Regional 4DVar Assimilation Studies on Weather Systems Over India Using the WRF Model

A thesis submitted
in partial fulfillment for the award of the degree of

Doctor of Philosophy

by

Deepak Gopalakrishnan

Department of Earth and Space Sciences
Indian Institute of Space Science and Technology
Thiruvananthapuram, India

April 2019

Certificate

This is to certify that the thesis titled ***Regional 4DVar Assimilation Studies on Weather Systems Over India Using the WRF Model*** submitted by **Deepak Gopalakrishnan**, to the Indian Institute of Space Science and Technology, Thiruvananthapuram, in partial fulfillment for the award of the degree of **Doctor of Philosophy**, is a bona fide record of the original work carried out by him under my supervision. The contents of this report, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

Prof. A. Chandrasekar	**Dr. Samir Mandal**
Supervisor & Outstanding Professor	Associate Professor & Head
Department of Earth & Space Sciences	Department of Earth & Space Sciences

Thiruvananthapuram
April 2019

Declaration

I declare that this thesis titled ***Regional 4DVar Assimilation Studies on Weather Systems Over India Using the WRF Model*** submitted in partial fulfillment for the award of the degree of **Doctor of Philosophy** is a record of original work carried out by me under the supervision of **Prof. A. Chandrasekar**, and has not formed the basis for the award of any degree, diploma, associateship, fellowship, or other titles in this or any other Institution or University of higher learning. In keeping with the ethical practice in reporting scientific information, due acknowledgments have been made wherever the findings of others have been cited.

Thiruvananthapuram
April 2019

Deepak Gopalakrishnan
(SC14D003)

Acknowledgements

After a fairly long academic life that spanned over nearly twenty-five years, I am landing-up here with a feel of contentment. There are efforts of many people, including my teachers- right from the schooling - that brought up me up to here. Remembering my family with love, particularly for the freedom given to me to follow my choices. It would be my pleasure to express my heartfelt gratitude to all, who have encouraged and supported me in the academic and non-academic life.

At the outset, I express my profound gratitude to Prof. Anantharaman Chandrasekar for his able guidance and valuable support that enabled me to grow as a researcher. I feel so fortunate to have such a great teacher, who has in-depth knowledge in the subject as my research supervisor. I am particularly indebted to him for offering me the freedom in choosing the research problems to work on, and designing the experiments. Despite his busy engagements, he always tried to find sufficient time for academic discussions and to provide insightful feedback.

I would like to thank all my doctoral committee members, Dr. Anandmayee Tej, Dr. Ravi S. Nanjudiah, Dr. Pradip K. Pal, Dr. Govindan Kutty, and Dr. Raju K. George for their constant support and valuable suggestions during the annual review meetings. Particularly, I must acknowledge the constructive comments and thoughtful ideas from Dr. Govindan Kutty, that have greatly helped in a better realization of my thesis. I am also thankful to Dr. M. V. Ramana, my former doctoral committee member, for his support and motivation. I would like to acknowledge Dr. V. K. Dadhwal, the present Director as well as Dr. K. S. Dasgupta, the former Director of IIST for providing the research facilities and research fellowship.

I am extremely grateful to Director, Indian Institute of Tropical Meteorology (IITM), Pune for providing the access to utilize the high-performance computing system, Aaditya to conduct the numerical simulations, without which the thesis could not have been realized. Furthermore, I am thankful to the entire Aaditya team for their prompt technical supports.

My sincere gratitude to Dr. Amit Kesarkar, Dr. Jyoti Bhate, Dr. Vikas Singh and all other members of Weather and Climate Research Group, National Atmospheric Research Laboratory, Tirupati for their valuable support during my visit to National Atmospheric

Research Laboratory. I have greatly benefited from the fruitful discussions I had with them.

I am highly indebted to Dr. Dhanya M. who has been a constant source of motivation and support. She extended her wholehearted support in making me learn the basics of data assimilation and Weather Research and Forecasting model. Moreover, the discussions I used to have with her would have definitely influenced me to approach the research problems in a better way.

I must acknowledge the support from several agencies for model and data. Thanks to National Center for Atmospheric Research for the WRF model, National Centers for Environmental Prediction for Global Forecast System data, National Aeronautics and Space Administration for Tropical Rainfall Measurement Mission rainfall estimates, Indian Space Research Organization (MOSDAC) for Scatsat-1 data, India Meteorological Department and Joint Typhoon Warning Center for tropical cyclone best track observations. Technical assistance from the WRF team in sorting out the model related issues is also acknowledged. I would like to thank the support from computer system group, IIST in providing necessary computational facilities. Also, I extend my sincere thanks to all the administrative staffs and library staffs.

The non-academic life I had at IIST was incredible. Remembering all my friends (obviously, the list is too big) at IIST who made the campus life more lively and memorable. Thanks to all my lab-mates for their encouragement and support throughout the PhD life. I am grateful to Rekha Gogoi for for the fruitful discussions regarding the assimilation of observations from Scatsat-1. Also, thanks to Mukul Pateria for his support in making figures with NCAR Command Language. I would like to acknowledge the sincere efforts of Sarath Babu and Abhishek Chakraborty in creating an excellent LaTeX thesis template.

Encouragement and motivation from my teachers at the Department of Atmospheric Sciences, CUSAT were unparallel. Remembering my all friends from CUSAT for their love and warm friendship. I am also grateful to Dr. E. I. Anila and Dr. K. K. Leelamma, my undergraduate teachers at Union Christian College, Aluva, who initially motivated me towards the scientific research.

Deepak Gopalakrishnan

Abstract

The present study aims to investigate the performance of state-of-the-art Weather Research and Forecasting (WRF) model in the simulation of weather systems over India through regional variational data assimilation (DA) methods. Initially, the present study examines the performance of the four-dimensional variational (4DVar) DA system over the three-dimensional variational (3DVar) DA system for the simulation of a few tropical cyclones (TCs) that formed over the Bay of Bengal region. The improved performance of the 4DVar experiments over the 3DVar counterpart has been quantified for the simulation of cyclone intensity and track by generating a large number (51) of analysis/forecast samples for the TCs investigated. Cyclic assimilations were carried out for the entire lifespan of the TCs at 6h interval, and short-range (48h) free forecasts were initiated from each of the analysis fields. The 4DVar analyzed fields were found to reproduce the initial structure of the TC vortex realistically well, as compared to the 3DVar analyzed fields. The study suggests that, on an average the 4DVar runs can contribute to an improvement of 17-50% in the intensity simulation and 22-57% in the track simulation of the TCs, at different forecast lead times, the above results are being significant at 99% confidence level. The improvement in the rainfall simulation with the 4DVar experiments are attributed to the improved representation of the humidity fields due to the model physics involved in the 4DVar minimization process.

The sensitivity of the 4DVar DA system to the different background error covariance (BEC) formulations have also been investigated in the present study. Three different BECs, which employ three different sets of control variables are utilized. Two of them employ stream function and velocity potential (ψ and χ) as momentum variables, whereas the third one utilizes horizontal wind components as momentum variables (uv-BE). Among the two BECs which employ ψ and χ as momentum variables, the $\psi\chi$-BE method treats humidity as univariate, while the other method ($\psi\chi$-MBE) treats humidity as multivariate. Three heavy rainfall events, that occurred over north Indian region are chosen as case studies to investigate the sensitivity of the 4DVar DA system to different BECs. Five cyclic assimilation were performed for each of the three rainfall cases by utilizing conventional surface and upper air observations together with the satellite derived winds. The analysis

fields obtained using uv-BE experiment were consistently found to be more closer to the radiosonde observations. The quantitative verification of rainfall forecast for 24h and 48h accumulated precipitation indicates that, in general the $\psi\chi$-MBE experiment has better skill in reproducing the observed rainfall. Furthermore, it is noted that, employing the humidity variable in the multivariate form successfully suppresses the overestimation associated with rainfall forecast.

Further experiments were conducted to examine the sensitivity of the 4DVar DA system to different BECs during a fortnight of the Indian summer monsoon. Short-range rainfall forecasts (24h and 48h) were examined for the first 15 sample days during the month of July, 2017. Results from the above study were consistent with the earlier results for the heavy rainfall events; the analysis fields from the uv-BE experiments were found to be the closest to the observations. The rainfall forecast verification revealed a marginal improvement with the $\psi\chi$-MBE experiment.

Lastly, the impact of assimilating ocean surface winds from scatterometer on board Scatsat-1, the Indian scatterometer in the simulation of Indian summer monsoon circulation and associated rainfall has been examined for the entire month of July, 2017. 4DVar assimilations (SCATSAT run) were performed once for each sample days (at 00Z) utilizing the Scatsat-1 winds, conventional surface and upper air observations in combination with the atmospheric motion vectors. Results of this experiment are then compared with two other runs, such as (i) run that assimilated the ocean surface winds from ASCAT (ASCAT run) instead of Scatsat-1 and (ii) the run that did not assimilate any scatterometer winds (CTRL run). Both the 24h and 48h rainfall forecasts for SCATSAT experiments were found to be consistently better as compared to the CTRL run for most of the rainfall thresholds. The comparative study of the results for SCATSAT experiment with the ASCAT experiment shows that, both the scatterometer runs have similar skill in reproducing the 24h rainfall. However, at 48h lead time, ASCAT experiment shows better rainfall forecast skills. Furthermore, the results indicated that the SCATSAT experiment has better probability of detection for most of the rainfall thresholds, as compared to the ASCAT run.

Contents

List of Figures

List of Tables

Abbreviations

3DVar	3-Dimensional Variational
4DVar	4-Dimensional Variational
AAE	Average Absolute Error
AD	Adjoint model
AE	Absolute Error
AMV	Atmospheric Motion Vector
ASCAT	Advanced Scatterometer
ARW	Advanced Research WRF
BEC	Background Error Covariance
BE	BEC Experiment
BES	Background Error Statistics
DA	Data Assimilation
DAR	Domain Averaged Rainfall
EnKF	Ensemble Kalman Filter
ETS	Equitable Threat Score
FAR	False Alarm Ratio
GFS	Global Forecast System
GTS	Global Telecommunication System
IMD	India Meteorological Department
IP	Improvement Parameter
JTWC	Joint Typhoon Warning Center
MetOp	Meteorological Operational Satellite Program of Europe
MBE	Multivariate BEC Experiment
MSLP	Minimum Sea Level Pressure
MWS	Maximum surface Wind Speed
NASA	National Aeronautics and Space Administration
NCAR	National Center for Atmospheric Research

NCEP	National Centers for Environmental Prediction
NMC	National Meteorological Center
NMM	Non-hydrostatic Mesoscale Model
NOAA	National Ocean and Atmospheric Administration
NWP	Numerical Weather Prediction
POD	Probability of Detection
TC	Tropical Cyclone
TL	Tangent Linear model
TRMM	Tropical Rainfall Measurement Mission
VAR	Variational data assimilation system
WRF	Weather Research and Forecasting

Nomenclature

$\mathbf{B}$	Background error covariance matrix
$\mathbf{J}$	Cost function
H	Non-linear observational operator
$\mathbf{H}$	Tangent linear observational operator
M	Non-linear forward model
$\mathbf{M}$	Tangent linear forward model
Q	Water vapor mixing ratio
$\mathbf{R}$	Observation error covariance matrix
T	Air temperature
u	Zonal component of velocity
v	Meridional component of velocity
w	Vertical component of velocity
$\mathbf{x}$	State vector
$\mathbf{x}_\mathrm{a}$	Analysis field
$\mathbf{x}_\mathrm{b}$	Background field
$\mathbf{y}$	Observation
ψ	Stream function
χ	Velocity potential
ζ	Relative vorticity
δ	Divergence

Chapter 1

Introduction

> *"Perhaps some day in the dim future it will be possible to advance the computations faster than the weather advances and at a cost less than the saving to mankind due to the information gained. But that is a dream."*
>
> – L. F. Richardson

Lewis Fry Richardson, the first to advance as well as to implement 'Numerical Weather Prediction'(NWP), nearly a century ago had dared to dream big as the above-mentioned quote indicates. However, NWP is now a highly advanced and complex procedure that is instrumental in predicting the future state of the atmosphere by employing numerical approximations to the dynamical equations that govern the evolution of the atmosphere. A very large number of numerical operations are essential for the realization of NWP, hence it demands enormous amount of computational resources. NWP is conceived as an initial value problem (also a boundary value problem in regional modeling framework) and thus the model requires an accurate knowledge of the initial state of the atmospheric system. The complexities of the physical processes within the atmospheric system and representation of such processes makes NWP a significantly challenging endeavour and therefore, realizing a useful weather forecast is not an easy task. With efficient and vast computing facilities becoming easily available, the accuracy of numerical weather forecasts have improved remarkably, mainly through the increase in the number and variety of observational platforms, better understanding of the physics of the atmospheric system, increase in the model resolutions, and development of sophisticated approaches that create the initial conditions for NWP models.

1.1 Historical overview of weather prediction

Historically, during the initial years, the weather forecasts were presented as a sequence of maps that showed the atmospheric conditions for the next few days. More systematic approaches for preparing the weather forecasts were introduced by Heinreich Wilhelm Brandes in 1816, who collected sub-daily meteorological observations, mostly for Europe, and identified severe storms, that are having low pressure centers. A group of Norwegian meteorologists, led by Vilhelm Bjerknes, was actively involved in developing weather prediction methods that were widely adopted in United States during the World War II [1].

Bjerknes was the first to initially advocate the idea of NWP in 1904. Bjerknes suggested that the weather prediction is an initial value problem, where the basic prognostic equation of atmospheric evolution are integrated forward in time from a known initial state of the atmosphere [2]. L. F. Richardson was the first to actually implement the idea of numerical weather prediction in 1922 with an experimental design, that is outlined in his book *Weather Prediction by Numerical Process* [3]. Unfortunately, Richardson's numerical model results predicted unrealistic and fallacious changes in atmospheric pressure ($\sim$150 hPa in 6h), leading to an extremely poor forecast. Richardson opioned that the extreme errors in the initial wind information was responsible for his poor forecast[1]. Later, it was recognized that, together with the lack of adequate observations, Richardson's numerical model was also not computationally stable.

It took a further couple of decades to demonstrate a successful numerical weather forecast, thanks to the pioneering work of John von Neuman, Jule Charney, and Ragnor Fjortoft in 1950 [4]. Instead of utilizing the complete set of hydro-dynamical equations as employed by Richardson, Charney's team solved [5] the barotropic vorticity equation using the digital computer ENIAC. Encouraging results from the above work paved the way for further advancements in NWP, and the growth of the latter was largely supported by improved computational resources. By the late 1950s a research group in Princeton University (at The Institute of Advanced Study) initiated attempts to develop multi-level computer models and Norman Philips successfully developed a two-layer quasi-geostrophic atmospheric general circulation model in 1956 [6, 7]. Eventually, active research groups from University of California, Los Angeles and National Center for Atmospheric Research, Boulder were involved in developing more sophisticated, primitive equation models. An important development of NWP was the launch of the first artificial satellite *Sputnik* in

1957, which provided an excellent platform to bridge the gaps in observation over the oceans (data sparse regions). Presently, the Earth's atmosphere is being routinely and continuously monitored by several satellites, which have become an integral part of the global observational system. Today, with the help of powerful supercomputers, reasonably good and reliable weather forecasts are available worldwide - thanks to the sophisticated NWP models. Surely, all the above initiatives have transformed Richardson's dream to reality [8].

1.2 Data problem: Emergence of data assimilation

One of the reasons for the disappointing failure of Richardson's efforts of NWP was the limited number of atmospheric observations available at that time, particularly, observations above the surface as well as over the oceans. As proposed by Bjerknes, an accurate knowledge of the initial state of the atmosphere i.e, the *diagnostic* part in the words of Bjerknes, is very critical in NWP. Evidently, preparation of an accurate initial condition to the numerical weather model is a challenging task.

Manually prepared initial conditions were used in the experiments of Richardson and Charney [9]. However, this was a highly tedious task. Gradually methods that fit the uneven observations to regular grid points came into use. Based on two-dimensional polynomial interpolation, R. A. Panofsky [10] introduced a novel method to carry out the *objective analysis* in 1949 [9]. The method was devised in such a manner that, greater 'weights'would be assigned to the observations that have higher accuracy. Moreover, Panofsky imposed a dynamic constraint based on the geostrophic balance, that would relate the wind and mass fields. By 1954, B. Gilchrist and G. P. Cressman [11] proposed a modified version of Panofsky's method, with local polynomial fit, by defining a quadratic polynomial in x and y at each grid points, where the coefficients of the polynomial were determined by minimizing the mean square difference between the polynomial and the observation within a pre-defined radius of influence [9]. More importantly, they opioned that an improved analysis is realizable, if a preliminary information of the state of the atmospheric system is obtained from a short-range forecast [12]. The above mentioned short-range forecast is referred as *first guess (background) field* and the above idea provided the first insight to the modern *data assimilation* (DA) techniques.

With the emergence of sophisticated primitive equation models, it was not meaningful to employ simple hand-interpolation exercises. The total number of grid points multiplied by the independent variables or the *number of degrees of freedom* for a typical NWP model is of the order of 10^8, whereas the number of observation at any one time is far less ($\sim 10^5$). Also, a very large number of the available observations, including those from remote sensing platforms, are irregularly distributed in space and time, necessitating the paramount requirement of the background information. Such a background field or the first guess, which can be derived from a short-range forecast or even climatology, would provide a reasonable estimate of the state of the atmosphere. All the available observations are then combined *optimally* with the first guess to yield the best possible estimate of the state of the atmospheric system, which is referred as the *analysis*. The technique of optimally combining the observations with background field is known as data assimilation. Typically in a cyclic assimilation, a short-range forecast started from the analysis field, at the analysis time will be used as the first guess for the next cycle of assimilation. Thompson P. D. suggested that, information would be transported from data rich regions to data sparse regions through continuous data assimilation cycles, which ultimately results in better and improved representation of data void regions [12].

1.3 Methods of data assimilation

Olivier Talagrand defined the data assimilation as a method of *"using all the available information, to determine as accurately as possible the state of the atmospheric (or oceanic) flow"* [13]. Earlier versions of DA were of the simpler empirical type, that were similar to the objective analysis techniques. In the *successive correction method (SCM)* proposed by Bergthorsson and Doos in 1955, the difference between the first guess and observations (the so-called observational increment) will be estimated and will be used to find the analysis increments. The final analysis would be obtained by adding the analysis increments to the first guess. Observations were weighted based on their distance to model grid points; less the distance, more will be the weights [12]. Various methods (Cressman in 1959, Barnes in 1964 etc.) were developed based on the SCM method. Due to its reasonable performance, SCM methods were widely used. Another empirical method that was popular is *nudging or Newtonian relaxation* method[9]. In nudging method, an additional term is added in each of the prognostic equations, which would "nudge" or "adjust" the solution towards the observations. This additional term will be proportional to the difference between the model variable and the observation. Yet another well known method based on statistical estimation theory was the *Optimum Interpolation (OI)* [14, 12]. The analysis would be obtained as a weighted linear combinations of the first guess and the observations that are available in a "radius of influence".

Inspired from the philosophy of calculus of variational mechanics, Yoshikazu Sasaki in 1958 proposed a novel method called *variational data assimilation scheme*. Despite its early introduction, the method could not be adopted for operational purpose, due to the lack of adequate computational resources. In general, the variational method seeks an analysis field by minimizing a scalar function called *cost function*, the latter being a measure of the misfit between the model state and the observations. In the variational method, the background error statistics (BES) is estimated as time-average difference between two short-range model forecasts of different forecast lengths, valid at the same time [15, 9], thus providing an "approximate" structure of the model errors. BES estimated as indicated above would be homogeneous and isotropic in nature and hence will not account for the evolution of model errors. It is to be noted that the variational framework allows for direct assimilation of non-conventional observations such as satellite radiance, radar reflectivity etc. A popular variational assimilation method is the *three-dimensional variational* (3DVar) method, where the first guess and observations are available only for a single time. Even if the observations are spread over a particular time duration, in 3DVar, all the observations at different times would be assumed to correspond to the time of analysis. A popular variant of the 3DVar method is the 3DVar with first guess at appropriate time (FGAT), where first guess at multiple times would be available so that the closest first guess will be chosen for each observations. The variational procedure is considerably complex if observations that are discretized over a time window are to be assimilated at the appropriate times. In this case the misfit between the temporal sequence of model states and observations will be minimized. The above procedure is referred as the *four-dimensional variational* (4DVar) method. In order to enable assimilation at the exact time of the observations, the forward model has to be integrated up to the time of observation (within the observation window). In the 4DVar algorithm, a linearised version of the non-linear forward model (a *tangent linear model*) and its transpose (an *adjoint model*) are required to do the forward and backward integrations within the assimilation time window. The above approach of employing the dynamics of tangent-linear and adjoint model, implicitly accounts for the evolution of model errors over the assimilation window [16, 17]. Furthermore, since the 4DVar method uses the forward model as a constraint, it will enhance the dynamic balance in the analysis fields. It is indeed true that, the computational costs and complexities involved in the 4DVar method are considerably high, as compared with the 3DVar method.

Yet another promising method for data assimilation is the Kalman filter based technique. The chief advantage of the above method is that the BES is propagated explicitly in time. Thus, the Kalman filter based technique does take care of *"error of the day"* [9]. Geir Evensen applied the Kalman filter method in 1992 to a multi-layered quasi-geostrophic ocean model [18]. Evensen realized that implementing the full Kalman filter algorithm to a weather model with a very large number of grid points ($\sim 10^7$) would be impossible. Subsequently, Evansen proposed the *ensemble Kalman filter* (EnKF) technique for high dimensional systems[19]. In EnKF, the full BES will be replaced by a sample BES, the latter being computed from an ensemble of model states and observations. Hybrid DA versions, that employ the combinations of variational methods and ensemble methods are also found to yield good results [9]. Ensemble based data assimilation methods are presently considered to be one of the highly promising approaches for data assimilation.

1.4 Studies on data assimilation: An overview

In the last couple of decades, the quality and overall skill of numerical weather forecasts have achieved notable and worthy improvement. A prime reason for the improved forecast skill is attributed to the improvements in DA capabilities [20]. With the establishment of variational DA methods (4DVar) (eg: [16, 21, 22]), significant efforts were undertaken to upgrade from the optimal interpolation techniques to the DA procedures based on the variational methods. The National Meteorological Center (NMC) adopted variational DA technique for operational NWP forecasts as early as 1992 [15]. The technique developed at NMC named as "Spectral Statistical Interpolation", was similar to the 3DVar technique. The European Centre for Medium-Range Weather Forecasts (ECMWF) also implemented 3DVar based operational NWP in January 1996 [23, 24, 25]. Other operational weather forecasting agencies such as UK Met Office, Canadian meteorological services etc. also adopted the variational DA approaches in the late 1990s [26, 27]. The 3DVar DA system was conceived as the first step towards achieving an operational 4DVar DA system on a global scale. As mentioned earlier, the 4DVar DA method was initially tested on simplified models including the primitive equation models [16, 21, 28, 29, 30]. In 1997, ECMWF implemented its operational 4DVar analysis system based on the incremental strategy [31, 32, 33]. In the incremental approach developed at ECMWF, a low-resolution model will be employed with simplified physics schemes during the minimization of the cost function to reduce the computational cost [31]. The results from a 12 week global experiments conducted at ECMWF reported significant improvement with the 4DVar experiment over the southern hemisphere and the mid-latitude oceanic areas of northern hemisphere, as compared to the results from the 3DVar experiment. Further, the 4DVar experiment was found to successfully reproduce the rapid cyclogenesis phases more accurately than the 3DVar experiment. Subsequently, the other operational centers from UK, France, Canada and Japan also upgraded to 4DVar based DA system [34, 35, 36, 20].

Parallel to the advancements in DA capabilities within the global NWP models, DA systems were introduced in regional models as well [37, 38, 39, 40, 41]. The 4DVar algorithm was implemented on the primitive equation-regional model at NMC by utilizing the tangent linear model and its adjoint in early 1990s [37]. The above study reported encouraging results with the 4DVar method as compared to the OI method. The adjoint for a primitive equation spectral High-Resolution Limited Area Model (HIRLAM) was developed and was found to provide better short-range forecasts [38]. Further, assimilating the observations using adjoint based DA systems when applied to baroclinically developing systems yielded marked positive impact. Later, the 3DVar system was implemented on the full-physics HIRLAM, by replacing the OI based DA method [42]. The HIRLAM-3DVar was also based on the incremental formulation. The assimilation experiments conducted with HIRLAM utilizing the 3DVar DA system resulted in significant improvements when compared with the statistical interpolation scheme [43]. The major improvements with the 3DVar system were the use of multilevel observation reports and variational quality control procedure[43]. The 4DVar algorithm was implemented on HIRLAM in 2003; however, the results indicated negative impacts with the newly introduced 4DVar DA system as compared with the existing 3DVar system, due to errors in the adjoint model [44]. Later the HIRLAM-4DVar was successfully implemented in 2006 [44, 45]. Another limited area model, the fifth-generation Mesocale Model (MM5) developed by Pennsylvania State University and NCAR [46] also implemented the 3DVar DA method [47]. The initial studies utilizing the MM5-3DVar showed that truncating the vertical eigenmodes reduces the computational cost considerably. Similar 3DVar DA system was implemented on the Weather Research and Forecasting (WRF) model [48]. The WRF model was developed with collaboration from various institutes including NCEP, NCAR etc. The variational DA within the WRF model was extended to 4DVar system following the incremental formulation [49]. A simplified, linear version of the WRF model is utilized for inner-loop minimization. The results from the above study revealed that the WRF-4DVar successfully produced flow-dependent analysis increments, due to the implicit evolution of background error covariance [49]. Real data assimilation experiments with the WRF-4DVar were found to outperform the WRF-3DVar for the simulation of two severe low pressure systems. A stand-alone tangent linear/adjoint model - WRFPLUS - was developed for the WRF-4DVar in 2013, that accounts for simplified physics schemes for representing the surface drag, cumulus convection and large-scale condensation [50, 51]. Cloud resolving numerical experiments were conducted with multi-incremental WRF-4DVar to simulate hurricane Sandy (2012) and results reported encouraging positive impact with the 4DVar assimilation system [51].

With the advancements in satellite technology, the extent of information available to the NWP centers have increased tremendously. Hence, optimal utilization of all the available observations became a challenging task. Earlier studies on DA also focused on the effective utilization of satellite derived information into the NWP models [52]. One of the major advantages with the variational technique is its ability to assimilate observations such as satellite radiance, radar reflectivity etc. that are not model variables. Since the retrieval of the meteorological variables from satellite radiance involve inevitable errors, direct assimilation of satellite radiance is found to have better impact than assimilating retrieved meteorological observations (eg: [53, 54, 9]). The radiance from the vertical sounder on board Television Infra-Red Observation Satellite (TIROS) of NOAA was assimilated at ECMWF by utilizing a one-dimensional variational technique [53]. The simulation of hurricane Danny was investigated by Chen *et al.* using MM5 and its DA system (3DVar) by assimilating the humidity information from Special Sensor Microwave/Imager (SSM/I) [55]. Results from the above study indicated that assimilating the moisture observation from SSM/I provided for improved moisture content over Gulf of Mexico and lower level convergence which contributed to better simulation of hurricane Danny. The above study also noted marked differences in results with assimilating the brightness temperature from SSM/I and retrieved data from SSM/I. An extensive study was carried out by Fan and Tilley [56] on developing dynamical assimilation of cloud properties and moisture observations derived from Moderate Resolution Imaging Spectrometer (MODIS) on the MM5 model. The assimilation of humidity profiles from MODIS resulted in accurate simulation of the model's humidity fields. The improved simulation of humidity fields in turn led to better rainfall forecasts with the MM5 model [56]. Analysis nudging experiments were conducted for the Indian Ocean tropical cyclones [57] with the assimilation of temperature and humidity profiles from Advanced Microwave Sounding Unit (AMSU). The results from the above study revealed that ingestion of AMSU observations provided for wind asymmetries and warm temperature anomalies more accurately. The assimilation of SSM/I observation was found to improve the simulation of Indian Ocean tropical cyclones as well

[58]. The total precipitable water obtained from MODIS was assimilated into the WRF model for the simulation of severe weather systems in a study by Chen *et al.*[59] reported improved results with the assimilation of total precipitable water observations.

Several studies have been reported, which investigate the impact of assimilating satellite observations on the simulation of monsoon depressions over India (eg: [60, 61, 62, 63, 64, 65]). Numerous studies were carried out on the simulation of Indian Ocean tropical cyclones as well. A study conducted by Singh *et al.* [58], utilizing the MM5 model with 3DVar assimilation of SSM/I total precipitable water and ocean surface winds from QuickSCAT also reported that the assimilation of SSM/I observations resulted in improved lower level moisture fields. The track forecast of the Orissa super cyclone (1999) also improved with the assimilation of QuickSCAT winds observations in combination with the SSM/I observations [58]. Simulation of four Indian Ocean tropical cyclones were studied by Singh *et al.* [66] utilizing WRF model with the assimilation of QuickSCAT winds, SSM/I total precipitable water, and atmospheric motion vectors from Meteosat-7. The study revealed negative impact of assimilating AMVs on track simulation of three cyclone cases. In addition, assimilation of SSM/I observations were also found to have negative impact on the intensity simulation. However, assimilation of QuickSCAT near-surface winds impacted positively on track and intensity simulation of the cyclones [66]. Similarly, assimilation of near-surface winds over the ocean from Indian scatterometer on board Oceansat-2 was also found to impact positively on the track simulation of tropical cyclone Phet [67]. However, the assimilation of Oceansat-2 winds degraded the intensity simulation of cyclone Phet. In a detailed study [68], the simulation of eight Indian Ocean tropical cyclones were undertaken using the WRF model and its 3DVar DA system with different sets of observations including AMSU radiance. The results from the above study suggested that the experiments with assimilation of conventional observations with satellite based wind observation yielded the largest impact while the ingestion of AMSU radiance degraded the forecasts of Indian Ocean tropical cyclones considered in the study. Radiance from Sounder for Probing Vertical Profiles of Humidity (SAPHIR) on board Megha-Tropiques satellite was assimilated with WRF-3DVar to study the impact on simulation of few tropical cyclones over India and reported moderate improvement in track and intensity forecasts [69]. Even though one of the above-mentioned studies ([68]) resulted in negative impacts on the simulation of Indian Ocean tropical cyclones due to the ingestion of AMSU radiance, another recent study [70] showed considerable improvement in simulating the features of cyclone Thane and cyclone Jal due to the assimilation of radiance from different satellites.

Almost all the above-mentioned studies utilized the 3DVar DA systems within a regional modeling framework. Studies that utilize 4DVar DA method are considerably fewer as compared with the studies with the 3DVar DA system. The assimilation of Spinning Enhanced Visible and Infra-Red Imager (SEVIRI) radiance with 4DVar DA system on HIRLAM resulted in improvement in mid-tropospheric humidity fields [45]. Satellite rainfall assimilation experiments were conducted over the Indian region by Kumar *et al.*[71] utilizing the 4DVar DA system within the WRF model. Results from the above study indicated that assimilating rainfall estimates from satellites have yielded positive impacts on temperature, humidity, wind, and domain average rainfall. Cloud-resolving numerical experiments were conducted by Ban *et al.*[72] with the WRF model by assimilating NCEP stage IV precipitation data utilizing 4DVar DA system. The results from the study suggested that the precipitation assimilation led to improvement in all the model fields, especially for the lower level humidity fields. In addition, the rainfall forecasts were better after assimilating the rainfall observation [72].

In the DA system, the error statistics of both observations and the model are specified in terms of respective error covariance matrices: the background error covariance (BEC, **B**) matrix and the observation error covariance matrix (**R**). The spread of any observation assimilated is directly related to the nature of BEC, through the error relations defined in the BEC matrix. Hence accurate and realistic representation of BEC matrix is of extreme importance [73, 15]. However, computation of a 'true'BEC is not computationally feasible, owing to the large size of the state space of the NWP model. Hence, in practice, an approximated BEC is estimated by employing the BEC modeling techniques [73]. There are only very few studies exist that addressed the impact of employing different BECs on the simulation of weather systems, particularly over the Indian region. The DA system within the WRF model employs either a global BEC from NCEP or a region specific BEC that has to be modelled for different region of interest. The impact of employing global and regional BEC on the simulation of heavy rainfall event over the Indian region was studied by Rakesh and Goswami [74] with the WRF model and reported clear advantage of employing the regional BEC with 3DVar DA system. A similar study was performed for a set of tropical cyclones over the Indian Ocean region [75], and results from the study were also found to be in agreement with the results of [74]. Experiments with the regional BEC yielded significantly improved results for the simulation of track and intensity [75]. In another study, Routray *et al.*[65] investigated the impact of regional and global BECs on the simulation of monsoon depressions over India and reported positive impact of utilizing a region specific BEC in simulating the location and intensity of rainfall.

The BEC modeling involves a *Control Variable Transform (CVT)*, in which the model variables will be transformed to appropriately chosen *control variables* [73]. The CVT is also responsible for imposing various balance relations among the analysis variables through linear regression. Different set of control variables are being employed by different DA systems [76]. The momentum variables employed in different operational DA systems are one of the following: (i) stream function and velocity potential (ψ and χ), (ii) vorticity and divergence (ζ and δ), and (iii) zonal and meridional wind components (u and v). Fewer studies have addressed the impact of different control variables on the simulation of weather systems. Based on the analytical studies that compared the impacts of the above three momentum variables on the 3DVar analysis, Xie and MacDonald [76] concluded that the DA system employing ψ and χ does introduce analysis errors and the horizontal velocity components are the preferred choice of momentum variables. The above study attributed the analysis error associated with the $\psi\chi$-based DA system to its inherent property of conserving the integral values of wind. A couple of studies ([77, 78]) have compared the impact of two BECs that are based on ψ and χ momentum variables with 3DVar DA system within the WRF model. Among the two BECs employed in the above studies, one was a fully multivariate BEC. Results from these studies indicated that the fully multivariate BEC has moderate positive impact on rainfall simulation. Another study ([79]) compared the impacts of $\psi\chi$-based BEC and uv-based BEC by assimilating radar winds utilizing 3DVar method on the simulation of a squall line and reported that the experiment with $\psi\chi$-based BEC generated discontinuities in the wind fields and contributed to degradation in the precipitation forecast. The study conducted by Sun *et al.*[80] also reported similar findings and observed that the uv-based analysis fields are more closer to radar observations. Here also, all the above mentioned studies are based on a 3DVar DA framework. Since advanced DA methods such as 4DVar are capable of providing an implicit flow-dependence when compared to the 3DVar method, the impact of different BECs on forecast fields may vary considerably within a 4DVar framework.

1.5 Motivation and objectives of the thesis

There are several studies that have investigated and demonstrated the positive impact of assimilating various observations (both conventional and satellite based) using variational DA technique for weather systems over India. However, most of the above studies are based on the 3DVar technique. Very few studies exist in the literature that have quantified and documented the improved performance of advanced DA systems like 4DVar over the performance of the 3DVar system for the simulation of severe weather systems like tropical cyclones over India. Since the implicit flow dependence realizable for the 4DVar system may differ from global model to regional one, it is indeed very important to quantify and document the improvements of employing the 4DVar scheme over its 3DVar counterpart in a regional model.

Studies which investigate the sensitivity of the 4DVar DA system to the control variables chosen for BEC modeling over India do not exist in the literature. Although there exists few studies that compare the relative performance of two BECs based on $\psi\chi$ formulation, they were investigated within the 3DVar framework ([77, 78]). Moreover, results from the above studies did not show any substantial improvement in the analysis and forecast fields with the use of the multivariate BEC. No study has been reported in the literature, which investigates the performance of three different BECs (two $\psi\chi$ BECs and one uv BEC) within a 4DVar framework. More importantly, with its inherent ability to account for the model errors and multicorrelations in an implicit way, the 4DVar method is expected to yield better results with the univariate uv based BEC than with a 3DVar system for the same BEC.

Assimilating ocean surface winds from scatterometers have yielded improved NWP forecasts. India had launched its new scatterometer on board Scatsat-1 as the successor of Oceansat-2 on 26 September, 2016. There are no studies reported in the literature, that have examined the impact of assimilating Scatsat-1 derived ocean surface winds on the simulation of weather systems over India in a regional modeling framework. Also there is no reported study that has successfully assimilated Scatsat-1 ocean surface winds over the Indian region using the 4DVar method. Hence it is important to investigate the performance of a regional model during the Indian summer monsoon period with the assimilation of Scatsat-1 winds using the 4DVar method. It would be worthwhile to compare the above results with the results of 4DVar assimilation of other scatterometers like Advanced Scatterometer (ASCAT), the latter having been recognized as an important part of the current observational system.

Accordingly, the objectives of the present study are:

- Analyze and quantify the improved performance of regional 4DVar system over the 3DVar system for the simulation of a few tropical cyclones that formed over India by generating a large number of analysis/forecast samples.

- Investigate the sensitivity of the 4DVar DA system to the control variables in the BEC formulation with emphasis on (i) heavy rainfall events over India and (ii) the Indian summer monsoon circulation and rainfall.

- Examine the impact of Scatsat-1 scatterometer observations on the simulation of Indian summer monsoon rainfall for entire month of July in 2017 using the 4DVar DA method.

A detailed account of the model and the DA systems employed, observations used for assimilation and forecast verification are provided in Chapter 2. Results from a preliminary study on the performance of 3DVar and 4DVar DA system within the WRF model on the simulation of two Indian Ocean tropical cyclones are presented in Chapter 3. The detailed study on the improved performance of the 4DVar DA system in the simulation of four Indian Ocean tropical cyclones together with the quantification of improvement with the 4DVar method as compared with the 3DVar method is described in Chapter 4. The next chapter (Chapter 5) examines the results from the experiments that investigated the sensitivity of the 4DVar DA system to the three different BEC formulations in simulating heavy rainfall events over the Himalayan region, India. The BEC experiments have been extended to the simulation of Indian summer monsoon rainfall for a period of 15 days in July, 2017. Results from the above experiments are also included in Chapter 5. The impact of assimilating ocean surface winds from Scatsat-1 is investigated by utilizing the WRF-4DVar DA system on the simulation of rainfall during Indian summer monsoon during the month of July, 2017 and the above results are presented and duly compared with the 4DVar assimilation of ASCAT and are outlined in Chapter 6. Major findings from the present study are summarized with concluding remarks in Chapter 7.

Chapter 2

Data and Experimental Methods

A detailed account on the materials and methods adopted in the present study is given in this chapter. Details of the model, observations utilized for assimilation, data used for analysis/forecast verification are briefly described.

2.1 The WRF model: An overview

The Weather Research and Forecasting (WRF) model [48] is a state-of-the-art numerical weather prediction (NWP) system, which is widely recognized as an ideal tool for limited area NWP applications. The WRF system is a collaborative project by many leading agencies and is maintained as a community model. Even though the WRF model is being utilized primarily for limited area applications, it supports global scale simulations as well. In addition, it can serve as regional climate model also. The components of the WRF modeling system are illustrated in Fig. 2.1 and consists of a pre-processing system, dynamic solvers, physics interface, initialization schemes, data assimilation system, and a chemical module.

