NMRF/TR/06/2019

TECHNICAL REPORT

New Background Error Statistics for Regional NCUM 4DVAR Data Assimilation System

A. Routray, V. P. M. Rajasree, Devajyoti Dutta, and John P. George

May 2019

National Centre for Medium Range Weather Forecasting
Ministry of Earth Sciences, Government of India
A-50, Sector-62, NOIDA-201309, INDIA

सत्यमेव जयते

New Background Error Statistics for Regional NCUM 4DVAR Data Assimilation System

A. Routray, V. P. M. Rajasree, Devajyoti Dutta,
and John P. George

May 2019

National Centre for Medium Range Weather Forecasting
Ministry of Earth Sciences
A-50, Sector 62, NOIDA-201309, INDIA

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	NMRF/TR/06/2019
3	Date of Publication	May 2019
4	Title of the document	New Background Error Statistics for Regional NCUM-4DVAR Data Assimilation System
5	Type of Document	Technical Report
6	No. of pages &Figures	23 Pages; 10 Figures and 3 Tables
7	Number of References	21
8	Author (s)	A. Routray, V. P. M. Rajasree, Devajyoti Dutta and John P. George
9	Unit	NCMRWF
10	Abstract	The domain-specific background error statistic (BES) has been computed for the high resolution regional NCUM-4DVAR analysis system. The CVT module used for calculation of BES is successfully installed in the Mihir HPCS of NCMRWF. BES were calculated using one month's forecast from the 4.4 km resolution regional model for Indian region. The diagnostics of BES calculations shows that they are reasonably well matched with the patterns of tropical domains BSE's of other NWP centres.. Several single observation (pseudo-obs) tests at various locations (latitude, longitude and model level) are performed over Indian region using the newly calculated BES. This is done to understand the response of BES to the assimilation system over the Indian region. The single-observation perturbation tests are applied for the temperature and u-wind component, of a 1K and 1 m/s innovation. The analysis increments of temperature and wind components obtained from single-observation tests using the newly calculated domain specific BES are showing isotropic in nature. To understand the impact of default and newly generated domain specific BES, a study of analysis-forecast has been carried out using the high resolution regional NCUM 4DVAR analysis system on the genesis of the tropical cyclone (TC) Titli, that formed over east-central BoB during 08-13 October 2018. The results suggest that the use of new domain-dependent BES in the NCUM 4DVAR system improves the analysis and forecast
11	Security Classification	Non-Secure
12	Distribution	Unrestricted Distribution
13	Key Words	Background Error Statistic, NCUM, 4DVAR, control variable, Innovation, Tropical Cyclone

Table of Contents

Abstract

The domain-specific background error statistic (BES) has been computed for the high resolution regional NCUM-4DVAR analysis system. The CVT module used for calculation of BES is successfully installed in the Mihir HPCS of NCMRWF. BES were calculated using one month's forecast from the 4.4 km resolution regional model for Indian region. The diagnostics of BES calculations shows that they are reasonably well matched with the patterns of tropical domains BSE's of other NWP centres.. Several single observation (pseudo-obs) tests at various locations (latitude, longitude and model level) are performed over Indian region using the newly calculated BES. This is done to understand the response of BES to the assimilation system over the Indian region. The single-observation perturbation tests are applied for the temperature and u-wind component, of a 1K and 1 m/s innovation. The analysis increments of temperature and wind components obtained from single-observation tests using the newly calculated domain specific BES are showing isotropic in nature. To understand the impact of default and newly generated domain specific BES, a study of analysis-forecast has been carried out using the high resolution regional NCUM 4DVAR analysis system on the genesis of the tropical cyclone (TC) Titli, that formed over east-central BoB during 08-13 October 2018. The results suggest that the use of new domain-dependent BES in the NCUM 4DVAR system improves the analysis and forecast

1. *Introduction*

Background Error (BE) statistics is important in Data Assimilation (DA) since it provides the error information of the background fields. Quality of the assimilation largely depends on the representation of error covariance. Several methods are available for the computation of background error statistics such as Innovation method, National Meteorological Centre (NMC) method (Parrish and Derber 1992), Ensemble method (Fisher 2003) and a few adaptive methods. Innovation method is an observation-space based method whereas NMC method is a model-space based method. Since the NMC method is easy to implement, it is/was used for Background Error computations at many NWP centres as described by Pereira (2006) e.g. at the European Centre for Medium-Range Weather Forecasts (ECMWF; Rabier et al. 1998), at the Canadian Meteorological Centre (CMC) in Canada (Gauthier et al. 1999), at Météo-France (Desroziers et al. 1995), and at the UK Met Office (UKMO; e.g., Lorenc et al. 2003). The Ensemble method uses the ensemble of assimilation experiments. In this method the observations as well as the physical parameterizations are perturbed for each ensemble member.

The data assimilation demands the explicit knowledge of BE statistics since it is one of the important components in the assimilation algorithm. However, the size of BE matrix is very large and it is very difficult to deal with such large matrices. Therefore, the calculation of BE is simplified using a technique called Control Variable Transform (CVT). There are several studies that clearly explain the choice of control variables selected in atmospheric and oceanic data assimilation (Bannister, 2008, 2017; Parrish and Derber, 1992; Derber and Bouttier, 1999; Berre, 2000; Weaver et al., 2005; Ménétrier et al. 2015). The size of BE matrix is reduced by choosing the control variables appropriately in such a way that cross covariance between these variables are minimum.

An effort is made to compute new BE statistics for the regional assimilation-forecast system of 4.4 km resolution over the Indian monsoon region. For this purpose the CVT suite of Unified Model (UM) is implemented in the high performance computer (HPC) Mihir. BE statistics using one month forecast is computed using the CVT system. This document is organized as follows; Section 2 describes the Preparation of Training Data and Methodology used in detail. How to set up CVT module is described in Section 3. The CVT module structure is explained in Section 4. Section 5 explains the BE diagnostics. Section 6 describes the results of Pseudo observation test, followed by Impact of BE Statistics in Section 7 and summary of the report presented in Section 8.

2. *Preparation of Training Data and Methodology*

The first step is the creation of training data from the forecast (short forecast) files i.e selection of linearization states (LS i.e ca* files). The NCUM-PS40 standard research suite has been used for the creation of linearization states from the short range forecasts in all synoptic hours (00, 06, 12 and 18 UTC) over regional domain. The regional domain is used for creation of linearization states is shown in Figure 1. The domain covers the area 63-107°E longitude and 0-40°N latitude and the horizontal resolution of the model is ~4.4 km with 1200 x 1200 grid points in the domain. The model has 80 sigma levels reaching upto ~38.5 km height. The CVT suite considers the NMC method (Parrish and Derber 1992) and can deal with different forecast lengths. The NMC method is one of the commonly used methods for the computation of BE statistics. It is based on the difference between pairs of forecasts of different lead times, but valid at the same time, to evaluate short-range forecast errors. When CVT is run on a limited area domain, the increments generated from the "CvtProg_CalibTp1" (discuss in later section) step need to have a zero row or column around the edge for model variables. Otherwise the results from CVT will be erroneous. Therefore, the same lateral boundary conditions (LBCs) are used in the CVT module at different cycle times (Table 1).

