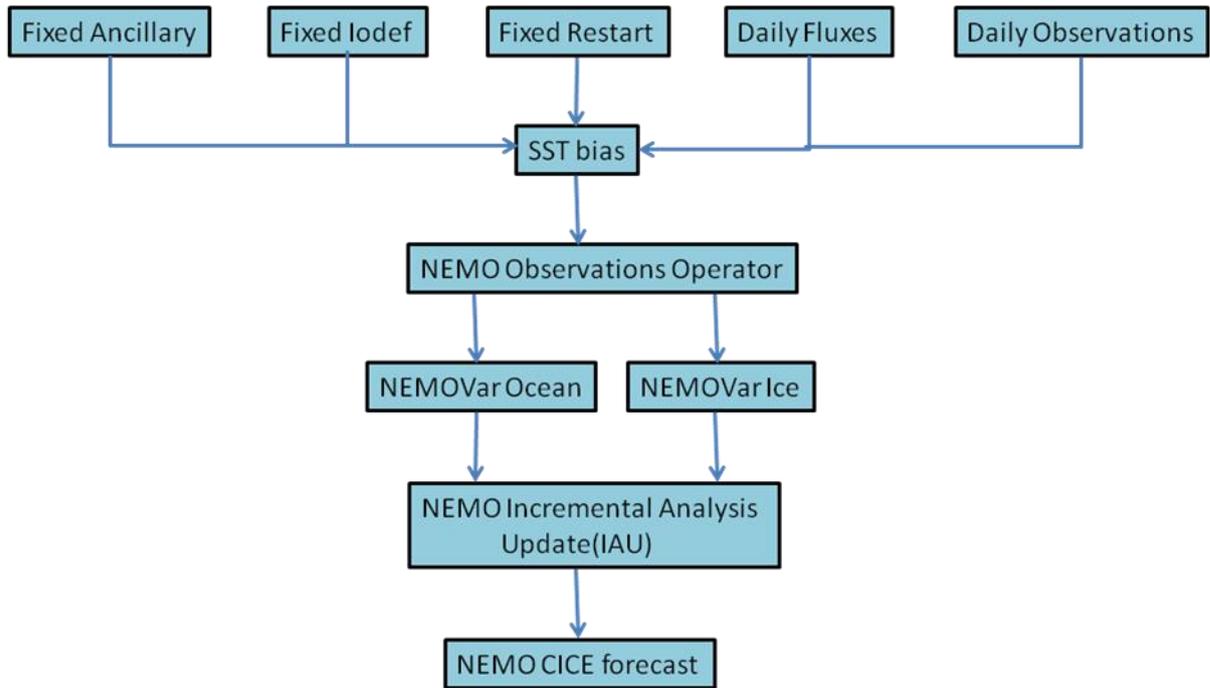


Figure 4: Flow chart of NEMOVar suite run.

The NEMO based ocean assimilation requires the ancillary files such as bathymetry, river runoff, grid information etc for physical ocean model while seasonal dependent background error covariance matrix, ratio of synoptic to meso-scale variability, Rossby radius, and mean dynamic topography for the assimilation purpose. The initialization of NEMOVar is carried out through the restart file. Some executables were precompiled before running the assimilation. Table-2 shows the list of executables used by the assimilation and its functions.

In NEMOVar suite, each task runs on either single or multi-processors. Table-3 shows the number of processors, time, and maximum memory used to complete each task for Bhaskara and Mihir HPCs. The `nemo_cice_obsoper` and `nemo_cice_iau` tasks uses the NEMO global model to produce the background information and increment analysis update respectively. They require ~5 minutes with 192 processors to finish task. Similarly, the ocean and ice assimilation requires the 20 minutes and 4 minutes to complete task with 192 processors. Table-3 shows the technical details of various tasks including the run time, no of processors, and memory usage at “*Bhaskara*” and “*Mihir*” HPCs.

**Table-2: List of executables used assimilation system and its functions**

| <i>Index</i> | <i>List of executables</i>                     | <i>Task Used</i> | <i>Function</i>   |
|--------------|--|------------------|---|
| 1            | interp.exe                                     | Preprocess       | To interpolate the UM grid to NEMO grid resolution  |
| 2            | Fbcomb.exe                                     | NEMOVar          | To combine the observations and first guess available on multiple processors (192 processors)                   |
| 3            | matchup.exe                                    | NEMOVar          | To find the collocated point (matchup point) between satellite observations and in-situ observations            |
| 4            | NemoQcProg_ExtractAndProcess.exe               | NEMOVar          | Apply basic quality control such as spatial and temporal thinning and rewrite it in feedback format (in NetCDF) |
| 5            | nemo_orca025.exe and nemovar_inner_orca025.exe | NEMOVar          | To perform observational operator/incremental analysis update and variational assimilation                      |
| 6            | rebuild_nemo                                   | Post processing  | To combine the restart files available on multiple processors (192 processors)                                  |