The WRF pre-processing system (WPS) consists of three programs, namely *geogrid, ungrib,* and *metgrid*. The purpose of WPS is to generate three-dimensional fields by combining the terrestrial information and meteorological data for a user-defined domain. The geogrid utility will interpolate the terrain information, soil categories, and land use categories, etc. on to the grid-points. The ungrib utility is responsible for extracting the meteorological fields from the gridded binary (GRIB) form to an intermediate format. The meteorological fields are typically obtained from global models such as GFS. Finally the

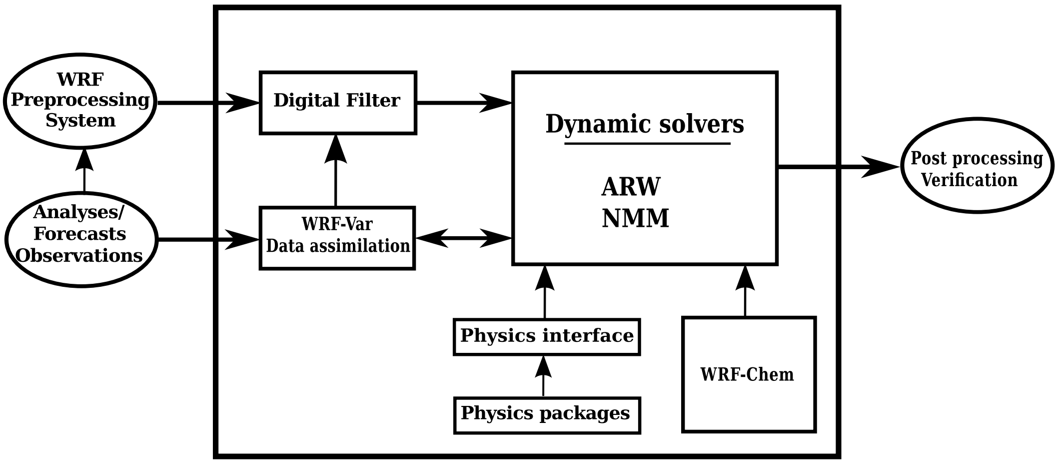

Fig. 2.1: Components of the WRF system. *Source: Skamarock et al. [48]*

metgrid utility will interpolate the meteorological fields "degribbed" by the ungrib program to user-defined experimental domain (generated by geogrid). The output from metgrid can then be utilized to perform numerical simulation using the dynamical solvers. The data flow and program utilities for WRF system are illustrated in Fig. 2.2

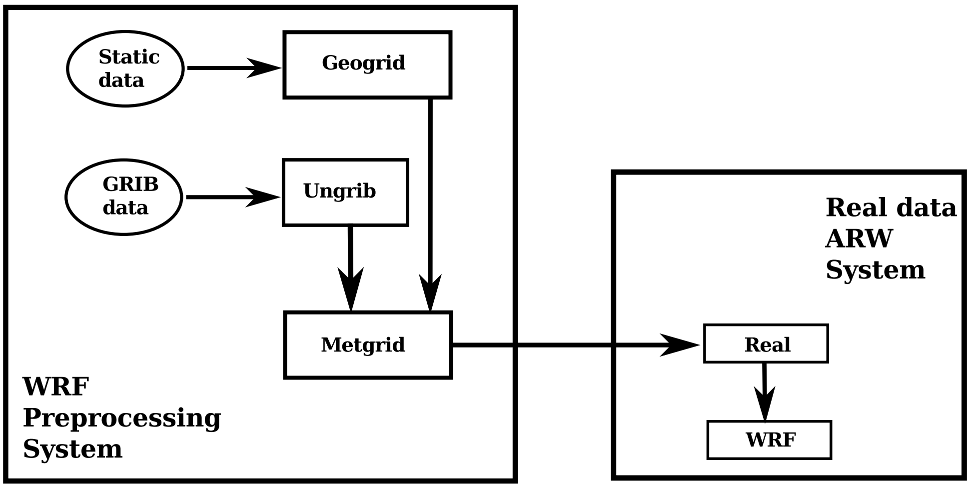

Fig. 2.2: Schematic of the WRF programs and data flow. *Source: Skamarock et al. [48]*

There are two variants of WRF solvers: (i) the Advanced Research WRF (ARW), and (ii) the Non-hydrostatic Mesoscale Model (NMM). The ARW core was mainly developed by NCAR, while the NMM core development was carried out by NCEP. Both the ARW and NMM cores are based on non-hydrostatic, primitive equations with terrain following vertical coordinates. The present study has utilized the ARW core of the WRF model. It employs the horizontal and vertical wind components, perturbation potential temperature, perturbation geopotential and the perturbation surface pressure of dry air as prognostic variables. Also, turbulent kinetic energy, water vapor mixing ratio, rain/snow mixing ratio, cloud water/ice mixing ratio etc. can also be used as optional variables. The vertical coordinates are terrain following eta-coordinates, which provide for vertical stretching[81]. The horizontal grids are staggered with Arakawa C-grid. Spatial discretization is done with 2nd-6th order advection options in both horizontal and vertical directions. Runge-Kutta 3rd order scheme[82] is employed for time integration. The high-frequency acoustic and gravity waves are taken care by utilizing smaller time-step for integration. The model employs physical/free slip bottom boundary condition along with a rigid lid condition for upper boundary, the latter assumed to be fixed at a constant pressure level.

2.1.1 Governing equations

The prognostic equations employed in the ARW solver are formulated on terrain following hydrostatic-pressure vertical coordinates, known as eta (η) coordinate system. If p_h is the hydrostatic component of the pressure and p_{hs}, p_{ht} are the pressure values along the surface and top boundaries, η can be defined as:

$$\eta = \frac{(p_h - p_{ht})}{(p_{hs} - p_{ht})} \tag{2.1}$$

At surface η will be 1 and 0 at the upper boundary of the atmosphere. The η coordinate system allows for vertical stretching that enables the lower atmospheric processes to be very finely resolved as compared to upper atmosphere.

Given that $\mu = (p_{hs} - p_{ht})$, $\mu(x, y)$ would indicate the mass per unit area within the column at (x, y) on the model domain. Then, the flux form of the prognostic variables can be written as:

$$\begin{aligned} \mathbf{V} &= \mu\mathbf{v} = (U, V, W) && (2.2) \\ \Omega &= \mu\dot{\eta} && (2.3) \\ \Theta &= \mu\theta && (2.4) \end{aligned}$$

here, $\mathbf{v} = (u, v, w)$ are the covariant velocities, $\omega = \dot{\eta}$ is the contravariant 'vertical' velocity and θ is the potential temperature. The Euler equations in the flux form are written in terms of the above defined variables as:

$$\begin{aligned}
\partial_t U + (\nabla \cdot \mathbf{V}u) - \partial_x(p\partial_\eta \phi) + \partial_\eta(p\partial_x \phi) &= F_U && (2.5)\\
\partial_t V + (\nabla \cdot \mathbf{V}v) - \partial_y(p\partial_\eta \phi) + \partial_\eta(p\partial_y \phi) &= F_V && (2.6)\\
\partial_t W + (\nabla \cdot \mathbf{V}w) - g(\partial_\eta p - \mu) &= F_W && (2.7)\\
\partial_t \Theta + (\nabla \cdot \mathbf{V}\theta) &= F_\Theta && (2.8)\\
\partial_t \mu + (\nabla \cdot \mathbf{V}) &= 0 && (2.9)\\
\partial_t \phi + \mu^{-1}[(\mathbf{V} \cdot \nabla \phi) - gW] &= 0 && (2.10)\\
\partial_\eta \phi &= -\alpha\mu && (2.11)\\
p &= p_0(R_d\theta/p_0\alpha)^\gamma && (2.12)
\end{aligned}$$

Here, ϕ is geopotential ($\phi = gz$), p is pressure and α is the specific volume (inverse density). γ is the ratio of specific heats of dry air ($\gamma = C_p/C_v$=1.4). Eqn. 2.11 is the diagnostic relationship for the specific volume (hydrostatic balance) and Eqn. 2.12 is the equation of state. The subscripts $x,\ y$ and η indicate respective derivatives. $F_U,\ F_V\ F_W$, and F_Θ are the forcing terms.

The coordinate for dry-air mass can be written as,

$$\eta = \frac{(p_{dh} - p_{dht})}{(\mu_d)} \tag{2.13}$$

Here, μ_d is the mass of the dry-air in the vertical column. Similarly, in the case of dry-air, Eqn. 2.2–2.4 become,

$$\begin{aligned}
\mathbf{V} &= \mu_d \mathbf{v} = (U, V, W) && (2.14)\\
\Omega &= \mu_d \dot{\eta} && (2.15)\\
\Theta &= \mu_d \theta && (2.16)
\end{aligned}$$

Now the modified Euler equation can be written as:

$$\begin{aligned}
\partial_t U + (\nabla \cdot \mathbf{V}u) + \mu_d \alpha \partial_x p + (\alpha/\alpha_d)\partial_\eta(p\partial_x \phi) &= F_U && (2.17)\\
\partial_t V + (\nabla \cdot \mathbf{V}v) + \mu_d \alpha \partial_y p + (\alpha/\alpha_d)\partial_\eta(p\partial_y \phi) &= F_V && (2.18)\\
\partial_t W + (\nabla \cdot \mathbf{V}w) - g[(\alpha/\alpha_d)\partial_\eta p - \mu_d)] &= F_W && (2.19)
\end{aligned}$$

$$\begin{aligned}\partial_t\Theta + (\nabla\cdot\mathbf{V}\theta) &= F_\Theta && (2.20)\\ \partial_t\mu_d + (\nabla\cdot\mathbf{V}) &= 0 && (2.21)\\ \partial_t\phi + \mu_d^{-1}[(\mathbf{V}\cdot\nabla\phi) - gW] &= 0 && (2.22)\\ \partial_t Q_m + (\nabla\cdot\mathbf{V}q_m) &= F_{Q_m} && (2.23)\\ \partial_\eta\phi &= -\alpha_d\mu_d && (2.24)\\ p &= p_0(R_d\theta_m/p_0\alpha_d)^\gamma && (2.25)\end{aligned}$$

In the above equation, α_d is the specific volume for dry-air, $\theta_m = \theta(1 + (R_v/R_d)q_v)$ and $Q_m = \mu_d q_m$.

The WRF-ARW supports both isotropic and anisotropic map projections (projection to sphere). For an isotropic projection, $(\Delta x/\Delta y)|_{earth}$ will be a constant on the grid. Lambert conformal, polar stereographic and Mercator projections are isotropic and the latitude-longitude projection is anisotropic in nature. In orthogonal projections, the physical distance between the grid points will vary with position. However, in computational space the distance between two grid points (dx and dy) are constants. Accordingly, *map scale factors* (m_x and m_y) are defined as the ratio of distance in computational space to the corresponding distance on the earth's surface. Finally, the map-factors can be incorporated in the governing equations with a new set of momentum variables:

$$U = \mu_d u/m_y,\ \ V = \mu_d v/m_x,\ \ W = \mu_d w/m_y,\ \ \Omega = \mu_d\dot{\eta}/m_y$$

The forcing terms in the Euler equations contain the curvature and Coriolis terms together with physical forcing terms. For isotropic projections, where $m_x = m_y = m$, the Coriolis and curvature terms are employed in the following form.

$$\begin{aligned}F_{U_{cor}} &= +\Big(f + u\frac{\partial m}{\partial y} - v\frac{\partial m}{\partial x}\Big)V - eWcos\alpha_r - \frac{uW}{r_e} && (2.26)\\ F_{V_{cor}} &= -\Big(f + u\frac{\partial m}{\partial y} - v\frac{\partial m}{\partial x}\Big)U - eWsin\alpha_r - \frac{vW}{r_e} && (2.27)\\ F_{W_{cor}} &= +e(Ucos\alpha_r - Vsin\alpha_r) + \Big(\frac{uU + vV}{r_e}\Big) && (2.28)\end{aligned}$$

Here, $f = 2\Omega_e sin\phi$ is the Coriolis parameter, $e = 2\Omega_e cos\phi$, Ω_e is the angular velocity of the earth, and r_e is the radius of the earth. In this study, the earth is considered to be a perfect sphere of radius r_e.

2.1.2 Temporal and spatial discretization

The governing equations in NWP are highly non-linear and coupled partial differential equations and hence cannot be solved analytically. Finite difference method or spectral methods are routinely employed to solve the NWP equations. The ARW solver uses finite difference method to solve the NWP equations that are discretized in both space and time.

Spatial discretization in WRF-ARW follows the Arakawa C grid [83] design, which is known to provide for better geostrophic adjustment [83, 9]. Fig. 2.3 illustrates the distribution (staggered) of prognostic variables on Arakawa C grid as employed in the ARW solver. From Fig. 2.3, it is clear that the velocity components are staggered one-half grid length from the potential temperature (at *mass points*). Pressure (p) and specific volume (α) are computed at mass points. Even though the horizontal model grid lengths (Δx and Δy) are fixed in the model formulation, the changes in physical grid lengths are taken into account using the map factors. The vertical grid spacing $\Delta\eta$ is not a constant in model configuration and is specified during the model initialization.

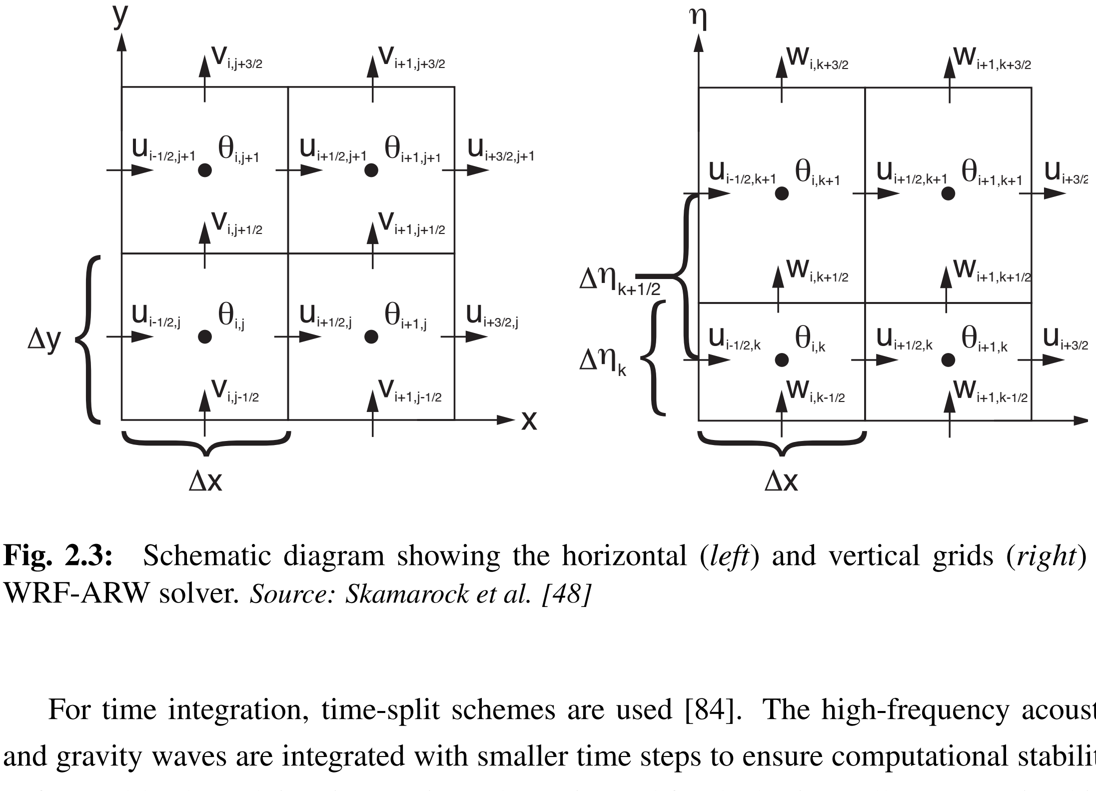

Fig. 2.3: Schematic diagram showing the horizontal (*left*) and vertical grids (*right*) in WRF-ARW solver. *Source: Skamarock et al. [48]*

For time integration, time-split schemes are used [84]. The high-frequency acoustic and gravity waves are integrated with smaller time steps to ensure computational stability. A forward-backward time integration scheme is used for the horizontally propagating high frequency modes, while a vertically implicit scheme is being employed for the integration of vertically propagating acoustic waves and buoyancy oscillations. However, the low frequency modes are integrated over relatively higher time steps using the third-order Runge-Kutta (RK3) scheme [82]. The RK3 method follows a predictor-corrector approach. If the prognostic equations are defined as $\Phi = (U, V, W, \Theta, \phi', \mu', Q_m)$ and the model equations as $\Phi_t = R(\Phi)$, the solution $\Phi(t + \Delta t)$ will be arrived at by marching in the following 3 steps (Δt is the model time step). * and ** indicate intermediate steps.

$$\Phi^* = \Phi^t + \frac{\Delta t}{3} R(\Phi^t) \tag{2.29}$$

$$\Phi^{**} = \Phi^t + \frac{\Delta t}{2} R(\Phi^*) \tag{2.30}$$

$$\Phi^{t+\Delta t} = \Phi^t + \Delta t R(\Phi^{**}) \tag{2.31}$$

2.1.3 Representation of sub-grid scale processes

In the NWP models, the atmosphere/ocean will be represented as a set of discrete three dimensional arrays, which have a specific grid spacing. Hence, the atmospheric or oceanic processes which have a horizontal scale smaller than the grid spacing of the NWP model, will not be represented explicitly. However, these smaller, *sub-grid scale* physical processes will have critical impacts on the evolution of large-scale weather phenomena. Hence, in order to take these sub-grid scale processes into account, their approximated effects will be formulated in terms of the large-scale, model variables and this procedure is known as *parameterization*. Thus, the processes that are being resolved by the model will be evolved by the *model dynamics* and the processes that need to be parameterized are taken care of through the *model physics*.

Major physical processes that are parameterized with the help of model physics are (i) cloud microphysics, (ii) cumulus convection, (iii) planetary boundary layer, (iv) land surface, and (v) radiation. The physics package in ARW utilizes physics drives and is designed in such a way that, it is independent of the main dynamical solver part. A pre-physics step will prepare the necessary variables in un-staggered space and ingest to the physics utility. The physics packages will then compute the tendencies for all the physics variables, which will be brought back to the staggered space by a post-physics step. Finally, the tendencies computed by the physics utilities will be coupled with the dynamical solver by the post-physics step. The tendencies for radiation, surface, boundary layer and cumulus convection are calculated in the first step of the time integration and will be held fixed throughout the Runge-Kutta step. To maintain the saturation condition at the end of the integration step, the microphysics will be computed only at the last Runge-Kutta step.

2.1.3.1 Atmospheric radiation

Electromagnetic radiation from the Sun is the driving force for all kinds of circulatory systems on the Planet Earth. Radiation parameterization will account for the total radiative flux at any given location and time. Basically, it provides the atmospheric heating due to the radiative flux divergence. The shortwave radiation primarily consists of visible spectrum together with the neighbouring frequencies. Shortwave radiative parameterization schemes would account for the interaction of radiation with the atmosphere in the form of scattering, reflection, and absorption. The longwave radiation consists of the infrared spectrum, which manifests as absorbed/emitted radiation by different atmospheric constituents and earth's surface. The upward longwave flux of the model atmosphere is essentially a function of surface temperature, that also depends on the land-use type.

The radiation parameterization schemes employed in WRF-ARW are one-dimensional schemes. The shortwave radiation parameterization scheme that is widely utilized is the one proposed by Jimy Dudhia in 1989 [85]. It determines the shortwave radiation reaching the surface by integrating the solar flux downward by accounting for the interactions such as scattering, water vapor absorption [86], reflection and absorption by the clouds etc. The rapid radiative transfer model (RRTM) scheme is the most commonly used longwave parameterization scheme [87]. The RRTM scheme is a spectral-band scheme and utilizes correlated-k approach. It provides for the longwave interaction corresponding to carbon dioxide, water vapor etc. with the help of look-up table. The scheme accounts for cloud optical depth as well.

2.1.3.2 Land surface process

A realistic representation of surface-atmosphere interaction is necessary for an accurate numerical weather forecast, since the former determines the exchange of mass, energy and momentum upward from the surface to the atmosphere. Moreover, all these fluxes may vary considerably with the nature of the terrain, vegetation, soil characteristics etc. In NWP models, land surface models (LSMs) are being utilized to simulate the surface heat, moisture and momentum fluxes. The LSM is integrated together with the main forward model simultaneously. The LSM employed in ARW is a one-dimensional column model and it yields heat and moisture fluxes based on the information from other physics schemes. Commonly used LSMs in WRF are 5-layer thermal diffusion model, Noah LSM, Rapid update cycle model etc. The thermal diffusion model is a 5-layer soil temperature model, with the soil layers centered at 1, 2, 4, 8, and 16 cm. Soil moisture field is taken as a season dependent constant and has fixed values depending on the land use type. Being a simpler model, it does not account for any vegetation effects explicitly. The Noah LSM [88] is a 4-layer model with soil layers having thicknesses of 10, 30, 60, and 100 cm, respectively which was developed at NCEP. The soil fields in Noah LSM are time-dependent and consistent with the analysis data set. The Noah LSM accounts for evepotranspiration, root zone, soil drainage, soil texture, vegetation categories etc. The rapid update cycle LSM is a 6 layer soil model which also has a multi-layer snow scheme. The above LSM seeks solution for heat diffusion and Richardson transfer equation [89, 90].

2.1.3.3 Planetary boundary layer

Planetary boundary layer (PBL) is the lowest portion of the atmosphere that constantly interacts with the surface of the planet earth. The PBL region is highly turbulent with associated eddies that are responsible for transporting heat and moisture flux upward from the surface. A realistic representation of these sub-grid scale eddies and the resulting vertical transports can be achieved using PBL parameterization. The PBL parameterization scheme takes care of the vertical diffusion not only in the boundary layer, but in the entire atmospheric column. The land-surface schemes and the surface layer provide the surface fluxes to the PBL model from which the flux profiles will be derived for the entire atmosphere together with temperature, moisture and momentum tendencies. Like the LSMs, the PBL schemes are also one-dimensional in nature.

Popular PBL schemes in WRF include the medium range forecast model (MRF) scheme, Yonsei university (YSU) scheme, Mellor-Yamada-Janjic (MYJ) sheme etc. The MRF scheme [91] uses counter-gradient flux for heat and moisture in unstable conditions and the entrainment is represented as part of the PBL mixing. The PBL height is determined using critical bulk Richardson number (Ri). The YSU scheme [92] is an updated version of MRF scheme. Fluxes are represented as counter-gradient terms. Entrainment is treated explicitly, and is considered proportional to the surface buoyancy flux. The PBL top in the YSU scheme is determined by the buoyancy profile (at the maximum entrainment layer).

The MYJ scheme [93, 94] follows the so-called 1.5-order closure scheme with prognosis equation for turbulent kinetic energy (TKE). In the MYJ scheme, entrainment is not treated explicitly and thus is considered as a part of PBL mixing. The PBL height is dependent on TKE and buoyancy and shear of the primary flow.

2.1.3.4 Cloud microphysics

Moist convection will result in formation of clouds. The physical processes that govern the formation and growth of cloud particles and their fallout as precipitation are being accounted by microphysics parameterization schemes in the NWP models. The length-scale of these cloud processes are in the range of 10^{-6}m to 10^{-3}m. There are two fundamental approaches to parameterize the cloud microphysics: *bin models* and *bulk models*. In bin modeling, the particle size distribution would be determined explicitly whereas in the bulk approach, the same is represented by a distribution functions such as exponential or gamma distributions for each hydrometeors[95, 96]. In general, the single-moment bulk parameterization schemes predict the particle mixing ratio or specific humidity only, while the double-moment schemes predict particle number concentration in addition to the mixing ratio.

The WRF-ARW supports a variety of microphysics parameterization schemes namely, Kessler [97], Purdue Lin [98, 99], WRF single moment 3-class (WSM3) [100], WSM5 [100, 101], WSM6 [98, 100], Thompson [102] etc. The above schemes differ from one another in the manner they handle the mass mixing-ratio variables. The various cloud mass variables are the cloud water mixing ratio (q_c), rain water mixing ratio (q_r), ice mixing ratio (q_i), snow mixing ratio (q_s), graupel mixing ratio (q_g), total ice mixing ratio (q_t), mixing ratio for hail (q_h). Among the above schemes, the Kessler scheme is a simple warm cloud scheme that does not provide for ice-phase/mixed-phase processes. All the other schemes indicated above do provide for both ice-phase processes and mixed-phase processes, except WSM3 and WSM5 schemes that do not account for mixed-phase processes.

2.1.3.5 Cumulus convection

In general, moist convection is associated with updrafts and downdrafts of relatively smaller length-scales, which are not resolved by the NWP model explicitly in general. The cumulus parameterization is responsible for representing the effects of smaller-scale deep/shallow convection in terms of the resolved variables. The cumulus schemes essentially compute the changes in temperature and moisture fields due to the convection [96] for individual columns and provides for vertical profiles of temperature and humidity. In addition, a few cumulus schemes also provide for cloud and precipitation tendencies as well. Cumulus schemes are generally employed for coarser horizontal resolution models (>10 km) where implicit representation of convection and resultant latent heat release are important. Though the cumulus parameterization schemes are found to have positive impact in triggering the convection in models with higher resolution (5-10 km), convective parameterization schemes are not applied for a cloud-resolving scale applications (< 5 km). There are few newer schemes that are "scale-aware schemes", which can be employed in smaller-scale applications as well.

Major cumulus schemes available in the WRF-ARW are Kain-Fritsch (KF), Simplified Arakawa Schubert (SAS), Betts-Miller-Janjic (BMJ), Grell-Devenyi ensemble (GD) scheme etc. The KF scheme [103, 104, 105] follows the mass-flux approach and uses a simple cloud model with moist updrafts and downdrafts. The KF scheme accounts for the effects of entrainment and detrainment as well. The BMJ scheme [106, 107] is a convective adjustment scheme which accounts for both deep and shallow convection. In the BMJ scheme, the temperature and humidity fields are adjusted to observed quasi-equilibrium thermodynamic structures. The GD scheme [108] is an ensemble cumulus parameterization scheme that follows mass-flux approach. In the GD scheme, multiple cumulus schemes are run and the ensemble mean would be fed back to the model. Finally, the SAS scheme [109, 110, 111] also follows the mass-flux approach to adjust the temperature and humidity fields by employing quasi-equilibrium assumption for closure.

2.2 Assimilation methodology

The primary goal of any data assimilation technique is to obtain a best or *optimal* state of the atmospheric (or oceanic) system by combining all the available observations with a background field. The *Best Linear Unbiased Estimator (BLUE)* for multi-dimensional problem can be written as:

$$\mathbf{x_a} = \mathbf{x_b} + \mathbf{K}(\mathbf{y_o} - H(\mathbf{x_b})) \tag{2.32}$$

$$\mathbf{K} = \mathbf{B}\mathbf{H^T}(\mathbf{H}\mathbf{B}\mathbf{H^T} + \mathbf{R})^{-1} \tag{2.33}$$

where, $\mathbf{x_a}, \mathbf{x_b}, \mathbf{y_o}$ H are analysis field, background field, observation, and observation operator respectively. **K** is a linear operator called *gain* or *weight matrix*. $\mathbf{B}, \mathbf{R}, \mathbf{H}$ are the background error covariance matrix, observation error covariance matrix, and linearized observation operator respectively. The best estimate or the *analysis* can be obtained by minimizing the analysis error variance, that is, by finding the *optimal weights* through a least squares approach. An alternate approach is the *variational method*, in which the analysis will be obtained by minimizing a *cost function* that measures the distance between the analysis and background and the analysis and observations. In fact, both the above approaches have yielded equivalent solutions [112, 9]. The present study utilizes the variational methods. Generally, two approaches are followed in variational data assimilation system: the three-dimensional variational (3DVar) method and four-dimensional variational (4DVar) method.

2.2.1 Three-dimensional variational approach

In the variational method, an optimal analysis is sought as an approximate solution by minimizing the cost function (J) instead of computing the gain, **K** (Eqn. 2.33). The cost function is defined as:

$$2J(\mathbf{x}) = (\mathbf{x}-\mathbf{x_b})^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x_b}) + [\mathbf{y_o}-H(\mathbf{x})]^T\mathbf{R}^{-1}[\mathbf{y_o}-H(\mathbf{x})] \tag{2.34}$$

As per definition, the analysis $\mathbf{x_a}$ is obtained when $\nabla_{\mathbf{x}}J(\mathbf{x_a})$ =0. Expanding the second term of Eqn. 2.34 by linearizing the observation operator H around the background $\mathbf{x_b}$ yields:

$$\mathbf{y_o}-H(\mathbf{x}) = \mathbf{y_o}-H[\mathbf{x_b}+(\mathbf{x}-\mathbf{x_b})] = \mathbf{y_o}-H(\mathbf{x_b})-\mathbf{H}(\mathbf{x}-\mathbf{x_b}) \tag{2.35}$$

Employing Eqn. 2.35 in Eqn. 2.34 gives:

$$\begin{aligned}
2J(\mathbf{x}) &= (\mathbf{x}-\mathbf{x_b})^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x_b}) && (2.36)\\
&+ [\{\mathbf{y_o}-H(\mathbf{x_b})\}-\mathbf{H}(\mathbf{x}-\mathbf{x_b})]^T\mathbf{R}^{-1}[\{\mathbf{y_o}-H(\mathbf{x_b})\}-\mathbf{H}(\mathbf{x}-\mathbf{x_b})]\\
2J(\mathbf{x}) &= (\mathbf{x}-\mathbf{x_b})^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x_b}) + (\mathbf{x}-\mathbf{x_b})^T\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}(\mathbf{x}-\mathbf{x_b})\\
&- \{\mathbf{y_o}-H(\mathbf{x_b})\}^T R^{-1}\mathbf{H}(\mathbf{x}-\mathbf{x_b}) - (\mathbf{x}-\mathbf{x_b})^T\mathbf{H}^T\mathbf{R}^{-1}\{\mathbf{y_o}-H(\mathbf{x_b})\}\\
&+ \{\mathbf{y_o}-H(\mathbf{x_b})\}^T\mathbf{R}^{-1}\{\mathbf{y_o}-H(\mathbf{x_b})\} && (2.37)
\end{aligned}$$

That is, from the above equation, the cost function is a quadratic function of the analysis increment $(\mathbf{x} - \mathbf{x_b})$. For a quadratic function of the form $F(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{d}^T\mathbf{x} + c$, its gradient will be $\mathbf{A}\mathbf{x} + \mathbf{d}$. Therefore,

$$\nabla J(\mathbf{x}) = \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x_b}) + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}(\mathbf{x} - \mathbf{x_b}) - \mathbf{H}^T\mathbf{R}^{-1}\{\mathbf{y_o} - H(\mathbf{x_b})\} \tag{2.38}$$

When $\nabla J(\mathbf{x_a}) = 0$, J will be minimum and in that case,

$$\mathbf{x_a} = \mathbf{x_b} + (\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{R}^{-1}\{\mathbf{y_o} - H(\mathbf{x_b})\} \tag{2.39}$$

The Eqn. 2.39 is the solution to the 3DVar analysis problem. The solution is usually achieved using minimization methods such as conjugate gradient method. The 3DVar system within the WRF model follows the implementation by Barker *et al.* [47, 113].

2.2.2 Four-dimensional variational approach

Major limitations with the 3DVar approach are (i) it treats all the observations available in an observation window as observations at a single time (i.e, the analysis time), and (ii) it uses a completely static background error covariance matrix (**B**). The 4DVar method accounts for the above-mentioned limitations of 3DVar [16, 22, 114]. The 4DVar method accounts for observations that are distributed over a time window and hence allows for observations to be assimilated at the exact time of the observation [17, 32, 35, 115]. In the case of 4DVar method, the cost function takes the form as given below:

$$\begin{aligned} J[\mathbf{x}(t_0)] &= \frac{1}{2}[\mathbf{x}(t_0) - \mathbf{x^b}(t_0)]^T\mathbf{B}_0^{-1}[\mathbf{x}(t_0) - \mathbf{x^b}(t_0)] \\ &+ \frac{1}{2}\sum_{i=0}^{N}[H(\mathbf{x}_i) - \mathbf{y}_i^o]^T\mathbf{R}_i^{-1}[H(\mathbf{x}_i) - \mathbf{y}_i^o] \end{aligned} \tag{2.40}$$

Here, the total number of observations (N) over a time window, say 6h, shows up as the second term in RHS with i varying from 0 to N. Thus, for 4DVar, the cost function essentially measures the distance between the model trajectory and the observations as well as the first guess at the beginning of the assimilation time window ($\mathbf{x}(t_0)$) [16, 20, 32]. Thus, it demands by definition that the sequence of model states $\mathbf{x}_i$ should be a solution of the (non-linear) forward model M. That is,

$$\forall i, \mathbf{x}_i = M_{0\to i}(\mathbf{x})$$

With the causality hypothesis, one can express the non-linear forward model as a sequence of products of intermediate steps [116].

Therefore, the above expression can be written as:

$$\mathbf{x}_i = M_i M_{i-1} ... \, M_1 \mathbf{x} \tag{2.41}$$

Moreover, the non-linear operator M can be linearized together with H by applying the tangent-linear hypothesis, yielding,

$$\mathbf{y}_i - H_i M_{0\to i}(\mathbf{x}) \approx \mathbf{y}_i - H_i M_{0\to i}(\mathbf{x_b}) - \mathbf{H}_i \mathbf{M}_{0\to i}(\mathbf{x} - \mathbf{x_b}) \tag{2.42}$$

where, $\mathbf{M}$ is the *tangent linear model*, which is essentially the differential of the forward model M. The 4DVar minimization problem becomes quadratic in nature with the help of above two hypotheses [116, 9]. The first part of the cost function, J_b is basically similar to the 3DVar minimization problem; however, the second part J_o demands n integrations of the forward model from the analysis time to the observations times i.

Following Kalnay (2003), the gradient of the cost function for the 4DVar minimization can be obtained as:

$$\left[\frac{\partial J_o}{\partial \mathbf{x}(t_0)}\right] = \sum_{i=0}^{N} \mathbf{M}(t_i, t_0)^T \mathbf{H}_i^T R_i^{-1} \left[H(\mathbf{x_i}) - \mathbf{y}_i^o\right] \tag{2.43}$$

where, M^T is the *adjoint operator*, which is the transpose of the tangent linear operator. Therefore, the minimization of 4DVar cost function involves a direct model integration of the tangent linear model and a backward integration of the adjoint model to the beginning of the assimilation window [116, 9]. This forward marching of the tangent linear model and the backward integration of the adjoint model in each iteration during the minimization process lead to evolution of model errors in an implicit way, which is a salient feature of the 4DVar procedure. Moreover, as the model dynamics is involved in the minimization procedure, the solution may achieve better dynamical balance. However, major limitations of the 4DVar method are the complexity in implementation, especially, the modeling of the adjoint operator and the huge computational cost.

The 4DVar method has been implemented in the WRF variational assimilation system in 2009 [49]. It follows the incremental formulation [31, 50, 51]. The evolution of the background state is predicted utilizing the full-physics of the non-linear model, whereas the inner-loop integrations are performed using the tangent linear/adjoint model, which is a simplified version of the actual forward model.

The tangent linear/adjoint model in WRF

The tangent linear/adjoint model in the WRF variational system (WRFPLUS) is developed following the procedure proposed by Xiao *et al.* in 2008 [117]. The formulation of WRFPLUS makes use of a source-to-source automatic differentiation tool, TAPANADE [118, 50]. The linear model takes care of major physical processes that have more critical impact on the forecasts, following the approach of Xiao et al [117]. The cumulus convection, surface drag, and microphysics processes are being incorporated in the WRFPLUS in a simplified form. In the cumulus convection scheme, it is assumed that the parcels having buoyant updrafts will be associated with a cloud-top with undiluted ascent to the equilibrium level. Condensates from all the parcels will be summed down as rainfall and the latter may be allowed to evaporate below the cloud base. The surface drag is approximated with a vertical diffusion code. The surface stress is computed with constant drag coefficients. The simplified microphysics scheme is based on simple large-scale condensation approach. The water vapor is allowed to condense when it reaches a threshold level, which then contributes to latent heat release. As an alternate to the condensation scheme, a simplified Kessler microphysics is also implemented in the linear model.

2.2.3 Modeling of errors

Both the model state and the observation have associated uncertainties. The errors in observation and the model are usually represented by their respective probability density function, from which one can compute average and variance. In variational assimilation procedure, the error statistics of the model and the observations are generally prescribed in the form of square matrices called background error covariance matrix (**B**) and observation error covariance matrix (**R**).

The observation error covariance matrix is given by:

$$\mathbf{R} = \overline{(\varepsilon_\mathbf{o} - \bar{\varepsilon_\mathbf{o}})(\varepsilon_\mathbf{o} - \bar{\varepsilon_\mathbf{o}})^\mathbf{T}} \qquad \text{where, } \varepsilon_\mathbf{o} = \mathbf{y} - H(\mathbf{x_t})$$

Observational errors arise from instrumental error, representative error and the errors associated with the observational operator. If $\mathbf{x_b}$ is the background state and $\mathbf{x_t}$ is the true state, the background error for a scalar system can be expressed as $\varepsilon_\mathbf{b} = \mathbf{x_b} - \mathbf{x_t}$. Then **B** will take the following form[116]:

$$\mathbf{B} = \overline{(\varepsilon_\mathbf{b} - \bar{\varepsilon_\mathbf{b}})(\varepsilon_\mathbf{b} - \bar{\varepsilon_\mathbf{b}})^\mathbf{T}}$$

If the total number of degrees of freedom of the NWP model is n, the size of **B** would be $n \times n$. In atmospheric models, n is of the order of 10^7 and hence direct calculation of **B** is not feasible. Thus, the **B** matrix is approximated by employing modeling techniques. One method of estimating BEC is based on analyzing the 'innovation', that is, the observation minus background fields [119]. In the above method, the forecast errors of well-observed quantities are estimated by determining the corresponding innovation fields, which can then be related to background errors. Though this method was found to be successful, its performance was much dependent on the number of *in situ* observations. A widely adopted method for the estimation of **B** matrix is the National Meteorological Center (NMC) method [15]. In this procedure, the model errors are approximated as the difference between two different forecasts that are valid at same time, typically with 12h and 24h forecast lead time. Short-range forecasts for one month will be used for estimating the average error statistics. As the **B** matrix yielding from this particular method would be homogeneous in nature, it will not provide for any kind of flow dependency. Another popular method for estimating the BEC is the ensemble perturbation method [120]. In the ensemble method, the ensemble mean will be used as the 'truth' and covariance statistics will be derived from the difference of the each member from the ensemble mean. The ensemble method is found to yield improved representation of the BEC matrix as the same does not have the problems related to the regions of less observations [73]. However, the above method is computationally more costlier than the NMC method. The NMC method has been employed for the estimation of **B** matrix in all the data assimilation experiments performed in the present thesis. The BEC matrices for respective experiments have been generated by considering 60 samples for the corresponding season of the meteorological phenomena studied.

2.3 Data used

Various data have been utilized in the present study for model initialization, assimilation, forecast verification etc. A brief mention of all the data sets used in the present study is given below.