Table 1: Suite set up for creating forecast pairs

Cycle Time	LBC used	Forecast Length	Different LS files
00 UTC	00 UTC	12 hrs	Pair-1 valid at 12 UTC
06 UTC		06 hrs	
12 UTC	12 UTC	12 hrs	Pair-2 valid at 00 UTC
18 UTC		06 hrs	

The schematic diagram to generate forecast difference for the NMC method is provided in Figure 2. The major role of the background error is to spread out the information from the observations. It provides statistically consistent increments at the neighbouring grid points and levels of the model and to ensure that observations of one model variable (e.g. temperature) produce dynamically consistent increments in the other model variables (e.g. vorticity and divergence). We used the period of 01-31 July 2018 to generate 62 forecast pairs for the creation of BE statistics.

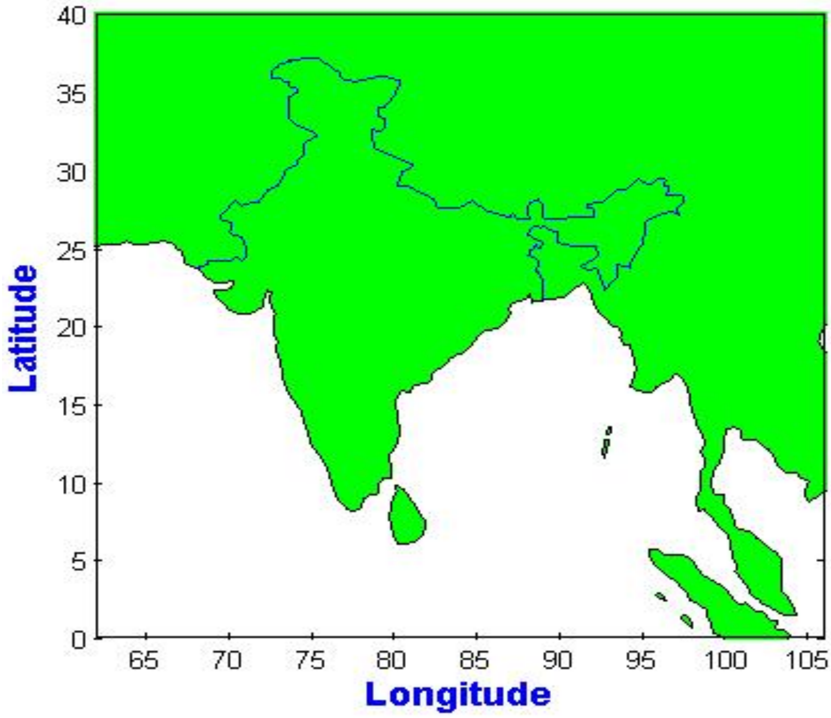


Figure 1: NCUM regional Model domain used for CVT

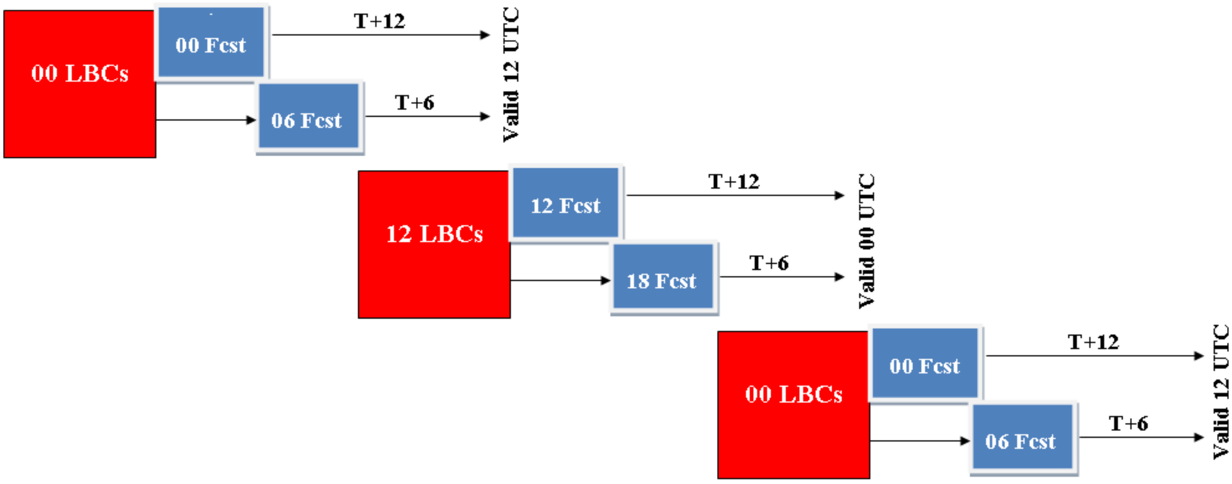


Figure 2: Schematic of NMC method used for generating forecast differences (*Source: McMillan, Martin. 2015* (http://code.metoffice.gov.uk)

The CVT module calculates a set of key statistical quantities which represent static background. The choice of key statistical quantities depends on the choice of control variables i.e. variables derived from the original model variables and are assumed to be uncorrelated with respect to each other, and are Gaussian. The control variables used here are stream function, velocity potential, atmospheric pressure and a moisture variable. The means of getting the control variables is done through a mixture of dynamical relationships and statistics.

Once the training data is transformed into the control variable space and appropriate statistical quantities are derived, then it is possible to represent an approximation to the horizontal structure, vertical structure and the variance of each of the control variables. The use of separate horizontal and vertical transforms strictly limits the degree to which one can approximate the spatial representation of the true unknown background error covariance. In this study, we applied vertical and horizontal transforms for regional domain as follows:

Control variables in grid point space -> vertical transform -> horizontal transform

Many operational centres adopted the above vertical and horizontal transforms for calculation of the covariance matrix. In this case, for each control variable an averaged vertical covariance matrix is generated that used to derive a projection matrix through the solution of a generalized eigen value problem. The vertical projection transforms the data from model levels to vertical modes as well as removes the variance. In the global domain, the homogeneous and isotropic power spectrum for each vertical mode is derived from the horizontal power spectra. However, in a regional domain a homogeneous and isotropic "Second Order Auto-Regression" (SOAR) is used to describe horizontal structure. The user can define the horizontal length scale for each vertical mode in the namelist as a variable. The CVT module also supports the vertical and horizontal transforms for global domain in the following order;

Control variables in grid point space -> Horizontal transform -> Vertical transform.

3. *How to set up CVT module*

Before copying the rose suite from the UK Met office repository, the user needs to make sure the accessibility to the following website, http://code.metoffice.gov.uk. Then the user can download the suitable CVT suite from the UK Met office repository by typing "**rosie u-av047**" (current stable version of CVT). The directory contains various suites and sub-directories as shown in Table 2. The "site/direct-elogin01.rc" suite contains the HPC Mihir related information and configuration (Host name; NPROC; NPROCX; NPROCY; Wall clock time, etc) required to run the CVT module.