**Table-3: Details of NEMOVar Computation Tasks**

| <i>Index</i> | <i>Task Name</i>   | <i>No of Processors</i> | <i>Time taken in Bhaskara/Mihir</i> | <i>Maximum Memory used in Bhaskara/Mihir</i> |
|--------------|--------------------|-------------------------|-------------------------------------|--|
| 1            | fixed_ancillary    | 1                       | 61 Sec.<br>10 Sec.                  | 17 MB  |
| 2            | fixed_iodef        | 1                       | 61 Sec.<br>10 Sec.                  | 17 MB  |
| 3            | fixed_restart      | 1                       | 298 Sec.<br>223 Sec.                | 19 MB<br>37 MB                               |
| 4            | daily_fluxes       | 1                       | 94 Sec.<br>88 Sec.                  | 19 MB  |
| 5            | forecast_fluxes    | 1                       | 119 Sec.<br>85 Sec.                 | 23 MB  |
| 6            | daily_observations | 1                       | 12 Sec.<br>5 Sec.                   | 14 MB  |
| 7            | nemo_cice_obsoper  | 192                     | 306 Sec.<br>244 Sec.                | 150 GB<br>43 GB                              |
| 8            | nemovar_ice        | 192                     | 252 Sec.<br>244 Sec.                | 1 GB   |
| 9            | nemovar_ocean      | 192                     | 1184 Sec.<br>1217 Sec.              | 150 GB<br>44 GB                              |
| 10           | nemo_cice_iau      | 192                     | 338 Sec.<br>266 Sec.                | 150 GB<br>43 GB                              |
| 11           | nemo_cice_forecast | 192                     | 2169 Sec.<br>1937 Sec.              | 190 GB<br>221 GB                             |

## 6. Ocean Analysis Outputs

The output of the ocean assimilation is interpolated from tripolar grid to regular grid resolution using the climate data operator package. The NEMO based analysis of temperature and salinity profiles are compared with the Research Moored Array for African–Asian–Australian Monsoon Analysis and Prediction (RAMA) buoy observations on daily time scale for the study period. RAMA is a key element of the Indian Ocean Observing System (IOOS) programme designed for the study of the large-scale ocean–atmosphere interactions, ocean circulation, and mixed-layer dynamics (McPhaden et al., 2009). Before computation of various statistics such as bias, standard deviation and RMSE for temperature and salinity profile, we first remove data from NEMO analysis corresponding to data gap region of buoy observations. Figure 5 (a)-(c) shows the daily mean temperature profiles of NEMO analysis and RAMA buoy, and its difference at three continuous buoy locations ( $4^{\circ}$  S;  $80.5^{\circ}$  E,  $12^{\circ}$  N;  $90^{\circ}$  E and  $15^{\circ}$  N;  $90^{\circ}$  E) for a long period daily (2016-2018). The temperature profile of RAMA buoy observations is available up to 200m depth. Comparison result shows that the NEMO analysis captures the daily mean temperature very well especially over the upper ocean up to 80 m. At buoy location ( $12^{\circ}$  N;  $90^{\circ}$  E), the higher systematic bias observed over the sub-surface compare to other two buoy observations. Various error metrics such as Mean, standard deviation, root-mean-square error (RMSE), and correlation coefficient are computed for NEMO Sea Surface Temperature (SST) analysis at seven different buoy observations on daily and 5-days timescales for 2016-2018 periods (Table 4 and Table-5). Both timescale, the NEMO analysis captures the mean and variability of SST very well with a small systematic underestimation. It also shows that the correlation of SST is more than 0.9 for all buoy observations except  $12^{\circ}$  N;  $90^{\circ}$  E.

Similar comparison of daily mean salinity profiles of NEMO analysis and RAMA buoy, and its difference at three different buoy locations ( $4^{\circ}$  S;  $80.5^{\circ}$  E,  $12^{\circ}$  N;  $90^{\circ}$  E and  $15^{\circ}$  N;  $90^{\circ}$  E) is shown in Figure 6(a)-(c). The result shows that the NEMO analysis captures the vertical distribution of salinity very well. For  $12^{\circ}$  N;  $90^{\circ}$  E and  $15^{\circ}$  N;  $90^{\circ}$  E buoy observations, the low salinity were examined near the surface. Table-6 and Table-7 represents the mean, standard deviation, RMSE, and correlation coefficient for NEMO Sea Surface Salinity (SSS) analysis at seven different buoy observations on daily and 5-days timescales (2016-2018). The NEMO analysis overestimates/underestimates the mean and variability for northern BOB and southern Indian Ocean region respectively. The maximum RMSE were observed at northern BOB with poor correlation. The various statistics are improved in 5-day average timescale as compare to daily time scale.

The surface zonal and meridional current from NEMO analysis was also compared against the RAMA buoy and satellite estimates current. Figure 7 shows the temporal variation of zonal and meridional current (m/s) with respect to RAMA buoy over the BOB region. It shows that the NEMO analysis captured zonal and meridional current very well. Further, the spatial distribution of NEMO analysed mean surface current (m/s) and its variability with respect to Ocean Surface Current Analysis Real-time (OSCAR; Bonjean, F., and G. S. E. Lagerloef, 2002) is shown in figure 8(a)-(b) during southwest monsoon-2016. NEMO captures the mean and its variability very well spatially over the off coast of Sri Lanka and east coast of India.

The warm sea surface temperature (SST) plays a major role in weather and climate change. The Indo-Pacific warm pool is a large region of warm water with SSTs above 28°C lies in the equatorial Indian and western Pacific Oceans while the Indian Ocean Warm Pool (IOWM) is a region of highest temperature and large spread over the Arabian Sea before the onset of summer monsoon (Vinayachandran and Shetye, 1991). We have computed the heat content up to 80 m (J/m<sup>2</sup>) from the NEMO ocean analysis at daily time scale to see the propagation of heat over the Arabian Sea mini warm pool region. The upper ocean heat content (80 m) is defined as

$$HC = \rho C_p \int_0^{80} T dz$$

Where,  $\rho$  is the density of the water and  $C_p$  is specific heat capacity of the sea water at constant pressure. T is the ocean temperature at each layer of ocean thickness dz. Figure 9 shows the upper ocean heat content up to 80m (J/m<sup>2</sup>) average over the latitude 5-15° N from the daily ocean analysis. It shows the gradually increases of upper ocean heat content from February average over the latitude 5-15° N. It starts decreasing during the onset of summer monsoon (late May to early June period).