NCEP-GFS forecast fields: All the numerical experiments conducted in the present thesis use 24h forecast fields of NCEP Global Forecasting System (GFS) as initial and boundary conditions for the WRF model. The above forecast fields are available at half-a-degree horizontal resolution and have 27 vertical levels. Data source: `ftp://nomads.ncdc.noaa.gov/GFS/Grid4/`

Surface and upper-air observations: The Global Telecommunication System (GTS) provides surface and upper-air observations from conventional platforms such as radiosonde, pilot balloon, buoy, ship, ground stations etc. All the conventional observations are utilized for all the data assimilation experiments in the thesis. Data source: `https://rda.ucar.edu/datasets/ds337.0/index.html#sfol-wl-/data/ds337.0`

Atmospheric motion vectors: Meteosat-7, on board Metop provides wind observations over the Indian region as atmospheric motion vectors. Being on a Geostationary platform, Meteosat-7 provides continuous monitoring of wind fields over the Indian domain. The atmospheric motion vectors derived from Meteosat-7 are assimilated in all the experiments performed in this study. Data source: `https://rda.ucar.edu/datasets/ds337.0/index.html#sfol-wl-/data/ds337.0`

ASCAT winds: Ocean surface winds from Advanced scatterometer (ASCAT) are utilized in the study. ASCAT is on board the European weather satellite on board Metop. It operates at C-band (at 5.255 GHz) and uses vertically polarized fan beam antenna. The satellite is at an altitude of 840 km. It carries 2 swaths of 500 km width and provides ocean surface wind observations at a spatial resolution of 0.25°. Data source: `https://rda.ucar.edu/datasets/ds337.0/index.html#sfol-wl-/data/ds337.0`

Scatsat-1 winds: Scatsat-1 is an Indian weather satellite, that carries a scatterometer operating at Ku-band (13.52 GHz). This was launched as the successor to Oceansat-2. It orbits at an altitude of 720 km and carries a pencil beam antenna. Scatsat-1 has a swath width of 1400 km and provides ocean surface winds at a spatial resolution of 0.25°. Data source: `ftp://scatsat1@ftp.mosdac.gov.in/`

AMSU radiance: The Advanced Microwave Sounding Unit (AMSU) A and B are widely recognized instruments, on board NOAA polar orbiting satellites, that provide for vertical temperature and humidity profiles. AMSU-A has 15 channels between 23.8 GHz and 89 GHz and provides for temperature profiles from 3 hPa to surface at a horizontal resolution of 48 km at nadir. The AMSU-B provides humidity profiles at a horizontal resolution of 15 km at nadir. Out of the five microwave channels, three are at 183 GHz water vapor line and they provide the moisture profiles within troposphere and lower stratosphere. AMSU-A radiance from MetOp-B, NOAA-15, NOAA-16, and NOAA-18 and AMSU-B radiance from NOAA-15, NOAA-16, and NOAA-17 are used in the study. Data source: `https://rda.ucar.edu/datasets/ds735.0/index.html#sfol-wl-/data/ds735.0?g=1`

MHS radiance: The Microwave Humidity Sounder (MHS) has five channels between 89 GHz and 190 GHz. It is a cross-track line scanned instrument. MHS instruments are on board NOAA-18 and NOAA-19 satellites. In conjunction with AMSU-A, MHS also provides for emissivity and the surface temperature. MHS is similar to AMSU-B in its design and both have almost identical field of view with a horizontal resolution of 16 km at nadir. MHS radiance from NOAA-18 and NOAA-19 are utilized in the present study. Data source: `https://rda.ucar.edu/datasets/ds735.0/index.html#sfol-wl-/data/ds735.0?g=5`

HIRS radiance: The High-resolution Infrared Sounder (HIRS-4) operates at the infrared band, which is primarily meant to provide for temperature and moisture profiles. HIRS-4 has 20 spectral bands comprising of: one visible channel and 19 infrared channels. With a field of view of about 1.3° near 833 km altitude, HIRS gives temperature and humidity information at a horizontal resolution of 20 km at nadir. HIRS radiance from NOAA-17 and NOAA-18 are utilized in the study. Data source: `https://rda.ucar.edu/datasets/ds735.0/index.html#sfol-wl-/data/ds735.0?g=4`

TRMM rainfall: The Tropical Rainfall Measurement Mission (TRMM) is a collaborative project of National Aeronautics and Space Administration (NASA) and Japan Aerospace Exploration Agency (JAXA) that was launched in 1997. TRMM-3B42 algorithm provides for rainfall estimates over the tropical belt (35°S to 35°N) around the glob by combining the microwave and infrared rainfall information [121]. It also provides 3-hourly high-resolution (0.25° X 0.25°) rainfall estimates [122]. The TRMM rainfall estimates are used in this study for validating model precipitation forecasts. Data source: `http://apdrc.soest.hawaii.edu/las/v6/constrain?var=13166`

2.4 Forecast verification methods

Model forecasts for all the numerical experiments have been validated with reference to corresponding observations. The metrics utilized for quantitative verification of model forecasts are briefly described here.

- **Root Mean Square Error (RMSE):** For any parameter x, whose forecast fields and observed fields are x^o and x^f respectively, the RMSE for N samples is given by,

$$RMSE(x) = \left[\frac{1}{N} \sum_{i=1}^{N} (x_i^f - x_i^o)^2 \right]^{\frac{1}{2}} \tag{2.44}$$

- **Improvement Parameter (IP):** It is essentially the difference in RMSE between a control run (C) and experiment run (E) with respect to observation (O). IP is given by,

$$IP = \left[\frac{1}{N} \sum_{i=1}^{N} \left(O_i - C_i\right)^2 \right]^{\frac{1}{2}} - \left[\frac{1}{N} \sum_{i=1}^{N} \left(O_i - E_i\right)^2 \right]^{\frac{1}{2}} \tag{2.45}$$

  As evident from the expression, positive values of IP indicate that the RMSE of the control run dominates over the RMSE of the experiment run and hence experiment run is more closer to observations.

- **Skill scores for rainfall forecast:** Statistical skill scores are utilized for quantitative verification of rainfall forecast with respect to TRMM rainfall estimates. *Equitable threat score* (ETS) is a measure of forecast skill excluding the events that happened merely by chance. The *bias score* indicates whether the model shows any tendency to overestimate or underestimate the observed rainfall. The *probability of detection* (POD) gives an estimate of model's skill in capturing the actually observed rainfall events. The fraction of false predictions, that is, the events where the model simulated rainfall is not observed, by the model is measured by the *false alarm ratio* (FAR). All the quantitative scores can be computed by preparing a *contingency table* for the rainfall event as shown in Table. 2.1.

Table 2.1: Contingency table for calculating the statistical skill scores.

	Observed Yes	**Observed No**
Forecast Yes	hits	false alarms
Forecast No	misses	correct negatives

$$hits_r = \frac{(hits + misses)(hits + false\ alarms)}{(hits + misses + false\ alarms + correct\ negatives)} \tag{2.46}$$

$$ETS = \frac{(hits - hits_r)}{hits + misses + false\ alarms - hits_r} \tag{2.47}$$

$$Bias = \frac{(hits + false\ alarms)}{(hits + misses)} \tag{2.48}$$

$$FAR = \frac{false\ alarms}{hits + false\ alarms} \tag{2.49}$$

$$POD = \frac{hits}{(hits + misses)} \tag{2.50}$$

Chapter 3

Performance of the 4DVar DA System on Tropical Cyclones: A Preliminary Study

The performance of the 4DVar and 3DVar DA systems has been examined for two TCs - cyclone Thane and Hudhud - that formed over the Bay of Bengal region. Cyclic 3DVar and 4DVar assimilations were performed utilizing the conventional surface and upper-air observations together with the satellite derived wind information. The 4DVar experiments are found to yield improved track and intensity simulations for the first 24h of free forecast.

3.1 Introduction

Accuracy of the numerical weather forecasts have improved remarkably in the last few decades. Enhancing the skill of the NWP model for simulating the intensity and movement of tropical cyclones is still a challenging task, owing to the uncertainties in the initial condition. With the advent of advanced remote sensing observational systems, oceans are now better observed with satellites and this has contributed to much improved initial conditions for the NWP models. Several studies have been conducted across the world[123, 57, 124, 58, 125, 66, 126, 69, 70] that have demonstrated the positive impact of assimilating both conventional and satellite observations on the simulation of tropical cyclones. All the studies mentioned above however have utilized the 3DVar DA framework. For severe weather systems such as tropical cyclones that intensify very rapidly in time, it is desirable to employ more advanced and improved DA system like 4DVar that allows for observations to be assimilated at the exact time of observation. Moreover, the implicit flow-dependence inherent with the 4DVar DA system may contribute to better representation of the error statistics of the forward model, that can in turn result in improved simulation of the tropical cyclone features[31, 17, 35, 127, 128, 129, 115]. Huang *et al.* [49] implemented the 4DVar DA system in the WRF model in 2009 and carried out preliminary studies for two tropical storms. Even though there are numerous studies that have investigated the impact of 3DVar DA method for tropical cyclones over the Indian Ocean basins, similar studies that have employed the 4DVar DA methods are very limited. Since the evolution of model errors, even in a 4DVar DA system, will be sensitive to the flow regimes, the impact of 4DVar DA methods may differ for different domains in the regional modeling framework.

This chapter presents the results of a preliminary study that examined the relative performance of the 3DVar and 4DVar DA systems within the WRF model in the simulation of two tropical cyclones, namely, cyclone Thane and cyclone Hudhud, that formed over the Bay of Bengal.

3.2 Case description

Cyclone Thane

On 25 December 2011, a depression formed over the southeast Bay of Bengal. While moving in the north-northwestwards direction, the system developed to a deep depression and subsequently intensified to a cyclonic storm by the midnight of 26 December, 2011. The storm then moved in the west-northwest direction and developed as severe cyclonic storm by the evening of 28 December, 2011. The cyclone experienced landfall between 00Z and 03Z on 30 December, 2011 near Tamil Nadu and Puducheri coast. The system then weakened to a severe cyclonic storm and subsequently to a deep depression by evening of the same day. The system continued to weaken and was finally observed as a well marked low pressure area near the northern Kerala on 31 December, 2011.

Cyclone Hudhud

The very severe cyclonic storm Hudhud formed during October, 2014, near North Andaman Sea, from a low pressure area. The system gradually strengthened to a depression on 07 October, 2014 and moved in the west-northwest direction. The depression further developed as a cyclonic storm and crossed the Andaman Islands on 08 October, 2014.

While continuing to move in the west-northwest direction, the storm intensified first as a severe cyclonic storm and then as a very severe cyclonic storm and reached its maximum intensity by 12 October, 2014 morning. The storm continued to move towards the Andhra Pradesh coast and experienced landfall near Vishakhapatnam between 06Z and 08Z of 12 October, 2014. The cyclone continued to move northwestwards and started to weaken. By same day evening the cyclonic storm weakened to deep depression. Subsequently, the system weakened to a depression on 13 October, 2014. Moving northward, the system weakened gradually to a low pressure area over Uttar Pradesh on 14 October, 2014 evening.

3.3 Model configuration

Numerical experiments were carried out with the WRF model and its DA modules (version 3.6.1) by employing two nested domains (with two way nesting) as shown in Fig. 3.1. The parent (outer) domain and the inner domain were configured with 27 km and 9 km horizontal resolutions, respectively. The parent domain was having 350 grid cells in both east-west and north-south directions. Both the domains have 36 levels in vertical direction (up to 50 hPa).

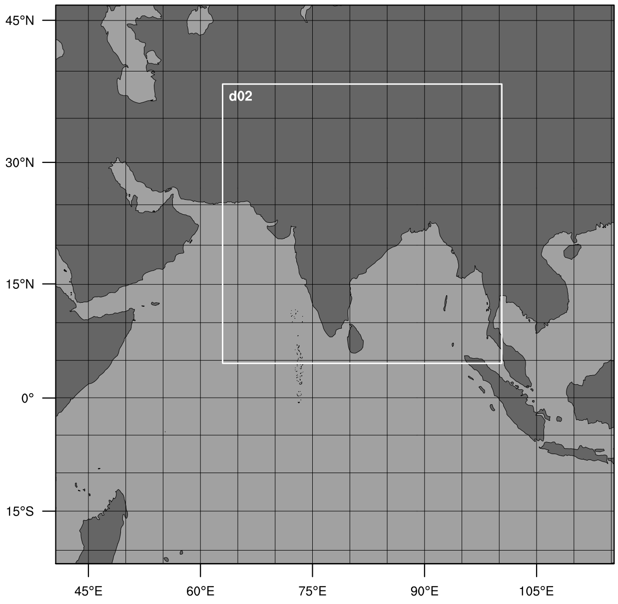

Fig. 3.1: The model domain used in the study.

The convective processes were represented using Kain-Fritsch scheme, which accounts for deep and shallow convection using mass flux approach. The boundary layer turbulence was specified following the Yonsei University scheme, while the microphysical processes follow the WRF single moment five-class scheme, that allows for the mixed-phase processes and super-cooled water. The Noah land surface model provides for the land surface physics. The radiative processes are parameterized following the rapid radiative transfer model and Dudhia scheme for longwave and shortwave radiations respectively. More details on the model configurations are given in Appendix A.

3.4 Experimental design

Cyclic assimilation runs were designed for both the TCs utilizing the 3DVar and 4DVar DA methods (WRFDA v3.6.1). Surface and upper air observations from conventional platforms were assimilated in combination with satellite derived wind information. A 6h long assimilation window was chosen for both the 3DVar and 4DVar experiments: t-3h < t < t+3h for 3DVar run and t < t+6h for the 4DVar run (tindicates the analysis time). Total 11 and 13 assimilation cycles were performed for cyclone Thane and Hudhud respectively. A schematic of the experimental design in given in Fig. 3.2.

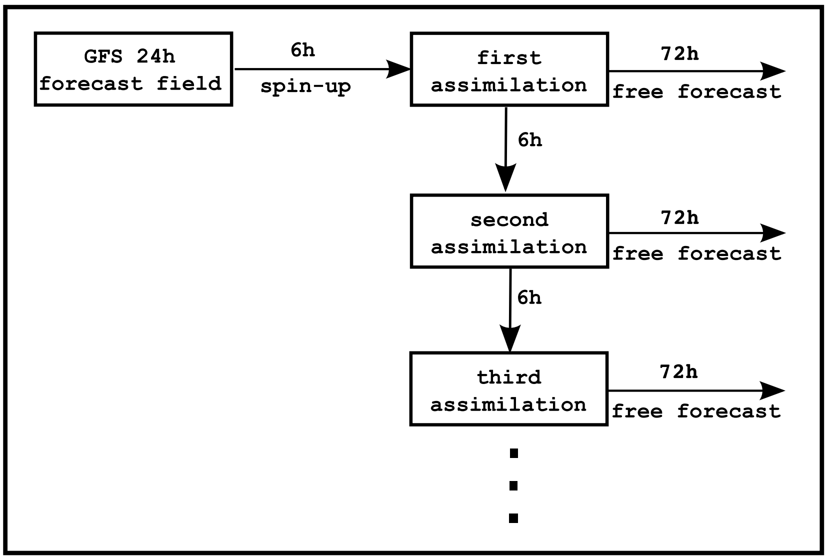

Fig. 3.2: Illustration of the experimental design.

In the case of cyclone Thane, after a spin-up run for 6h, 11 cyclic assimilations were carried out during the period 12Z 25 to 00Z 28 December 2011. A free forecast for 72h was initiated from each of the analysis fields available at 6h interval. Similarly for cyclone Hudhud, 13 cyclic assimilations were performed during 00Z 08 to 00Z 11 October 2014, after a short (6h) spin-up run. Also, short-range forecast for 72h was launched from each of the analysis fields. NCEP-GFS 24h forecast fields at the horizontal resolution of 0.5°×0.5°were used as initial and boundary conditions. The distribution of various observations that were assimilated are given in Fig. 3.3 and typical number of observations used for assimilation at 0Z 08 October 2014 are indicated in Table 3.1.

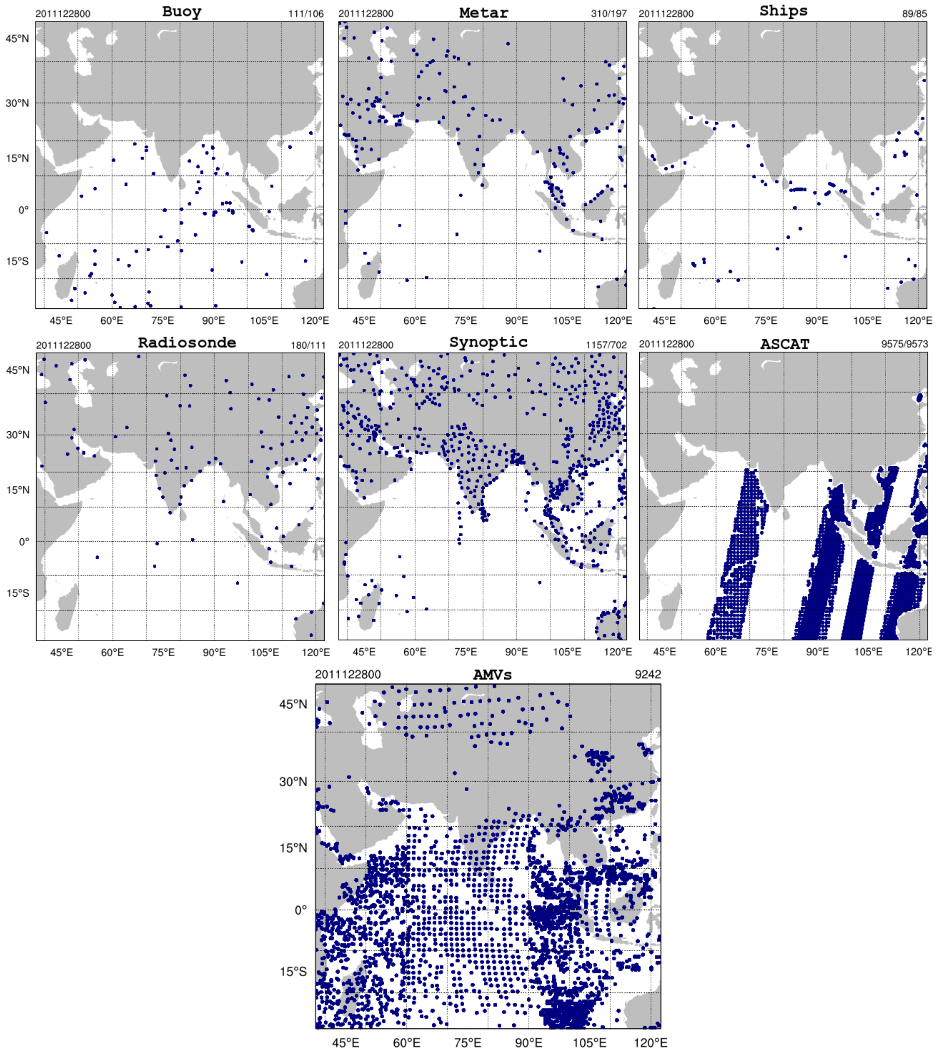

Fig. 3.3: Distribution of various observations over the experimental domain that were assimilated on 00Z 28 December 2011.

Table 3.1: Typical number of observations used for assimilation at 0Z 08 October 2014.

Observation type	Number of observations used
Synoptic	1060
Soundings	9221
Metar report	218
Pilot	207
Buoy	24
Ship	80
AMVs	11065
Scatterometer	10970

3.5 Results and discussion

3.5.1 Improvement in the analysis fields

Improvement in the analysis of ocean surface winds obtained after employing the 3DVar and 4DVar DA systems at the first analysis time were examined by comparing the analysis results with the observed ocean surface winds from Advanced Scatterometer (ASCAT) for both the TCs Thane and Hudhud and are shown in Fig. 3.4.

On an average, both the experiments have simulated the near-surface winds over the ocean as seen in the ASCAT observations. However, a close examination would reveal notable differences in the wind speed between the two experiments. For instance, for TC Thane, the west coast of Kerala is characterized with weaker surface winds in the 3DVar analysis field (Fig.3.4 a), whereas, both ASCAT observation (Fig.3.4 c) and 4DVar analysis field (Fig.3.4 b) do not show the above-mentioned pattern. Similarly, a zone of weaker wind is seen over the eastern side of the 3DVar simulated vortex of cyclone Thane. The ASCAT winds however suggests that the above-mentioned feature is not observed. The 4DVar run for cyclone Thane simulates stronger winds over the central and eastern Bay of Bengal, which is at variance with the ASCAT observations. In the case of cyclone Hudhud, there are no notable differences between the 3DVar and 4DVar analyzed surface winds (Fig.3.4 d and e). However, when compared with the ASCAT observed winds (Fig.3.4 f), both the experiments show significant overestimation in the wind speed near the vortex.

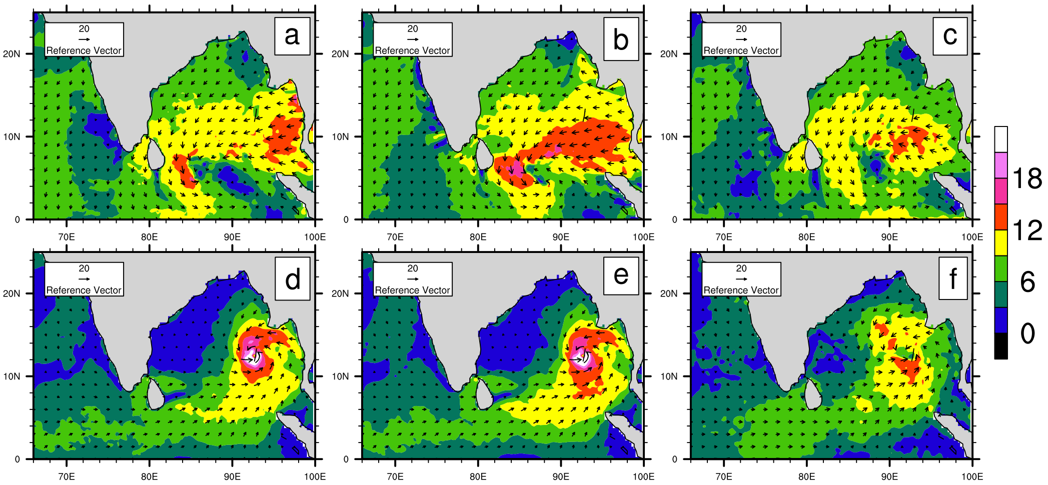

Fig. 3.4: Ocean surface winds (at 10 m height) for initial analysis fields for 3DVar run (a, d), 4DVar run (b, e) and ASCAT observations (c, f) TC Thane [(a-c) valid at 12Z 25 December 2011] and for TC Hudhud [(d-f) valid at 00Z 08 October 2014].

A quantitative measure on the impact of both the assimilation methods on the ocean surface winds can be obtained by estimating the Improvement Parameter (IP) for surface winds, which is defined as in Eqn. 2.45. Here, IP is being computed with respect to the ASCAT observations and is shown if Fig. 3.5.

Since the IP is calculated by considering all the analysis samples available for both the cyclones, it would provide for more robust quantitative impacts of the 4DVar DA system. From Eq. 2.45, it is evident that positive values of IP indicate greater departure for the 3DVar fields (as compared to the 4DVar fields) from the ASCAT observations and hence provides a quantitative improvement for the 4DVar DA system. Fig. 3.5 depicts IP for zonal (a,c) and meridional (b,d) wind components for cyclone Thane (a,b) and cyclone Hudhud (c,d). In the case of cyclone Thane, a clear positive impact due to the 4DVar assimilation can be observed, especially for the zonal wind component. However, for the cyclone Hudhud (Fig. 3.5 c), 3DVar zonal winds are found to be closer to the ASCAT winds over the Bay of Bengal region. On the other hand, the meridional wind component (Fig. 3.5 d) for cyclone Hudhud indicates that the 4DVar run impacted positively over most of the grid points. While analyzing the first analysis fields (Fig. 3.4), the results indicated that the positive impact due to the 4DVar run was not very significant. This suggests that the cyclic or repeated 4DVar assimilation will improve the analysis fields at later times. The classical adjustment theory indicates that the wind information is more crucial in describing

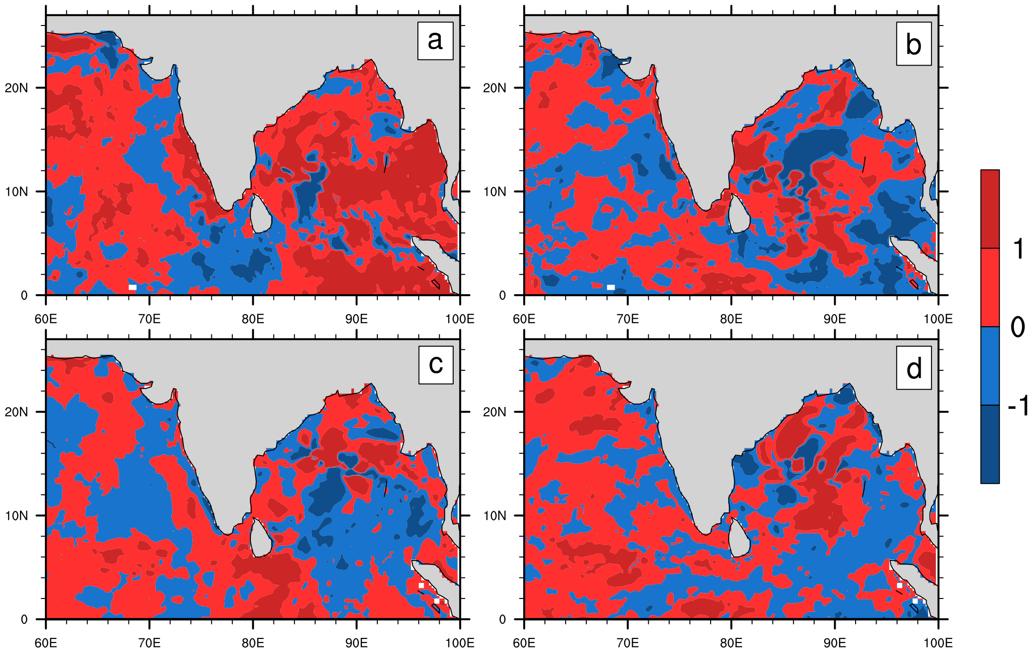

Fig. 3.5: IP for ocean surface winds estimated for zonal (a, c) and meridional (b, d) wind components for cyclone Thane (a, b) and cyclone Hudhud (c, d).

the evolution of tropical weather systems, particularly if the length scale is smaller than the Rossby radius of deformation [130, 131]. Hence, an improved wind fields associated with the 4DVar run would contribute to better simulation of cyclone features. Furthermore, a realistic representation of surface turbulent fluxes is highly desirable in achieving a better simulation of tropical cyclone features [132]. Since the 4DVar run accounts for more realistic representation of surface winds as compared to the 3DVar run, the former may contribute to improved simulation of tropical cyclone features through improved realization of surface turbulent fluxes.

3.5.2 Intensity of the simulated cyclones

The intensity of the simulated cyclones were examined by comparing the time evolution of the model minimum sea level pressure (MSLP) and model maximum surface wind speed (MWS) from the Joint Typhoon Warning Center (JTWC) best track data. Fig. 3.6 depicts the time evolution of average absolute error in MSLP (a) and MWS (b) for both the cyclones.

It can be observed that the initial error in MSLP simulation has reduced significantly with the 4DVar analysis, especially for cyclone Hudhud. To be specific, the initial MSLP error with the 3DVar run is 11.5 ms^{-1}, whereas, the initial MSLP error for the 4DVar analysis is only 4.5 ms^{-1}. Up to 42h, the 4DVar run for cyclone Hudhud indicated improved MSLP simulation, however, beyond 42h of free forecast, the improvement in the 4DVar results are not observed. For cyclone Thane, even though the reduction in initial MSLP error with the 4DVar run is not as marked as with the cyclone Hudhud, the same is associated with better MSLP simulation up to 24h free forecast, beyond which the 3DVar run shows lower MSLP errors. Towards the end of free forecast (>54h), a steady increase in the MSLP error is observed for cyclone Thane in both the 3DVar and 4DVar runs.

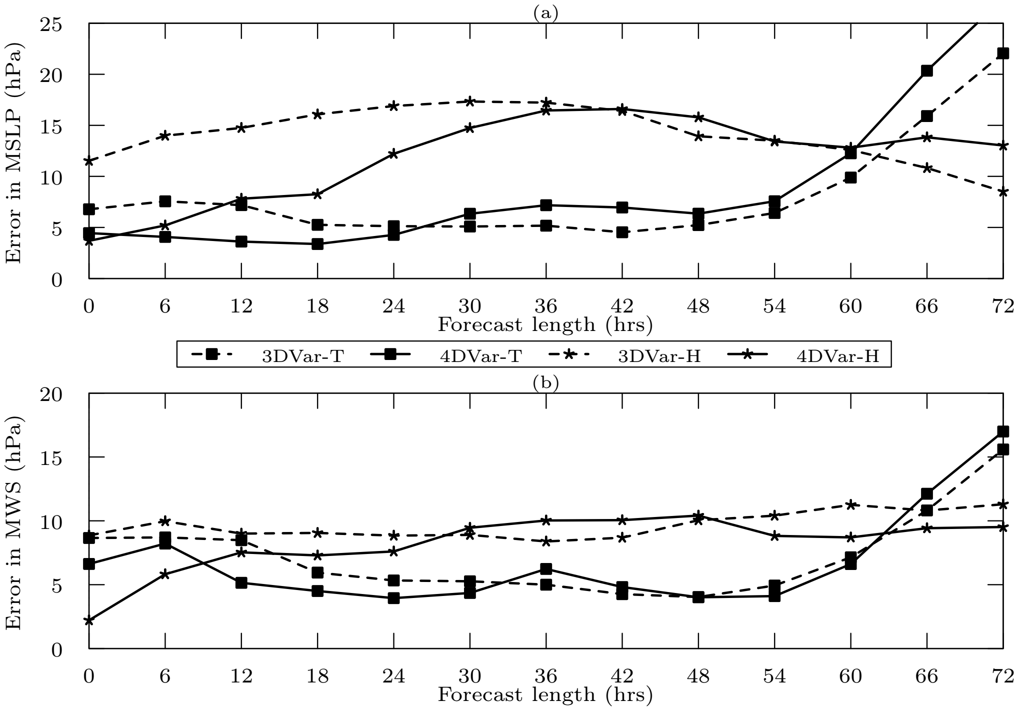

Fig. 3.6: Time series of average absolute error in MSLP (a) and MWS (b) for cyclone Thane and cyclone Hudhud. 'T'and 'H'in legends stand for 'Thane'and 'Hudhud'respectively.

Almost similar inferences can be drawn from the analysis of the intensity simulation in terms of MWS (Fig. 3.6 b). For cyclone Hudhud, the average MWS error of the 4DVar run is $\sim$60% lower as compared with its 3DVar counterpart. In the first 28h of free forecast for cyclone Hudhud, the 4DVar run has lower MWS error and towards the end of the free forecast, the errors are nearly the same for both DA systems. Likewise, for cyclone Thane also, the average MWS errors are on the lower side for the 4DVar run (up to 32h). Consistent with the MSLP simulation, the average MWS errors also show a steady rise beyond 54h free forecast for cyclone Thane for both DA systems. The skill of the model in the intensity simulation for both the cyclones in terms of MWS are compared with the operational forecasts from India Meteorological Department (IMD). IMD forecasts show MWS error of 5.4 to 8.9 ms^{-1} for forecasts ranging from 12 to 48h for cyclone Thane. The 4DVar experiments from the present study show lower MWS errors, ranging from 4.0 to 7.5 ms^{-1} for forecasts up to 48h. In the case of cyclone Hudhud, results from the 4DVar experiments are found to have comparable MWS errors with reference to IMD forecasts up to 24h (4-11 ms^{-1}). However, towards the end of the forecasts, the MWS errors associated with the 4DVar run are found to be slightly higher. Nevertheless, the results suggest that the intensity of the TCs can be simulated more accurately when the model is initialized with 4DVar analysis fields, particularly for the first 24h of forecast.

3.5.3 Rainfall simulation

The skill of the model in simulating the rainfall associated with the TCs is analyzed by examining the 24h accumulated rainfall with the Tropical Rainfall Measurement Mission (TRMM) observed rainfall at the landfall phase for both the cyclones and are depicted in Fig. 3.7.

Fig. 3.7 confirms that the model has captured the general pattern of rainfall reasonably well as seen from the TRMM observations. However, in general the model shows a notable degree of overestimation in simulating the rainfall. For cyclone Thane, 24h accumulated rainfall valid at 06Z 30 December 2011 is depicted in Fig. 3.7 (a-c). It can be noted that the rainfall simulated with the 3DVar run shows significant overestimation and the location of maximum rainfall is away from the observed location. On the other hand, the 4DVar run has successfully suppressed the overestimation of 3DVar run rainfall to a great extent. Moreover, the location of maximum rainfall of the 4DVar run also matches with the location of maximum TRMM observed rainfall. The 24h accumulated rainfall for cyclone Hudhud, valid at 00Z 13 October 2014 is depicted in Fig. 3.7 (d-f). It is worthwhile to note that the 3DVar run simulated two distinct rainfall maxima, while, the results of the 4DVar run shows an intense rainfall pattern along the path of the cyclone. Both the above features are however not observed in the TRMM rainfall. The extended rainfall band on the eastern side

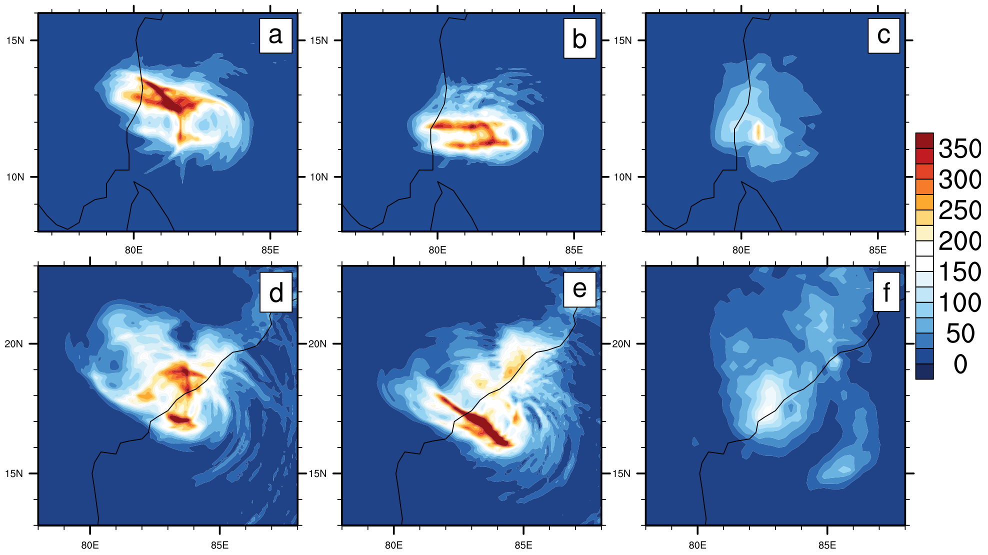

Fig. 3.7: 24h accumulated rainfall from 3DVar (a,d) and 4DVar (b,e) experiments and TRMM observation (c,f) at the landfall phase of cyclone Thane (a-c) and Hudhud (d-f).

of the vortex, as seen in the observed rainfall pattern has been simulated by the 4DVar run fairly well. The overall rainfall distribution during the landfall phase of both the cyclones are reproduced reasonably well with 4DVar run when compared to its 3DVar counterpart.

Further, the rainfall simulation with the 3DVar and 4DVar experiments were examined quantitatively by comparing the skill scores such as ETS, bias, FAR for both the model runs. Fig. 3.8 shows the time series of all the three skill scores for both the cyclones for both the experiments. The skill scores were computed for 72h accumulated rainfall for (i) 100 mm and (ii) 200 mm rainfall thresholds with respect to the TRMM observation.

The most striking observation from Fig. 3.8 is the general and gradual improvement in most of the skill scores for both the model runs with time. The above inference is mostly true for both the rainfall thresholds and for both the cyclones. For cyclone Thane (Fig. 3.8 a-c), the quantitative skill scores for both the 3DVar and 4DVar runs are nearly comparable during the initial hours. However, eventually with time, all the skill scores show marked improvement for the 4DVar run (for both the rainfall thresholds). As discussed earlier (Fig. 3.7), the higher overestimation with the 3DVar run is evident from higher bias score particularly for the 200 mm threshold. It can be noted that the bias score for the 4DVar run is fairly closer to unity, except during the final hours for the 200 mm threshold. The ETS for cyclone Thane also shows that the 4DVar run provides for significantly improved

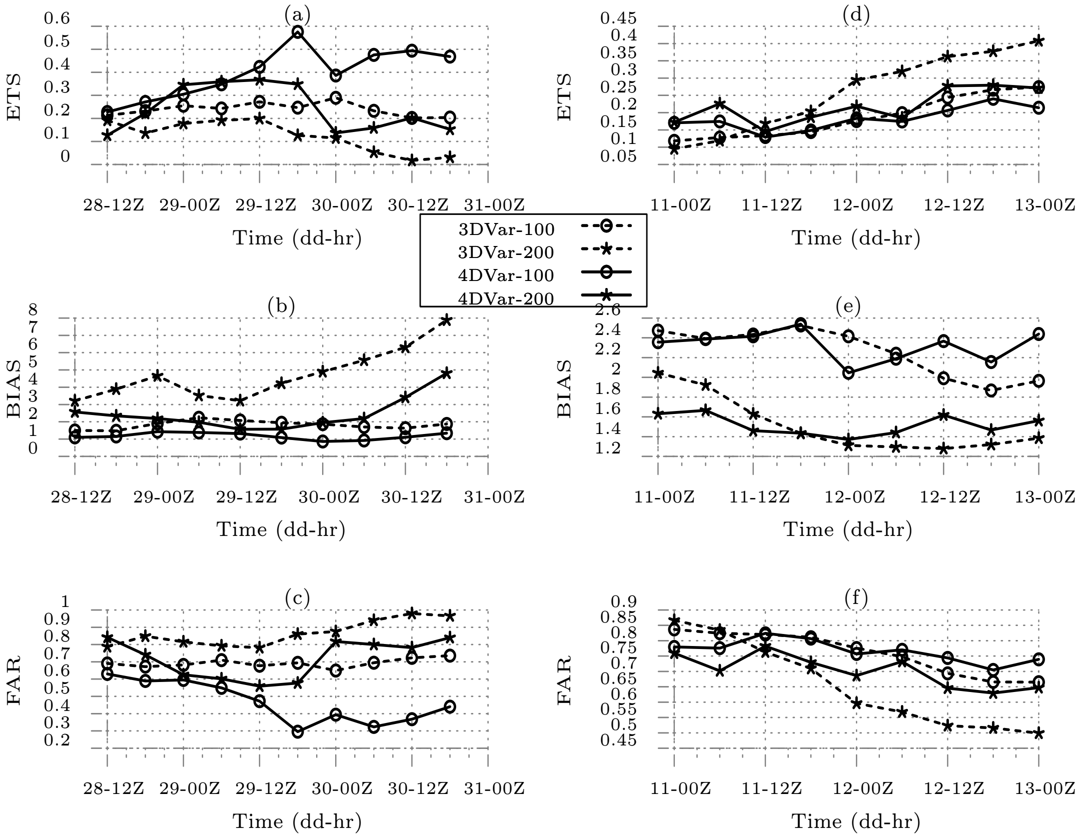

Fig. 3.8: Time series of statistical skill scores for 100 mm and 200 mm rainfall threshold for cyclone Thane (a-c) and cyclone Hudhud (d-f) based on 72h accumulated rainfall.

simulation of the rainfall as compared to the 3DVar run. Similar inferences can be drawn from FAR also, since the 4DVar run has significantly reduced the number of false alarms. The above feature is more evident for the lower rainfall threshold (100 mm).

In the case of cyclone Hudhud (Fig. 3.8 d-f) also, the rainfall forecast skill scores show improvement with time. However, the 4DVar run shows improved skill scores only for the initial hours, and the skill scores of the 3DVar runs are better at later times, particularly for the 200 mm rainfall threshold. The bias score confirms the overestimation in rainfall associated with the 4DVar run, which was already pointed out earlier (Fig. 3.7 e). Both ETS and FAR indicate a slight degradation in the rainfall forecast skill for the 4DVar run, which is discernible at and after the landfall and especially for the higher rainfall threshold. Since the skill scores are sensitive to the spatial distribution of rainfall, it would be interesting to examine the track simulation to get a clear idea about the lower ETS and higher FAR associated with the 4DVar run.