Table 2: List of suites and sub-directories in the CVT module.

```
ashish@elogin01:~/roses/cvt_july> ls
ReadMe                 rose-suite.conf_org       suite-horizstruct.rc
app                    rose-suite.info           suite-innerprod.rc
bin                    site                      suite-removemean.rc
direct-elogin01.rc     suite-calibtp1.rc         suite-removemean_qc.rc
direct-local.rc        suite-calibtp2.rc         suite-spectralproj.rc
doc                    suite-calibtpmu.rc        suite-spectralvariance.rc
etc                    suite-defaults.rc         suite-variance.rc
meta                   suite-gathercovstats.rc   suite-verticalproj.rc
opt                    suite-graphics.rc         suite.rc
rose-suite.conf        suite-gridbinning.rc
ashish@elogin01:~/roses/cvt_july>
```

3.1. Setting up rose-suite-[Optional].conf:

The settings of variables in the "rose-suite.conf" are overwritten by the optional "rose-suite-[Optional].conf" (opt/rose-suite-[Optional].conf). The user can change the configurations of the important variables in the "opt" directory. The users are strongly suggested to create their own wrapper script for the CVT runs. Here, we used the wrapper "rose-suite-ncmrwf-62sample.conf" in the opt directory. There are a number of important variables that need to be set in the wrapper "rose-suite-ncmrwf-62sample.conf" as described below:

CVT_GLOBALHEADER_*: The global headers are key parameters that provide the detailed information about the covariance generation (i.e type of forecast differences; domain global or regional; training data used; organization; etc).

CVT_DATA: Path to directory prefix for training data (e.g. /home/ashish/cylc-run/PS40india6hrly/share/cycle).

***FILESUM*:** Name of file in etc/ directory contains LS files, which is used to create the forecast differences (e.g. etc/listing_ncmr_july.txt).

The text file is organised with 3 columns in one row as follows:
20180701T0000Z/umqvaa_ca0009   20180701T0600Z/umqvaa_ca0003   001
20180701T1200Z/umqvaa_ca0009   20180701T1800Z/umqvaa_ca0003   001
and so on.

First column: Path to the first full LS file; *Second column*: Path to the second LS have the same validity time as the first LS; *Third column*: Number of ensemble member (If an ensemble is not used then it should be "001"). The forecast difference is defined as first LS file minus the second LS file.

NFILESUM: Total number of rows in the FILESUM file.

CVT_PATH1: Path of parallel VAR Fortran executable e.g. CvtProg_CalibTp1.exe (*/home/umprod/PS40/SOURCE_DIR/old_fcm_make_var_xc40_haswell_ifort_opt/build/bin*)

CVT_PATH2: Path of serial VAR Fortran executable e.g. CvtProg_CalibTp2.exe (*/home/umprod/PS40/SOURCE_DIR/old_fcm_make_var_xc40_haswell_ifort_opt/build-serial/bin*)

CVT_VAR_LIST: "CvtOnly" (If "All" then it will calculate the statistics for both the control variables and variables in model space. If "CvtOnly" then it will generate statistics for the control variables. "DiagOnly" will only generate the diagnostic variables.

CVT_OPT_FileBatch: This is also set to an integer and defines the maximum number of forecast differences that are processed together within a rose task. Works in modules CALIBTP2, CALIBTPMU and CvtProg_SpectralProj.

CVT_OPT_Tp1BinningBatch: This limits the number of ROSE jobs running at a given time for the CvtProg_CalibTp1 step.

NCKS_PATH: Path to an "nco library" directory. This directory includes the ncks utility, which is used to combine the NetCDF files.

CVT_TransformOrder: "TvThTp" (calculating vertical covariances as a function of total wavenumber) or "ThTvTp" (calculating SOAR length scales for each vertical mode).

GraphicsOnly: True or False (Whether to generate graphics).

CVT_Overide: True [It allows just single modules to be run on their own. This capability is very useful when something goes wrong on a particular module].

The details of other variables can be found in the CVT user documents (VTDP9) of UK Met Office available in https://code.metoffice.gov.uk/trac/var/wiki/CVT.

4. *CVT Module Structure*

The detailed structure of the CVT modules is illustrated in Fig.-3 and Fig.-4. The first step (Fig.-3) transforms the training data into control variables (stream function, velocity potential, atmospheric pressure and moisture) and derives the necessary statistical variables.

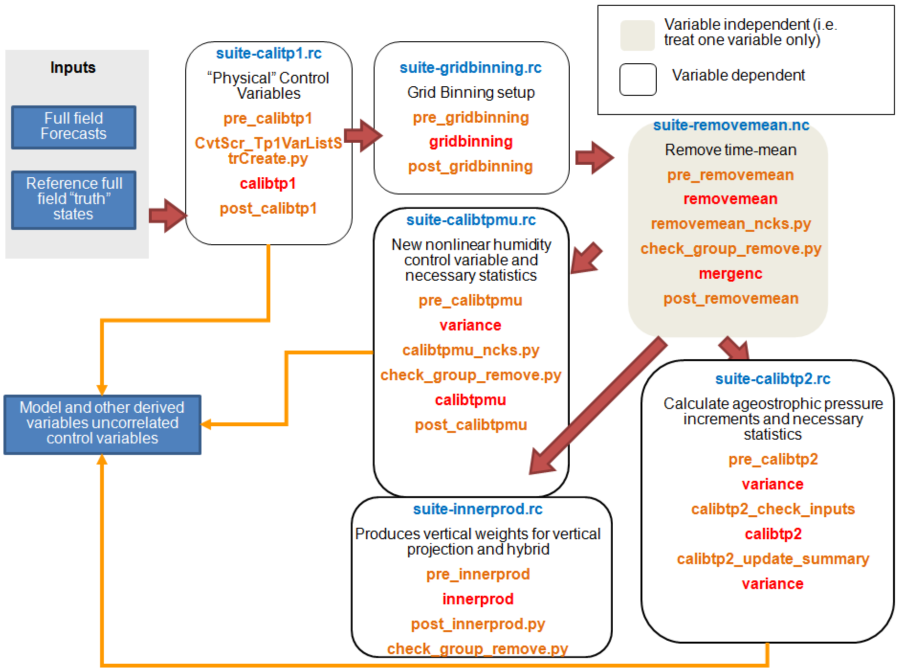


Figure 3: First step of CVT computation (*Source: CVT technical document VTDP9 of UK Met Office by Marek Wlasak;* https://code.metoffice.gov.uk/trac/var/browser/main/trunk/doc/technical/VTDP9)

The various stages evolved in the first step are "CalibTP1" (creates the control variables from the forecast files), "Gridbinning" (collect the information, group the information and provide weight), "Removemean" (remove the time-mean at each grid point for each variable), "InnerProd" (creates normalised vertical profiles of the model level difference in the pressure field and calculate the mean pressure vertical profile), "CalibTp2" (calculates vertical covariances and cross covariances needed to generate the unbalanced control variables) and "CalibTpMU" (calculates variances and cross-variances for moisture control variable).