### **Acknowledgements**

We would like to thank the UK Met Office for supporting in scientific and technical matters related to NEMOVar. We are thankful to NEMOVar consortium for their support. This work was carried out under the MoES-UKMO collaboration. We would also like to thank Ms Shivali Gangwar of CRAY India Ltd. for her technical support during the implementation. RAMA buoy data is obtained from the TAO Project Office of NOAA/PMEL. OSCAR current data is taken from [https://podaac.jpl.nasa.gov/dataset/OSCAR\\_L4\\_OC\\_third-deg](https://podaac.jpl.nasa.gov/dataset/OSCAR_L4_OC_third-deg).

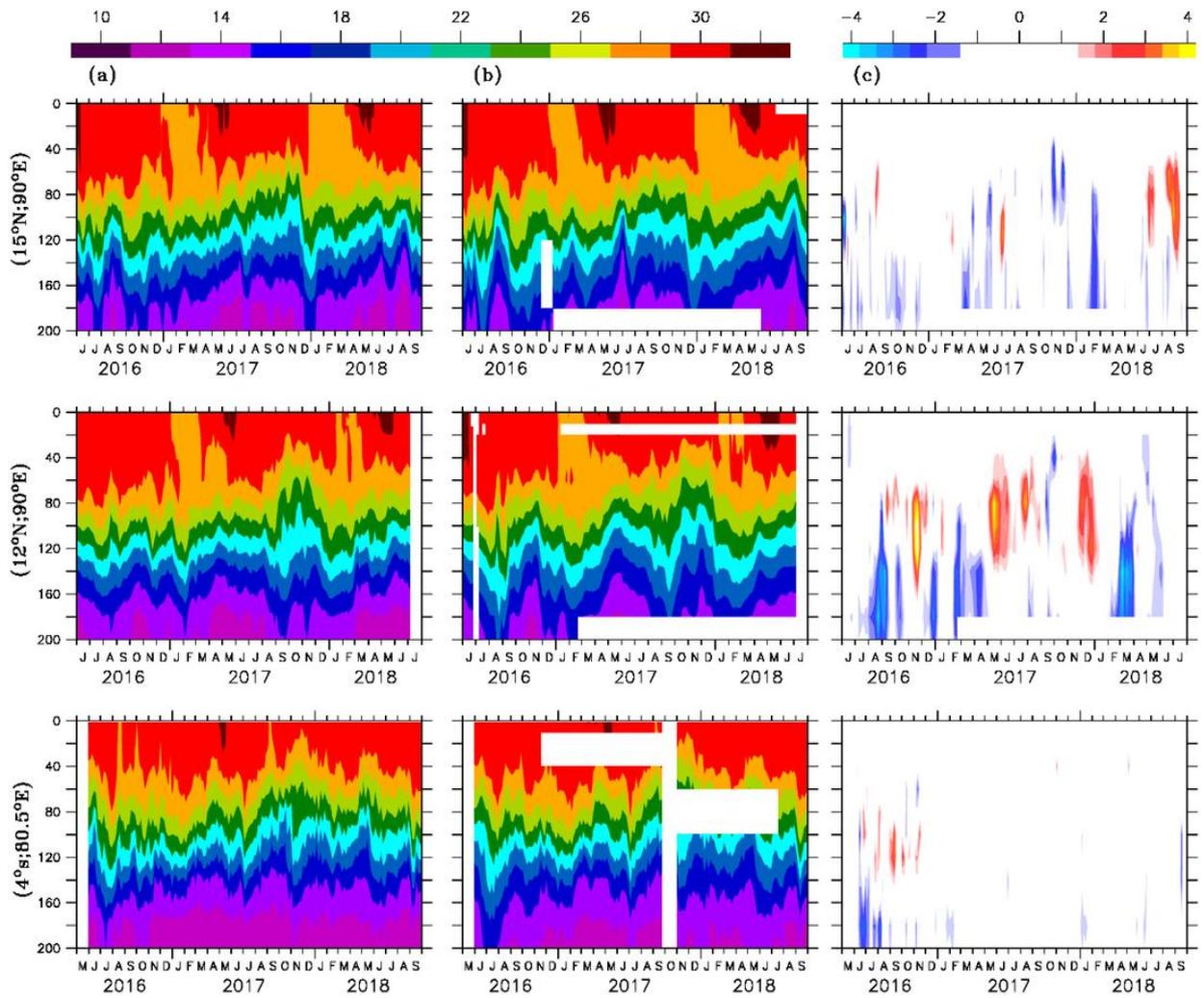


Figure 5 (a)-(c): Vertical distribution of daily mean temperature ( $^{\circ}$  C) profile from NEMO analysis, buoy observation and its difference at  $15^{\circ}$  N;  $90^{\circ}$  E (upper panel),  $12^{\circ}$  N;  $90^{\circ}$  E (middle panel), and  $4^{\circ}$  S;  $80.5^{\circ}$  E (lower panel).