3.5.4 Track of the simulated cyclones

The model simulated TC track has been analyzed with reference to the JTWC observations. Error in simulating the vortex position after each analysis for both the experiments are indicated in Fig. 3.9. The most striking feature in the case of cyclone Thane (Fig. 3.9 a) is the systematic improvement in the simulation of the cyclone position with the number of assimilation. Furthermore, the 4DVar simulated vortex position is more closer to the JTWC observed position for all the cycles. The positional error for the 4DVar run after the first analysis was about 320 km and the same reduced drastically to ∼100 km after the third assimilation cycle and gradually reduced to ∼50 km with seventh cycle. The larger positional error in the first cycle may be attributed to the fact that the assimilation started one day before the storm started its development to the depression stage.

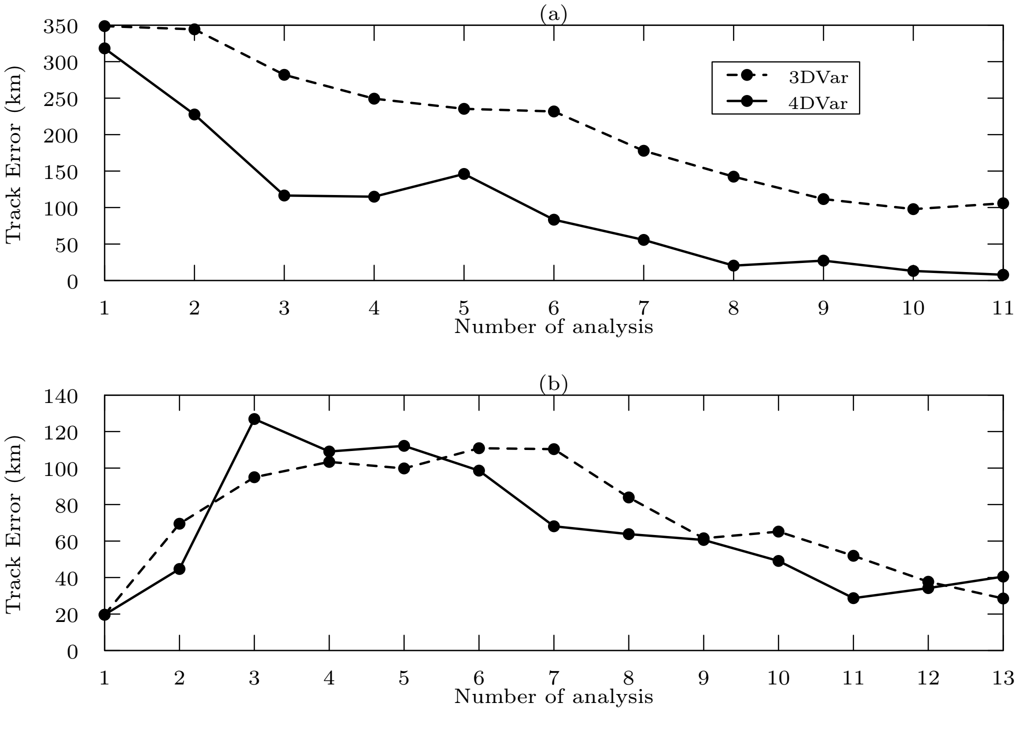

Fig. 3.9: Time series of vortex position error in the analysis fields with respect to JTWC observations for cyclone Thane (a) and cyclone Hudhud (b).

The error in the vortex position simulation from the analysis fields for cyclone Hudhud (Fig. 3.9 b), however, reveals slightly different behaviour from that of cyclone Thane. Even though, the positional error is found to increase for the first three analysis cycles, thereafter the positional error decreased with increase in analysis cycles. However, the improvement in reproducing the cyclone position for cyclone Hudhud with the 4DVar assimilation is not as marked as in the case of cyclone Thane, and the results are comparable for both experiments. Unlike for cyclone Thane, the magnitude of the positional errors for cyclone Hudhud is relatively lower, possibly because the assimilation started when the storm was already in deep depression stage.

The error in TC track forecast has been examined by estimating the average track error for all the forecast samples (with respect to the forecast length). Evolution of the average track error in time for cyclone Thane is shown in Fig. 3.10 (a). As discussed before, the track error at the initial time for the 4DVar run was about half of the average error associated with the 3DVar run. Up to 42h of free forecast, the 4DVar run showed lower track error, while at later times, both the experiments showed comparable track errors. The improved track forecast with the 4DVar experiment for cyclone Thane can be partly attributed to the fact that the 4DVar analysis could reproduce the initial position of the vortex more accurately than the 3DVar analysis.

The evolution of track error for cyclone Hudhud with respect to the forecast length is shown in Fig 3.10 (b). As seen from the Fig. 3.9 (b), the 4DVar run does not show any significant improvement in track forecast when compared to the 3DVar run for cyclone Hudhud. Instead, there is an apparent degradation in the 4DVar track forecast, which peaks at 42h of forecast lead time. The 4DVar run shows a maximum average track error of $\sim$150 km at 42h and 48h forecast length. Thereafter, the track error decreases to $\sim$100 km for the 4DVar run, whereas the track error for the 3DVar run increases from 50 km (42h) to 125 km (72h). Interestingly, the intensity errors for cyclone Hudhud also showed a notable degradation during 30h to 48h of the forecast (Fig 3.10 b).

The model simulated cyclone track from both the 3DVar and 4DVar runs are depicted in Fig. 3.11 together with corresponding best track (JTWC) data for TC Thane and TC Hudhud. For TC Thane, the 72h free forecasts initiated from 12Z 27 December 2011 are considered, while the model forecasts for TC Hudhud were initialized at 00Z 10 October 2014. As the results from Fig. 3.9 (a) indicated, the initial position of the vortex centre simulated by the 3DVar run is slightly away from the observed location of the vortex centre. On the other hand, the position of the vortex centre simulated by the 4DVar run at the

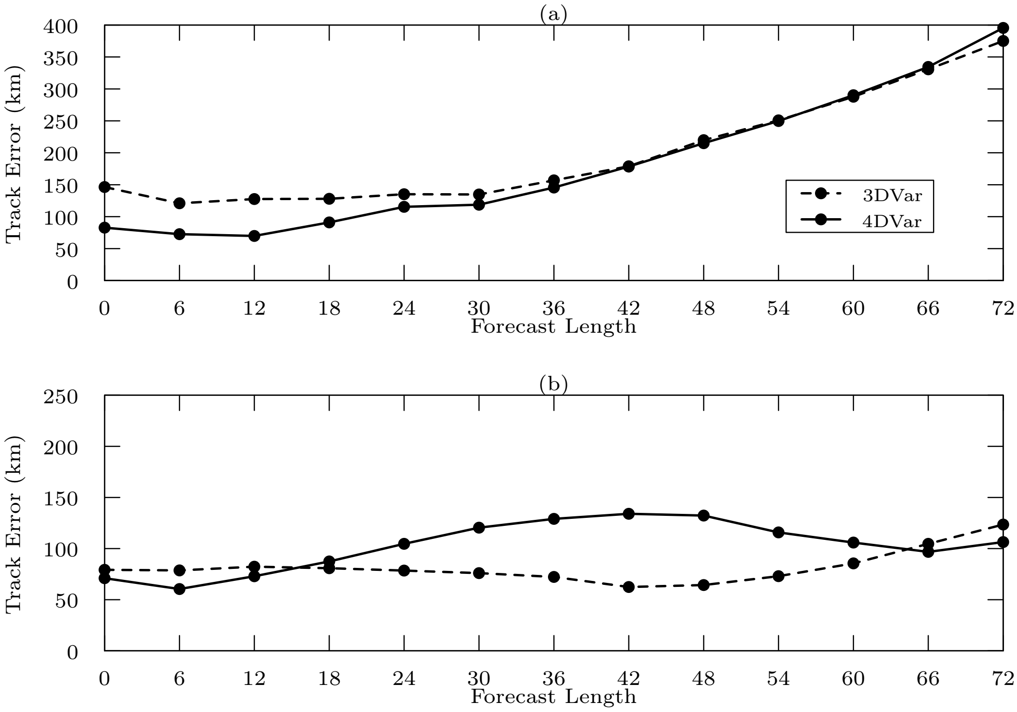

Fig. 3.10: Time series of average track error with respect to JTWC observations for cyclone Thane (a) and cyclone Hudhud (b).

beginning of the forecast is quite closer to the observed location. The 3DVar simulated cyclone moved in the north-west direction, at variance with the observed track of TC Thane, whereas the 4DVar run reproduced the observed track of TC Thane fairly well, and hence captured the location of landfall very closely. However, both the model runs show considerable differences in the speed at which the cyclone moved; the model simulated cyclones are moving faster as compared with the movement of the observed system.

It may be recalled that the track simulated by the 4DVar for TC Hudhud was associated with more error when compared with the corresponding 3DVar run, as evident from Fig. 3.9 (b). One can note from Fig. 3.11 that the position of the vortex centre simulated by the 4DVar run is relatively closer (as compared with the 3DVar run) to the observed position of the storm centre at the beginning of the forecast. However, the 3DVar run successfully reproduced the observed track of TC Hudhud, while the 4DVar simulated cyclone moved more northward after initial hours of the free-forecast and then recurved to the direction of the observed cyclone, after 42h of the free-forecast. It is interesting to note that, despite the 4DVar simulated cyclone moving in a direction at variance with

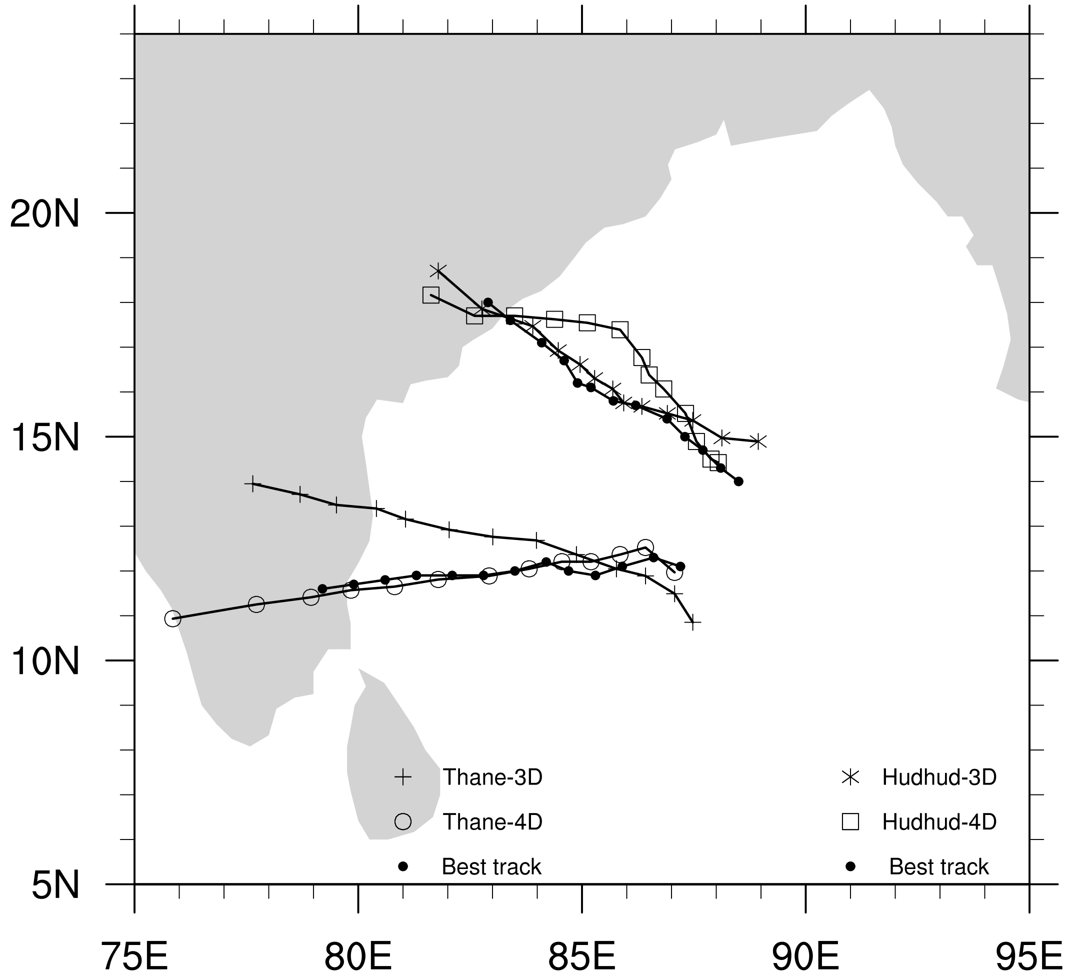

Fig. 3.11: Model simulated 72h tracks from 3DVar and 4DVar runs for TC Thane and TC Hudhud and corresponding best track observations. Model forecasts are initialized at 12Z 27 December 2011 for TC Thane and 00Z 10 October 2014 for TC Hudhud.

the track of the observed cyclone during 12h-42h of the free-forecast period, the system recurved to the observed direction and experienced landfall over a location very near to the observed landfall position. Results from a few forecast samples during mid-forecast period have been examined closely to provide a plausible explanation for the above discussed degradation for the 4DVar run in track simulation for TC Hudhud.

The "steering" flow for the cyclone Hudhud was analyzed by examining the upper level geopotential height and wind (averaged for 400-100 hPa) and are depicted in Fig. 3.12. Free forecasts at 42h and 66h that are initialized at 18Z 09 October 2014 are considered here, since the 4DVar run showed maximum track error of 128 km (28 km for 3DVar run),

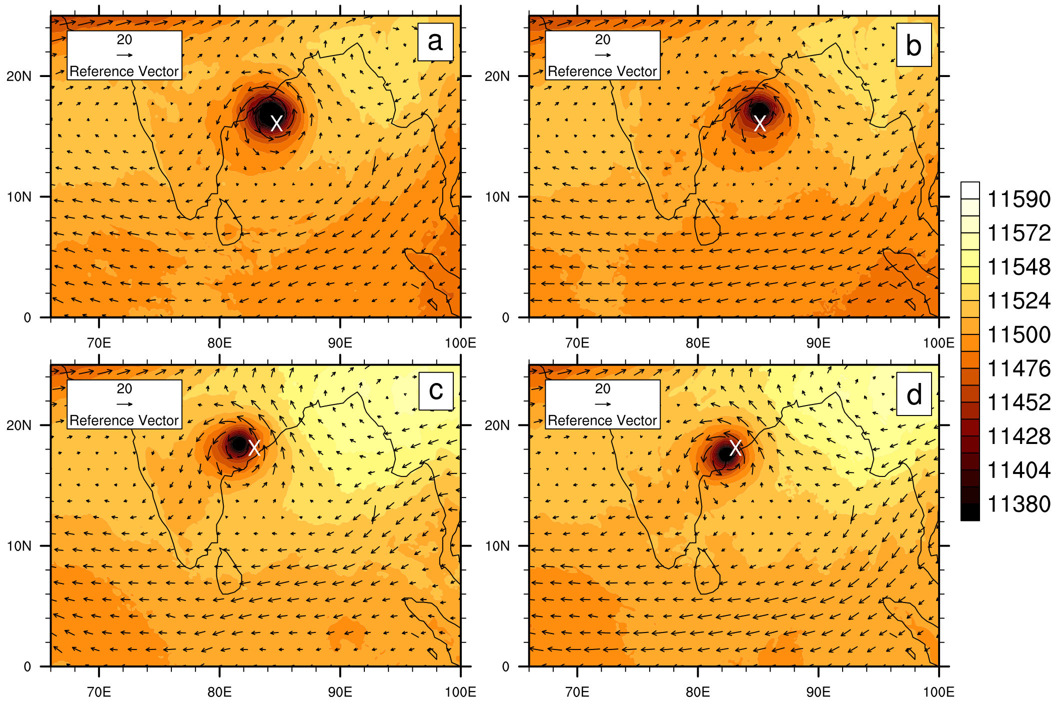

Fig. 3.12: Average geopotential height (in gpm) and wind vector (400-100 hPa) for cyclone Hudhud as revealed from 42h (a,b) and 66h (c,d) free forecast from initial condition valid at 18Z 09 October 2014 for 3DVar run (a,c) and 4DVar run (b,d). 'X' mark shows the observed position of the storm.

when the model was initialized with this particular initial condition. In Fig. 3.12, top panel (a and b) shows 42h forecasts and the bottom panel (c and d) shows 66h forecasts. It can be noted that the vortex simulated by the 4DVar run is weaker than that of the 3DVar run at 42h forecast. This early weakening of the vortex in the 4DVar run might have contributed to the higher intensity error during 36-48h forecast lead time. A closer examination also revealed that the vortex in the 42h forecast 4DVar field has moved slightly more northward when compared to the 3DVar forecast. In the 42h forecast field for the 4DVar run, the lower height fields are seen over greater extended regions over the southern Bay of Bengal and adjoining areas. As a result, stronger easterlies are observed in the 4DVar run over the same areas. In contrast with this, the 42h 3DVar forecast field shows slightly weaker northeasterlies over the south-eastern Bay of Bengal and adjoining areas. J. C. Chan in 1985 had reported that the cyclonic vortex may move right of the mean steering current in the northern hemisphere based on the results of his numerical studies [133]. Hence the anomalously stronger northeasterlies manifested in the 4DVar forecast could have contributed to the degradation of the observed vortex position. It was noted that the average track error for 4DVar run reached its peak at 48h of the forecast length and decreased thereafter. Hence, it would be interesting to analyze the forecast fields of the 4DVar run towards the end of the forecast length. Fig. 3.12 (c and d) illustrate the 66h forecasts from the 3DVar and 4DVar runs respectively. The prevailing winds over the southern Bay of Bengal is mostly northeasterlies in the 66h 4DVar forecast field and are similar to that of the 3DVar simulated winds except over the north-east sector of the vortex (over the head Bay of Bengal). While the 3DVar run simulated weaker southwesterly winds over the northern Bay of Bengal and adjoining area near the vortex, the 4DVar run simulated relatively stronger southwesterlies over the above area. Consequently, the position of the 3DVar simulated vortex was found to be more west of the observed position, in contrast, the 4DVar simulated cyclone was found to be more closer to the observed cyclone position. Accordingly the positional error at 66h of forecast was found to be 148 km for 3DVar run and 82 km for the 4DVar run. The above discussion suggests that the manifestation of stronger easterlies over the eastern side of Bay of Bengal by the 4DVar run (in contrast to the northeasterlies as in the 3DVar run) might have contributed to the degradation in track forecast with the 4DVar run. This has been further verified by analyzing the 66h forecast fields from both the model runs, which confirmed the change in the above easterlies to northeasterlies in the 4DVar run.

3.6 Summary

The present chapter investigated the performance of the 3DVar and 4DVar DA systems within the WRF model in the simulation of two TCs - cyclone Thane and cyclone Hudhud - through cyclic assimilation experiments. Assimilated observations include conventional surface and upper-air observations, AMVs from Meteosat and ASCAT ocean surface winds.

Results from the study suggest that the fields analyzed with 4DVar DA method can reproduce the initial structure of the cyclone vortex more realistically than the 3DVar DA method. The average intensity errors associated with the 4DVar run were found to be nearly half of that of the 3DVar run in the analysis fields (at 0h). Forecast fields for both the TCs clearly showed that the 4DVar runs can yield better simulation of TC intensity up to 24h of free forecast. The quantitative verification of rainfall forecast for TC Thane revealed significant improvement for the 4DVar experiment. For cyclone Thane, the 4DVar analysis fields were found to simulate the initial position of the vortex fairly close to the JTWC observed position. More importantly, the error in the initial position of the cyclone reduced significantly for the 4DVar run after each analysis cycle. Even though the initial position error for cyclone Hudhud also was lower for the 4DVar experiment, the improvement of the position error with time over the 3DVar run was not marked. While the 4DVar forecast fields for cyclone Thane showed an improvement in track forecast for the first 30h, the same for cyclone Hudhud did not show any significant improvement over the corresponding 3DVar run. The average track error for cyclone Hudhud indicated a slight degradation in track forecast for the 4DVar experiment as compared to the 3DVar experiment, which is being attributed to the manifestation of anomalous easterly winds simulated by the 4DVar experiment. Even though the 4DVar run showed slight degradation in the track forecast with few samples of TC Hudhud, the location of landfall was simulated reasonably well by the 4DVar run (as seen from Fig. 3.11).

The results from the preliminary study presented in this chapter that examined the relative performance of the 4DVar DA system in comparison to the 3DVar system yielded encouraging results that need to be pursued through more systematic studies over the Indian region involving additional satellite observations such as satellite radiances, which forms the results of next chapter.

Chapter 4

Quantifying the Improved Performance of the 4DVar DA System for Tropical Cyclones

The present chapter aims to examine and quantify the improved performance of the 4DVar DA system over the the 3DVar DA system on the simulation of four tropical cyclones that formed over the Bay of Bengal by generating a large number of analysis/forecast samples. On an average, an improvement of 22-57% in the simulation of cyclone track and 2-43% improvement in simulating cyclone intensity are observed for the 4DVar experiments.

4.1 Introduction

The initial study conducted in the previous chapter examined the relative performance of the 3DVar and 4DVar DA systems for two Indian Ocean TCs and yielded encouraging results. The 4DVar analysis fields were found to reproduce the position and intensity of the storms more accurately. The short-range forecasts, particularly the 24h track and intensity forecasts with the 4DVar experiments revealed clear positive impact.

The present chapter aims to provide a systematic analysis on the improved performance of the 4DVar DA system over the 3DVar system for the simulation of four TCs by generating a large number of analysis/forecast samples. In addition to TC Thane and TC Hudhud studied in the previous chapter, TC Phailin and TC Vardah are also considered in the present chapter. Apart from the conventional surface and upper air observations and satellite wind observations that were utilized in the previous chapter, the present study also ingests satellite radiance observations from different instruments for assimilation.

4.2 Case description

Cyclone Thane

On 25 December 2011, a depression had formed over the southeast Bay of Bengal. While moving in the north-northwestward direction, the depression developed to a deep depression and subsequently intensified to a cyclonic storm by the midnight of 26 December, 2011. The storm then moved in the west-northwest direction and strengthened as severe cyclonic storm by the evening of 28 December, 2011. The cyclone made landfall between 00Z and 03Z on 30 December, 2011 near Tamil Nadu and Puducheri coast. The system then weakened to a severe cyclonic storm and subsequently to a deep depression by evening of the same day. The system continued to weaken and finally manifested as a well marked low pressure area near northern Kerala in the morning of 31 December, 2011.

Cyclone Phailin

The very severe cyclonic storm Phailin developed over South China Sea on 06 October 2013 from a remnant cyclonic system. The low pressure area intensified into a depression over North Andaman Sea by 08 October 2013. The system moved west-northwestwards and strengthened into a cyclonic storm on the evening of 09 October 2013. The storm further moved northwestward and developed into very severe cyclonic storm on 10 October 2013 and finally made landfall on 12 October 2013 near Gopalpur. TC Phailin had a rapid intensification phase that was observed during 00Z10-00Z11 October 2013 which led to an increase in wind speed from 23 ms^{-1} to 60 ms^{-1}.

Cyclone Hudhud

The very severe cyclonic storm Hudhud originated from a low pressure area during October, 2014, near North Andaman Sea. The system gradually strengthened to a depression on 07 October, 2014 and moved in the west-northwest direction. The depression further developed as a cyclonic storm and crossed the Andaman Islands on 08 October, 2014.

While continuing to move in the west-northwest direction, the storm intensified as a severe cyclonic storm after moving to the Bay of Bengal. The system further concentrated to a very severe cyclonic storm and reached its maximum intensity by 12 October, 2014 morning. The storm continued to move towards the Andhra Pradesh coast and made landfall near Vishakhapatnam between 06Z and 08Z of 12 October, 2014. The cyclone continued to move northwestwards and started to weaken. By the same day evening, the cyclonic storm weakened to deep depression. Subsequently, the system became as a depression on 13 October, 2014. Moving northward, gradually the system weakened to a low pressure area and lay centered over Uttar Pradesh by the evening of 14 October 2014.

Cyclone Vardah

Over south Andaman Sea, a low pressure area had formed on 04 December 2016. The very severe cyclonic storm Vardah intensified from a well marked low pressure area near north Andaman Sea on 05 December 2016. The low pressure system moved westwards and developed into a depression on 06 December 2016 and further strengthened to deep depression on 07 December 2016. The system continued to move northwestward and concentrated into a severe cyclonic storm on 09 December 2016. The storm subsequently developed as a very severe cyclonic storm on 10 December 2016 and made landfall over Chennai on 12 December 2016. The storm continued to move west-southwestwards and weakened to a cyclonic storm. The system subsequently weakened to a depression by 13 December 2016 and lay centered as a well marked low pressure region afterwards.

4.3 Model configuration

The model configuration in this chapter is similar to the configuration employed in the previous chapter (as in Fig. 3.1). Two nested domains were employed with two way nesting enabled. The outer domain had a horizontal resolution of 27 km and the inner domain was configured with 9 km horizontal resolution. Both the domains had 36 vertical levels.

The convective processes are represented using the Kain-Fritsch scheme and the WSM 5-class scheme for parameterizing the microphysical processes. The Yonsei University scheme is employer for the boundary layer turbulence. The Noah land surface model provides for the land surface physics. The radiative processes are parameterized following the RRTM scheme and Dudhia scheme for longwave and shortwave radiations respectively.

Numerical simulations were conducted employing the 3DVar and 4DVar DA methods. The experimental design is similar as in the previous chapter (Fig. 3.2), except the free forecast length is 48h. For all the four TCs, cyclic DA experiments were performed by assimilating conventional surface and upper air observation, satellite radiance and satellite derived wind observations. Radiance from AMSU-A, AMSU-B, MHS, and HIRS-4 were assimilated. As in the previous set of experiments, a 6h long assimilation window was chosen for both the 3DVar and 4DVar experiments: t-3h < t < t+3h for 3DVar run and t < t+6h for the 4DVar run ('t'indicates the analysis time). The present study has utilized the community radiative transfer model (CRTM) for simulating the brightness temperature from the model variables. The selected satellite instrument channels corresponding to the radiance observations assimilated are indicated in Table 4.1. The radiance observations within the range of 150-350 K only assimilated in the study. Further, the observations that are being affected by rain or cloud are also removed. The bias correction for the radiance observations was also carried out by employing the variational bias correction method within the WRFDA [134].

Table 4.1: Satellite sensors and channels selected for assimilation

Satellite instrument	Satellite and channels used
AMSU-A	NOAA-15, 5-9 NOAA-16, 5-9 NOAA-18, 5-9 MetOp-2, 5-9
AMSU-B	NOAA-15, 3-5 NOAA-16, 3-5 NOAA-17, 3-5
HIRS	NOAA-17, 2-15 NOAA-18, 2-15
MHS	NOAA-18, 3-5 NOAA-19, 3-5

For TC Phailin, 12 cyclic assimilations were carried out while 13 assimilation cycles were performed for other three cyclones, after a short spin-up run (6h). 48h free forecasts were initiated from each of the analyses fields that were available at 6h interval for all the four TCs. The period of spin-up and assimilation are mentioned in the Table 4.2.

NCEP-GFS 24h forecast fields at the horizontal resolution of 0.5° × 0.5° were used as initial and boundary conditions.

Table 4.2: Details of the experiments carried out in the study

Case	Category (IMD)	Spin-up period	Period of assimilation	No. of cycles
Thane	VSCS	06Z-12Z 25 December 2011	12Z 25-12Z 28 December 2011	13
Phailin	VSCS	00Z-06Z 08 October 2013	06Z 08-00Z 11 October 2013	12
Hudhud	VSCS	18Z 07-00Z 08 October 2014	00Z 08-00Z 11 October 2014	13
Vardah	VSCS	18Z 07-00Z 08 December 2016	00Z 08-00Z 11 December 2016	13

Typical spatial distribution of observations from different platforms that are available for assimilation (valid for 00Z 28 December 2011) are shown in Fig. 4.1 and 4.2. Satellite radiance from different instruments (AMSU, MHS, HIRS) are shown in Fig. 4.1 and all other observations are depicted in Fig. 4.2. The typical number of observations available for assimilation at 00Z 08 October 2014 is indicated in Table 4.3.

Table 4.3: Typical number of observations used for assimilation at 00Z 08 October 2014.

Observation type	**Number of observations used**
Synoptic	1060
Soundings	9221
Metar report	218
Pilot	207
Buoy	24
Ship	80
AMVs	11065
Scatterometer	10970
AMSU-A	13580
HIRS-4	10046
MHS	12062

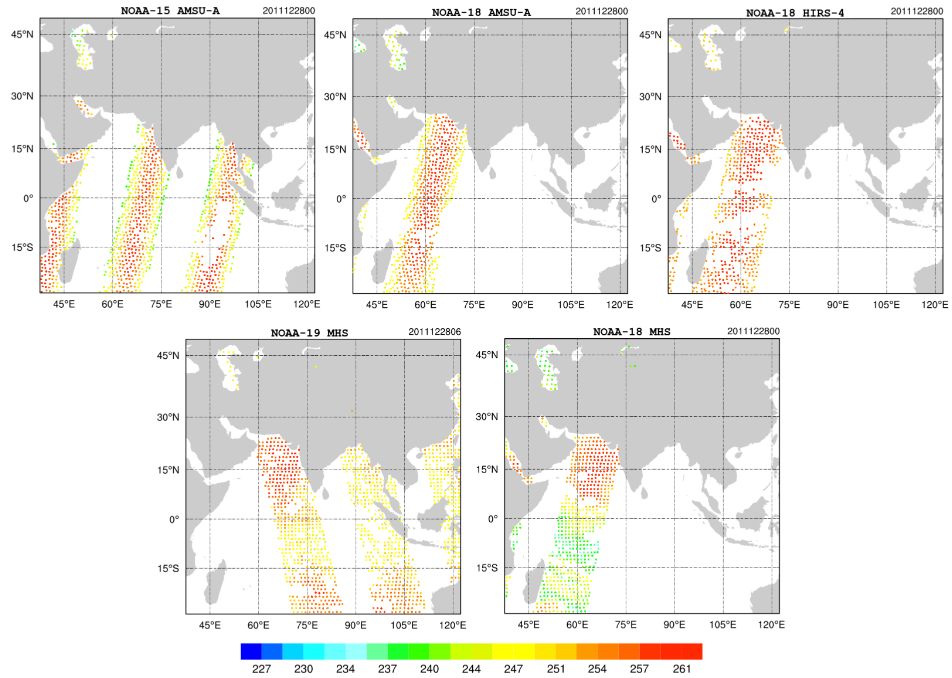

Fig. 4.1: Spatial distribution of satellite radiance observations from different platforms over the experimental domain available on 00Z 28 December 2011.

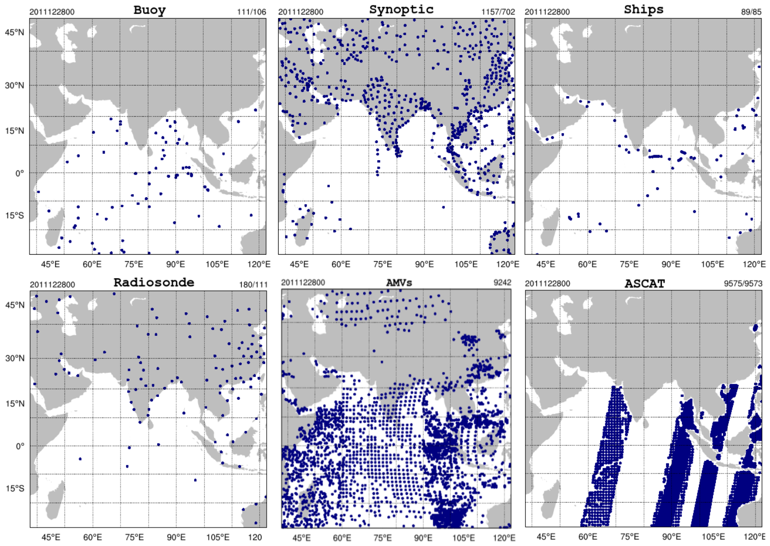

Fig. 4.2: Spatial distribution of various observations other than satellite radiance over the experimental domain available 00Z 28 December 2011.

4.4 Improvement in the analysis fields

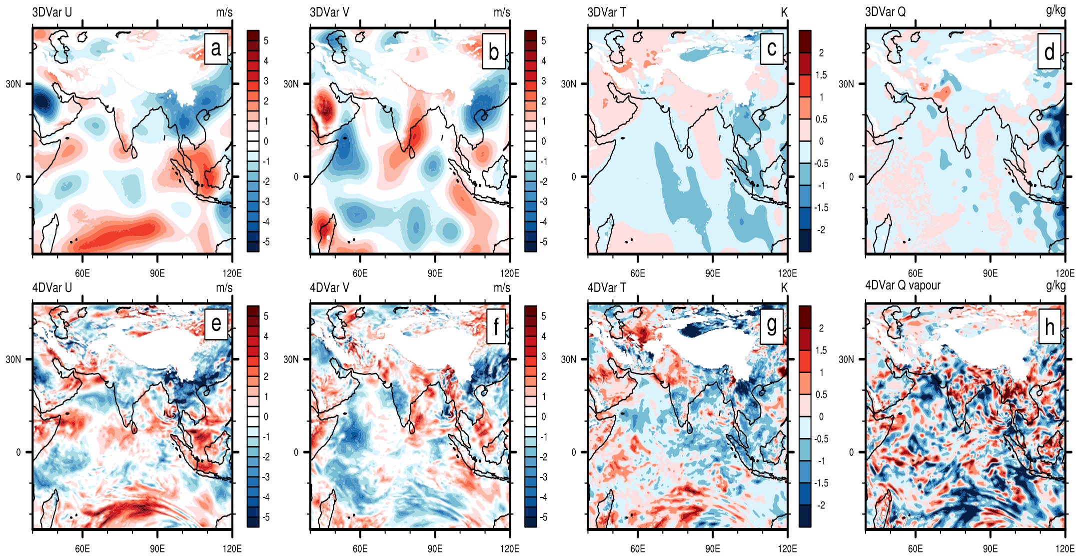

Fig. 4.3: Analysis increment in u-wind (a,e), v-wind (b,f), temperature (c,g), and water vapor mixing ratio (d,h) at 850 hPa for 3DVar run (a-d) and 4DVar run (e-h). Fields are valid at 06Z 08 October 2013.

The impact of DA can be assessed by the examining magnitude and spatial pattern of the analysis increments (analysis minus first guess) due to the assimilation. Fig. 4.3 shows the analysis increments in horizontal wind components, temperature, and mixing ratio due to the 3DVar assimilation (top panel) and the 4DVar assimilation (bottom panel). It is known that the 3DVar DA method utilizes a climatological BEC and hence the increments produced with the same would be homogeneous and isotropic in nature. However, the dynamics of the tangent linear and adjoint models associated with the 4DVar minimization process can account for evolution of model errors in an implicit manner, which in turn may contribute to flow-dependent analysis increment structures. The above observation is clearly evident from Fig. 4.3. In the case of horizontal wind components, the magnitude of increments are almost similar for both 3DVar and 4DVar experiments, whereas the flow-dependent structures are not seen in the 3DVar increment fields. Specifically, the flow pattern over the southern region of Indian Ocean is seen clearly in the increments for both the wind fields, while the same is not at all observed from the 3DVar increments. Also for the temperature and moisture fields, there are significant difference in the increment's magnitude among the 3DVar and 4DVar experiments. Interestingly, the flow pattern that is observed in the horizontal wind fields (in 4DVar increments) is consistently seen in the 4DVar increments in temperature and moisture fields also. The increments in temperature fields are more higher for the 4DVar experiments than that for the 3DVar experiments. The most striking difference is the adjustments in the moisture fields due to the 4DVar assimilation. In the 3DVar experiment, apart from the high negative increments in the humidity fields over the eastern boundary of the domain, no significant increments are seen. On the other hand, the refinements in the moisture field due to the 4DVar assimilation are more significant, especially over the Indian Ocean region. A closer examination reveals marked localized increments over the northern Bay of Bengal and northern Arabian Sea.

The tangent linear/adjoint model employed in the 4DVar assimilation does account for the major physical processes. Thus, the refinements in the temperature and humidity fields due to the 4DVar assimilation can be attributed to the impact of the simplified physics schemes incorporated in the tangent linear model. Since the TCs are convectively driven systems, the adjustments in the thermodynamic fields may impact the forecasts significantly. It is worthwhile to note that, unlike the increments in temperature and moisture fields, the magnitude and location of increments in the horizontal wind components are almost similar except for the flow-dependent structure in the 4DVar increments. The similar pattern of analysis increments in the wind components due to 3DVar and 4DVar assimilation are possibly because of the dense coverage of wind observations assimilated in terms of AMVs.

Fig. 4.4 shows the root mean square (RMS) fit error with respect to the radiosonde observations for horizontal winds, air temperature and humidity fields. RMS fit is computed by averaging all the analysis fields for all the four TCs. Fig. 4.4 suggests considerable improvement in both the wind components (with respect to the radiosonde observed winds) due to the 4DVar assimilation as compared to the 3DVar counterpart. However, neither the temperature field nor the humidity fields show reduction in RMS-fit error for the 4DVar analyzed fields. The reason for the above result may be the increments in both temperature and humidity fields with the 4DVar assimilation are much more localized in nature ($\sim$160 radiosonde samples only are available for verification over the entire domain).

Apart from the comparison with radiosonde observations, the RMSE and standard deviation in observation minus analysis (O-A) fields with respect to the surface synoptic observation are also verified. Fig. 4.5 depicts the scatter plot of O-A for u-wind (a,e), v-wind (b,f), temperature (c,g), and water vapor mixing ratio (d,h) for 3DVar run (a-d) and 4DVar run (e-h). The RMSE and standard deviation (given on left top of the plots) indicate that the 4DVar analysis fields are relatively closer to the observed fields, as revealed from

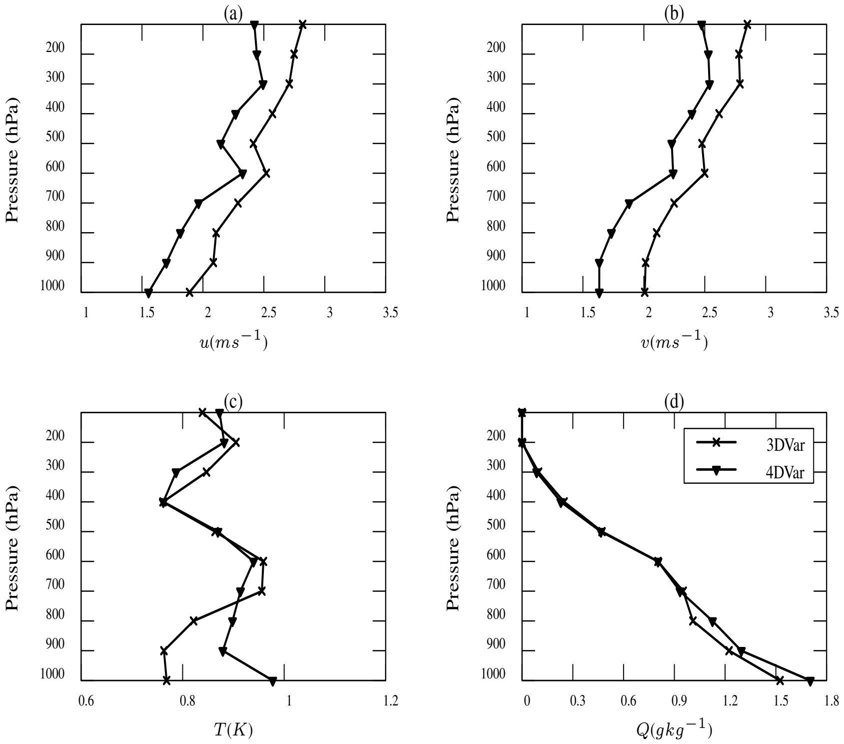

Fig. 4.4: RMS-fit error for u-wind (a), v-wind (b), temperature (c), and water vapor mixing ratio (d) for 3DVar run and 4DVar run. Fields are valid for TC Phailin.

lower RMSE for all the four fields. It is worthwhile to note that, the reduction in RMSE is more significant in the case of horizontal wind components, rather than for the temperature and humidity fields. Since the analysis increments and RMS-fit to radiosonde observation for horizontal wind components, temperature, and humidity fields for both the 3DVar and 4DVar experiments are found to have similar pattern for all the cases, figures of RMS-fit and analysis increments corresponding to TC Phailin are only shown here.