The second part of the CVT module is as shown in the Fig.-4 and it creates a spatial parameterization of the horizontal and vertical structure. This part of the CVT module calculates the "Variance" (variance and covariances are binned in their spatial grid point position), "VerticalProj" (forecast error samples transformed from model levels to vertical modes), "HorizStruct" (calculates the horizontal length scale for each variables).

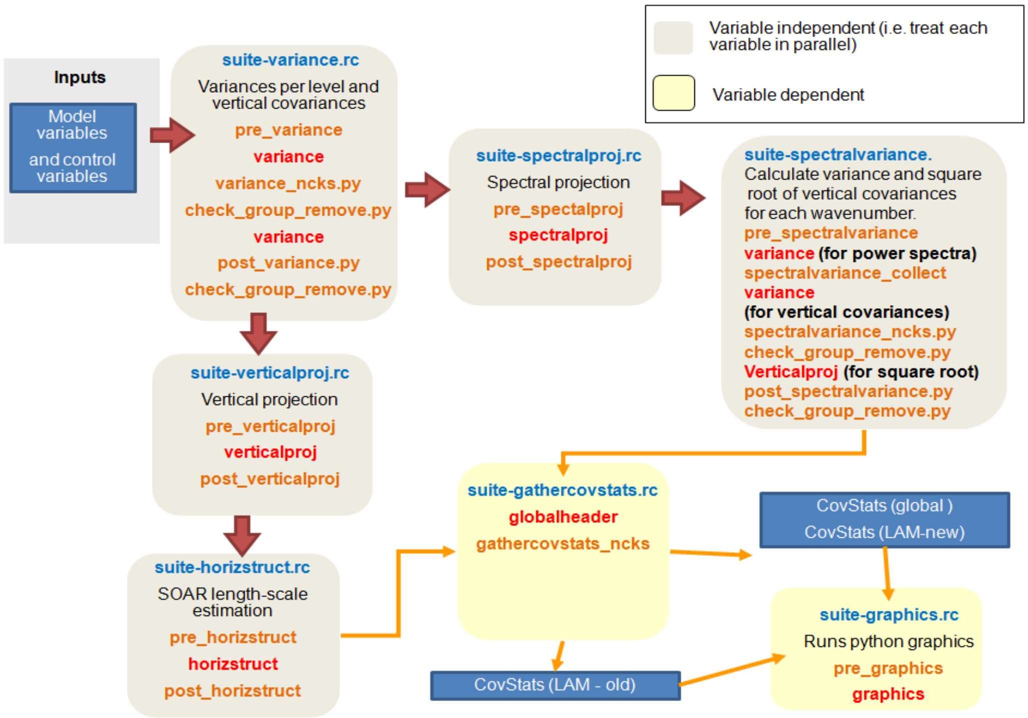


Figure 4: Second step of CVT computation (*Source: CVT technical document VTDP9 of UK Met Office by Marek Wlasak;* https://code.metoffice.gov.uk/trac/var/browser/main/trunk/doc/technical/VTDP9).

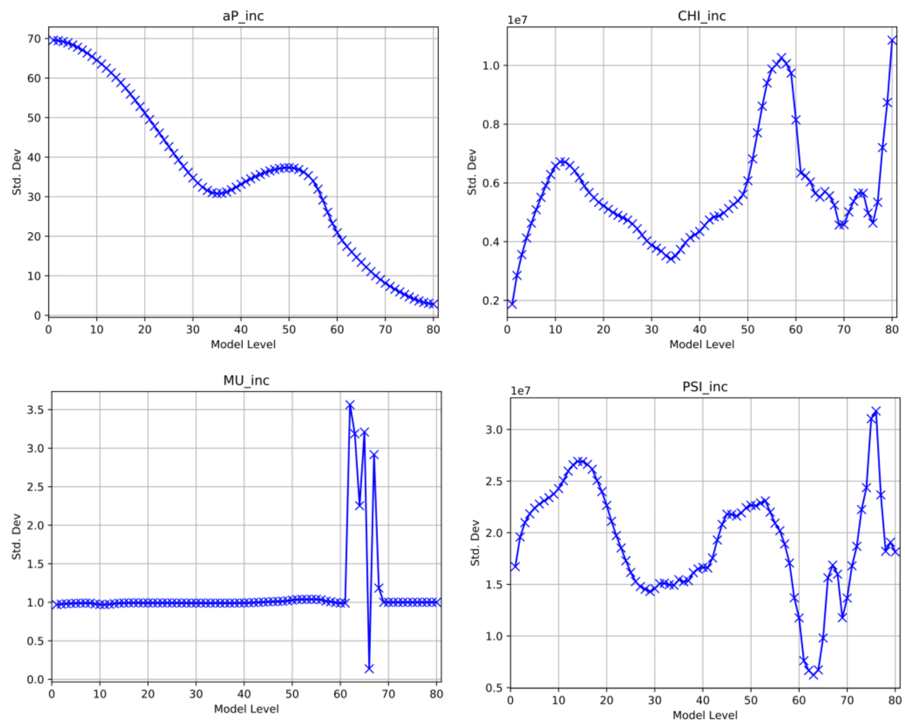


Figure 5: Averaged standard deviation of the control variables

5. *BE Diagnostics:*

The CVT module produces various diagnoses of the control variables through graphical packages which are mainly developed in python scripting language. Figure-5 illustrates the averaged profile of standard deviation of the control variables - unbalanced atmospheric pressure (aP); velocity potential (χ); moisture (Mu) and stream function (ψ). The standard deviation of aP is higher at lower levels and it gradually decreases as model level is increased. It indicates the high variation in the aP control variable at lower model levels. However, the standard deviation of other control variables is less throughout all model levels. The variation of the control variables is more clustered around the mean value.