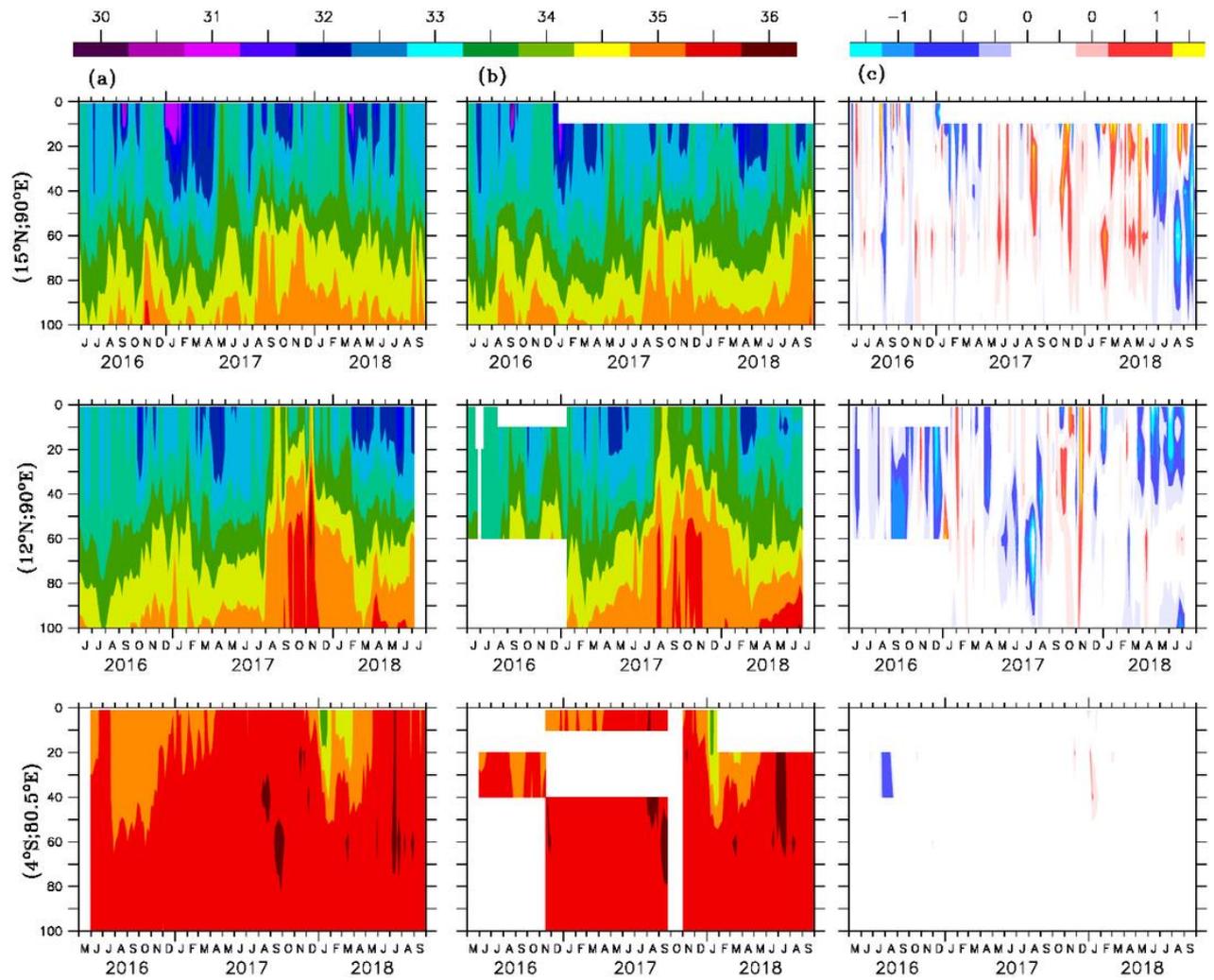


Figure 6(a)-(c): Vertical distribution of daily mean salinity profile from NEMO analysis, buoy observation and its difference at 15° N; 90° E (upper panel), 12° N; 90° E (middle panel), and 4° S; 80.5° E (lower panel)