The impact of assimilation on the simulation of surface winds over the ocean has been examined by estimating the improvement parameter (IP, as defined in Eqn. 2.45). The IP is computed for the analysis surface winds (u-wind and v-wind) with respect to the scatterometer winds observed by ASCAT. Fig. 4.6 depicts the IP for u-wind (a-b) and v-wind (e-h) for TC Thane (a,b), Phailin (c,d), Hudhud (e,f), and Vardah (g,h). As per the definition of IP, positive values of IP indicate the improvement with the 4DVar assimilation, as compared to the 3DVar assimilation. It is observed that, for all the four TCs, of most of the grid points (>70%) are positively impacted by the 4DVar assimilation.

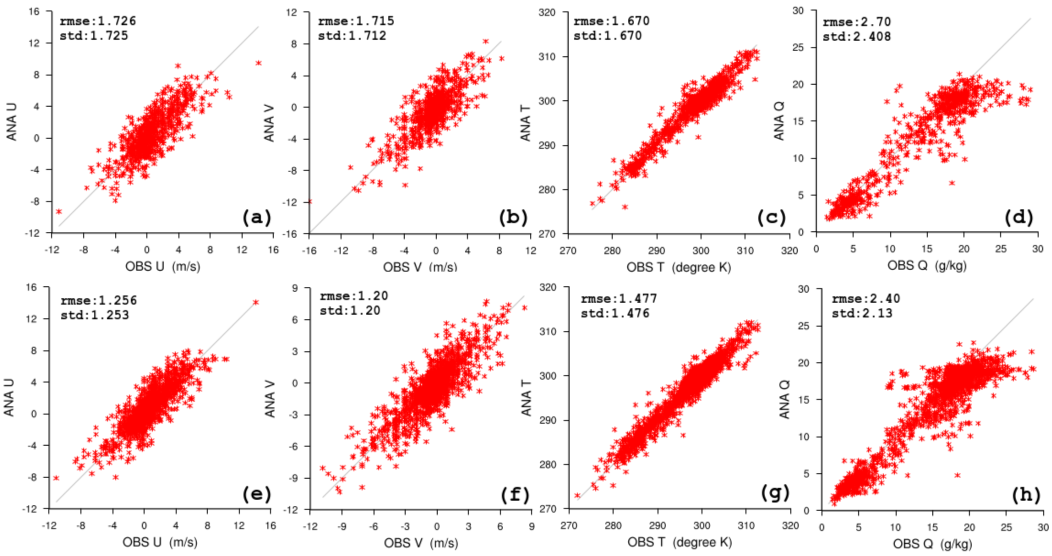

Fig. 4.5: Scatter plot of O-A (observation minus analysis) for u-wind (a,e), v-wind (b,f), temperature (c,g), and water vapor mixing ratio (d,h) for 3DVar run (a-d) and 4DVar run (e-h). Fields are valid at 12Z October 2013.

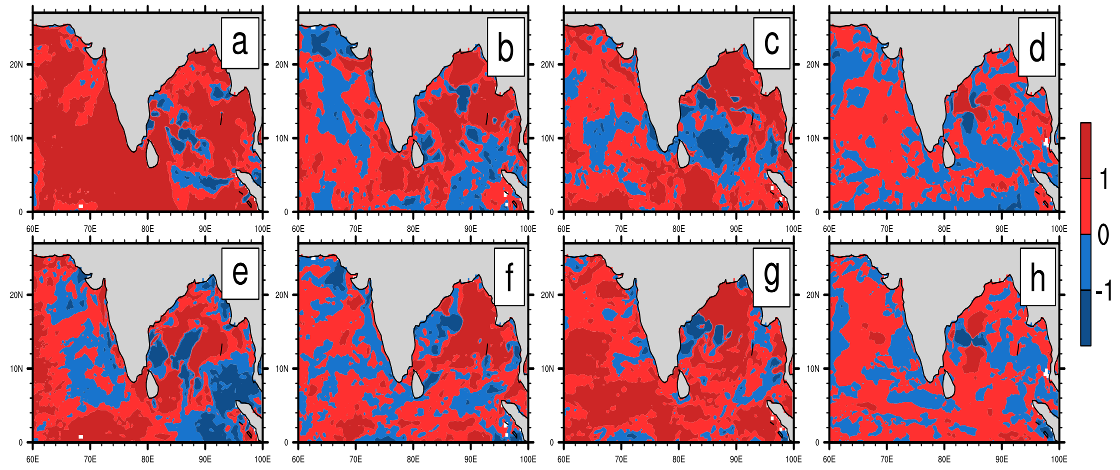

Fig. 4.6: Improvement parameter for u-wind (a,c,e,g) and v-wind (b,d,f,h) over the ocean for TC Thane (a,b), Phailin (c,d), Hudhud (e,f), and Vardah (g,h) with respect to ASCAT observations.

The ocean surface winds play an important role in the transport of mass, momentum and energy through turbulent fluxes. Thus, an accurate representation of surface winds over the oceans would account for improved turbulent fluxes and transport. Since the turbulent processes over the ocean surface plays a pivotal role in the development of TCs, better representation of surface winds over the oceans can hopefully contribute to improved simulation of tropical cyclone features [132].

4.5 Intensity of the simulated cyclones

At the end of each cyclic assimilation, the analyses fields were subjected to 48h free forecasts and the features of the TCs simulated during the free forecast period were verified against corresponding observations. The intensity of the simulated cyclones in terms of MWS and MSLP were validated with respect to the IMD best track observations. The average absolute error in both MWS and MSLP were computed by averaging the absolute error in MWS and MSLP for all the samples for each TCs, with respect to the forecast length (up to 48h). Fig. 4.7 shows the time series of average absolute error in MWS for the 3DVar and 4DVar experiments for all the four TCs.

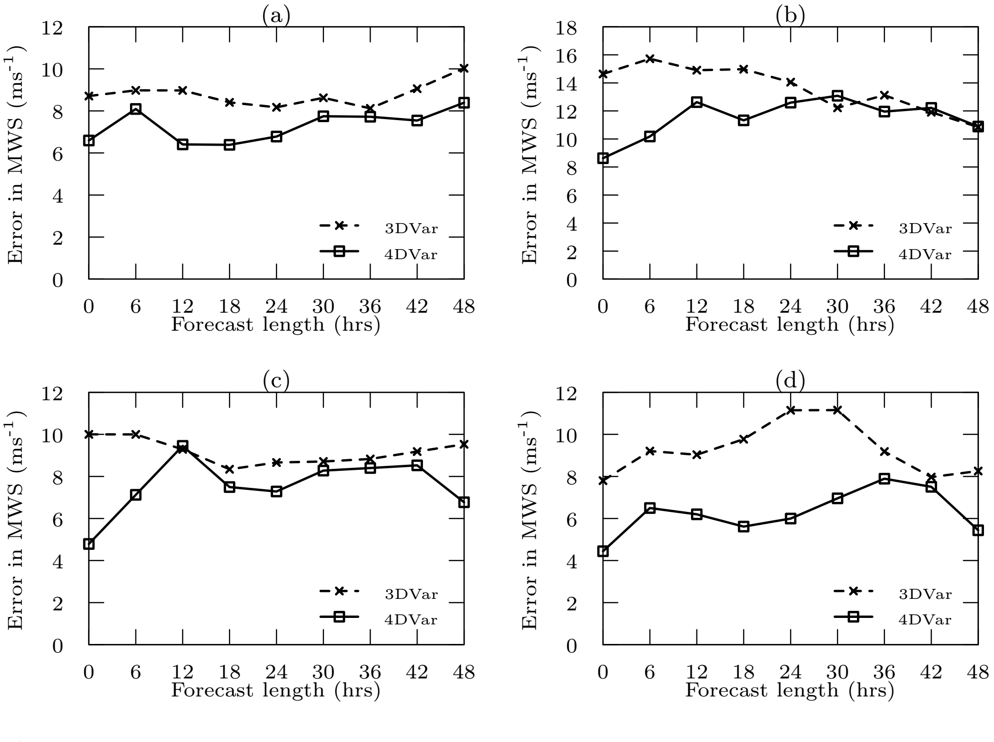

Fig. 4.7: Time evolution of average error in MWS for 3DVar and 4DVar experiments for TC Thane (a), Phailin (b), Hudhud (c), and Vardah (d).

The average error associated with the 4DVar run is consistently lower when compared with the error in the 3DVar run. In general, the errors are growing with time, irrespective of the DA method employed. However, for TC Phailin simulated by the 3DVar run, the average error in MWS is decreasing with time, while the MWS average error for the 4DVar run is increasing with forecast lead time (Fig. 4.7b). It is interesting to note that the 4DVar analysis fields (0h forecast) shows significant reduction in the average MWS error for all the cyclones as compared to their 3DVar counterparts. On an average, the 4DVar analyzed fields (at 0h forecast length) are found to reduce the MWS error up to $\sim$ 50% (in the case of TC Hudhud) with respect to the 3DVar analysis fields.

The intensity simulation in terms of time evolution of MSLP errors are also examined for all TCs and are depicted in Fig. 4.8. The 4DVar experiments clearly have better skill in reproducing the MSLP associated with the TCs throughout the forecast length. Also, unlike in the case of MWS simulation, the MSLP simulation shows better impact due to the 4DVar assimilation for all the TCs. Consistent with the MWS errors, the MSLP errors associated with the 4DVar analysis fields (at 0h forecast length) have reduced significantly when compared to the 3DVar analysis fields for all TCs.

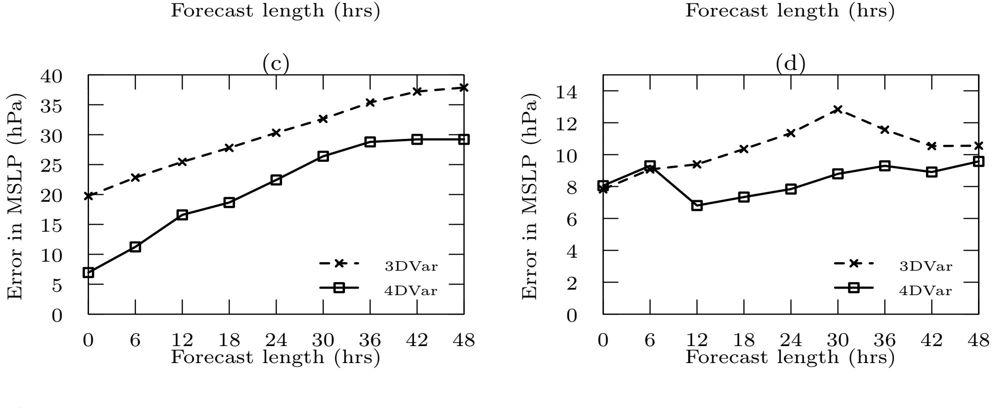

Fig. 4.8: Time evolution of average error in MSLP for 3DVar and 4DVar experiments for TC Thane (a), Phailin (b), Hudhud (c), and Vardah (d).

For TC Phailin, the evolution of MSLP error shows somewhat different behaviour, with the initial 3DVar error being much higher and the same decreases with time (as noted in the evolution of MWS errors for TC Phailin). Fairly large overestimation in the simulated MSLP has been noted in the case of TC Hudhud with both the experiments, which contributed considerably large error in MSLP simulation of TC Hudhud. However, the corresponding MWS errors for TC Hudhud did not reveal such large overestimation. The IMD best track observations for TC Hudhud indicate that the lowest MSLP is 950 hPa. The 3DVar run has simulated the cyclone with lowest MSLP of 929 hPa while the 4DVar run simulated the lowest MSLP of 937 hPa. However, the IMD observed highest MSW for TC Hudhud is 51 ms^{-1} while the highest MWS simulated by the 3DVar and 4DVar run are 61 ms^{-1} and 53 ms^{-1} respectively. Despite the general overestimation in the MSLP simulation by the model for TC Hudhud, the results of the 4DVar run are relatively closer to the IMD observations. It can be recalled that the 4DVar analysis fields accounted for more realistic surface winds over the ocean, that will contribute to better representation of turbulent transport. Hence improvement in the intensity simulation with the 4DVar experiments may be attributed to improved simulation of ocean surface winds. It is also important to recall the improvements in moisture fields with the 4DVar assimilation, which is attributed to the moist physics processes incorporated in the linear model employed in 4DVar minimization.

4.6 Rapid intensification of TC Phailin

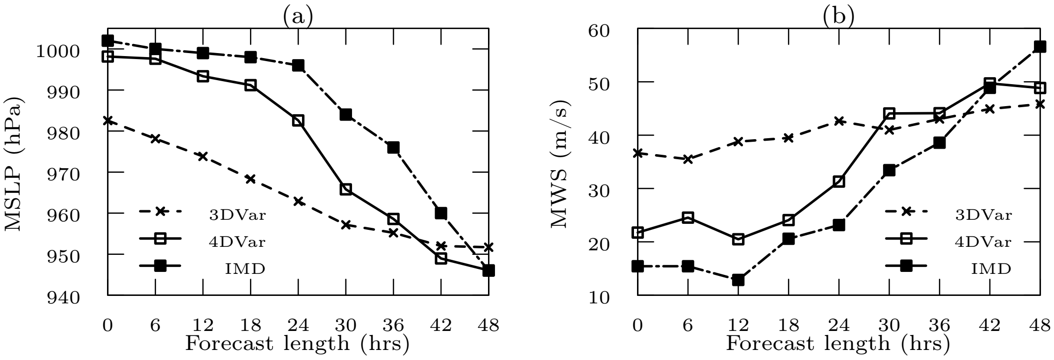

Fig. 4.9: Time evolution of MSLP (a) and MWS (b) for 3DVar and 4DVar experiments for TC Phailin. Forecast was initialized with analysis at 00Z 09 October 2013.

TC Phailin experienced a rapid intensification during 10-11 October 2013 that led to a rapid decrease in MSLP from 996 hPa to 940 hPa and an associated increase in maximum wind speed from 23 ms^{-1} to 60 ms^{-1}. The model runs simulating TC Phailin have been closely analyzed to investigate the model's skill in reproducing the above rapid intensification phase by examining a single deterministic forecast initialized at 00Z 09 October 2013. Fig. 4.9 shows the evolution of MSLP (a) and MWS (b) for both the model runs (3DVar and 4DVar) during 48h free forecast period (00Z 09-00Z 11 October 2013). The observed MSLP and MWS at 00Z 09 October 2013 are 1002 hPa and 15 ms^{-1} respectively. The 4DVar run simulated a vortex with MSLP of 999 hPa and MWS of 22 ms^{-1}, which is much closer to the observed cyclone vortex intensity features. However, the vortex simulated by the 3DVar run was found to be excessively stronger than the observed one, with MSLP of 982 hPa and MWS of 36 ms^{-1}. Also, the 3DVar simulated cyclone intensified slowly with time throughout the forecast duration. However, the results of the 4DVar run revealed very little intensification up to 18h of free forecast and showed rapid intensification at later times. Overall the 4DVar run reproduced the rapid intensification phase very realistically and is in good agreement with the observed intensification. Fig. 4.9 also provides an explanation for the 3DVar results that showed a decrease in intensity errors with forecast length for TC Phailin. The 3DVar run simulates a very intense storm that contributes to large error at the beginning of the forecast. Since the 3DVar run is not reproducing the intensification process accurately, at the end of the forecast, the simulated intensity of the 3DVar run is very nearly the same as the observed intensity. The above resulted in larger error in the initial time of the forecast and much lower error in the end of the forecast for the 3DVar run for TC Phailin.

The vertical wind shear (VWS) plays a key role in modulating the intensification of tropical storms. Strong wind shear can inhibit the development of vertically growing clouds which in turn limit the intensity of TCs. Thus, in general lower wind shear always favours intensification of the storms. The VWS (the vector difference of the 850 hPa and 200 hPa wind) is examined for TC Phailin to understand the intensification process. Fig. 4.10 shows the VWS corresponding to analysis fields (a,b) and 48h forecast fields (c,d) for 3DVar (left panel) and and 4DVar (right panel) runs that were initialized at 00Z 09 October 2013. The 3DVar analysis field (Fig. 4.10a) indicates the presence of stronger shear fields over the position of the vortex and in the immediate neighbourhood. On the other hand, for the 4DVar run, VWS has relatively lower magnitude over the southern sector of the vortex and much lower magnitude over the position of the vortex. The 48h model forecasts indicate that, the weaker shear zones in the 4DVar run have expanded to regions that are in the north-northeast sector of the TC over time. However, the immediate vicinity of the 3DVar storm is characterized by very large VWS. Hence, it is inferred that, the presence of lower magnitude of the vertical wind shear as observed in the case of the 4DVar run has favoured the rapid intensification phase in the case of TC Phailin. Furthermore, the relatively higher magnitude of VWS may have restricted the further intensification of the 3DVar storm to an extent.

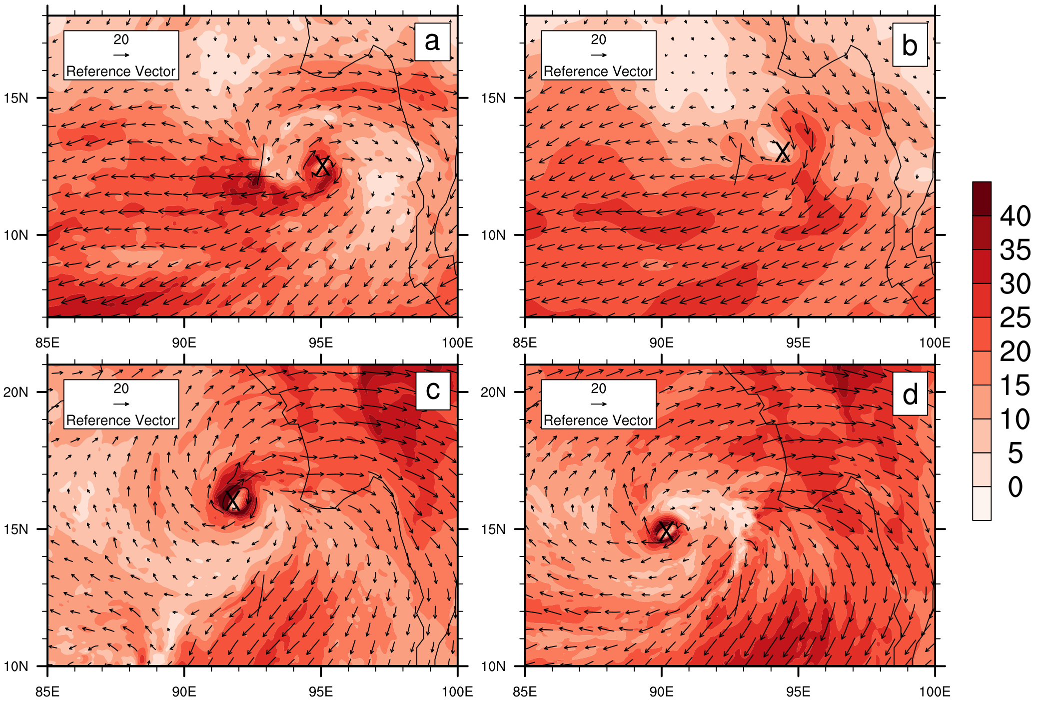

Fig. 4.10: Vertical wind shear vector and magnitude corresponding to analysis fields (a,b) and 48h forecast fields (c,d) for 3DVar (a,c) and 4DVar (b,d) experiments for TC Phailin. Forecast was initialized with analysis at 00Z 09 October 2013. Black cross indicates the position of the storm.

4.7 Rainfall simulation

Tropical storms are usually associated with torrential rainfall. The impact on rainfall simulation is examined by verifying the model accumulated rainfall with estimates from TRMM. Fig. 4.11 shows the 48h accumulated rainfall from 3DVar run (a, d, g, j), 4DVar run (b, e, h, k), and TRMM rainfall estimates (c, f, i, l) for TC Thane (a-c), Phailin (d-f), Hudhud (g-i), and Vardah (j-l).

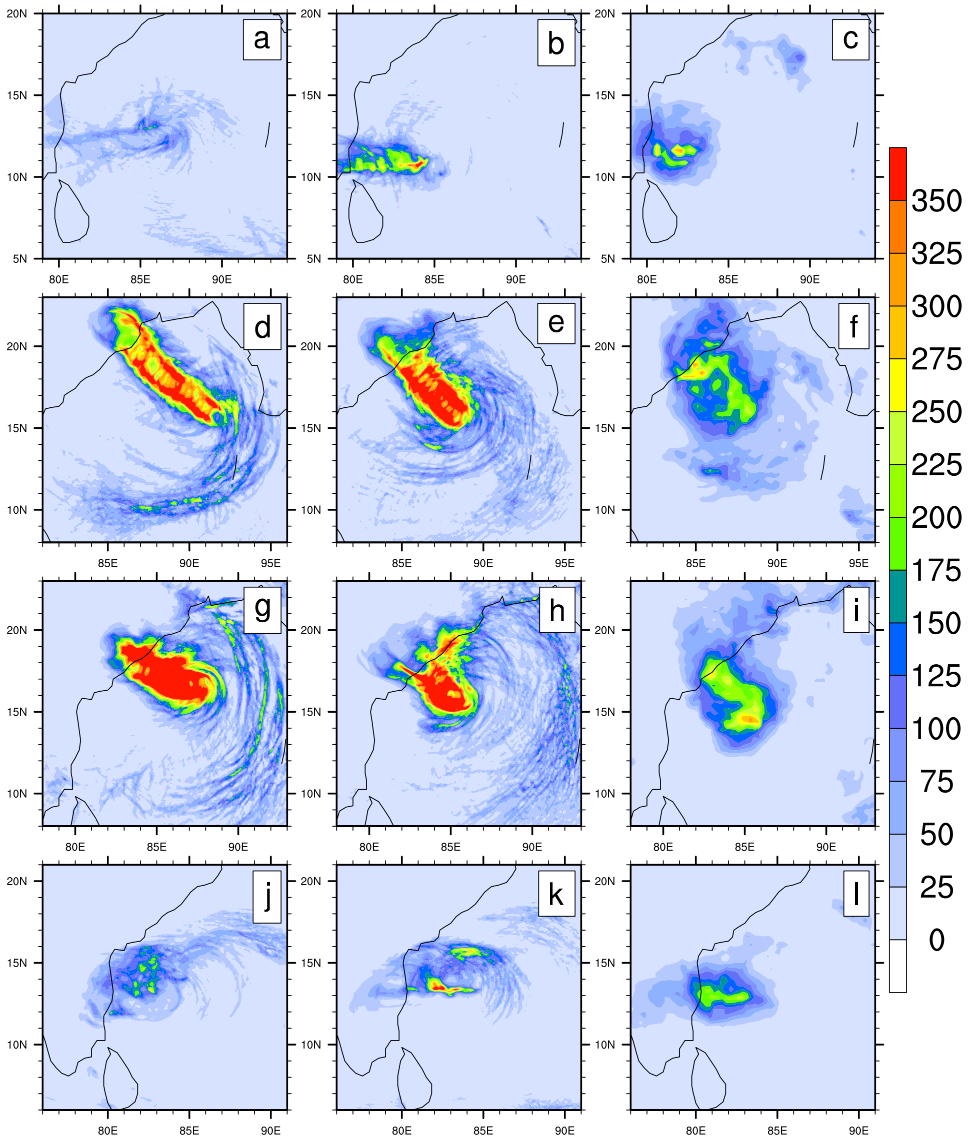

Fig. 4.11: 48h accumulated rainfall from 3DVar run (a,d,g,j), 4DVar run (b,e,h,k), and TRMM rainfall estimates (c,f,i,l) for TC Thane (a-c), Phailin (d-f), Hudhud (g-i), and Vardah (j-l).

The results shown here correspond to 48h accumulated rainfall when initialized from 12Z 27 December 2011 for TC Thane, 00Z 11 October 2013 for TC Phailin, 00Z 11 October 2014, TC Hudhud, and 00Z 11 December 2016 for TC Vardah. The above time periods correspond to the maximum intense rainfall phase of each cyclones. Even though both the model runs have reproduced the overall general feature of rainfall associated with all the four TCs with reasonable accuracy, there are marked difference in the location and intensity of the intense rainfall. Both model results show significant overestimation in the rainfall for TC Phailin and Hudhud, while, the magnitude of maximum rainfall has been reproduced fairly well by both the model runs in the case of TC Thane and Vardah. It is to be noted that, TCs Phailin and Hudhud are relatively stronger systems, whereas TCs Thane and Vardah are relatively weaker storms.

For TC Thane (Fig. 4.11 a-c), the 4DVar run has reproduced the rainfall distribution and magnitude very realistically, with respect to the observed TRMM estimates. In addition to the above, the location of maximum rainfall is also reproduced well by the model when initialized with 4DVar analysis. However, the model could not reproduce the location and structure of the rainfall bands realistically with the 3DVar run. Furthermore, a closer examination would reveal a slight underestimation in rainfall with the 3DVar run. The disagreement in the simulation of the location of maximum rainfall can be attributed to the different speeds at which the system advanced, i.e, the 3DVar simulated storm has advanced slowly as compared to the 4DVar storm and the observed system. Interestingly, the rainfall pattern associated with the 3DVar storm appears to have a structure that is not fully organized. Also, the intensity of rainfall is lower when compared to the observed system and the 4DVar storm. The above observation provides an indication of an early weakening of the 3DVar simulated storm. Hence, it is concluded that the 4DVar run simulated the life cycle of TC Thane more realistically than the 3DVar run did.

The 48h accumulated precipitation for TC Phailin is shown in Fig. 4.11 (d-f). It is clearly evident that both the model runs have overestimated the rainfall considerably. While the TRMM rainfall indicates more wide-spread rain, both the model runs show more localized, intense rainfall patterns. The results for the 3DVar storm show much elongated rainfall pattern, along the track of the cyclone, while, the rainfall distribution simulated with the 4DVar run resembles fairly well with the observed pattern, except for the magnitude of the rainfall intensity. The 3DVar run has simulated a wider rain band on the southern sector of the vortex, which is not seen in the TRMM rainfall and the 4DVar experiment. Also, the TRMM rainfall pattern indicates that the system is very close to the coast. A closer analysis will reveal that the position of the 4DVar storm is in good agreement with the TRMM observed position, while the 3DVar simulated rainfall pattern is slightly away from the coast and also from the observed location.

Fig. 4.11 (g-i) shows the 48h accumulated rainfall for TC Hudhud. As in the case of TC Phailin, both the model runs have overestimated the rainfall intensity when compared with the TRMM rainfall estimate. Among the two model runs, the 3DVar run has significantly larger overestimation of rainfall as compared to the 4DVar run. Nevertheless, both the 3DVar and 4DVar experiments have reproduced the rainfall pattern reasonably well as observed in the TRMM estimate. The 4DVar run has reproduced the location of the maximum rainfall, which is closer to the maximum rainfall location as seen from the TRMM rainfall. The TRMM rainfall indicates that the location of maximum rainfall is centered near 15.0°N, 85.0°E. The location of maximum rainfall with the 4DVar run is near 15.2°N, 85.0°E, whereas the maximum rainfall location simulated with the 3DVar is slightly north of the observed position and is at 15.5°N, 85.2°E.

Being a relatively weaker system, the magnitude of the rainfall associated with TC Vardah is not very high. Fig. 4.11 (j-l) shows the 48h accumulated rainfall in the case of TC Vardah. Both the model runs have reproduced the rainfall distribution fairly well. However, as mentioned earlier, the 3DVar run has slightly underestimated the rainfall intensity while the 4DVar run has slightly overestimated the rainfall intensity. Also, the 3DVar run did not simulate any distinct rainfall maxima. The 4DVar simulated rainfall pattern is characterized with two rainfall maxima, which is however, not observed in the TRMM rainfall and the 3DVar run. Nonetheless, the location of maximum rainfall is very much closer to the observed location of maximum rainfall.

The quantitative verification of the rainfall simulation has been performed by computing the statistical skill scores such as ETS, bias, POD, and FAR for both the model runs. Fig. 4.12 shows the time series of ETS (a), bias (b), FAR (c), and POD (d) for TC Thane, calculated for two rainfall thresholds (75 mm and 150 mm) by considering 48h accumulated rainfall with reference to TRMM rainfall estimate. For lower and higher thresholds, significant improvement in rainfall simulation can be noted in terms of higher ETS, lower FAR and higher POD for the 4DVar run. As seen in Fig. 4.11 (a-c), the skill scores for the 3DVar run also confirm the underestimation of the higher rainfall threshold (150 mm). The results indicate that the 4DVar run tends to overestimate the high-threshold rainfall. In addition, low ETS for higher threshold beyond a particular time (for the 3DVar run) indicates early weakening of the cyclone, as mentioned earlier.

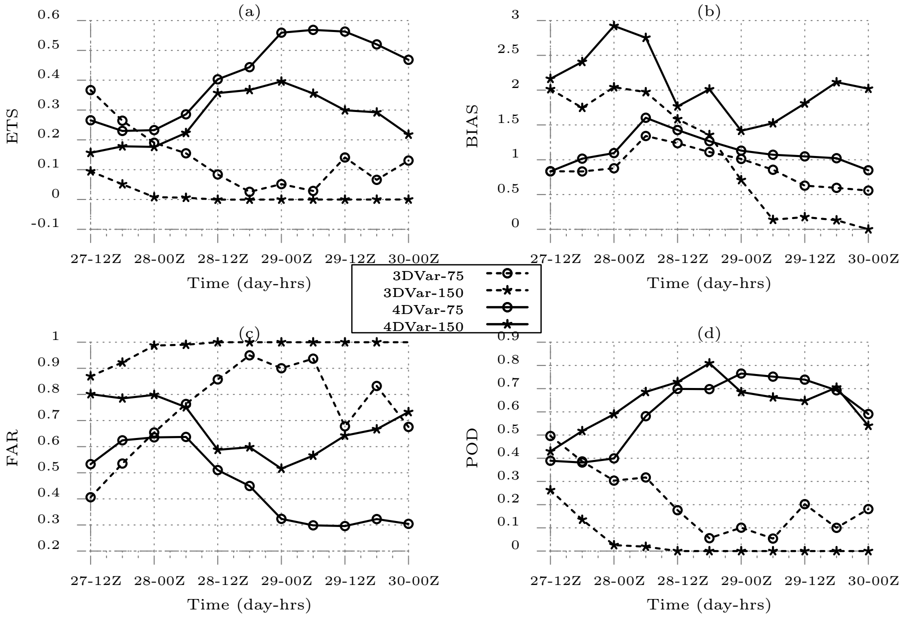

Fig. 4.12: Quantitative skill scores for TC Thane, computed for 48h accumulated rainfall from 3DVar run and 4DVar run with respect to TRMM rainfall for 75 mm and 150 mm rainfall thresholds.

Interestingly, all the 4DVar skill scores for TC Thane show improvement consistently with respect to time. This is partly due to the gradual refinement in the moisture fields after every 4DVar assimilation. Another possible reason is the improvement in simulating the position of the storm with the analysis cycle for the 4DVar run. Better simulation of storm track would result in accurate simulation of location of maximum rainfall, which in turn leads to improvement the in rainfall forecast skill by the model.

The skill scores for TCs Phailin and Hudhud are shown in Fig. 4.13 and Fig. 4.14 respectively. Almost similar inferences can be drawn for TCs Phailin and Hudhud also. The skill scores show appreciable improvement with the 4DVar experiments over the 3DVar counterparts. Furthermore, the rainfall forecast skill is continuously getting improved with number of analyses, as seen in the case of TC Thane. The rainfall distribution for TC Phailin and Hudhud (Fig. 4.11d-f, g-i) had revealed considerable overestimation in rainfall with the 3DVar run. It is important to note that the overestimation in the rainfall associated with the model (for 3DVar run) has reduced significantly with the 4DVar run for both the TCs Phailin and Hudhud, resulting in improved rainfall forecast with the 4DVar run.

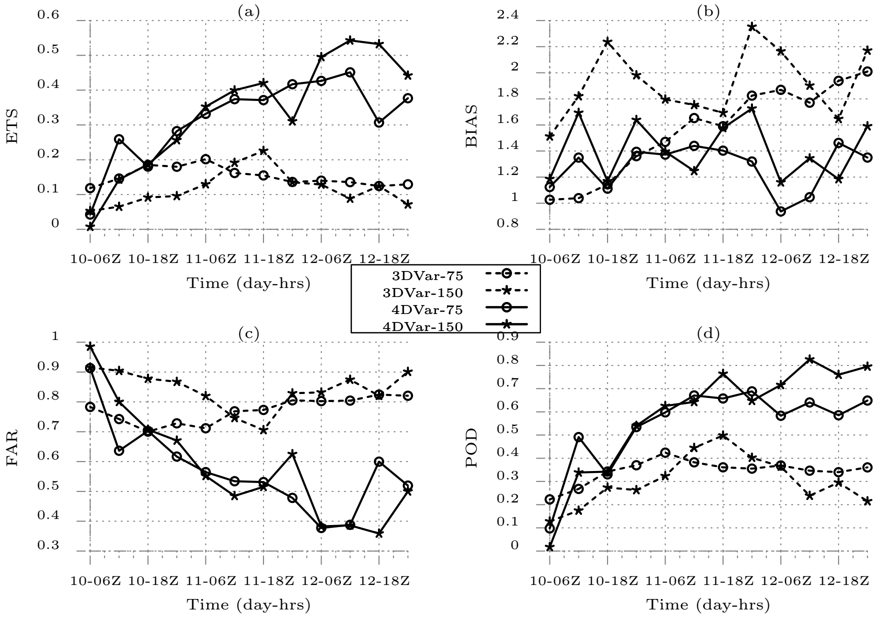

Fig. 4.13: Same as Fig. 4.12 but for TC Phailin.

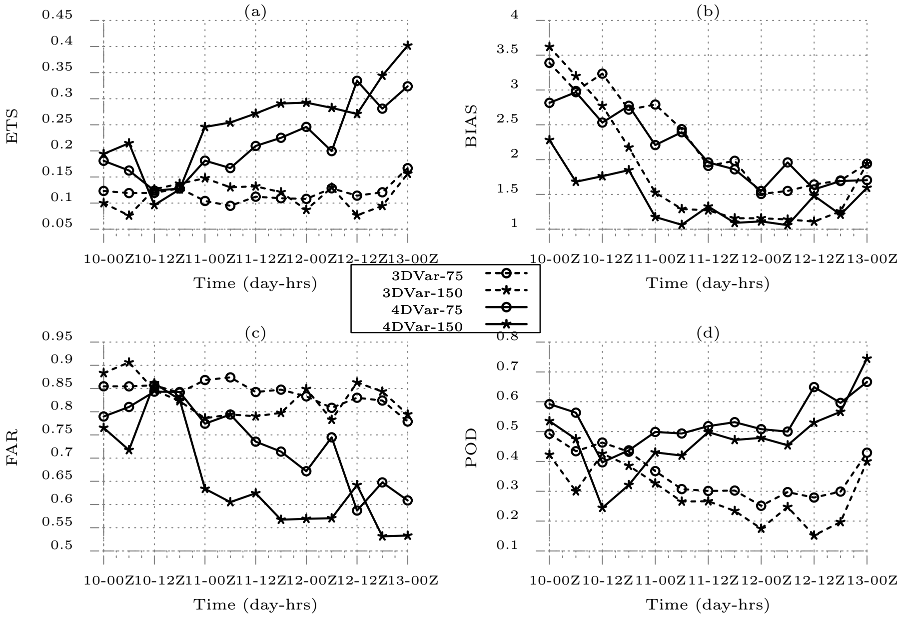

Fig. 4.14: Same as Fig. 4.12 but for TC Hudhud.

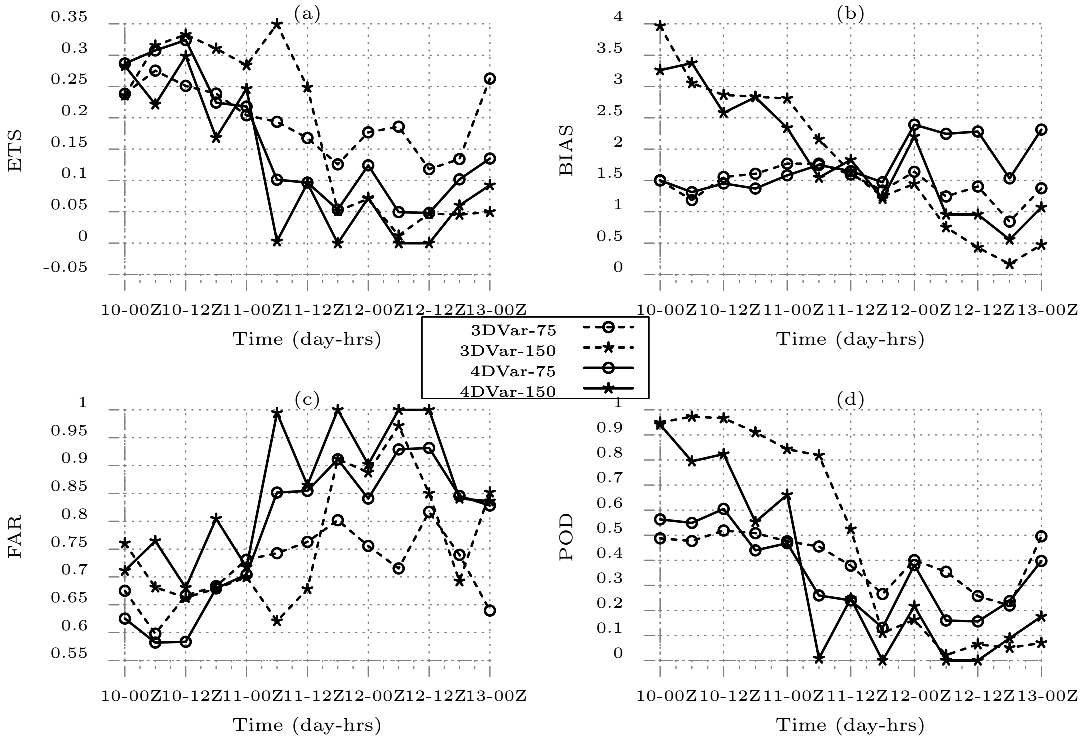

Fig. 4.15: Same as Fig. 4.12 but for TC Vardah.

The rainfall forecast skills for TC Vardah show slightly different behaviour from the other three TCs and the 4DVar run results do not show marked improvement over the corresponding 3DVar experiment. In fact, the 3DVar run has slightly better skill scores for most of the times. It may be recalled that the rainfall distribution for TC Vardah (Fig. 4.11j-l) revealed the 4DVar run simulated the location of maximum rainfall slightly north of the observed location. Also, the 4DVar simulated rainfall was characterized with two rainfall maxima, which was not seen in either the TRMM rainfall or in the 3DVar run. The above might have contributed to lower skill scores for the 4DVar run as compared to the 3DVar run. Unlike in the other three TCs, the skill scores for TC Vardah decreased with time. This may be due to the discrepancies in simulating the cyclone track with the analysis cycles, i.e, after each analysis cycle, the positional error of the cyclone center may be increasing with time. The above observation will be verified by analyzing the track of the TCs.