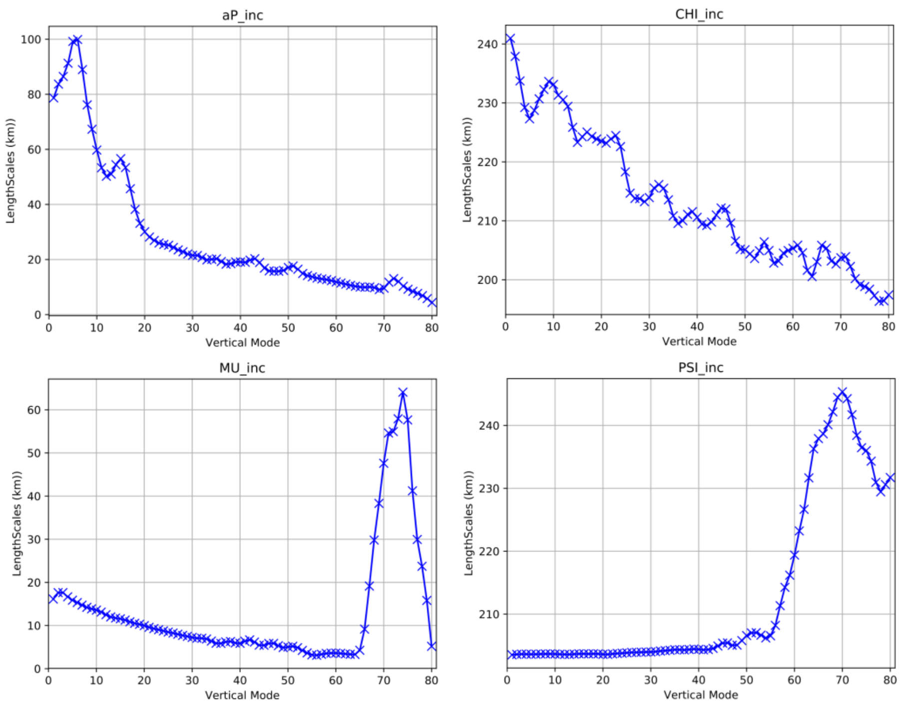


Figure 6: Averaged SOAR length scales of control variables

Figure-6 shows the SOAR horizontal length-scales associated with the vertical modes. It is seen from the Fig.-6 the SOAR length-scale decreases with height, both for Psi and Chi. In the case of aP and Mu variables there is not much change in the SOAR length scales till vertical mode 60. However, thereafter both the control variables show sudden spike and then decreases. The pattern of the control variables is reasonably well matched with the other tropical region's SOAR length scales (i.e Singapore and Queensland; McMillan, 2015). Similarly, the auto and cross-correlation matrix of various control variables (Fig-7) are well matched with the pattern obtained from Singapore and Queensland domains (McMillan, 2015).

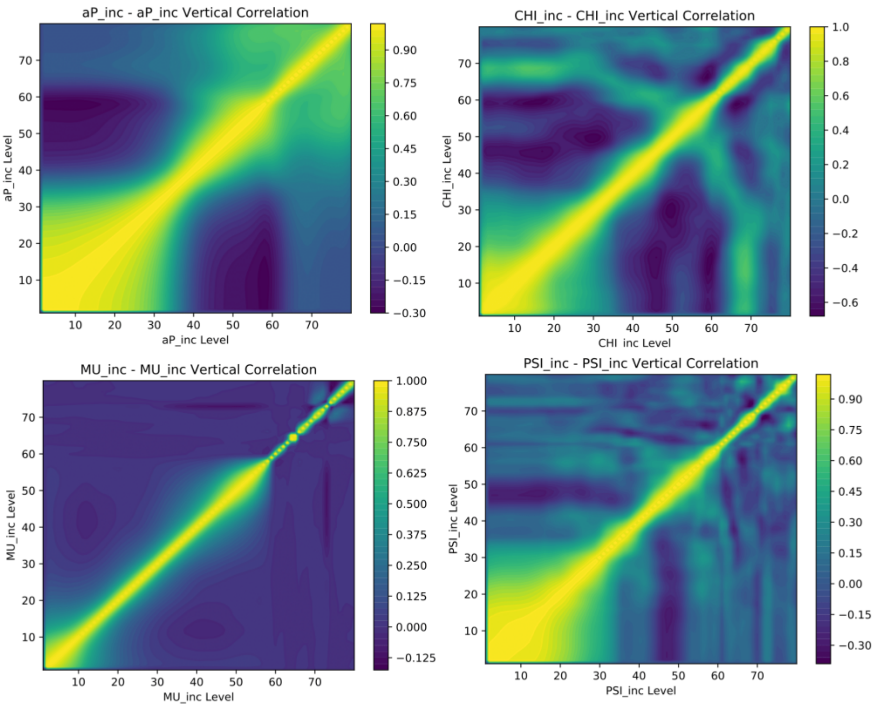


Figure 7: Vertical correlation of control variables

6. Pseuedo observation tests:

The single-observation (pseudobs) test is often used as a proof-of-concept test to determine how the observed entity spreads to its vicinity via the established correlations among analyses variables (Wu et al., 2002; Wang et al. 2014; Routray et al. 2010 and 2014). For this purpose, the pseudobs test has been carried out (Met Office suite u-au654 is successfully installed in HPC Mihir, Table 3). The user can modify required information in the rose-suite.conf file before running the pseudobs suite such as:

[jinja2:suite.rc]

COVTYPE_LIST=['ThTvTp'] **# Different TransformOrder i.e 'ThTvTp' or 'TvThTp'.**

CYCLE_TIME="20181010T00" **# Date tag yyyymmddThh.**

DET_MODEL="ukv" **# Model name.**

HPC_HOST="elogin01" **# HPC login node.**

PSEUDOOB_FIELD_LIST=['u','t'] **# Single obs for particular variable.**

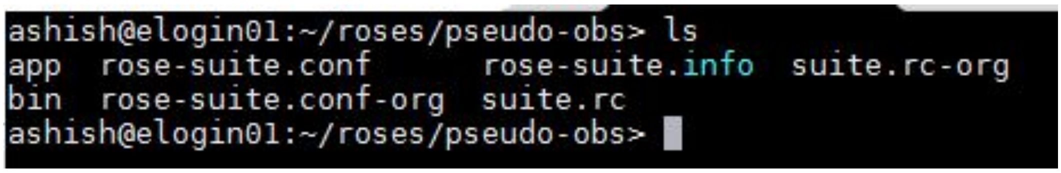
PSEUDOOB_LEVEL_LIST=[28] # **Model level where single obs perform.**

PSEUDOOB_TS_LIST=[0] # **By default zero no need to change.**

PSEUDOOB_XY_LIST=[[220,300]] # **Model points in x and y direction.**

TESTCASE_DIR="/home/ashish/cylc-run/cvt_july/share/data" # **Path of 'CovStats_Full', 'CovStats_Var' directory.**

Table 3: List of suites and sub-directories in the Pseudobs module.



The path of VAR executives is fixed in the app/install_cold/rose-app.conf file like [file:$ROSE_SUITE_DIR/share/fcm_make_var-opt]

mode=symlink

source=/home/umprod/PS40/SOURCE_DIR/old_fcm_make_var_xc40_haswell_ifort_opt

The detailed path of low and high resolution linearization files (LS, i.e *ca files) is fixed in app/ install_cycledata/rose-app.conf file like

[file:$CYLC_SUITE_SHARE_PATH/cycle/${ENS_CYCLE_LG00}/${DET_MODEL}_em
]

mode=symlink

source=/home/ashish/cylc-run/PS40india-6hrly/share/cycle/20181010T0000Z/ukv_ls # Low resolution *ca files.