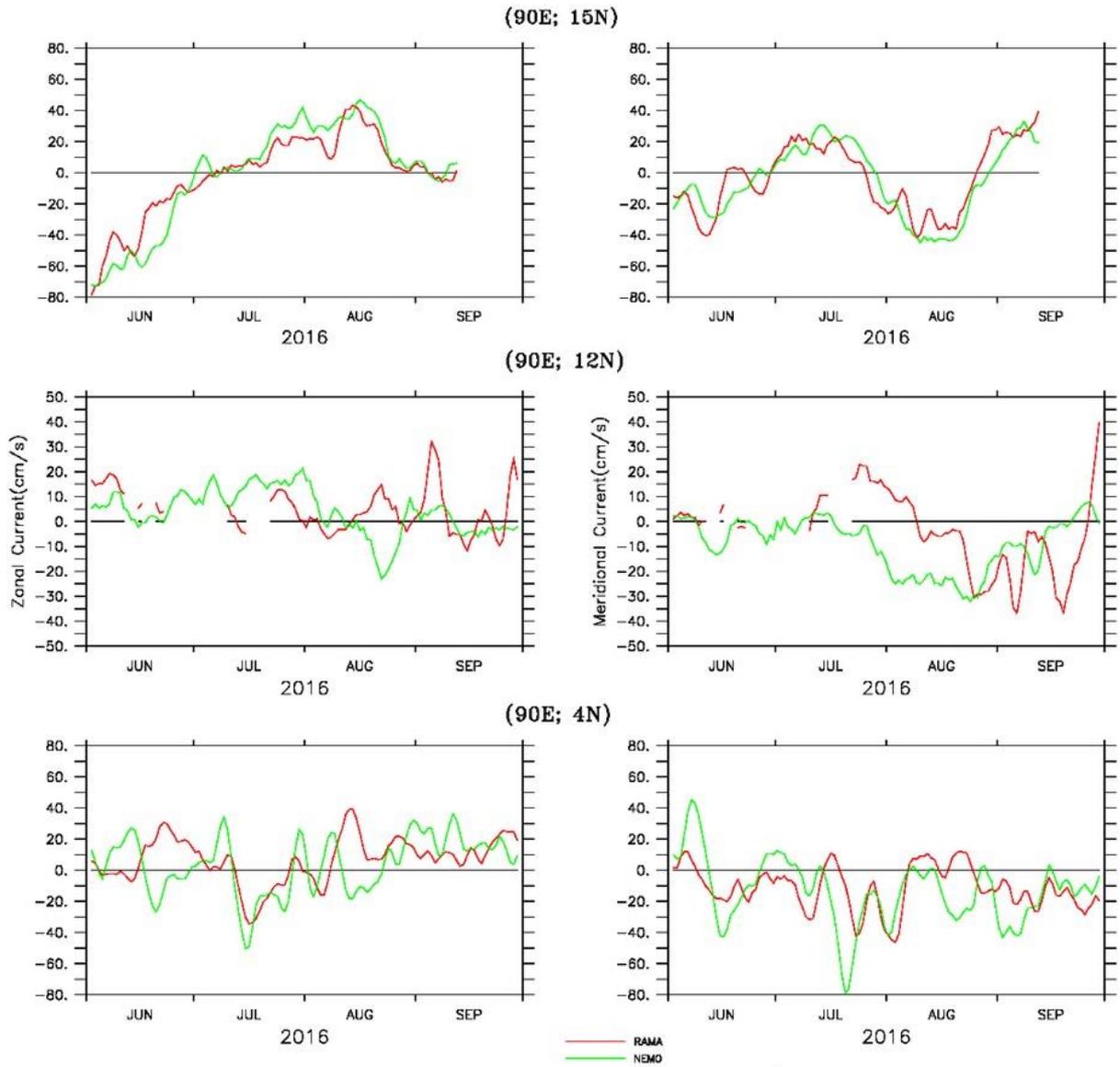


Figure 7(a)-(c): The temporal variation of surface zonal and meridional current (m/s) from NEMOVar analysis and RAMA buoy over the BoB during southwest monsoon-2016

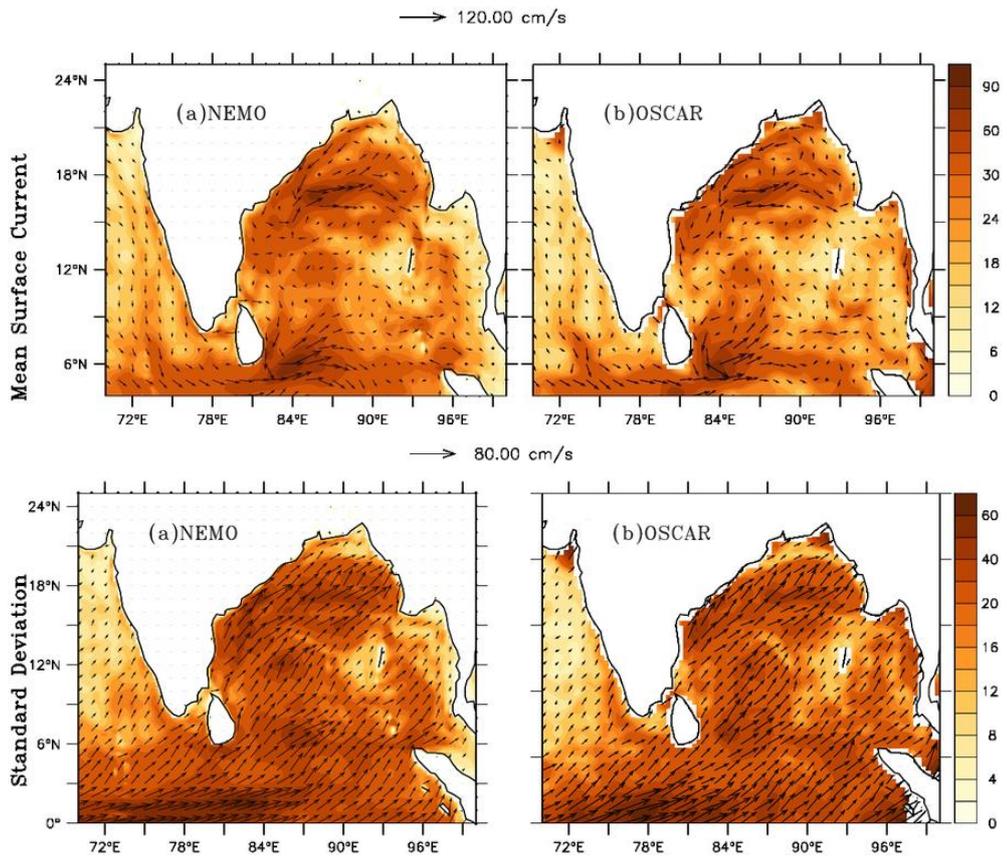


Figure 8: The spatial distribution of NEMO simulated mean surface current and its variability with OSCAR current during southwest monsoon-2016