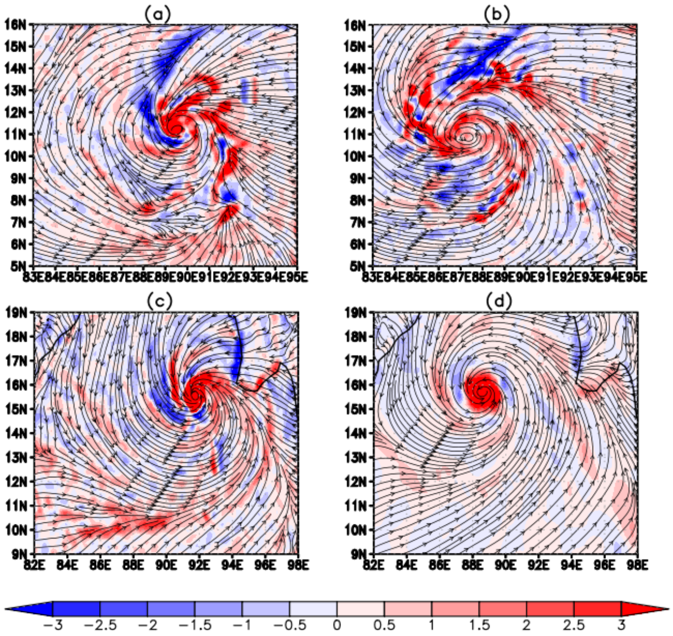

Fig. 4.16: Surface moisture convergence (in gkg^{-1}s^{-1}) and streamlines for 3DVar (a,c) and 4DVar (b,d) run for TC Thane (a-b) and Phailin (c-d).

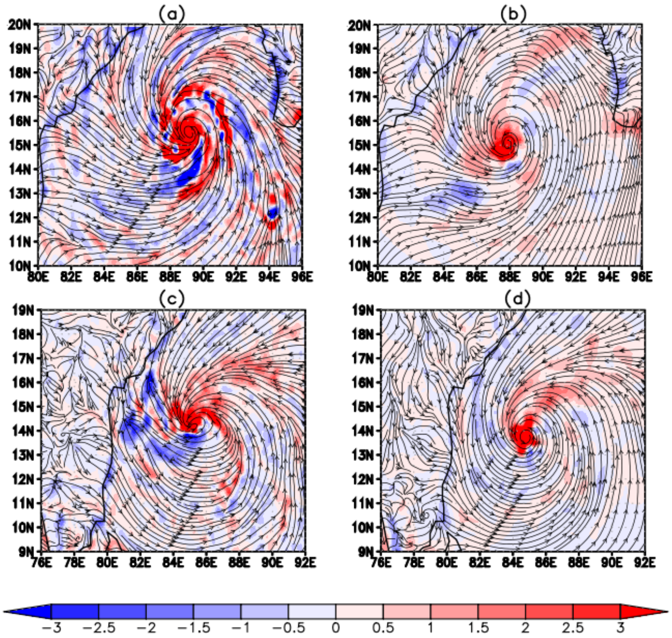

Fig. 4.17: Surface moisture convergence (in gkg^{-1}s^{-1}) and streamlines for 3DVar (a,c) and 4DVar (b,d) run for TC Hudhud (a-b) and Vardah (c-d).

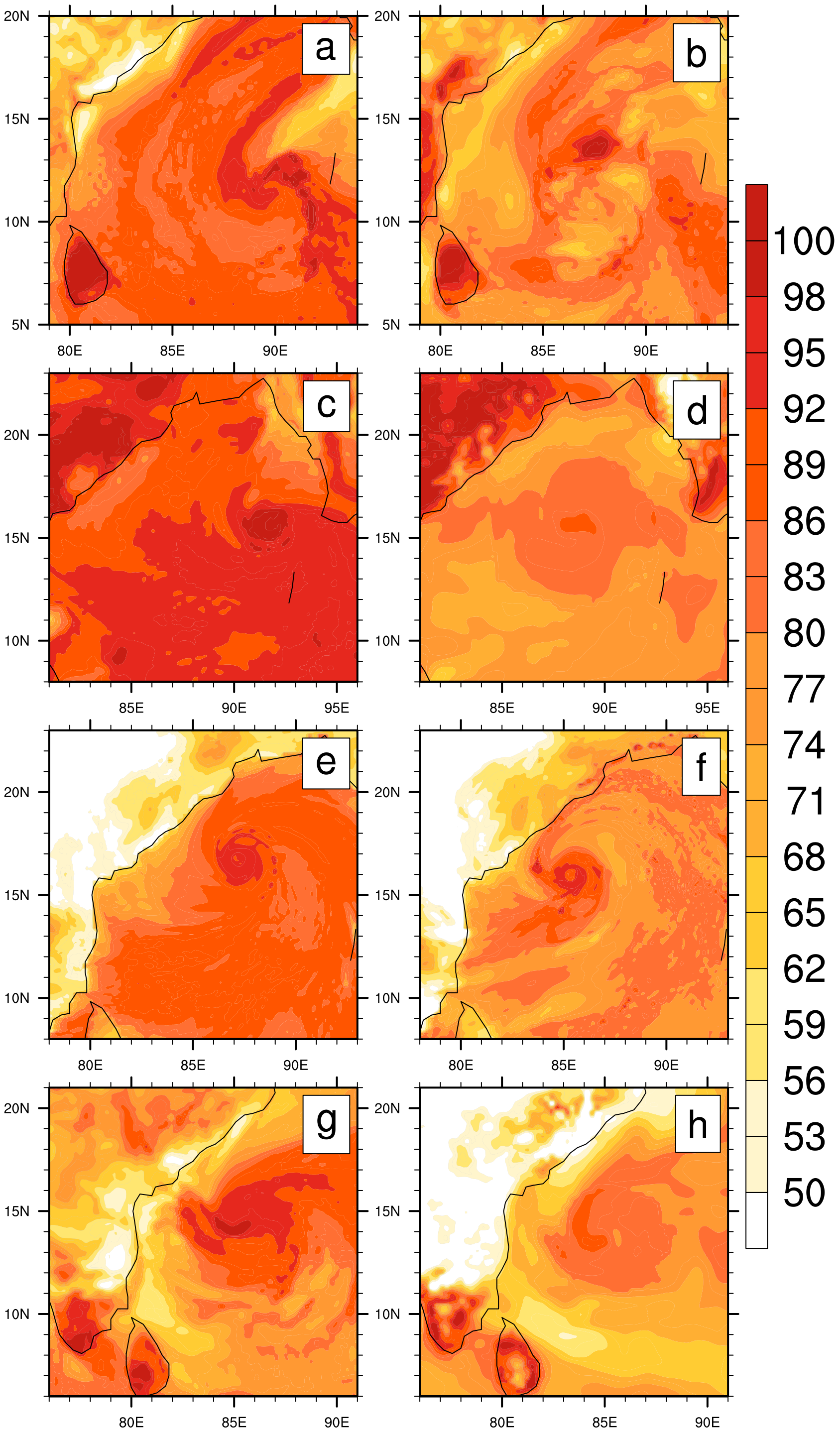

Fig. 4.18: Surface relative humidity for 3DVar (a,c,e,g) and 4DVar (b,d,f,h) run for TC Thane (a-b), Phailin (c-d) Hudhud (e-f), and Vardah (g-h).

The analysis of surface winds over the ocean has revealed that the 4DVar analyzed wind fields account for a more accurate and realistic surface winds as compared to the 3DVar winds. Hence, such an improvement in the surface winds would definitely impact the surface moisture convergence and in turn the rainfall intensity and its distribution. The surface moisture convergence associated with both the model runs have been examined for all the TCs during their maximum intensity phase and are shown in Fig. 4.16-Fig. 4.17. A general observation is that the 3DVar runs are characterized with slightly stronger moisture convergence as compared with the corresponding 4DVar runs. In the case of TC Phailin (Fig. 4.16c) and Hudhud (Fig. 4.17a), the above observation is more applicable, i.e, the moisture convergence for the 3DVar run is considerably stronger than the 4DVar run. Consequently, the intensified moisture convergence for the 3DVar run resulted in significant bias in the model simulated rainfall. For TC Phailin, the 4DVar run indicates a moisture convergence of 2.5-3 $gkg^{-1}s^{-1}$, while for the 3DVar counterpart simulated moisture convergence value is above 3 $gkg^{-1}s^{-1}$. This has resulted in higher bias score for the 3DVar simulated rainfall, particularly for the higher rainfall threshold. In the case of TC Hudhud, the overestimation in the moisture convergence by the 3DVar run is even more significant and has resulted in extreme overprediction of rainfall. However, the model runs do not show marked differences (among the 3DVar and 4DVar runs) in the moisture convergence simulation with the relatively weaker cyclonic systems, TC Thane or Vardah. The above observation provides an indication that the 3DVar runs have a tendency to overestimate the moisture convergence of stronger storms while the 4DVar runs are able to suppress the overestimation in the moisture convergence appropriately. Hence, it can be inferred that the 4DVar analysis accounts for accurate surface moisture convergence as compared to the 3DVar run, which in turn contributes to improved rainfall forecasts. The analysis of surface relative humidity fields for all the four TCs (Fig. 4.18) confirms that the 4DVar analyses simulate considerably drier atmosphere when compared with the 3DVar analyses fields. The above leads to excessive moisture convergence for the 3DVar fields as compared with the 4DVar analysis fields.

4.8 Track of the simulated cyclones

The skill of the model in simulating the TC track has been evaluated with respect to IMD observations. The positive impact of assimilation can also be examined in terms of how well the analyses fields have reproduced the position of the cyclone vortex. Fig. 4.19 shows the error in simulating the vortex position for 3DVar and 4DVar experiments as revealed from analyses fields.

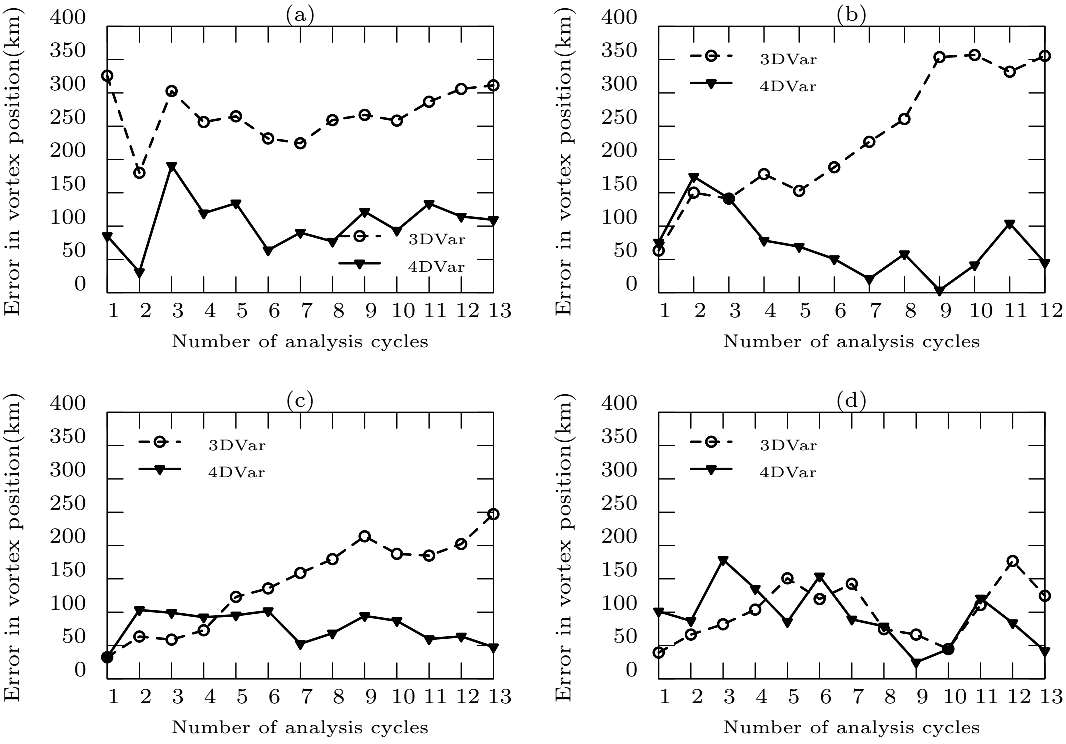

Fig. 4.19: Time series of error in vortex position for 3DVar and 4DVar runs for TC Thane (a) Phailin (b) Hudhud (c) and Vardah (d) with respect to IMD observations.

It is quite evident from Fig. 4.19 that the 4DVar analysis fields have successfully reproduced the position of the storm centre and the error in simulating the vortex position associated with the 4DVar run is significantly lower than the corresponding 3DVar run, with an exception for TC Vardah. In the case of TC Thane (Fig. 4.19 a), from the first analysis itself, the positional error with the 3DVar run is much higher (∼300 km) than the 4DVar run (∼100 km), indicating that the large initial error for the 3DVar run in the very first analysis is consistently observed throughout the experiment. For TCs Phailin (Fig. 4.19 b) and Hudhud (Fig. 4.19 c), the positional error in 3DVar and 4DVar runs are comparable for the first few cycles. However, after a few analysis cycles, the errors in the 3DVar run started to diverge. To be specific, after the third cycle in the case of TC Phailin, the error in the 3DVar experiment increased significantly, while the error in the 4DVar run started decreasing. Almost similar inferences can be drawn in the case of TC Hudhud also. The results for TC Vardah (Fig. 4.19 d) are different from other TCs and exhibits a mixed behaviour. For the first few cycles, the 3DVar analyses has slightly lower errors, however after 3-4 analyses, both the 3DVar and 4DVar runs are found to have comparable error values.

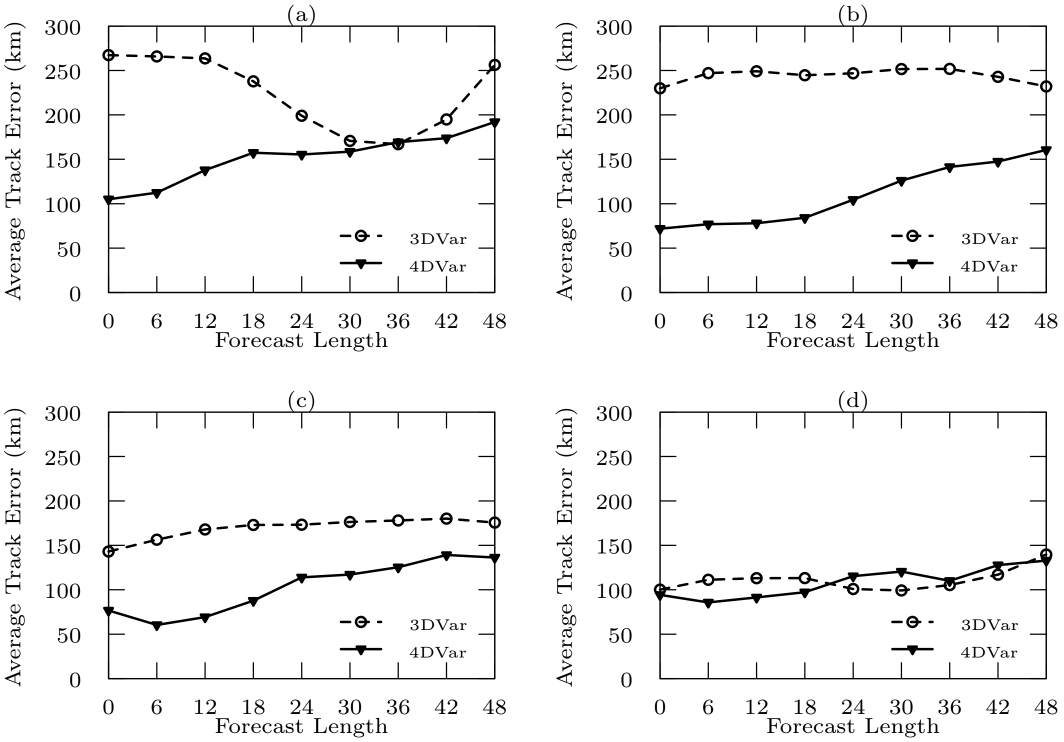

Fig. 4.20: Time series of average track error for 3DVar and 4DVar runs for TC Thane (a) Phailin (b) Hudhud (c) and Vardah (d).

The performance of the model in forecasting the cyclone track is examined by estimating the average track error, that is, by averaging for all the samples with respect to the forecast length. Fig. 4.20 illustrates the time evolution of the average track error for both the model runs and for all the TCs. As observed from Fig. 4.19, the average error for the 3DVar run at 00 forecast hour is quite large, particularly for TCs Thane and Phailin. The average track error associated with the 4DVar of all the TCs increased slightly with time. For TC Thane, the error with the 3DVar run decreases from 350 km to 150 km during the first 30h forecast and increases afterwards. It is observed that the 3DVar simulated storm tracks crossed the observed cyclone track (see next figure, Fig. 4.21). It is clear that as the 3DVar track approaches the actual observed track, the track error will decrease and subsequently increases as the simulated cyclone moves away from the observed system. The above is attributed for the decrease in 3DVar track error in the first 30h forecast. For TC Phailin and Hudhud, the 4DVar simulated track is consistently closer to the observed track and hence resulted in very low average track errors throughout the forecast length.

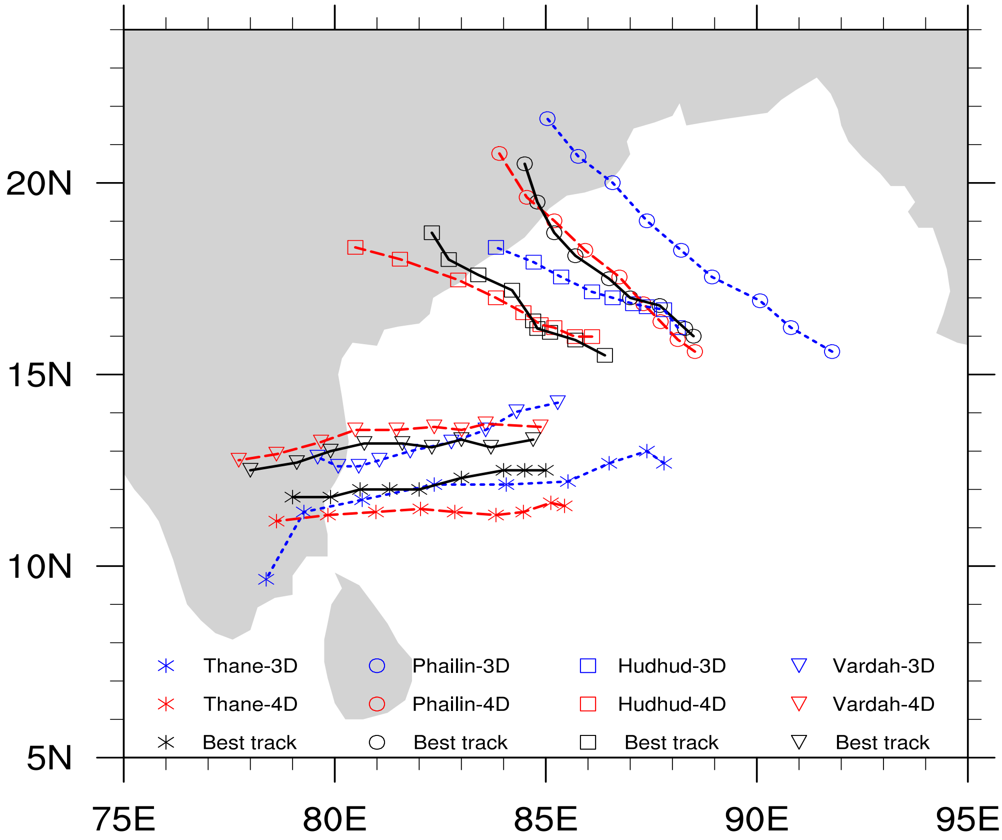

Fig. 4.21: Track of all the cyclones at their landfall phase for 3DVar and 4DVar runs with respect to IMD observations.

The average track error for TC Vardah (Fig. 4.20 d) shows entirely different behaviour from the other three cyclones and the results do not show any significant improvement with the 4DVar run. Up to 18h of free forecast, the 4DVar run is found to have slightly lower average error than that of the 3DVar error, and both the model results are nearly comparable towards the end of the free forecasts. While analyzing the model simulated track of TC Vardah, it is observed that, the 3DVar simulated track crossed the observed track (see Fig. 4.21). Such a crossing of the 3DVar track with the actual track has contributed to lower track errors for a few samples, and hence resulted in slightly lower average track error for the 3DVar run as compared to the 4DVar run. However, overall the 4DVar simulated track was consistently very close to the observed track.

To investigate how well the model could reproduce the landfall of the cyclones, the model simulated tracks (both 3DVar and 4DVar) are examined and the results are shown in Fig. 4.21. All the simulations are initialized from the last analysis cycle of all the respective cyclones. Apparently, the initial position of the TC Thane simulated by the 3DVar analysis is much farther from the observed position. while, the 4DVar analysis could reproduce the position of the storm centre quite well. As mentioned earlier, the 3DVar simulated track does cross the observed track after 24h of forecast and made landfall slightly closer than the 4DVar cyclone did. However, the rate at which the system moved differ significantly between the two model runs, with the 4DVar run being in better agreement with the IMD observations. Hence, the 4DVar simulated storm experienced landfall at exactly same time as that of the observed system. For TC Phailin and Hudhud, the positional error for the 3DVar run is much higher at the beginning of the forecast, while the same is much lower for the 4DVar run. The model simulated tracks for the 4DVar run, for both TCs Phailin and Hudhud, are very close to the observed tracks and the time of landfall is realistically close to the observed landfall time, resulting in much lower track errors. Unlike other three cyclones, the initial position simulated by the 3DVar analysis for TC Vardah is very close to the observed cyclone. Nevertheless, the 4DVar analysis has simulated the initial position more closely as compared with the observed initial position. As pointed out earlier, the 3DVar simulated track is found to cross the observed track after 12h forecast and moves away afterwards. On the other hand, the track simulated by the 4DVar run is found to be in good agreement with the IMD observed track consistently throughout the free forecast period and also experienced landfall, nearly at the observed location and time . A striking observation while analyzing the two model simulated tracks is that there are differences in the rate at which the system moves, in both the model runs. The 3DVar experiments fail to reproduce the actual speed of the movement of storms and thus contributes to large temporal error, even if they capture the location of storm centre at the initial time. In general, the 3DVar analysis fields also fails to reproduce the actual position of the storm centre, which will contribute to considerable track errors throughout the forecast. On the other hand, the 4DVar analysis is able to reproduce the initial position of the vortex fairly well and the forecasts utilizing the 4DVar analysis fields are simulating the speed of the storm also quite well, resulting in very low track forecast errors. The spatial and temporal error in the model simulated landfall of the storms are given in Table 4.4. It is quite evident that the forecasts from the 4DVar analysis fields have reproduced the landfall more accurately, both in space (location) and time (landing time). For all the TCs, the 4DVar forecasts have predicted the landfall time within 3h of the actual landing time.

Table 4.4: Error in landfall (LF) position and time for all the cyclones. +/- sign in temporal error indicates ahead/delay in landfall.

Cyclone and initial condition	Error in LF position (km)		Error in LF time (hrs)	
	3DVar	4DVar	3DVar	4DVar
Thane 12Z 28 December 2011	80	43	+ 3	0
Phailin 00Z 11 October 2013	196	28	+ 6	+ 3
Hudhud 00Z 11 October 2014	91	40	- 12	+ 3
Vardah 00Z 11 December 2016	63	36	- 6	+ 3

Among the four TCs examined in the study, the average track errors revealed significant error reduction for TC Phailin, having an average error of $\sim$100 km as compared to the 3DVar run which has an average track error of $\sim$250 km. Such a drastic reduction in cyclone track error indicates the potential positive impact of the 4DVar assimilation for cyclone forecasting. Furthermore, Fig. 4.19 (b) shows that the error in simulating the position of the vortex by the 3DVar analysis increased significantly after a few analysis cycles. Hence, the track forecast for TC Phailin has been examined closely to identify the reason behind such a higher track error for TC Phailin in the 3DVar run.

As pointed out earlier, the error in simulating the position of the vortex centre was comparable for the first three analyses (see Fig. 4.19 b). However, after the fourth analysis, the 3DVar positional error of the vortex centre started increasing while, that of 4DVar analysis started decreasing. Since the first guess or background field for the fourth analysis is based on the 6h forecast from the third analysis field, the analysis field and 6h forecast field from the third analysis have been examined closely. The wind field aloft the storm are known to influence the track of the cyclones and thus the 500 hPa wind fields corresponding to the third analysis field (valid at 18Z 08 October 2013) and 6h forecast from the same field (valid at 00Z 09 October 2013) are examined from both the model runs for TC Phailin and are depicted in Fig. 4.22.

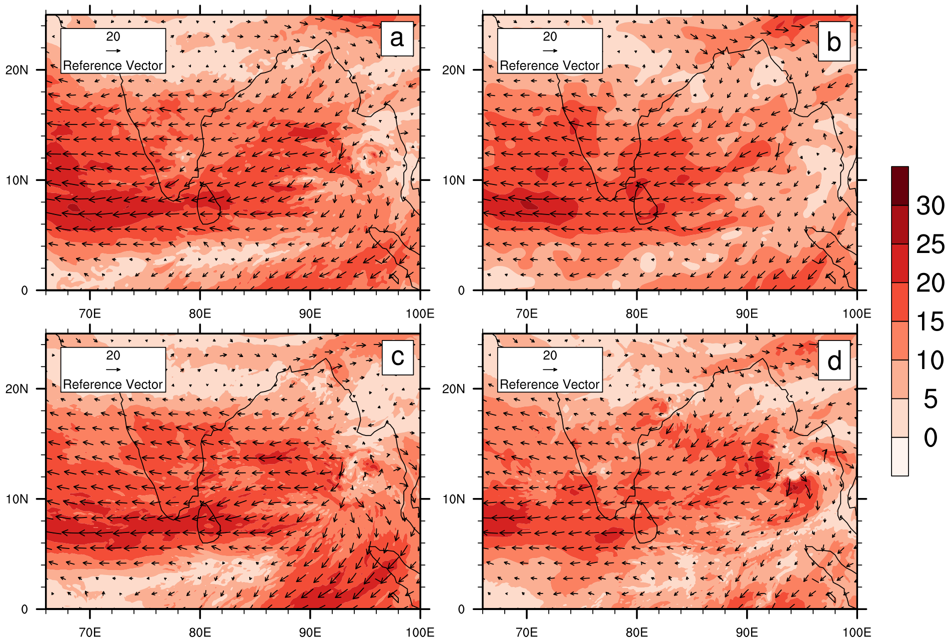

Fig. 4.22: 500 hPa wind fields from the analysis fields valid at 18Z 08 October 2013 (a,b) and 6h forecasts (c,d) from the same analysis field for 3DVar (a,c) and 4DVar (b,d) runs for TC Phailin.

Fig 4.22 depicts the 500 hPa wind fields corresponding to 3DVar (a,c) and 4DVar (b,d) analysis fields (a,b) and respective 6h forecast fields (c,d). The presence of north-easterlies over the south-eastern Bay of Bengal is seen clearly from both the 3DVar and 4DVar analysis fields (Fig 4.22 a and b). However, the above north-easterlies are considerably stronger in the 3DVar analysis field as compared to the 4DVar analysis field. Interestingly, these stronger north-easterlies in the 3DVar analysis field has further intensified after 6h, whereas, the weaker north-easterlies in the 4DVar analysis field weakened further in the next 6h (Fig 4.22d). It is postulated that the manifestation of these anomalous north-easterly winds in the 3DVar forecast fields would have contributed to the drift in the position of the storm away from the actual position, thus resulting in further degradation of the track forecasts from the subsequent 3DVar cycles.

4.9 Average improvement for all the cyclones

The present study has generated 12 analysis/forecast samples for TC Phailin and 13 each for the other three TCs, with a total of 51 analysis/forecast samples altogether for all the four TCs. A comprehensive analysis was carried out to get a statistics of, how well the 4DVar experiments can perform over their 3DVar counterparts in simulating the intensity and movement of the tropical cyclones.

Table 4.5: Absolute average error (AAE) in MSLP, MWS and track averaged for all the four TCs for 3DVar and 4DVar runs. Percentage of improvement with the 4DVar run over 3DVar run is given in brackets (significant at 99% confidence level.)

Forecast lead time	AAE in MSLP (hPa)		AAE in MWS (m/s)		AAE in Track (km)	
	3DVar	4DVar	3DVar	4DVar	3DVar	4DVar
00	12.75	6.34 (50)	11.77	6.62 (44)	185.1	86.9 (53)
06	14.45	8.45 (41)	12.66	8.49 (33)	195.1	83.8 (57)
12	15.20	10.14 (33)	12.03	10.22 (15)	198.4	94.1 (52)
18	15.98	11.16 (31)	12.01	8.94 (26)	192.2	106.6 (44)
24	16.98	12.59 (26)	11.98	9.61 (20)	179.9	122.2 (32)
30	17.86	13.66 (23)	11.08	10.35 (7)	174.5	130.5 (25)
36	18.27	14.27 (22)	11.06	10.05 (9)	175.5	136.5 (22)
42	18.07	14.16 (21)	10.24	10.11 (2)	222.1	147.0 (34)
48	17.41	14.47 (17)	9.87	8.49 (14)	234.3	155.4 (33)

The most striking feature is the improvement in the analysis fields (0h forecast). It is to be noted that approximately 50% improvement in reproducing accurately the intensity and position of the storm centre is achievable by employing the 4DVar DA methods for the simulation of TCs. However, as the forecast proceeds, the model errors are growing in time. The rate at which the model errors amplify is more for 3DVar experiments as compared to the 4DVar experiments. Interestingly, the intensity errors are growing at a higher pace than that of the track errors. The average track error for the 4DVar shows an improvement of ∼30% even for the 48h forecast. On the other hand, the MSLP error for the 4DVar run shows only ∼15% at 48h forecast lead time. Thus, the accurate prediction of intensity changes associated with TCs are a more challenging task even with the 4DVar DA methods, as compared with the track prediction. It is well recognized that the intensity simulation is more sensitive to smaller scale processes. The present study was performed in a domain having a 9 km horizontal resolution, which was nested over a parent domain of 27 km resolution. The errors associated with intensity simulation would have further decreased if the study had employed a domain with more finer horizontal resolution.

4.10 Summary

The studies conducted in the present chapter examined the relative performance of 3DVar and 4DVar DA methods in simulating the features associated with four TCs. Also, the improved performance of the 4DVar method over the corresponding 3DVar method is quantified by generating a large number of analysis/forecast samples for intensity and track simulation of the TCs. Cyclic assimilation were performed (at 6h interval) utilizing conventional surface and upper-air observations, satellite based wind observations and satellite radiance from AMSU, MHS and HIRS-4 for assimilation. The analyses fields available at 6h interval for all the TCs were subjected to short-range free forecasts (48h).

The study suggests that the 4DVar analysis fields perform significantly better than the 3DVar analysis fields by reproducing accurately the initial structure and intensity of the TCs. On an average, the 4DVar analysis (0h forecast) fields indicate significant (nearly 50%) improvement in simulating the position and intensity of the storm. The rate at which the model errors grow with time also decreases for the 4DVar experiments. It has been noted that the refinements in the moisture fields, due to the dynamics of moist-physics schemes incorporated in the tangent linear/adjoint models have a significant impact in simulating accurately the moisture convergence at the surface, which in turn contributed to improved rainfall simulation. Furthermore, the flow-dependent structure of analysis increments suggests that the implicit evolution of model errors due to the forward marching of tangent linear model and backward marching of the adjoint model contributes to a more realistic background error statistics. Interestingly, the 4DVar experiment realistically accounted for the rapid intensification phase of TC Phailin, during 00Z 10 to 00Z 11 October 2013, while the corresponding 3DVar run failed in simulating the intensification process appropriately. This suggests that, for rapidly evolving systems, the 3DVar DA method that treats all the available observation for a 3/6 hour window as observations at the analysis time, may be inadequate. The 3DVar method with FGAT may be able to provide improved results as compared with the simple 3DVar approach. Hence the asynoptic, satellite based observations can be appropriately utilized with the 4DVar DA method, which allows for observations to be assimilated at the exact time of the observations.

Chapter 5

BEC Sensitivity: Study on Heavy Rainfall Events

The present chapter investigates the sensitivity of the WRF-4DVar DA system to different formulations of background error covariance statistics. Three different BEC formulations are employed with different sets of control variables and the impact on the analysis and short-range forecasts for three heavy rainfall events over the Himalayan region, India are investigated. The results of the study shows that the multivariate BEC is performing better in simulating the heavy rainfall episodes.

5.1 Introduction

The error statistics of the forecast model is an integral component of the variational data assimilation system [73, 15]. Essentially, the background error covariance (BEC) matrix describes the nature of the forecast error corresponding to the non-linear model. The information assimilated at any point is *communicated* to the neighbouring grid points and other variables through the error relations described by the BEC matrix, thereby ensuring an improved dynamic balance in the analysis field. Hence, realistic and accurate representation of the model errors is extremely important. However, since the size of the state space of the NWP models is of the order of 10^7, calculation of an exact the BEC matrix (of size 10^{14}) is not computationally feasible. Hence, an approximated form of BEC will be estimated through the BEC modeling techniques [135].

Various BEC modeling methods utilize different control variables corresponding to different prognostic variables. Air temperature, relative humidity, surface pressure and a set of momentum variables are the commonly used control variables. For representing momentum, three sets of variables are employed by different NWP centers: (i) stream function and velocity potential, (ii) vorticity and divergence, and (iii) zonal and meridional wind components [136, 137, 23, 138, 139, 140, 141]. Among the three different choices, the commonly used momentum variables are the stream function and velocity potential (ψ and χ), primarily because the above is found to yield better performance for the large-scale DA applications [142]. However, studies have reported that there are potential disadvantages with the $\psi\chi$-based DA system. Since both ψ and χ have an inherent property of preserving the integral values of wind, the $\psi\chi$-based system may introduce analysis errors [76]. In addition, both the stream function and velocity potential are not usually the model variables. The conversion of model momentum variables (u, v) to stream function and velocity potential is a complex boundary value problem and there may be disagreements between the estimated wind field and actual wind field [76, 143]. Furthermore, for smaller scale DA applications the $\psi\chi$-based system is found to have demerits [76]. On the other hand, the BEC formulation based on uv-momentum variables is found to be more useful for high-resolution DA systems [80, 76].

In the real atmosphere, several balances exist among the variables (eg. the geostrophic balance). The most convenient way to incorporate these balances in BEC modeling is to approximate them in the form of linear regression relations, wherein the regression coefficients are estimated from an ensemble of forecast errors. One advantage of imposing the balance relations is that, the balanced part of the variable can be removed from their respective full field and thus one need to deal with the unbalanced part only [142, 80]. The DA system within the WRF model (WRFDA) supports the $\psi\chi$-based control variables. The balance relations are applied to velocity potential, temperature and surface pressure while, stream function and relative humidity are treated as full fields. An alternate BEC formulation was proposed by Krysta *et al.*[144] in 2009 by introducing balance relations for humidity fields and additional relations for temperature, surface pressure and velocity potential to yield a fully multivariate BEC [144, 77]. Wang *et al.* (2013) introduced an additional BEC formulation [145] by employing u and v as momentum variables (available from v3.7 onwards) instead of ψ and χ. The above BEC formulation utilizes air temperature, surface pressure and relative humidity as other control variables apart from the momentum variables. Unlike the other two formulations based on ψ and χ, the BEC formulation based on u and v treats all the variables as univariate.

Few studies have examined the relative performance of different BEC formulations in simulating various weather systems. It has been reported in the literature that the studies that have employed the multivariate $\psi\chi$-based BEC yielded improved rainfall forecasts [77]. A similar study performed to analyze the impact of two BECs ($\psi\chi$-based) on the simulation of monsoon depressions over India also found that the multivariate $\psi\chi$-based BEC provides better rainfall forecasts [78]. In a recent study conducted by Sun *et al.* (2016), the performance of $\psi\chi$-based BEC and uv-based BEC were investigated on the simulation of few heavy rainfall events that occurred over the Rocky mountain front range and the results of the study indicated that the uv-based BEC experiments yielded improved rainfall forecasts during 0-12h forecast period [80]. Most of the above mentioned studies are based on the 3DVar DA framework, that employs a static, homogeneous BEC. Utilizing advanced DA systems that can evolve the model errors such as 4DVar DA method to study the sensitivity of the WRFDA system to different formulations of BEC statistics may lead to interesting results. Hence, the present study aims to investigate the performance of the above-mentioned three BEC formulations on the simulation of three heavy rainfall events that occurred over the north Indian regions using the WRF model and its 4DVar DA module.

5.2 BEC formulation in WRFDA

The BEC matrix is usually estimated in a decomposed form $\mathbf{B} = \mathbf{U}\mathbf{U}^T$, to avoid the complexities in the inversion of the high-dimensional matrix. The analysis increment can then be represented in terms of the control variable as:

$$\mathbf{U}\mathbf{v} = \mathbf{x}^n - \mathbf{x}^b \tag{5.1}$$

where, $\mathbf{v}$ is the control variable. The operation $\mathbf{U}$ consists of three transforms and hence Eq. 5.1 can be written as:

$$\delta\mathbf{x} = \mathbf{x}^n - \mathbf{x}^b = \mathbf{U}\mathbf{v} = \mathbf{U}_p\mathbf{U}_v\mathbf{U}_h\mathbf{v} \tag{5.2}$$

where, $\mathbf{U}_h$ and $\mathbf{U}_v$ are horizontal and vertical transforms. The horizontal correlations are applied using recursive filters while the vertical correlations are applied by employing empirical orthogonal functions. The physical transform is designed to map the analysis increment in terms of control variables to the model variable space. The basis of the physical transform $\mathbf{U}_p$ is the regression coefficients between the control variables [47, 77].

In WRFDA, currently three sets of control variables for different BEC formulations are supported. For one $\psi\chi$-based BEC formulation ($\psi\chi$-BE), the control variables employed are stream function (ψ), unbalanced part of velocity potential (χ_u), unbalanced temperature (T_u), unbalanced surface pressure (Ps_u) and pseudo-relative humidity (rh_s). The balance relations for the unbalanced variables are:

$$\begin{aligned}
\chi_u(i,j,k) &= \chi(i,j,k) - \alpha_{\psi\chi}(i,j,k)\psi(i,j,k) && (5.3)\\
T_u(i,j,k) &= T(i,j,k) - \sum_{l=1}^{N_k} \alpha_{\psi T}(i,j,k,l)\psi(i,j,l) && (5.4)\\
Ps_u(i,j) &= Ps(i,j) - \sum_{l=1}^{N_k} \alpha_{\psi Ps}(i,j,l)\psi(i,j,l) && (5.5)
\end{aligned}$$

where, α is the regression coefficients between the variables indicated with subscripts, i, j represent horizontal grids while k, l represent vertical grids respectively.