[file:$ROSE_DATACPT1H/${DET_MODEL}_ls]

mode=symlink

source=/home/ashish/cylc-run/PS40india-6hrly/share/cycle/20181010T0000Z/ukv_umback # High resolution *ca files

After completion of all the above changes, the user can run pseudobs suite by ‘rose suite-run’. The output will be written in the path

$HOME/cylc-run/pseudo-obs/share/cycle/yyyymmddThh00Z/

Several single-observation tests at various locations (latitude, longitude and model level) are performed over the Indian region using the calculated domain dependent BES. This is done to understand the response of BES in the assimilation system over the Indian region. The single-observation perturbation tests are applied for the temperature and u-wind component, for a 1K and 1 m/s innovation [observation minus background (O - B)], respectively, at the middle of the domain (18.1N; 78.0E, model level 28). The analysis increments of temperature and wind components obtained from single-observation tests are shown in Fig.-8(a-c) and (d-f) respectively. The horizontal spreading of the increment of temperature (Fig-8a) is isotropic in nature. The features of the multivariate analysis increments of the u- and v-components of wind (Fig.-8b and c) are obtained from the isotropic temperature increment (Fig. 8a) through the balance response of the BES. It is suggested that the horizontal spreading of the increments of temperature and wind components are properly responded by the newly calculated BES. Similarly, the horizontal spreading of the increments of temperature (Fig-8d) and v-wind (Fig.-8f) is isotropic in nature in the response to the single u-wind component (1 m/s) innovation applied at the middle of the domain. The responses of the v increments (Fig.-8f) support cyclonic and anti-cyclonic circulation patterns to the north and south of the u-increment location, respectively. Both warm and cold temperature increments (Fig.-8d) response are found from the innovation of the single u-wind component. The symmetric responses in *C(u, T), C(u, u)* and *C(u, v)* etc. are consistent with the other studies (Daley,1993 and 1991; Routray et al. 2014). Therefore, the response of analysis increments from single innovation is propagated properly in the surrounding area by the BES. In our study, we found that the effect of a single wind observation is consistent with theoretically derived wind correlations for non-divergent flow (Daley, 1991 and 1996). However, the geostrophic coupling (mass–wind balance) decreases near the equator, which makes the circulation in the tropics different from other regions away for the tropics. Therefore, the calculation and tuning of the BES over tropical regions are very important to represent properly the tropical wind errors, mainly arising for divergent flow unlike other regions.

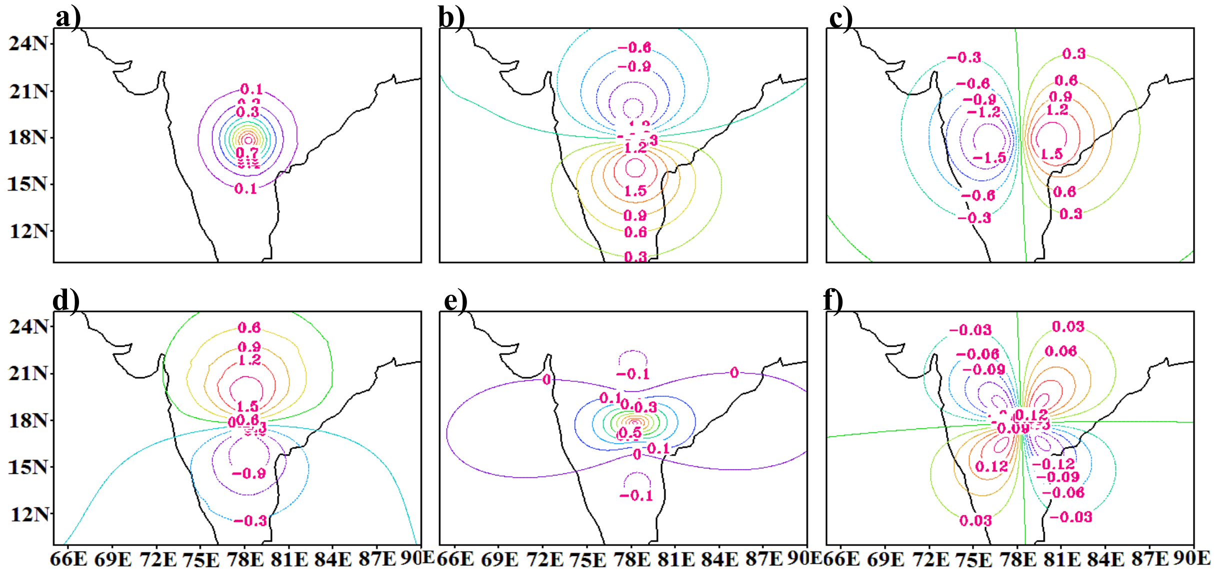


Figure-8: (a – c) Response of the analysis increments to a single temperature observation 1K at (18.1N; 78.0E; 28-model level). Similarly, (d – f) are the response of the analysis increments to a single u-wind perturbation of 1 m/s.

7. Impact of BE Statistics:

Two numerical experiments (assimilation-forecast) named as D-BES (experiments with default or old Background Error Statistics) and N-BES (experiment with New Background Error Statistics) are carried out to assess the impacts of default and newly calculated domain specific BES. The 4DVAR analyses are created with old and new BES to study the impact of new BES (N-BES experiment) on the genesis of the tropical cyclone (TC) Titli in NCUM-R model forecast. The very severe cyclonic storm (VSCS) Titli originated from a low pressure area (LPA) which formed over southeast Bay of Bengal (BoB) and adjoining north Andaman Sea in the morning (0830 IST) of 7th October and became a well marked low pressure area (WML) over the same region in the evening (1730 IST) of the same day. The system moved continuously northwestward intensified into a depression (D)- deep depression (DD)- cyclonic storm (CS) during 8th -9th October under favorable environmental conditions. The system further intensified into a severe cyclonic storm (SCS) to VSCS on 10th October. It crossed north Andhra Pradesh and south Odisha coasts (18.8N/84.5E) to the southwest of Gopalpur during 0430-0530 IST of 11th October as a VSCS with the wind speed of 140-150 gusting to 165 kmph. The details of the genesis, intensification and further movement of the VSCS Titli after landfall can be found from the IMD RSMC report of 2018.