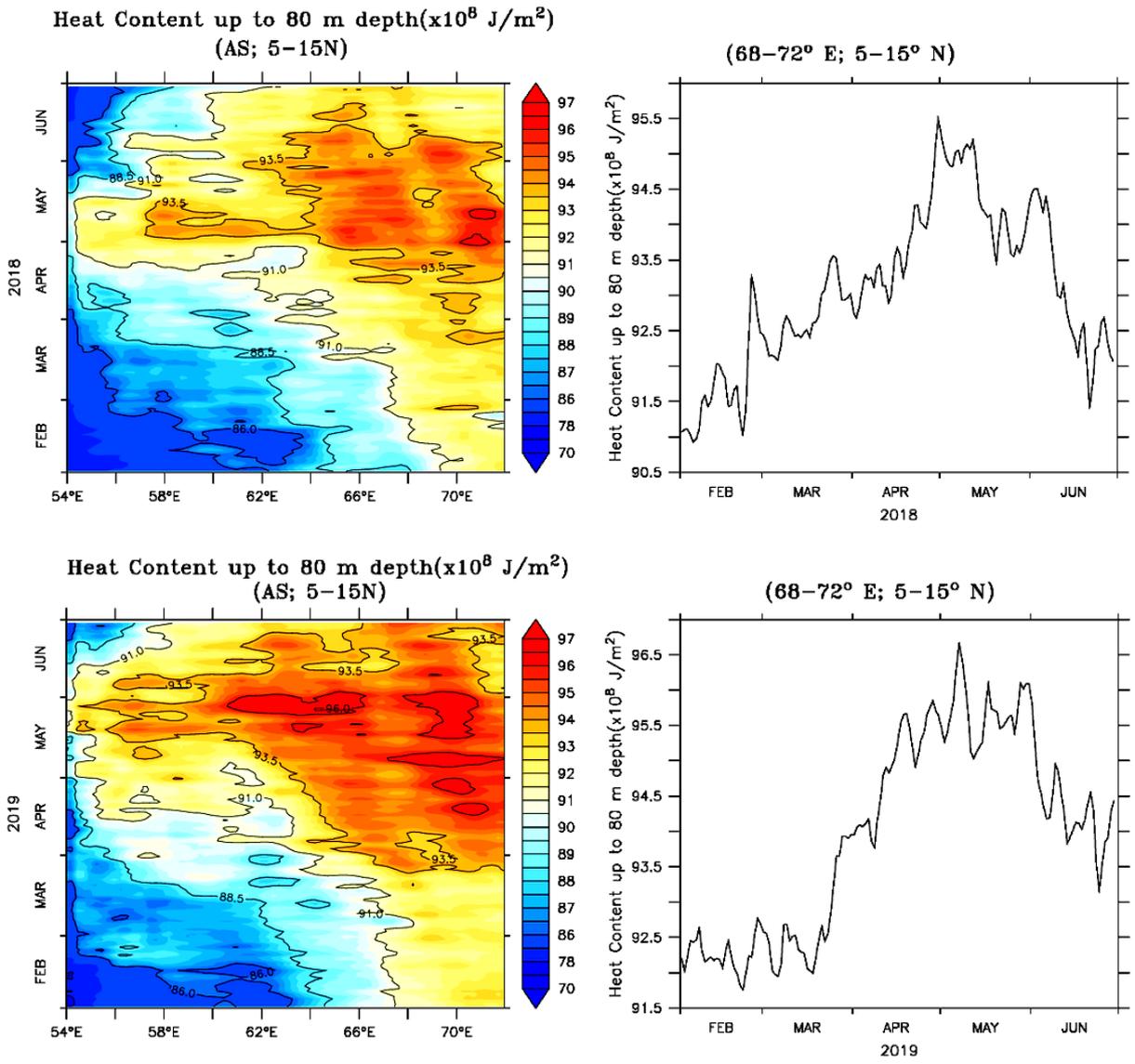


Figure 9: Upper Ocean Heat Content up to 80 m ( $\text{J/m}^2$ ) from NEMO ocean analysis for 2018 and 2019 periods

**Table-4: Mean, standard deviation, root-mean-square error (RMSE), and correlation coefficient for NEMO SST (° C) and RAMA buoys at daily time scale**

| <i>Buoy Locations</i> | <i>No. of Observations</i> | <i>Mean SST (° C)</i> |         | <i>Standard Deviation (° C)</i> |         | <i>Root Mean Square Error (° C)</i> | <i>Correlation</i> |
|-----------------------|----------------------------|-----------------------|---------|---------------------------------|---------|-------------------------------------|--------------------|
|                       |                            | OBS                   | NEMOVAR | OBS                             | NEMOVAR |                                     |                    |
| 12° S; 67° E          | 437.00                     | 27.38                 | 27.35   | 1.26                            | 1.15    | 0.20                                | 0.99               |
| 8° S; 67° E           | 439.00                     | 28.03                 | 28.14   | 0.95                            | 0.90    | 0.19                                | 0.98               |
| 8° S; 55° E           | 847.00                     | 27.43                 | 27.39   | 1.36                            | 1.30    | 0.25                                | 0.98               |
| 4° S; 80.5° E         | 813.00                     | 29.12                 | 29.00   | 0.54                            | 0.54    | 0.25                                | 0.91               |
| 0° S; 67° E           | 846.00                     | 29.50                 | 29.37   | 0.63                            | 0.58    | 0.30                                | 0.90               |
| 12° N; 90° E          | 767.00                     | 29.10                 | 28.90   | 0.90                            | 0.78    | 0.48                                | 0.88               |
| 15° N; 90° E          | 848.00                     | 28.94                 | 28.79   | 1.06                            | 1.01    | 0.36                                | 0.95               |