A different BEC formulation ($\psi\chi$-MBE), which is a modification of $\psi\chi$-BE utilizes a different set of control variables through introduction of additional regression relations to the unbalanced variables and defining balance relations for humidity variable [144, 77]. The balance relations for this multivariate BEC ($\psi\chi$-MBE) formulation are given by:

$$\begin{aligned}
rh_u(i,j,k) &= rh(i,j,k) - \sum_{l=1}^{N_k} \alpha_{\psi rh}(i,j,k,l)\psi(i,j,l)\\
&- \sum_{l=1}^{N_k} \alpha_{\chi_u rh}(i,j,k,l)\chi_u(i,j,l)\\
&- \sum_{l=1}^{N_k} \alpha_{T_u rh}(i,j,k,l)T_u(i,j,l)\\
&- \alpha_{Ps_u rh}(i,j,k)Ps_u(i,j) && (5.6)
\end{aligned}$$

$$\begin{aligned}
\chi_u(i,j,k) &= \chi(i,j,k) - \alpha_{\psi\chi}(i,j,k)\psi(i,j,k) && (5.7)\\
T_u(i,j,k) &= T(i,j,k) - \sum_{l=1}^{N_k} \alpha_{\psi T}(i,j,k,l)\psi(i,j,l)\\
&- \sum_{l=1}^{N_k} \alpha_{\chi_u T}(i,j,k,l)\chi_u(i,j,l) && (5.8)
\end{aligned}$$

$$\begin{aligned} Ps_u(i,j) &= Ps(i,j) - \sum_{l=1}^{N_k} \alpha_{\psi Ps}(i,j,l)\psi(i,j,l) \\ &- \sum_{l=1}^{N_k} \alpha_{\chi_u Ps}(i,j,l)\chi_u(i,j,l) \end{aligned} \tag{5.9}$$

From the above-mentioned balance relations pertaining to the $\psi\chi$-MBE formulation, it is evident that the temperature and surface pressure fields are related to the velocity potential field via the newly introduced regression coefficients. Hence, the wind field may be influenced by the assimilation of pressure or temperature or observations through the divergent part of wind in the above formulation. Since the relative humidity is represented in the multivariate form, assimilation of any other observation can impact the moisture field and vice versa [77].

The third formulation (uv-BE) utilizes zonal velocity (u) and meridional velocity (v) as momentum variables together with temperature (T), surface pressure (Ps) and pseudo-relative humidity (rh_s) as the control variables [145]. In this formulation, multivariate correlations among the variables are not considered and thus all the variables are completely univariate in nature. Accordingly, the physical transform $\mathbf{U}_p$ employed in estimating the BEC (see Eq. 5.2) will be devoid of any statistical balance relations [145]. More detailed description on the formulation of all the above BECs can be found in [47, 144, 77, 145].

5.3 Case description

The present study focuses on a few heavy rainfall events that occurred over the foothills of Himalaya. A comprehensive study was conducted by Vellore *et al.* (2016) by examining the characteristic dynamical features of the interaction between Indian summer monsoon and extratropical circulation that result in torrential rainfall episodes over the Himalayan region [146]. After performing a detailed empirical orthogonal function/principal component analysis, 34 heavy rainfall events with large normed metric of the major two principal components were chosen in the above study. Three extreme rainfall cases from the above chosen list of 34 events are selected for the present study. The three selected cases are: (i) Case 1: 16-18 June 2013 (Uttarakhand event), Case 2: 15-17 August 2011, Case 3: 19-21 September 2008. All the above three cases experienced extremely heavy rainfall resulting from the tropical-extratropical circulation interaction. For instance, during 14-18 June

2013, Uttarakhand state received torrential rainfall that led to unprecedented devastation which resulted in death toll of about 5000 persons. The Uttarakhand heavy rainfall event of 2013 resulted from an interaction between the southward moving extra-tropical systems and the tropical monsoonal circulation [146, 147]. The accumulated rainfall during 16-18 June 2013 was 492 mm over Dehradun and 485 mm over Nainital [147].

5.4 Model configuration

The WRF model (version 3.8.1) was configured with three nested domains having horizontal resolution of 27 km, 9 km, and 3 km and 36 terrain following vertical coordinates. The domain used in the study is illustrated in Fig. 5.1. The parent (outer) domain was configured with 375 grids in east-west direction and 250 grids in the south-north direction.

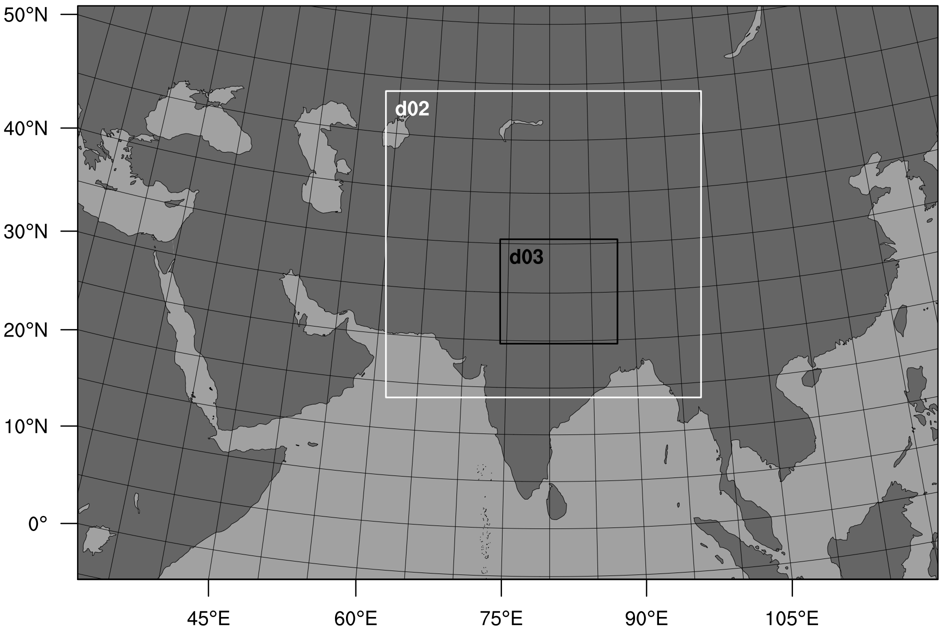

Fig. 5.1: The model domain used in the study.

The convective processes for the first and second domain are parameterized using the Kain-Fritsch scheme, that accounts for deep and shallow convection using mass flux approach. For the finer 3 km domain, explicit convection is considered and no convective scheme is employed. Since the detailed study by Dimri and Chevuturi in 2014 [148] on the simulation of western disturbances over Himalayan region has shown that the Eta-Ferrier scheme of microphysics parameterization provides improved rainfall forecasts, the present study also utilized the same parameterization scheme for representing the microphysical processes. The boundary layer turbulence is specified following the Yonsei University scheme, while the Noah land surface model is used for the land surface physics. The radiative processes are parameterized following the rapid radiative transfer model and Dudhia scheme for longwave and shortwave radiations respectively. More details on the model configurations are given in Appendix A.

5.5 Experimental design

Numerical simulations were performed for all the three cases by utilizing the 4DVar DA technique (WRFDA v3.8.1). Three experiments corresponding to three different BECs were performed for each rainfall case and are listed in Table 5.1.

Table 5.1: List of experiments and the corresponding control variables.

Experiment	**Control variables used**
$\psi\chi$-BE	$\psi, \chi_u, T_u, Ps_u, rh$
$\psi\chi$-MBE	$\psi, \chi_u, T_u, Ps_u, rh_u$
uv-BE	u, v, T, Ps, rh

Five cyclic assimilations at 6h intervals were carried out for each case after a short spin-up run for 6h. A free forecast lasting 48h was initiated from the last analysis. Initial and boundary conditions were derived from NCEP-GFS 24h forecast fields. Period of assimilation and duration of the free forecast run for each of the three case studies are given in Table 5.2.

Table 5.2: List of the cases studied and experimental details

Case	**Period of assimilation**	**Free forecast**
Case 1	00Z 15 - 00Z 16 June 2013	00Z 16 - 00Z 18 June 2013
Case 2	00Z 14 - 00Z 15 August 2011	00Z 15 - 00Z 17 August 2011
Case 3	00Z 18 - 00Z 19 September 2008	00Z 19 - 00Z 21 September 2008

Data assimilated include the surface and upper-air observations from all conventional platforms such as radiosonde, ship, buoy, pilot balloons, ground stations etc. In addition to the above conventional observations, atmospheric motion vectors from Meteosat-7 and ocean surface winds from Advanced Scatterometer are also utilized for assimilation. Typical number of observations utilized for assimilation over the study domain valid at 00Z 15 June 2013, which corresponds to Case 1 of the present study are given in Table 5.3.

Table 5.3: Typical number of observations used for assimilation at 00Z 15 June 2013.

Observation type	**Number of observations used**
Synoptic	3375
Soundings	4041
Metar report	2185
Pilot	10
Buoy	29
Ship	79
AMVs	8065
Scatterometer	8537

5.6 Single observation assimilation experiments

It was thought worthwhile and meaningful to perform simulations with the assimilation of single observation before conducting real data assimilation experiments. Single observation experiments provide a good understanding on the response of the assimilation system to the observations assimilated. More importantly, the effect of the BEC matrix in spreading the assimilated information can be well understood since the analysis increment is highly sensitive to the nature of the formulation utilized for estimating the BEC matrix. In the case of assimilation of a single observation y_i, the 4DVar solution [49] can be written as:

$$\mathbf{M}(x^a - x^b) = (\mathbf{MBM}^T)_i(\sigma_b^2 + \sigma_o^2)^{-1}(y_i - x_i) \tag{5.10}$$

where σ_b and σ_o are background and observation errors respectively. $\mathbf{M}$ is the tangent linear model and $\mathbf{B}$ is the BEC matrix, x^a is the analysis field and x^b is background field. The most striking point is that the analysis increment is directly proportional to the BEC matrix. In addition, the term $\mathbf{MBM}^T$ indicates that the model dynamics (in terms of the tangent linear model $\mathbf{M}$ and its adjoint $\mathbf{M}^T$) do play a key role in determining the analysis increment. The above discussion shows the role of the 4DVar DA system in accounting for evolution of the model errors.

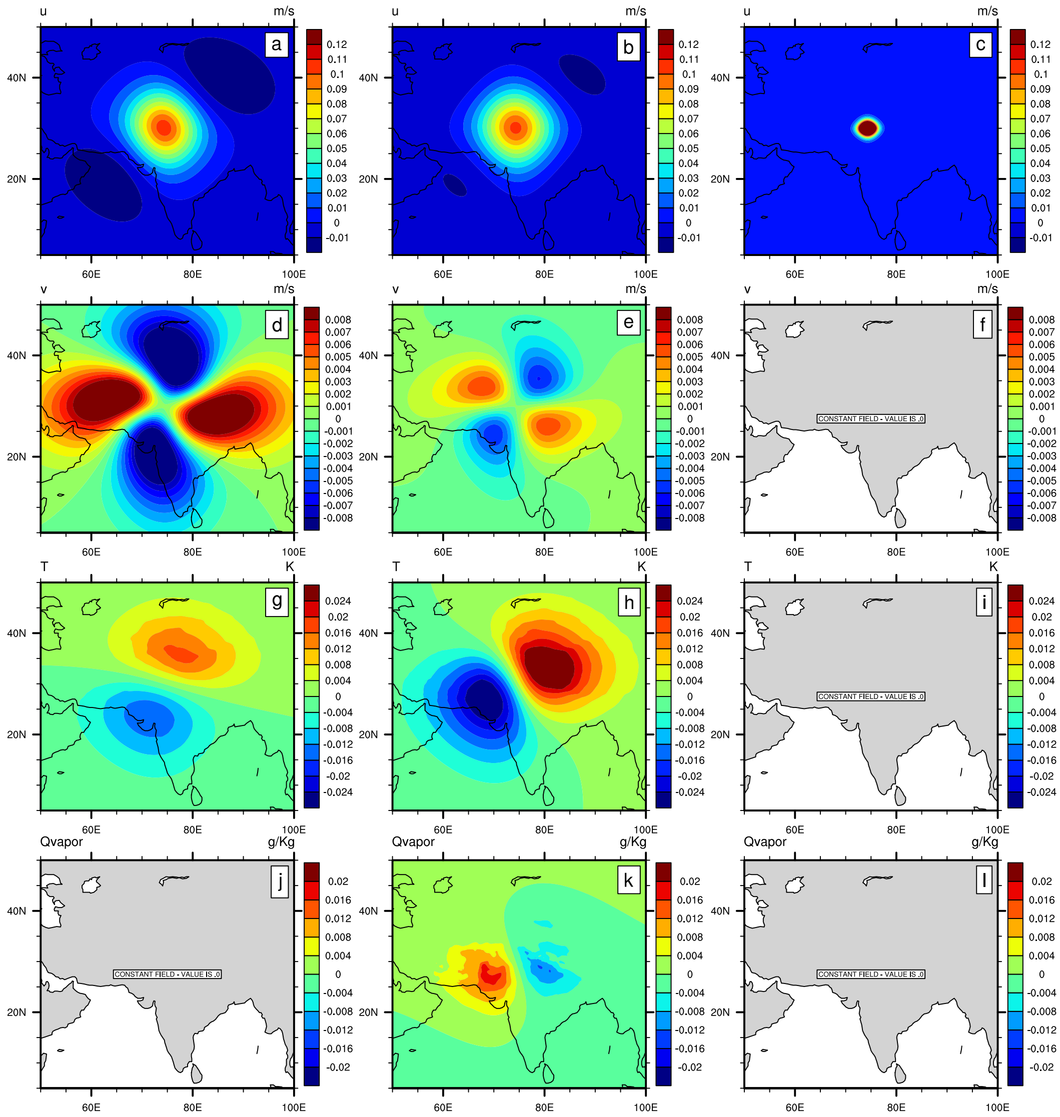

Fig. 5.2: Analysis increment corresponding to zonal wind (a-c), meridional wind (d-f), temperature (g-i) and water vapor mixing ratio (j-l) for single 'u' observation assimilation utilizing $\psi\chi$-BE (a,d,g,j), $\psi\chi$-MBE (b,e,h,k) and uv-BE (c,f,i,l).

Single observation assimilation studies were conducted by assimilating a single 'u' observation at the first model level at the center of the first domain. The assimilated 'u' observation was characterized with innovation and error statistics of 1 ms^{-1}. The analysis increments for single 'u' observation experiment are shown in Fig. 5.2. All the fields are valid at 00Z 16 June 2013. In the above figure, results for $\psi\chi$-BE experiments are shown in Fig. 5.2(a, d, g, j), which shows that the analysis increments in the zonal wind has slight negative increment values in the north-east and south-west sector of the observation location. Since only one observation with positive innovation is being assimilated, the negative increments could be attributed to the analysis errors. It is known that the analysis errors are inherent to the $\psi\chi$-based BEC formulation, since both the stream function and the velocity potential are integrals of the velocity fields [76]. The $\psi\chi$-based assimilation system attempts to maintain integrals of the analysis velocity fields with the same values as the background velocity fields. Hence, over the regions where the analysis wind fields depart from the background wind fields, the integrals will have to adjust themselves, causing the appearance of opposite analysis increments in the neighbouring wind fields [76]. Ideally, the BEC should take care of these negative analysis increments. In some previous studies with the 3DVar DA system, the negative increments in the wind fields were markedly dominant [77, 78]. Considering that the magnitude of negative analysis increments are very small, it can be inferred that the $\psi\chi$-BE formulation with the 4DVar DA system provides for improved adjustments to the analysis increments in the wind fields as compared to the 3DVar DA system.

Assimilation of single 'u' observation results in increments in zonal and meridional wind, and temperature fields in the $\psi\chi$-BE experiment. Since no balance relations are incorporated for the humidity variable in the $\psi\chi$-BE formulation, assimilation of wind observation did not impact the moisture fields in the $\psi\chi$-BE experiments (Fig. 5.2j). Positive and negative increments of equal magnitude can be observed in the meridional wind field (Fig. 5.2d) as a result of 'u' wind assimilation. The increments in the meridional wind field has a magnitude one order less than the respective increments in the zonal wind field. The analysis increments in the meridional wind fields have induced convergence/ divergence fields that has resulted in slight rotation about the location of the observation. Very weak analysis increments in temperature field are also seen due to the 'u' assimilation with $\psi\chi$-BE experiment (Fig. 5.2g).

In the case of $\psi\chi$-MBE experiments, the analysis increments in the zonal wind field (Fig. 5.2b) indicate that the negative increments observed over the north-east region and south-west region of the observation location for the $\psi\chi$-BE experiment are no longer dominant. The increments in the meridional wind field associated with the $\psi\chi$-MBE are much weaker in magnitude as compared to the $\psi\chi$-BE experiment. Also, the increment pattern in temperature field appears stronger with the $\psi\chi$-MBE as compared to the $\psi\chi$-BE experiment. More importantly, it is clear from Fig. 5.2(k) that the assimilation of 'u' wind has impacted the humidity field as well. It can be discerned that, when the zonal wind with an innovation of 1 ms^{-1} is assimilated as a single point observation, a maximum increment of magnitude 0.02 gkg^{-1} is produced in the moisture field due to the multivariate nature of the humidity variable in the $\psi\chi$-MBE experiment. It is interesting to note that the additional regression coefficients introduced in the balance relations pertaining to the temperature field in the $\psi\chi$-MBE formulation have resulted in significant adjustment in the temperature increment when compared with the $\psi\chi$-BE experiment.

Finally, the results for the uv-BE experiment reflects the univariate nature of the control variables. Since the uv-BE formulation does not account for any physical relationship among the control variables, assimilation single 'u' wind is expected to influence the zonal wind field only. It is therefore not surprising that assimilating a single 'u' observation employing the uv-BE experiment had no impact on the other variables (Fig. 5.2 f, i, l). It is important to note that, unlike in the case of other two BE experiments that produced somewhat weak and horizontally extensive patterns of increments, the uv-BE experiment yields very strong, localized increments. To be precise, while the $\psi\chi$-BE and $\psi\chi$-MBE experiments yielded maximum increments of 0.106 ms^{-1} and 0.103 ms^{-1} in the zonal wind field, the uv-BE experiment resulted in a maximum increment of 0.554 ms^{-1} in the zonal wind field. Similarly, the spatial extent of the spread of information for both $\psi\chi$-BE and $\psi\chi$-MBE are approximately $20^\circ \times 20^\circ$ (lat×long), whereas the spread for uv-BE is limited to approximately $5^\circ \times 5^\circ$. It is thus clear that there is a marked difference in the information spread among the $\psi\chi$-based BECs and the uv-based BEC. Also, the analysis increments observed in the present study with 4DVar DA system using the single observation assimilation are notably different in structure and in magnitude from such similar studies performed with 3DVar DA method [77, 78]. The implicit flow dependency, which is inherent to the 4DVar system could have contributed for the differences seen in the present single observation study.

Since the results of the single observation assimilation experiment revealed significant differences in the spread of information assimilated among the three formulations, the length scales of different variables as a function of vertical mode are analyzed and are illustrated in Fig. 5.3. The above length scales are inherent to the various BEC formulations and are expected to directly impact the spread of assimilated observation to the neighbouring locations. Vertical variation of momentum variables are shown in Fig. 5.3 (a and b). In the case of $\psi\chi$-BE and $\psi\chi$-MBE, momentum variables are ψ and χ, while for uv-BE, the momentum variables are u and v. The most striking observation is the large difference in

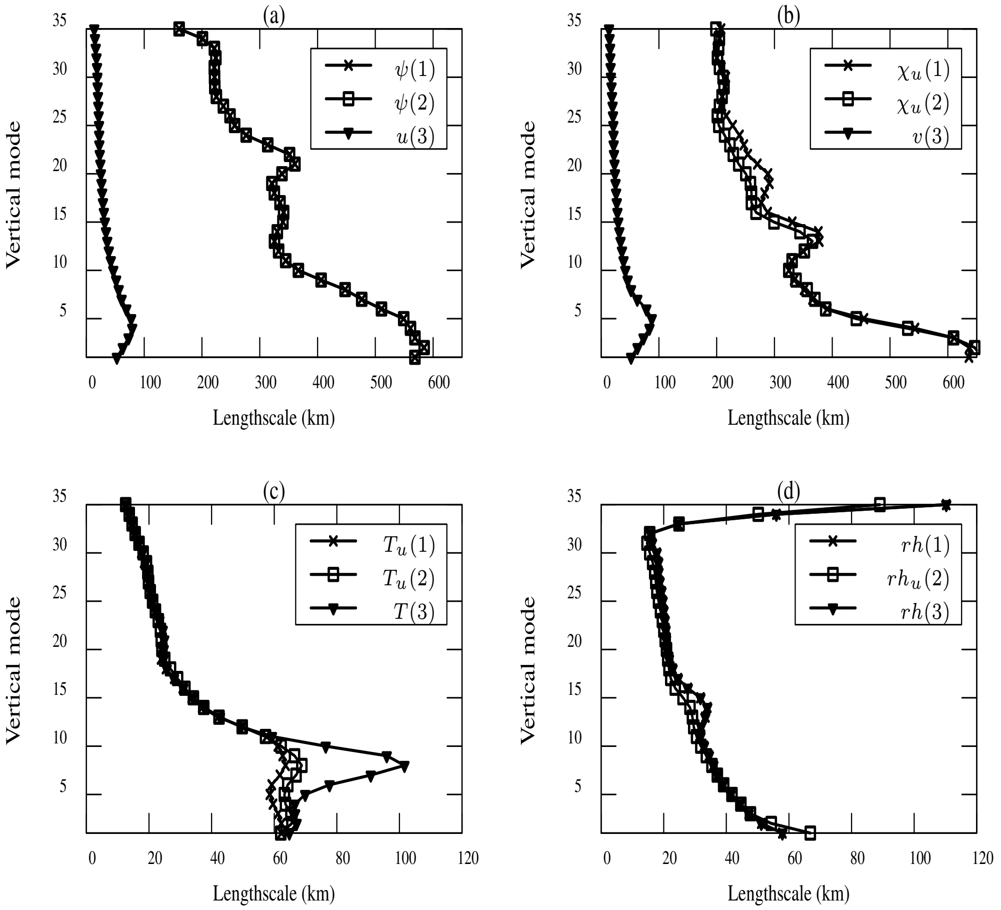

Fig. 5.3: Vertical variation of length scale (in km) for all control variables. 1, 2, 3 given in brackets for legends to indicate their respective BEC formulations; 1 for $\psi\chi$-BE, 2 for $\psi\chi$-MBE and 3 for uv-BE.

length scales of the momentum variables among the $\psi\chi$-based experiments and uv-based experiment. The length scales of ψ and $\chi_u$in $\psi\chi$-BE and $\psi\chi$-MBE peak near the surface do not show any differences and overlay over one another. The maximum length scale for both the momentum variables (ψ and χ) in $\psi\chi$-BE and $\psi\chi$-MBE is approximately 600 km, and is seen near the second vertical mode. On the other hand, the maximum length scale corresponding to the momentum variables with the uv-BE is $\sim$100 km, and the same is observed near the fifth vertical mode. The other two control variables, viz, temperature and humidity variables do not show any marked differences in their respective length scales with different BEC formulations, except for the temperature variable with uv-BE, which peaks at the eighth model level with a length scale of 100 km. Above the tenth vertical level, the length scales for temperature for all the experiments are almost identical. At the first level, the length scale for humidity is slightly higher for $\psi\chi$-MBE formulation as compared with the other two formulations. Hence it is interesting to note that the additional balance relations introduced in $\psi\chi$-MBE formulation does not alter the length scale of the original BEC formulation.

The momentum variables within both the $\psi\chi$-BE and $\psi\chi$-MBE formulations tend to propagate the assimilated information to a larger area, however the u and v variables within the uv-BE formulation produces very intense and sharp increments that are confined to a small neighbourhood of the location of assimilated observation. The above mentioned observation may be attributed to the inherent features of each of the BEC formulations and the above is expected to get reflected in the real data assimilation as well. Hence it is expected that the results for the the analysis fields which utilize each one of the above three BEC formulations may differ markedly and hence can thereby influence the forecasts.

5.7 Real data assimilation experiments

For each of the three heavy rainfall cases, five cyclic assimilations were performed within a 24h period. The evolution of cost function with iterations during the minimization for each of the three BE experiments is shown in Fig. 5.4 (for first assimilation cycle of Case 1).

The numerical value of the cost function for the uv-BE experiment differs significantly from the other two experiments and has a lower value than the cost function value for both $\psi\chi$-BE and $\psi\chi$-MBE experiments. However, there are no discernible differences in the cost function value among the $\psi\chi$-BE and $\psi\chi$-MBE experiments. Regarding the computation time, the uv-BE experiment required few more iterations than the $\psi\chi$-based experiments. The $\psi\chi$-MBE experiment converged relatively faster (26 iterations) while the $\psi\chi$-BE experiment converged in 28 iterations. Whereas, the uv-BE experiment converged only after 32 iterations. It is interesting to note that the additional balance relations imposed for humidity analysis in the $\psi\chi$-MBE formulation did not pose additional difficulties in the minimization process and in fact, the convergence occurred for the $\psi\chi$-MBE formulation, even faster as compared to the $\psi\chi$-BE formulation. The above-mentioned behaviour in the minimization process is consistently seen for all the assimilation cycles performed in the present study.

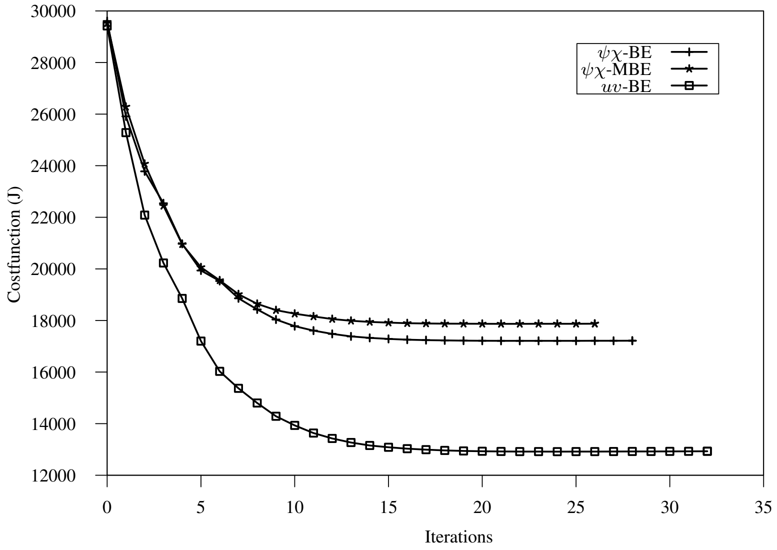

Fig. 5.4: Time evolution of cost function for $\psi\chi$-BE, $\psi\chi$-MBE and uv-BE experiments during minimization process for the first assimilation cycle of Case 1.

5.7.1 Improvement in the analysis fields

The impact of assimilation can be examined by analyzing the analysis increment (analysis minus first guess). Fig. 5.5 shows the analysis increment in zonal and meridional wind components, temperature, and water vapor mixing ratio at 850 hPa due to first assimilation cycle corresponding to $\psi\chi$-BE (left panel) $\psi\chi$-MBE (middle panel) and uv-BE (right panel) experiments for Case 1. It is evident that, there are no significant differences in the pattern of the analysis increment among the three experiment results, however, the magnitude of the increments do vary. For instance, the analysis increments in zonal and meridional wind components are mostly negative, and a closer examination will reveal that the magnitude of negative increments from the uv-BE experiment are relatively lower as compared to the increments from both the $\psi\chi$-BE and $\psi\chi$-MBE experiments. The above observation is true for temperature fields as well (Fig. 5.5 g-i). It is interesting to note that, although the BEC employed in the uv-BE experiment is completely univariate in nature, the overall pattern of analysis increments in all the four meteorological fields do not show notable differences from the analysis increment pattern of the experiment that utilized fully multivariate BEC ($\psi\chi$-MBE). The above finding shows that the 4DVar DA system is capable of account for the multivariate correlations among the analysis variables implicitly, owing to the dynamics of the tangent linear/adjoint model that is involved in the minimization process.

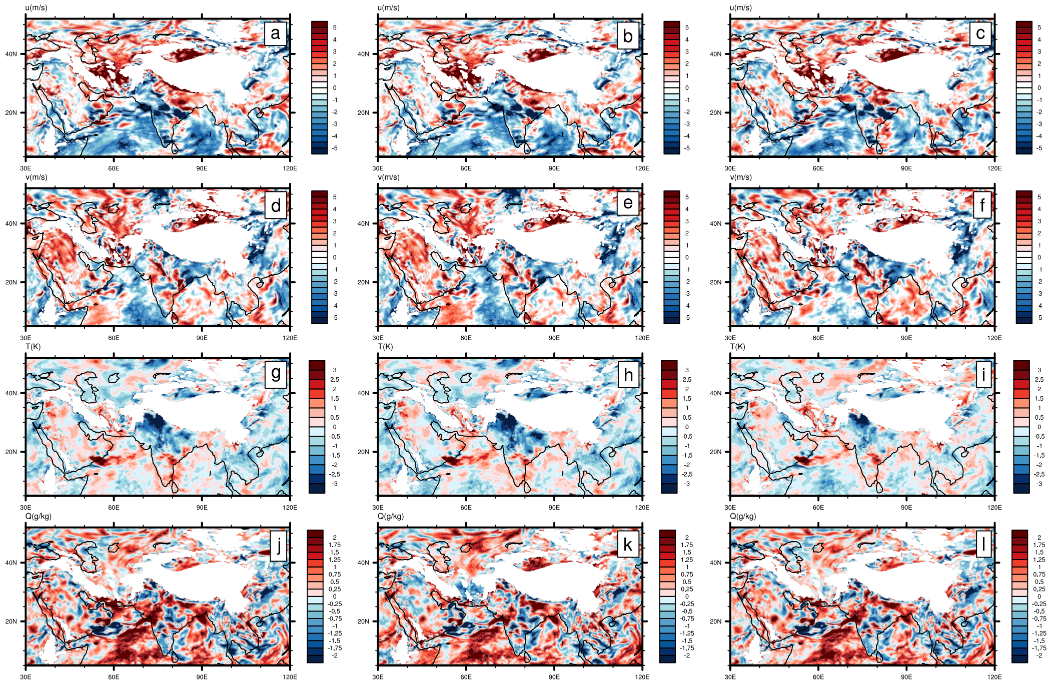

Fig. 5.5: Analysis increment in zonal wind (a,b,c), meridional wind (d,e,f), temperature (g,h,i), and water vapor mixing ratio (j,k,l) at 850 hPa for $\psi\chi$-BE (a,d,g,j) $\psi\chi$-MBE (b,e,h,k) and uv-BE (c,f,i,l) experiments corresponding to Case 1.

Fig. 5.6 depicts the increments in geopotential height fields at 200 hPa for the first assimilation cycle and for each of the three rainfall case. The analysis increments for Case 1 are shown in top panel, the same for Case 2 in middle panel and the bottom panel shows the increments for Case 3. In general, there are no marked differences in the geopotential height increments among the $\psi\chi$-BE experiments (a, d, g) and $\psi\chi$-MBE experiments (b, e, h). On the other hand, the impact on the geopotential height with the uv-BE experiments (c, f, i) are notably different from the other two experiments, with positive increments of larger magnitude are seen over most of the region. Over the northern region of the domain, negative increments with magnitude 300 m (blue and green shading) are seen for the $\psi\chi$-BE and $\psi\chi$-MBE experiments. However, the above region is characterized with very little analysis increments (-50 to 50 m, white shading) in the uv-BE experiment. Also, over the Bay of Bengal region, for both the $\psi\chi$-BE and $\psi\chi$-MBE experiments, mostly negative increments (green shading) are seen for Case 3. For the uv-BE experiment associated with Case 3, the negative increments over the Bay of Bengal have reduced considerably

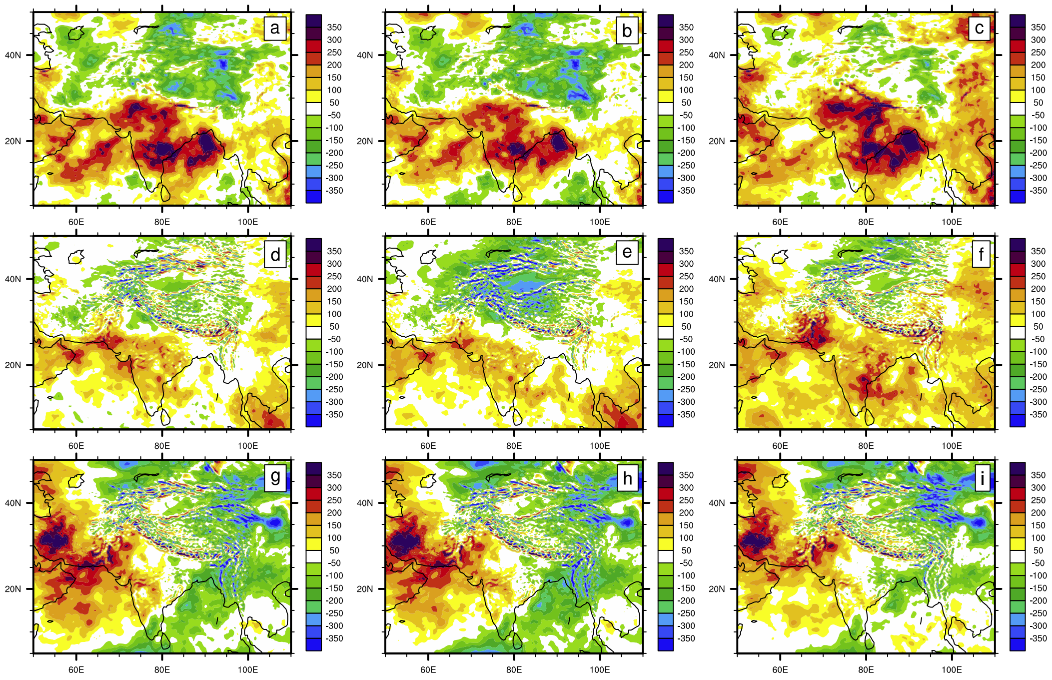

Fig. 5.6: Analysis increment in geopotential height at 200 hPa for $\psi\chi$-BE (a,d) $\psi\chi$-MBE (b,e) and uv-BE (c,f) experiments corresponding to Case 1 (a-c), Case 2 (d-f), and Case 3 (g-i).

and have been replaced by positive increments (white and yellow shading). Similarly for the other two cases also, the uv-BE experiment show a clear positive bias in geopotential height increments, i.e, the geopotential height fields simulated with the uv-BE experiments are quite higher than that of $\psi\chi$-BE and $\psi\chi$-MBE experiments. Differences in simulating height fields at 200 hPa will directly affect the simulation of the intensity and position of the subtropical jet stream. Since the trough associated with the jet streams are known to play a key role in modulating the mountainous heavy rainfall events, the differences in the height fields simulated by the three BEC experiments will definitely impact the forecasts.

The sensitivity of the different BEC formulations and hence the impact of assimilation can also be gainfully studied by examining the closeness of the analysis fields with the observations. Fig. 5.7 depicts the vertical variation in RMS-fit error in zonal wind (a), meridional wind (b), air temperature (c), and water vapor mixing ratio (d) with respect to radiosonde observations for Case 1. RMS-fit error is computed by averaging all the analysis fields for each rainfall case. Since results for all the three cases investigated in this study are found to be similar in nature, figures for RMS-fit error is shown only for Case 1 here.

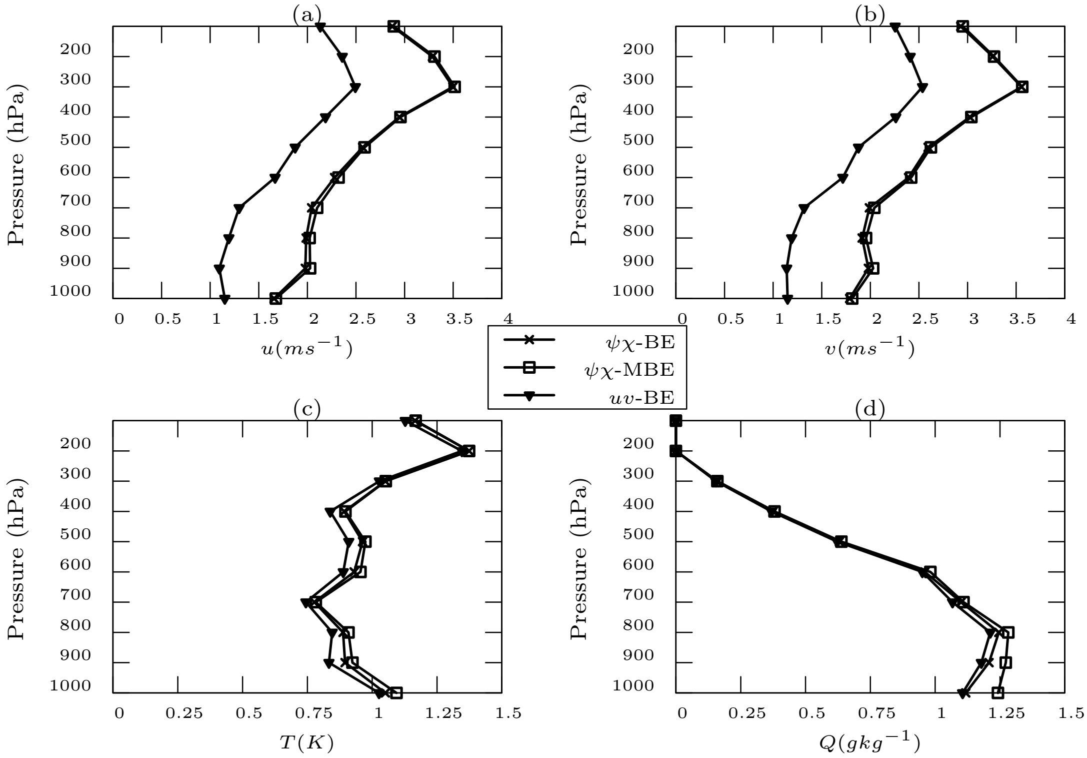

Fig. 5.7: RMS fit to radiosonde observations for u wind (a), v wind (b), temperature (c) and water vapor mixing ratio (d) for all the three BEC experiments.

Evidently, it is seen from Fig. 5.7 that the uv-BE analysis fields have consistently lower error value in all the four meteorological fields (u, v, T, and Q) and hence are closer to the radiosonde observations. The uv-BE analysis fields provides for a minimum error reduction of ~0.5 ms^{-1} in the zonal and meridional wind fields as compared with both $\psi\chi$-BE and $\psi\chi$-MBE analysis fields. Temperature field also shows slightly lower error for the uv-BE experiment. Below 500 hPa, the analysis humidity field for the uv-BE experiment shows lower error with respect to the radiosonde observed humidity profiles. Interestingly, there are no significant differences between the $\psi\chi$-BE analysis field and $\psi\chi$-MBE analysis fields in all the four meteorological variables. A closer examination reveals that the RMS-fit errors are slightly higher for the $\psi\chi$-MBE analysis fields as compared with the $\psi\chi$-BE analysis fields. Furthermore, the deviation of the analysis fields from the synoptic surface observations and satellite based wind observations (AMVs) have also been examined in terms of the O-A (observation minus analysis) fields for the first assimilation cycle of Case 1 and are shown in Fig. 5.8.