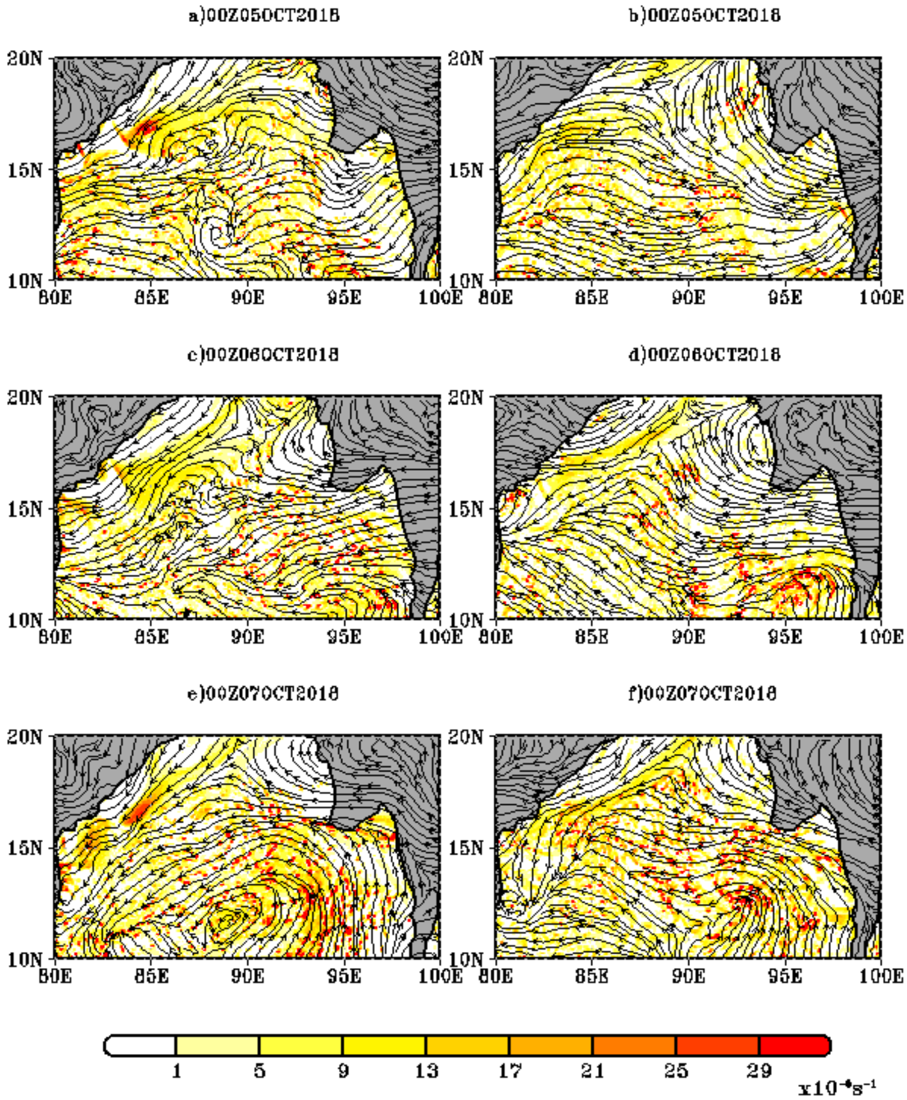


Figure-9: Left panel (a, c and e); relative vorticity (shaded) and streamlines at 850 hPa from domain-specific BES (N-BES) and right panel (b, d and f) from the default BES (D-BES).

Figure 9 shows the 850 hPa relative vorticity overlaid with streamlines from the both sets of experiments 3-days prior to the genesis of the TC. From the Fig.-9 (a and c), it is clearly seen that a small cyclonic circulation with scattered relative vorticity (~17-21 x 10^{-5} s^{-1}) is developed in the surrounding region during 5-6 October 2018. As per the India Meteorological Department (IMD) report a well marked low (WML) pressure area formed over southeast BoB and the adjoining regions of north Andaman Sea on morning 7th October

2018. The WML is clearly depicted in the N-BES analysis (Fig-9e) with strong vorticity (~21-25 x 10^{-5} s^{-1}) around the vortex. These features are not clearly depicted in the D-BES experiment [Fig-9 (b, d and f)] during genesis period. Particularly, in the Fig-9f a feeble closed circulation is noticed over southeast BoB on 7th October 2018 even if strong vorticity is seen as compared to the Fig-9e around the vortex. Figure-10 shows the 850 hPa relative humidity (RH) overlaid with streamlines from the both sets of experiments 3-days prior to the genesis of the TC. It is clearly seen that the RH is higher over BoB as well as the surrounding the vortex in N-BES [Fig-10 (a, c and e)] analyses as compared to D-BES analyses [Fig-10 (b, d and f)] during the genesis period of the TC. The results of this study suggested that the use of domain-dependent new BES in the NCUM 4DVAR analysis system improved the initial conditions (analysis) which leads to the improved forecast of the genesis of the storm reasonably well.

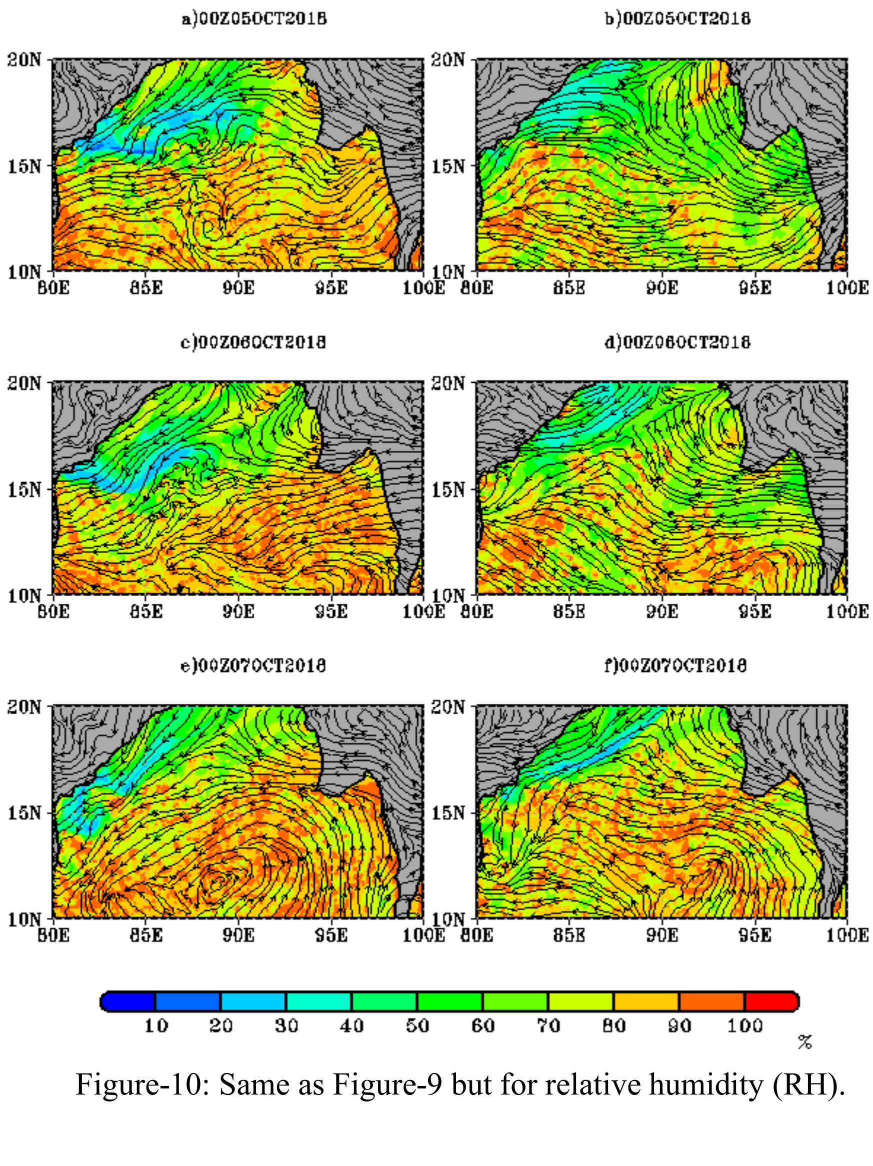


Figure-10: Same as Figure-9 but for relative humidity (RH).