**Table-5: Mean, standard deviation, root-mean-square error (RMSE), and correlation coefficient for NEMO SST (° C) and RAMA buoys at 5-day time scale**

| <i>Buoy Locations</i> | <i>No. of Observations</i> | <i>Mean SST (° C)</i> |         | <i>Standard Deviation (° C)</i> |         | <i>Root Mean Square Error (° C)</i> | <i>Correlation</i> |
|-----------------------|----------------------------|-----------------------|---------|---------------------------------|---------|-------------------------------------|--------------------|
|                       |                            | OBS                   | NEMOVAR | OBS                             | NEMOVAR |                                     |                    |
| 12° S; 67° E          | 88                         | 27.37                 | 27.36   | 1.25                            | 1.14    | 0.17                                | 0.99               |
| 8° S; 67° E           | 88                         | 28.03                 | 28.14   | 0.94                            | 0.88    | 0.14                                | 0.99               |
| 8° S; 55° E           | 171                        | 27.44                 | 27.38   | 1.35                            | 1.29    | 0.20                                | 0.99               |
| 4° S; 80.5° E         | 164                        | 29.12                 | 29.00   | 0.52                            | 0.52    | 0.20                                | 0.94               |
| 0° S; 67° E           | 170                        | 29.50                 | 29.37   | 0.61                            | 0.55    | 0.24                                | 0.94               |
| 12° N; 90° E          | 153                        | 29.10                 | 28.90   | 0.89                            | 0.76    | 0.44                                | 0.90               |
| 15° N; 90° E          | 170                        | 28.94                 | 28.79   | 1.05                            | 0.99    | 0.31                                | 0.97               |

**Table-6: Mean, standard deviation, root-mean-square error (RMSE), and correlation coefficient for NEMO SSS (psu) and RAMA buoys at daily time scale**

| <i>Buoy Locations</i> | <i>No. of Observations</i> | <i>Mean SSS (psu)</i> |         | <i>Standard Deviation (psu)</i> |         | <i>Root Mean Square Error (psu)</i> | <i>Correlation</i> |
|-----------------------|----------------------------|-----------------------|---------|---------------------------------|---------|-------------------------------------|--------------------|
|                       |                            | OBS                   | NEMOVAR | OBS                             | NEMOVAR |                                     |                    |
| 12° S; 67° E          | 395.00                     | 34.69                 | 34.79   | 0.45                            | 0.4     | 0.32                                | 0.71               |
| 8° S; 67° E           | 439.00                     | 34.87                 | 34.98   | 0.50                            | 0.46    | 0.14                                | 0.96               |
| 8° S; 55° E           | 640.00                     | 35.12                 | 35.11   | 0.28                            | 0.25    | 0.12                                | 0.92               |
| 4° S; 80.5° E         | 813.00                     | 35.00                 | 34.94   | 0.35                            | 0.36    | 0.16                                | 0.93               |
| 0° S; 67° E           | 847.00                     | 35.28                 | 35.28   | 0.29                            | 0.28    | 0.15                                | 0.87               |
| 12° N; 90° E          | 611.00                     | 33.05                 | 32.83   | 0.48                            | 0.58    | 0.47                                | 0.74               |
| 15° N; 90° E          | 646.00                     | 32.52                 | 32.56   | 0.61                            | 0.66    | 0.52                                | 0.66               |

**Table-7: Mean, standard deviation, root-mean-square error (RMSE), and correlation coefficient for NEMO SSS (psu) and RAMA buoys at 5-day time scale**

| <i>Buoy Locations</i> | <i>No. of Observations</i> | <i>Mean SSS (psu)</i> |         | <i>Standard Deviation (psu)</i> |         | <i>Root Mean Square Error (psu)</i> | <i>Correlation</i> |
|-----------------------|----------------------------|-----------------------|---------|---------------------------------|---------|-------------------------------------|--------------------|
|                       |                            | OBS                   | NEMOVAR | OBS                             | NEMOVAR |                                     |                    |
| 12° S; 67° E          | 80                         | 34.67                 | 34.79   | 0.45                            | 0.4     | 0.31                                | 0.75               |
| 8° S; 67° E           | 88                         | 34.87                 | 34.98   | 0.49                            | 0.45    | 0.11                                | 0.98               |
| 8° S; 55° E           | 129                        | 35.12                 | 35.11   | 0.27                            | 0.24    | 0.09                                | 0.96               |
| 4° S; 80.5° E         | 164                        | 35.00                 | 34.94   | 0.34                            | 0.36    | 0.14                                | 0.94               |
| 0° S; 67° E           | 170                        | 35.27                 | 35.28   | 0.28                            | 0.27    | 0.12                                | 0.90               |
| 12° N; 90° E          | 122                        | 33.05                 | 32.83   | 0.47                            | 0.56    | 0.44                                | 0.77               |
| 15° N; 90° E          | 129                        | 32.52                 | 32.56   | 0.56                            | 0.63    | 0.46                                | 0.71               |

## References

- Balmaseda M. A., K. Mogensen, F. Molteni and A. T. Weaver, 2010: The NEMOVAR-COMBINE ocean re-analysis, *COMBINE Technical Report No. 1*, [http://www.combine-project.eu/fileadmin/user\\_upload/combine/tech\\_report/COMBINE\\_TECH\\_REP\\_n01.pdf](http://www.combine-project.eu/fileadmin/user_upload/combine/tech_report/COMBINE_TECH_REP_n01.pdf)
- Balmaseda M. A., K. Mogensen, A. T. Weaver, 2013: Evaluation of the ECMWF ocean reanalysis system ORAS4, *Q. J. R. Meteorol. Soc.* 139: 1132–1161.
- Bell, M.J., M. Lefèvre, P.-Y. Le Traon, N. Smith, and K. Wilmer-Becker. 2009: GODAE: The Global Ocean Data Assimilation Experiment. *Oceanography* 22(3):14–21, <https://doi.org/10.5670/oceanog.2009.62>.
- Blockley E. W., M. J. Martin, A. J. McLaren, A. G. Ryan, J. Waters, D. J. Lea, I. Mirouze, K. A. Peterson, A. Sellar, and D. Storkey, 2014: Recent development of the Met Office operational ocean forecasting system: an overview and assessment of the new Global FOAM forecasts, *Geosci. Mod. Dev.*, 7, 2613–2638.
- Bonjean F. and G.S.E. Lagerloef, 2002: Diagnostic model and analysis of the surface currents in the tropical Pacific ocean, *J. Phys. Oceanogr.*, 32, 2938-2954.
- Daget, N., A. T. Weaver and M. A. Balmaseda, 2009: Ensemble estimation of background error variances in a three-dimensional variational data assimilation system for the global ocean. *Q. J. R. Meteor. Soc.*, 135, 1071-1094
- Delay R, 1991: Atmospheric data analysis. *Cambridge University Press: Cambridge, UK*.
- Evensen G., 1994: Sequential data assimilation with a non linear quasi geostrophic model using the Monte Carlo methods to forecast error statistics, *J. Geophys. Res.*, 99: 10143-10162.
- Lorenc A. C, 1986: Analysis methods for numerical weather prediction. *Q. J. R. Meteorol. Soc.* 112: 1177–1194.
- Madec G, 2003: Nemo ocean engine. In Note du Ple de modlisation, number 27.
- McPhaden, M. J., G. Meyers, K. Ando, Y. Masumoto, V. S. N. Murty, M. Ravichandran, F. Syamsudin, J. Vialard, L. Yu, and W. Yu. 2009: “RAMA: The Research Moored Array for African–Asian–Australian Monsoon Analysis and Prediction.” *Bulletin of the American Meteorological Society*, 90: 459–480. doi:10.1175/2008BAMS2608.1.
- Mogensen K. S., Balmaseda, M. A., and A. Weaver, 2012: The NEMOVAR ocean data assimilation system as implemented in the ECMWF ocean analysis for System 4, *ECMWF Tech. Memo.* 668.
- Oliver, H. J., Matt Shin, and Oliver Sanders, 2018: Cylc: A Workflow Engine for Cycling Systems, *Journal of Open Source Software*, 3(27), 737. DOI: 10.21105/joss.00737.
- Parrish D. F., and J.C. Derber, 1992: The National Meteorological Center’s spectral statistical interpolation analysis system, *Mon. Weather Rev.*, 120, 1747-1763.

- Rajagopal, E.N., G.R. Iyengar, John P. George, Munmun Das Gupta, Saji Mohandas, Renu Siddharth, Anjari Gupta, Manjusha Chourasia, V.S. Prasad, Aditi, Kuldeep Sharma and Amit Ashish, 2012: Implementation of Unified Model based Analysis-Forecast System at NCMRWF, *NMRF/TR/2/2012*, 45p. ([https://www.ncmrwf.gov.in/UM\\_OPS\\_VAR\\_Report.pdf](https://www.ncmrwf.gov.in/UM_OPS_VAR_Report.pdf))
- Rakhi R., A. Jaykumar, M. N. R. Sreevathsa and E. N. Rajagopal, 2016: Implementation and Up-gradation of NCUM in Bhaskara HPC, *NMRF/TR/03/2016*, 22p. ([https://www.ncmrwf.gov.in/NMRF\\_TR3\\_2016.pdf](https://www.ncmrwf.gov.in/NMRF_TR3_2016.pdf))
- Sumit Kumar, A. Jayakumar, M. T. Bushair, Buddhi Prakash J., Gibies George, Abhishek Lodh, S. Indira Rani, Saji Mohandas, John P. George and E. N. Rajagopal, 2018: Implementation of New High Resolution NCUM Analysis-Forecast System in Mihir HPCS, *NMRF/TR/01/2018*, 17p. ([https://www.ncmrwf.gov.in/NCUM-Report-Aug2018\\_final.pdf](https://www.ncmrwf.gov.in/NCUM-Report-Aug2018_final.pdf))
- Storkey, D., Blockley, E. W., Furner, R., Guiavarc'h, C., Lea, D., Martin, M. J., Barciela, R. M., Hines, A., Hyder, P., and Siddorn, J. R. 2010: Forecasting the ocean state using NEMO: the new FOAM system, *Journal of Operational Oceanography*. 3:3–15.
- Vinayachandran, P. N., and S. R. Shetye (1991): The warm pool in the Indian Ocean, *Proc. Indian Acad. Sci. (Earth Planet. Sci.)*, 100 (2), 165–175.
- Waters J., D. J. Lea, M. J. Martin, D. Storkey, and J. While, 2013: Describing the development of the new foam-NEMOVar system in the global 1/4 degree configuration, *Technical Report 578, Met Office*.
- Waters J., D. J. Lea, M. J. Martin, I. Mirouze, A. T. Weaver, and J. While, 2014: Implementing a variational data assimilation system in an operational 1/4 degree global ocean model, *Q. J. Roy. Meteor. Soc.*, doi:10.1002/qj.2388.
- Weaver A. T., C. Deltel, E. Machu, S. Ricci and N. Daget, 2005: A multivariate balance operator for variational ocean data assimilation, *Q. J. Roy. Meteor. Soc.*, 131, 3605-3625.