8. Summary

The background error statistic (BES) is computed over the regional domain using high resolution regional NCUM. The NCUM-PS40 standard research suite has been used for the creation of training data i.e linearization states from the short range forecasts for the period 1-31 July 2018 over the regional domain. The BES is calculated based on the NMC method which is one of the commonly used methods in the operational centres. It is based on the difference between pairs of forecasts of different lead times, but each valid at the same time, to evaluate short-range forecast errors. The CVT module (u-av047) is successfully installed in the Mihir HPC. The details of the namelist variables and structure of the CVT module is described in the earlier sections. The CVT module produced various diagnostics of the control variables. The various diagnostics obtained from the CVT module are analysed in this study and found that these are well matched with the patterns of other tropical centers' BES like Singapore and Queensland.

The single-observation (pseudobs) test is often used as a proof-of-concept test to determine how the observed entity spreads to its vicinity via the established correlations among analyses variables. For this purpose, the pseudobs suite (u-au654) is successfully installed. Several single-observation tests at various locations (latitude, longitude and sigma level) are performed over the Indian region using the calculated domain dependent BES. This is done to understand the response of BES in the assimilation system over the Indian region. The single-observation perturbation tests are applied for the temperature and u-wind component, of a 1K and 1 m/s innovation. The analysis increments of temperature and wind components obtained from single-observation tests using the newly calculated domain specific BES is isotropic in nature.

To study the impact of newly generated BES NCUM 4DVAR analysis-forecast system, assimilation-forecast experiment is carried out to study the genesis of the TC Titli formed over east central BoB during 08-13 October 2018. The results suggested that the use of new BES in the NCUM 4DVAR help to capture the genesis of the storm in the forecast.

Acknowledgements

The authors gratefully acknowledge scientists from UK Met Office mainly Dr. Richard Renshaw and Dr Marek Wlasak for their immense help to install the CVT module and clarify the technical as well as scientific doubts throughout the research period. The report used materials (text and figures) available with UK Met Office technical/scientific documents. We would like to thank Dr. E. N. Rajagopal, Head NCMRWF for his advice and encouragements. We would also like to acknowledge technical support provided by Cray HPC Team especially to Ms. Shivali Gangwar and Mr Virender Kumar.

References:

Bannister, R. N. (2008). A review of forecast error covariance statistics in atmospheric variational data assimilation. I: Characteristics and measurements of forecast error covariances. *Quarterly Journal of the Royal Meteorological Society*, *134*(637), 1951–1970. https://doi.org/10.1002/qj.339

Bannister, R. N. (2017). A review of operational methods of variational and ensemble-variational data assimilation: Ensemble-variational Data Assimilation. *Quarterly Journal of the Royal Meteorological Society*, *143*(703), 607–633. https://doi.org/10.1002/qj.2982

Berre, L. (2000). Estimation of Synoptic and Mesoscale Forecast Error Covariances in a Limited-Area Model. *Monthly Weather Review*, *128*(3), 644–667. https://doi.org/10.1175/1520-0493(2000)128<0644:EOSAMF>2.0.CO;2

Daley, R., 1991: Atmospheric Data Analysis. Cambridge University Press, 457 pp.

Daley, R., 1993: Atmospheric data assimilation on the equatorial beta plane. Atmos.–Ocean, 31, 421–450, doi:10.1080/ 07055900.1993.9649479.

Daley, R., 1996: Generation of global multivariate error covariances by singular-value decomposition of the linear balance equation. Mon. Wea. Rev., 124, 2574–2587, doi:10.1175/1520-0493(1996)124,2574:GOGMEC.2.0.CO;2.

Derber, J., & Bouttier, F. (1999). A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus A*, *51*(2), 195–221. https://doi.org/10.1034/j.1600-0870.1999.t01-2-00003.x

Desroziers, G., Mathiot, V., & Orain, F. (1995). A study of ARPEGE forecast error covariances. In *Proc. WMO Second Int. Symp. on Assimilation of Observations in Meteorology and Oceanography* (Vol. 1, pp. 263–268).

Fisher, M. (2003). Background error covariance modelling. In *Seminar on Recent Development in Data Assimilation for Atmosphere and Ocean* (pp. 45–63).

Gauthier, P., Charette, C., Fillion, L., Koclas, P., & Laroche, S. (1999). Implementation of a 3D variational data assimilation system at the Canadian Meteorological Centre. Part I: The global analysis. *Atmosphere-Ocean*, *37*(2), 103–156. https://doi.org/10.1080/07055900.1999.9649623

Lorenc, A. C. (2003). Modelling of error covariances by 4D-Var data assimilation. *Quarterly Journal of the Royal Meteorological Society*, *129*(595), 3167–3182. https://doi.org/10.1256/qj.02.131

McMillan, Martin. (2015). SINGV DA User Guide, UK Met Office, 33pp, http://code.metoffice.gov.uk.

Ménétrier, B., Montmerle, T., Michel, Y., & Berre, L. (2015). Linear Filtering of Sample Covariances for Ensemble-Based Data Assimilation. Part I: Optimality Criteria and

Application to Variance Filtering and Covariance Localization. *Monthly Weather Review*, *143*(5), 1622–1643. https://doi.org/10.1175/MWR-D-14-00157.1

Parrish, D. F., and J. C. Derber (1992), The National Meteorological Center's spectral statistical-interpolation analysis system, *Monthly Weather Review*, *120*(8), 1747–1763.

Pereira, M. B., & Berre, L. (2006). The Use of an Ensemble Approach to Study the Background Error Covariances in a Global NWP Model. *Monthly Weather Review*, *134*(9), 2466–2489. https://doi.org/10.1175/MWR3189.1

Rabier, F., McNally, A., Andersson, E., Courtier, P., Unden, P., Eyre, J., et al. (1998). The ECMWF implementation of three-dimensional variational assimilation (3D-Var). II: Structure functions. *Quarterly Journal of the Royal Meteorological Society*, *124*(550), 1809–1829.

Routray, A., U. C. Mohanty, S. R. H. Rizvi, Dev Niyogi, Krishna K. Osuri and D. Pradhan (2010). Impact of Doppler Weather Radar Data on Numerical Forecast of Indian Monsoon Depressions, *Quarterly Journal of the Royal Meteorological Society*, 136, 1836-1850.

Routray, A., S. C. Kar, P.Mali, and K. Sowjanya (2014). Simulation of Monsoon Depressions Using WRF-VAR: Impact of Different Background Error Statistics and Lateral Boundary Conditions, *Monthly Weather Review*, 142, 3586-3613. *DOI: 10.1175/MWR-D-13-00285.1*

Wang, Hongli, Xiang-yu Huang, Juanzhen Sun, Dongmei Xu, Man Zhang, Shuiyong Fan and Jiqin Zhong (2014). Inhomogeneous Background Error Modeling for WRF-Var Using the NMC Method, *Journal of Applied Meteorology and Climatology*, 53, 2287- 2309.

Weaver, A. T., Deltel, C., Machu, E., Ricci, S., & Daget, N. (2005). A multivariate balance operator for variational ocean data assimilation. *Quarterly Journal of the Royal Meteorological Society*, *131*(613), 3605–3625. https://doi.org/10.1256/qj.05.119.

Wu W-S, Purser RJ, Parrish DF. (2002). Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Weather Rev.* 130, 2905–2916.