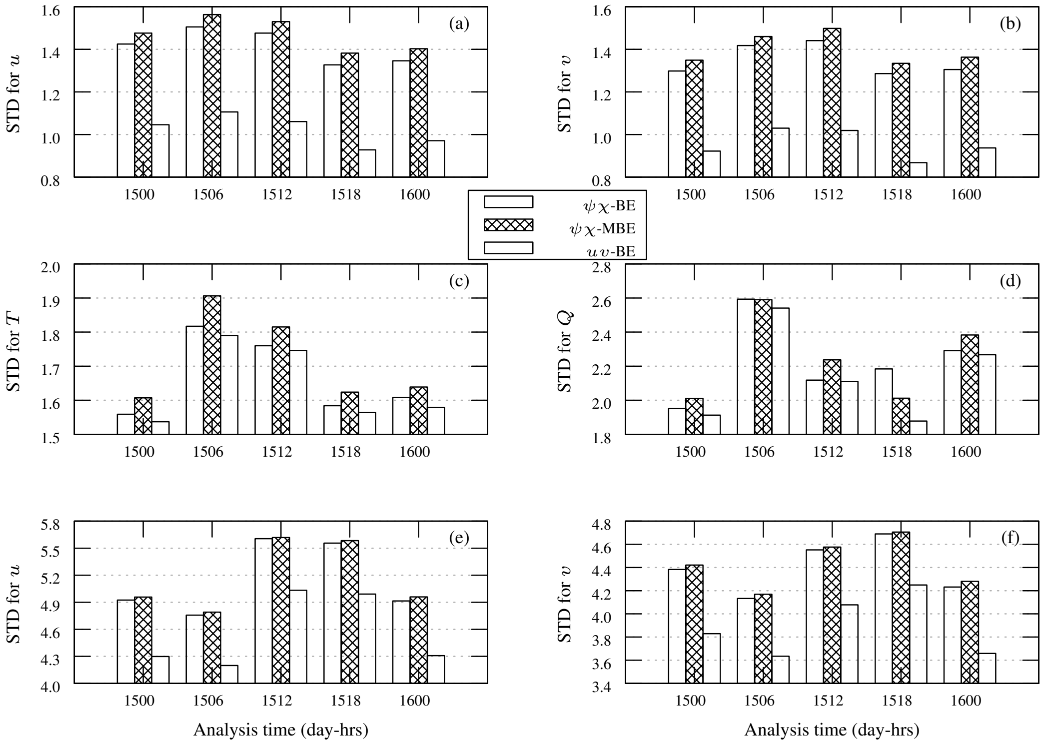

Fig. 5.8: Standard deviation for the analysis fields with respect to surface observations for u wind (a), v wind (b), temperature (c) and mixing ratio (d) for Case 1. Standard deviation with respect to AMVs for u wind (e) and v wind (f) are also shown for Case 1.

Standard deviation in the O-A fields corresponding to zonal wind (a), meridional wind (b), temperature (c), and water vapor mixing ratio (d) from surface observations for all the three experiments are given in Fig. 5.8 for Case 1. Additionally, the standard deviation in the O-A fields with respect to AMVs are also shown for Case 1 in the same figure: for zonal wind (e) and meridional wind (f). Consistent with the inferences drawn from the previous figure (Fig. 5.7), standard deviations for the uv-BE analysis field are considerably lower for all the four meteorological variables as compared with both $\psi\chi$-BE and $\psi\chi$-MBE analysis fields. The above results are found to be consistent throughout the entire analysis period. It is important to note that the reduction in standard deviation for the wind fields (Fig. 5.8 a, b, e, f) are more marked than that of temperature and humidity fields. Also, as compared to the results of $\psi\chi$-BE experiment, the standard deviation for $\psi\chi$-MBE experiment in all the meteorological variables are slightly higher, which indicates that the $\psi\chi$-MBE analysis fields are slightly away from the observed fields.

5.8 Forecast verification

The model's skill in accurately reproducing the rainfall associated with each convective events is examined by verifying the model simulated rainfall with respect to TRMM rainfall estimates (3B42).

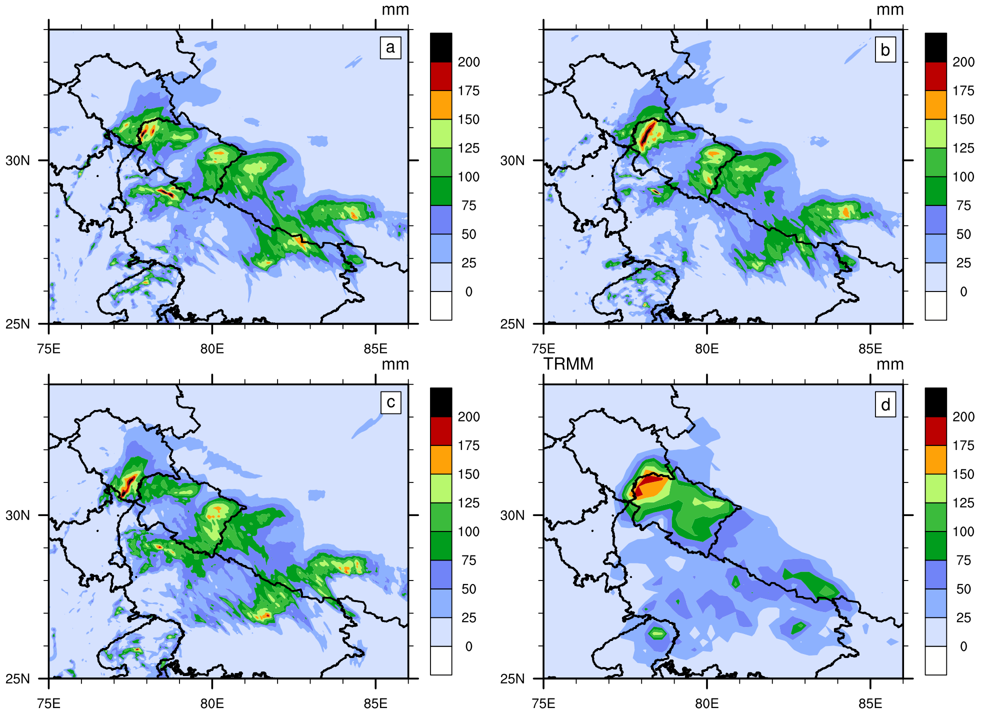

Fig. 5.9: 24h accumulated rainfall for $\psi\chi$-BE (a), $\psi\chi$-MBE (b), uv-BE (c), and TRMM estimate (d) for Case 1. Fields are valid at 00Z 17 June 2013.

In the case of Uttarkhand event (Case 1), larger number of casualties were associated with the intense rainfall that occurred on 16 June 2013. Fig. 5.9 shows the 24h accumulated rainfall for Uttarakhant heavy rainfall event, valid at 00Z 17 June 2013, corresponding to $\psi\chi$-BE run (a), $\psi\chi$-MBE run (b), uv-BE run (c), and TRMM estimate (d). From the TRMM estimate (Fig. 5.9 d), it is clear that the most intense downpour occurred over the state boundary of Uttarakhand state and Himachal Pradesh (near Kedarnath). Fig. 5.9 indicates that, both the $\psi\chi$-BE and $\psi\chi$-MBE runs have reproduced the location and intensity of maximum rainfall reasonably well. However, the uv-BE run simulated the maximum rainfall to the west of the actual maximum rainfall location.

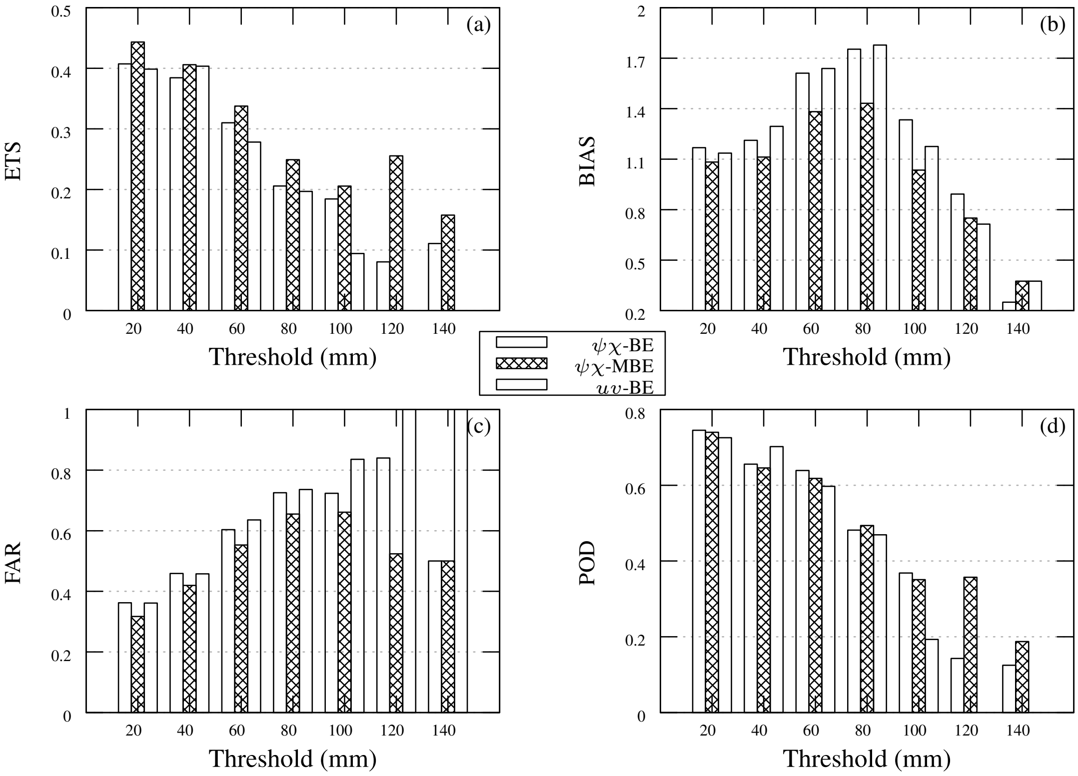

Fig. 5.10: Quantitative skill scores based on 24h accumulated rainfall for $\psi\chi$-BE (a), $\psi\chi$-MBE (b), uv-BE (c), and TRMM estimate (d) for Case 1. Fields are valid at 00Z 17 June 2013.

The quantitative verification of rainfall simulated by the model runs is performed by calculating the statistical skill scores such as ETS, Bias, FAR, and POD. Fig. 5.10 depicts the above-mentioned skill scores based on 24h accumulated rainfall with respect to TRMM rainfall estimates. As seen from Fig. 5.9, the uv-BE experiment did not accurately reproduce the location of the high-intensity rainfall for Case 1. It is clear from Fig. 5.10 that the skill of the model in reproducing the high intensity rainfall for the uv-BE experiment is considerably lower as compared with skill for the $\psi\chi$-BE and $\psi\chi$-MBE experiments. The $\psi\chi$-MBE experiment shows improved skill in reproducing the 24h accumulated rainfall as compared with the other two experiments. It is important to note that the uv-BE experiment shows lower bias scores for lower rainfall thresholds, however, beyond 100 mm threshold, all the three experiments show considerable underestimation in simulating the observed rainfall.

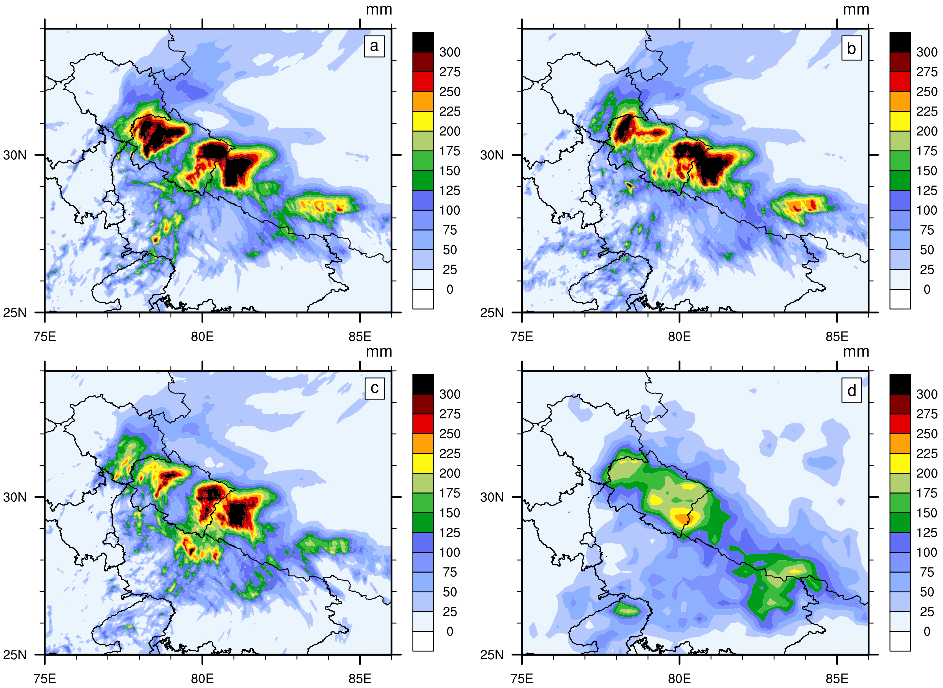

Fig. 5.11: 48h accumulated rainfall for $\psi\chi$-BE (a), $\psi\chi$-MBE (b), uv-BE (c), and TRMM estimate (d) for Case 1. Fields are valid at 00Z 18 June 2013.

The 48h accumulated rainfall (16-18 June 2013) from the three model runs and TRMM are shown for Case 1 in Fig. 5.11. As evident from the TRMM rainfall (d), the location of the maximum rainfall reported at 00Z 18 June 2013 was over the south-east region of Uttarakhand state (near Nainital). All the three model runs have reproduced the heavy rainfall over the said region fairly well. However, all the three model runs are having a tendency to overestimate the rainfall. Evidently, the overestimation is on higher side for the $\psi\chi$-BE as compared with the other two experiments. Particularly, the rainfall distribution over the Uttarakhand-Himachal Pradesh border is extremely overestimated by the $\psi\chi$-BE run. Although the overestimation in the rainfall simulation is lower for the uv-BE experiment, the above-said experiment has not reproduced the location of the maximum rainfall accurately.

The quantitative skill scores are computed for 48h accumulated rainfall for Case 1 and are shown in Fig. 5.12. The results from the above figure are similar to the inferences from Fig. 5.10 (for day 1). Clear improvement in all the skill scores can be seen for both $\psi\chi$-BE run and $\psi\chi$-MBE run. As the 48h accumulated rainfall distribution suggests (Fig. 5.11 a), the overestimation associated with the $\psi\chi$-BE experiment is quite high (as revealed from

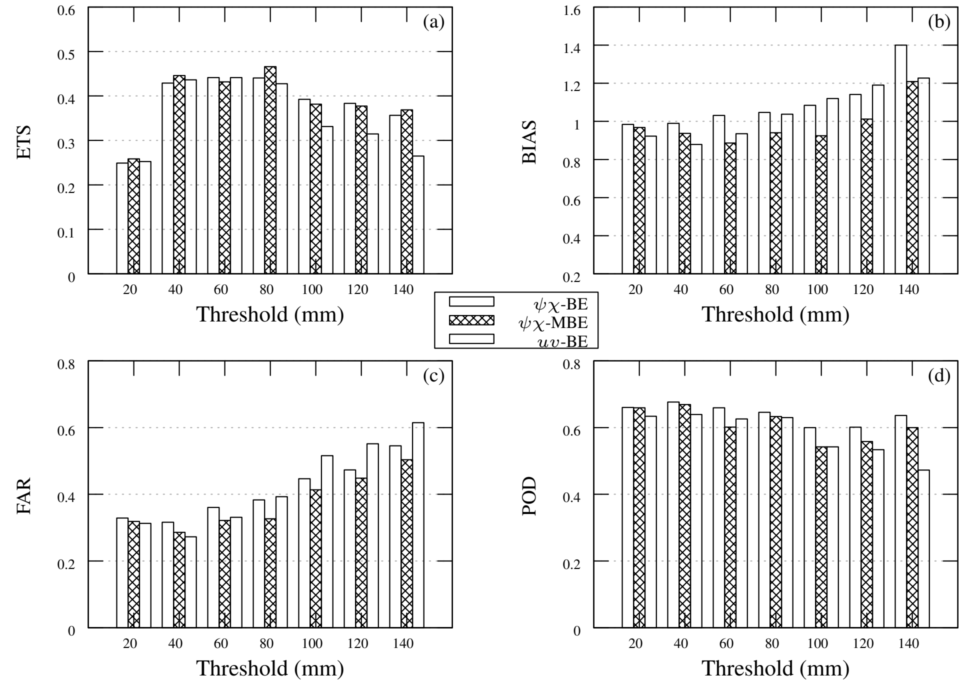

Fig. 5.12: Quantitative skill scores based on 48h accumulated rainfall for $\psi\chi$-BE (a), $\psi\chi$-MBE (b), uv-BE (c), and TRMM estimate (d) for Case 1. Fields are valid at 00Z 18 June 2013.

higher bias score) when compared with the other two experiments. The probability of detection (POD) values (Fig. 5.12 d) are consistently lower for the uv-BE experiment for all the thresholds. Although the ETS is comparable for both the $\psi\chi$-BE and $\psi\chi$-MBE runs, the number of false predictions (FAR) are consistently lower for the $\psi\chi$-MBE run (Fig. 5.12 c). Even though the $\psi\chi$-BE run is reproducing the observed rainfall (higher POD), it is associated with more number of false alarms also (as compared to the $\psi\chi$-MBE run). It is important to note that the high positive bias (overestimation) associated with the $\psi\chi$-BE experiment is somewhat lower with the $\psi\chi$-MBE experiment and the latter suggests that incorporating humidity as a multivariate variable in the $\psi\chi$-MBE formulation has a positive impact in accurately reproducing the observed rainfall distribution. The improved skill scores for rainfall simulation, especially at the higher thresholds, for the $\psi\chi$-MBE experiment suggests that introducing additional balance relations for the control variables in the $\psi\chi$-MBE formulation does provide an important advancement towards accurately simulating the high intensity rainfall events.

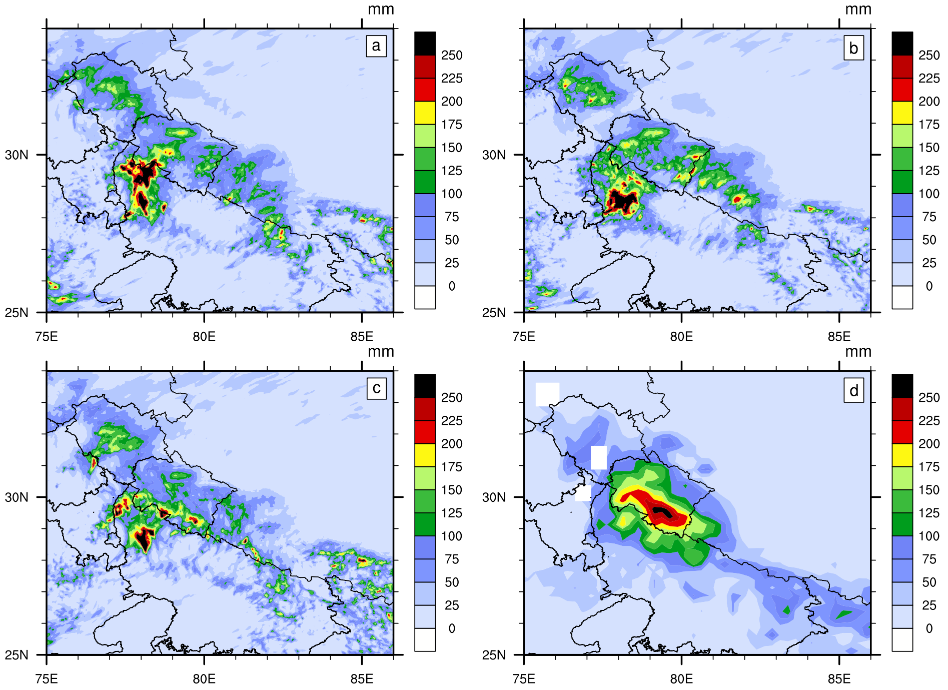

Fig. 5.13: 48h accumulated rainfall for $\psi\chi$-BE (a), $\psi\chi$-MBE (b), uv-BE (c), and TRMM estimate (d) for Case 2. Fields are valid at 00Z 17 August 2011.

The 48h accumulated rainfall from the three model runs and TRMM for Case 2 are shown in Fig. 5.13. As seen from the TRMM rainfall (Fig. 5.13 d), intense rainfall occurred mostly along the boundary of the Uttarakhand state. It is evident that all the model runs have failed to reproduce accurately the location of maximum rainfall. None of the three model runs did capture the elongated structure of the rainfall band along the boundary of the state. Model has simulated rainfall with magnitude >100 mm in the Uttarakhand state in all the three experiments. However, the location of strongest rainfall (>200 mm) in all the model runs is not coinciding with the observed location of maximum rainfall band. Both $\psi\chi$-BE and $\psi\chi$-MBE runs have realistically simulated the intensity of the heavy rainfall, while, the uv-BE experiment failed to reproduce the magnitude of rainfall as well.

A closer analysis was performed by calculating the skill scores for 48h accumulated rainfall for Case 2 and these are depicted in Fig. 5.14. It can be seen from Fig. 5.14 that, for the lower rainfall thresholds, the uv-BE run has a higher ETS, higher POD, and lower FAR. However, for rainfall thresholds beyond 60 mm, the uv-BE run show relatively lower skill, especially, the rainfall intensity is heavily underestimated by the uv-BE run. The $\psi\chi$-MBE run has relatively better skill in reproducing the high-intensity rainfall with higher POD and lower FAR. As indicated from the results of the previous case (Case 1), the $\psi\chi$-BE run for Case 2 is also characterized with relatively higher false alarms among all the three model runs. Results from Case 2 experiment also confirm that the $\psi\chi$-MBE run has better skill, particularly in reproducing the heavy to very-heavy rainfall intensities. While the uv-BE experiment has successfully reduced the overestimation in rainfall intensity, the same is not able to capture the location of intense rainfall accurately. Being an upgradation over the $\psi\chi$-BE formulation, the $\psi\chi$-MBE formulation has been somewhat successful in simulating the observed rainfall distribution.

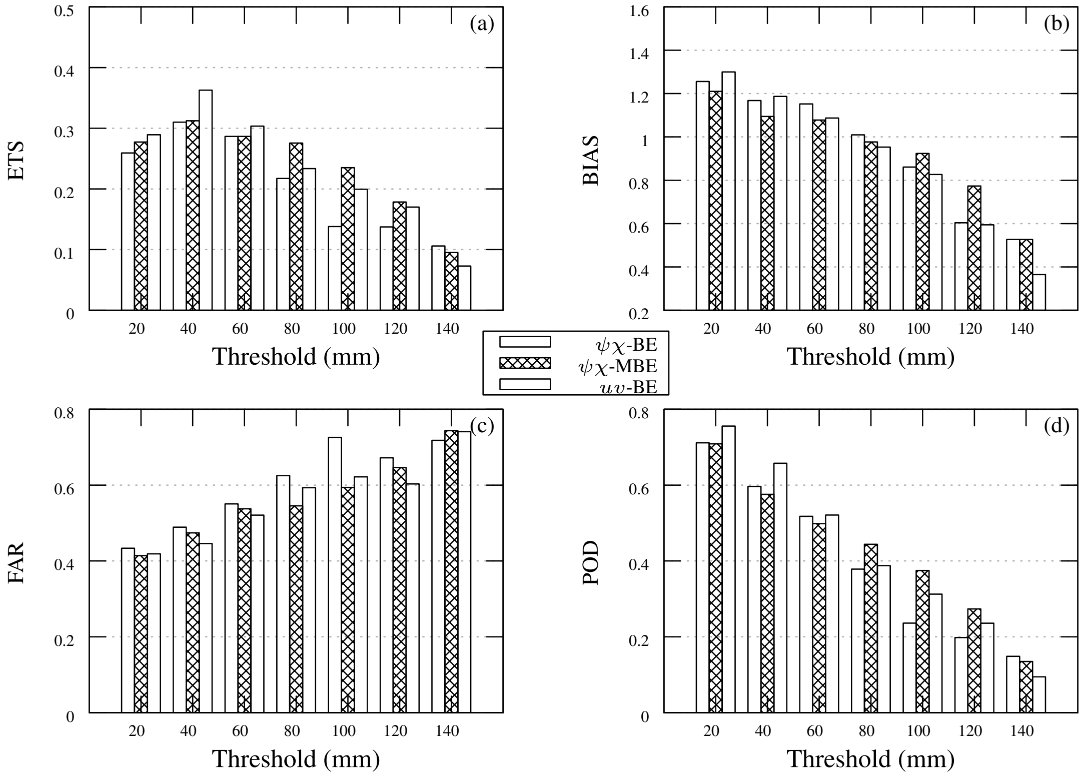

Fig. 5.14: Quantitative skill scores based on 48h accumulated rainfall for $\psi\chi$-BE (a), $\psi\chi$-MBE (b), uv-BE (c), and TRMM estimate (d) for Case 2. Fields are valid at 00Z 17 August 2011.

Finally, the rainfall distribution from the three model runs and TRMM, accumulated for 48h corresponding to Case 3 (19-21 September 2008) is shown in Fig. 5.15. All the three model runs have simulated two distinct rainfall maxima: one maximum over north-west Himachal Pradesh (77°E 32°N) and another maximum over the boundary of Uttarakhand and Nepal (80°E 29°N). The TRMM rainfall (Fig. 5.15 d) indicates that, the excessive maximum rainfall over the north-west region is somewhat absent. TRMM rainfall distribution indicates an extended rainfall maximum over the boundary of Uttarakhand and Nepal, and none of the model runs have reproduced the above rainfall maximum. All the three model runs have excessively over-predicted the northern maximum. While TRMM recorded maximum rainfall (over the above-mentioned area) is 100-125 mm, all the model runs have predicted a rainfall maximum of 250 mm and above. The primary maxima (over 80°E 29°N) is reasonably well simulated by all the three model runs, both in intensity and position. As seen in the previous two cases, the over-prediction is on a higher side for the $\psi\chi$-BE run. The rainfall band highlighted by the rectangular box in Fig. 5.15 is not simulated well by any of the three model runs. Both the $\psi\chi$-BE and $\psi\chi$-MBE runs have simulated a weaker rainfall band with moderate intensity (>150 mm) inside the box; however, such a rainfall band is not simulated by the uv-BE run. Also, the rainfall band in the $\psi\chi$-BE is extended over slightly larger area when compared with other two model runs.

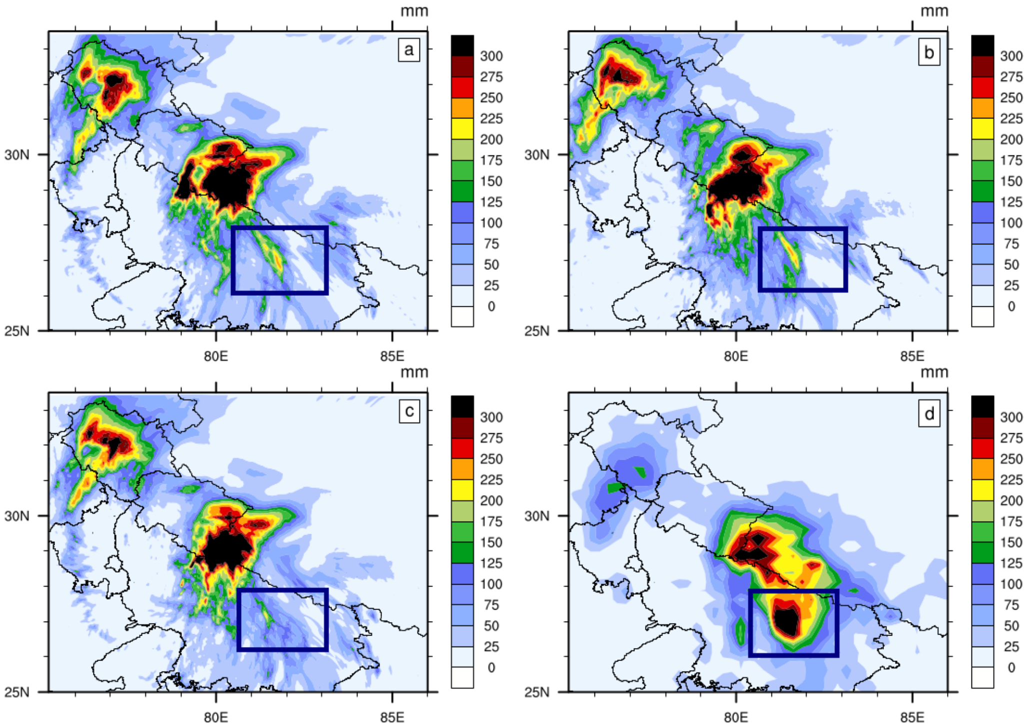

Fig. 5.15: 48h accumulated rainfall for $\psi\chi$-BE (a), $\psi\chi$-MBE (b), uv-BE (c), and TRMM estimate (d) for Case 3. Fields are valid at 00Z 21 September 2008.

Fig. 5.16 depicts the skill scores based on 48h accumulated rainfall corresponding to Case 3. It is noted that, except for the very low rainfall thresholds, the $\psi\chi$-BE run

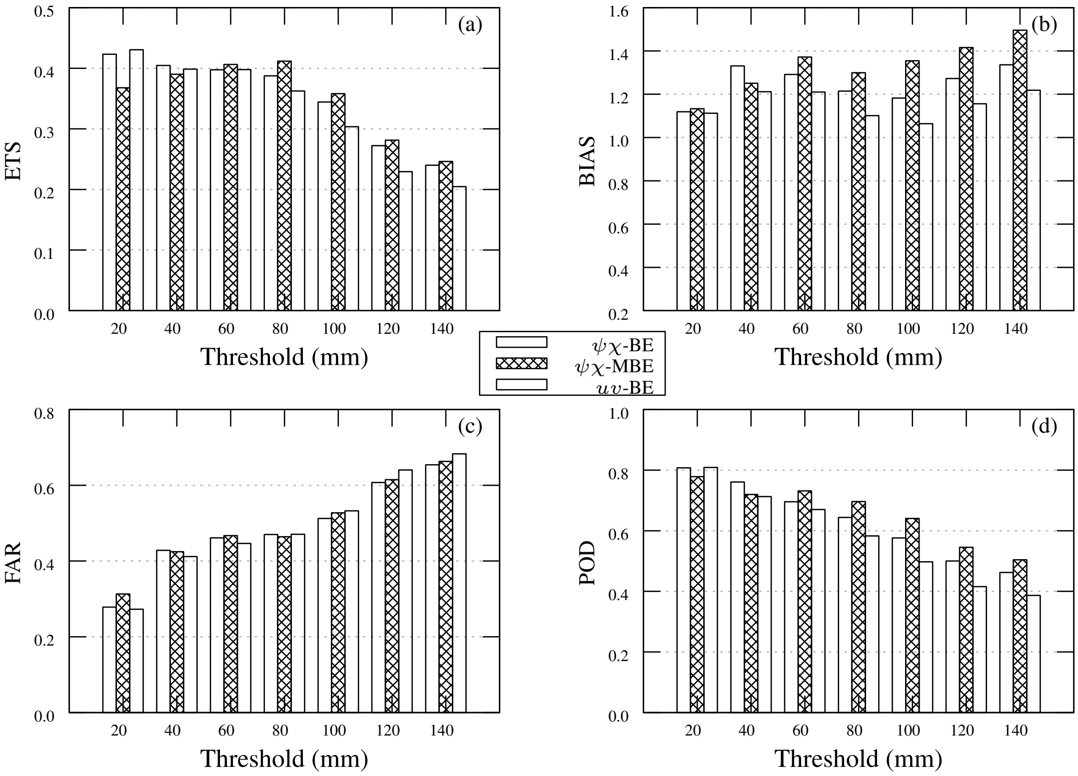

Fig. 5.16: Quantitative skill scores based on 48h accumulated rainfall for $\psi\chi$-BE (a), $\psi\chi$-MBE (b), uv-BE (c), and TRMM estimate (d) for Case 3. Fields are valid at 00Z 21 September 2008.

and $\psi\chi$-MBE are associated with higher ETS. Unlike the results for other two cases, the $\psi\chi$-MBE run is characterized with higher overestimation in the intensity of rainfall for Case 3. The uv-BE experiment shows a marginally higher number of false alarms. It is important to note that the $\psi\chi$-MBE run shows higher POD for higher rainfall thresholds, which can be partly attributed to the fact that the $\psi\chi$-MBE run accounted for the extended rainfall band (in the rectangular box) relatively well. Similarly, the POD for uv-BE is notably lower at higher thresholds; possibly due to the fact that the same did not capture the rainfall maximum intensities (indicated in the box) appropriately. Hence, the results from Case 3 also suggest that the $\psi\chi$-MBE experiment is associated with improved rainfall simulation, which is consistent with all the results from other three cases.

In the case of heavy rainfall events, the accurate simulation of the evolution of rainfall intensity is of utmost importance. Detailed analysis was performed by assessing ETS for the high-intensity rainfall of all the three cases by considering the accumulation of rainfall at different forecast lead times (i.e, the cumulative rainfall from 6h to 48h at 6h interval). Rainfall threshold for different lead times were chosen in such a way that, the same will be a representative figure of heavy rainfall for the given forecast lead time. 25 mm threshold was chosen for 6h and 12h accumulation, for 18h and 24h, 50 mm was considered, for 30h and 36h, 75 mm threshold was chosen and finally, 100 mm threshold was considered for accumulation of rainfall at 42h and 48h. In addition to ETS, the domain averaged rainfall (DAR) (accumulated for different lead times) was also analyzed for all the three cases and both ETS and DAR are shown in Fig. 5.17.

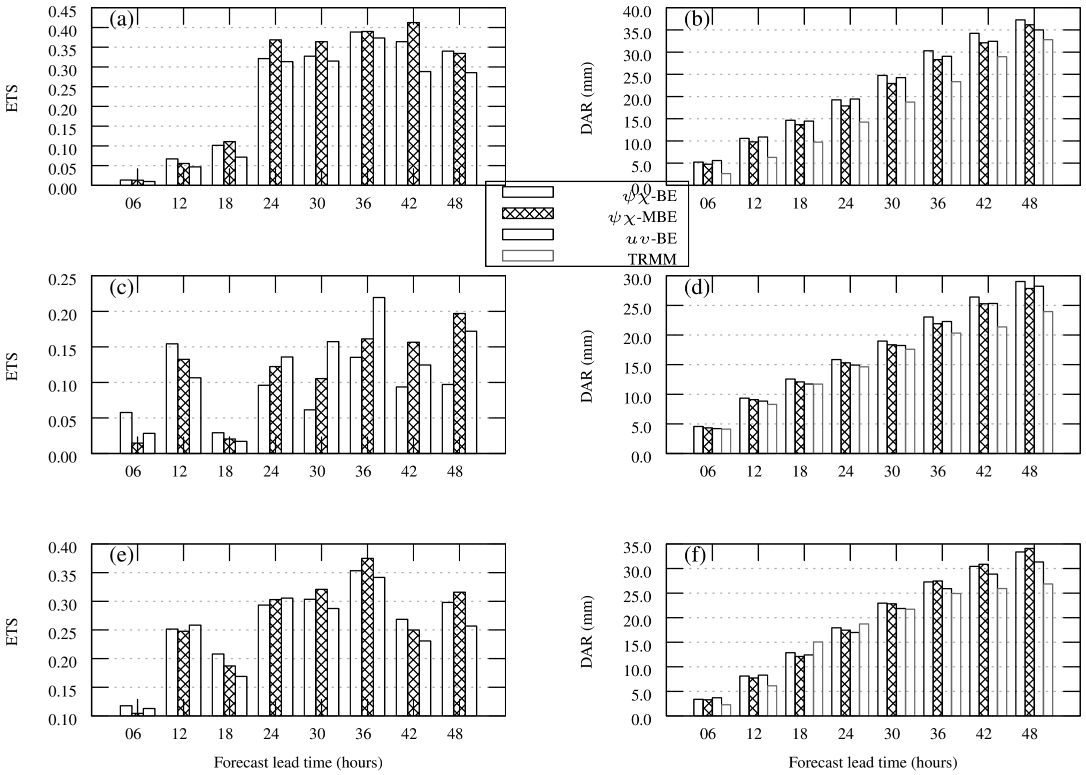

Fig. 5.17: Evolution ETS (a,c,e) and DAR (b,d,f) for 6-48h accumulation of rainfall for Case 1 (a, b), Case 2 (c, d), and Case 3 (e, f) for all the three experiments.

Fig. 5.17 provides important results on the ability of the three BEC experiments to accurately reproduce the temporal evolution of rainfall for all the three cases. Except for Case 2 (Fig. 5.17 c), high intensity rainfall simulation has been reproduced reasonably well by both the $\psi\chi$-BE run and $\psi\chi$-MBE run, in terms of higher ETS values as compared to the uv-BE experiment. It is interesting to note that the uv-BE experiment shows better rainfall forecast skill primarily during the mid-forecast hours (24h-36h). For Case 1 (Fig. 5.17 a), the performance of the uv-BE run is consistently lower as compared to the other two runs. Analysis of DAR (Fig. 5.17(b, d, f)) confirms the positive bias of all the three model runs in the rainfall simulation. The $\psi\chi$-BE is consistently associated with significant overestimation in rainfall with respect to the TRMM rainfall. Incorporation of additional regression relations along with the multivariate nature of humidity variable with $\psi\chi$-MBE experiment helps in successfully suppressing the excessive rainfall (as simulated by the $\psi\chi$-BE run) and provides more realistic rainfall forecast.

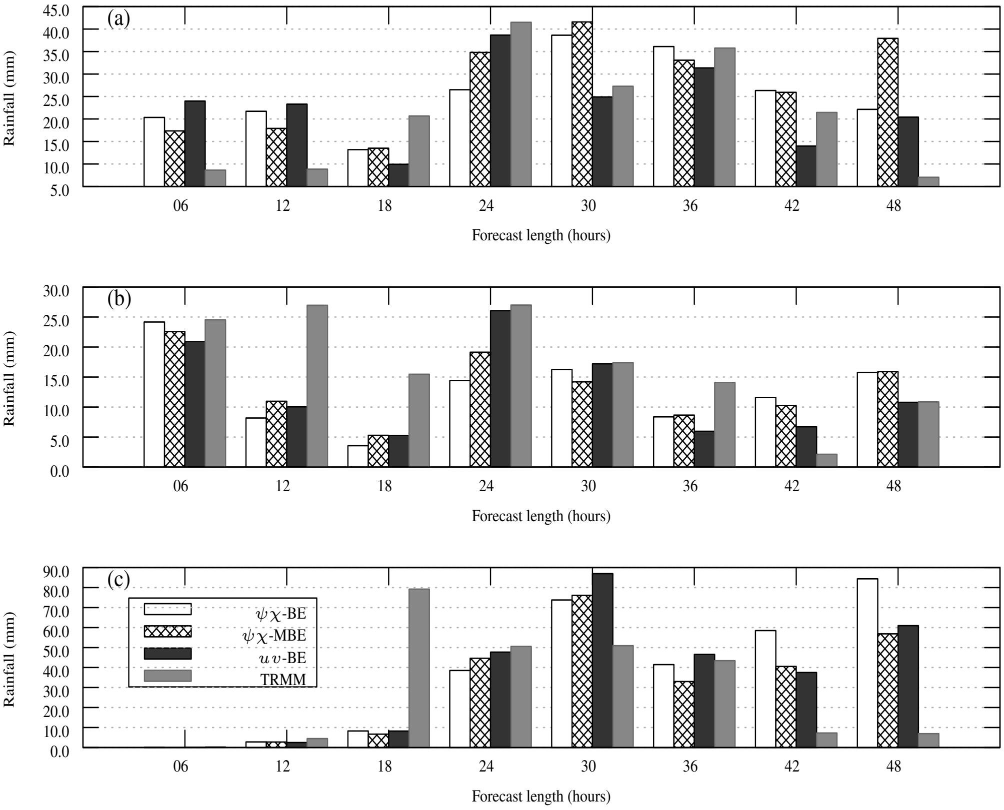

Fig. 5.18: Time evolution rainfall over the location of maximum rainfall for each 6h period Case 1 (a), Case 2 (b), and Case 3 (c).

Further, the time evolution of rainfall over the location of maximum rainfall for all the three cases has been analyzed and is shown in Fig. 5.18. The above rainfall represents the spatially averaged precipitation over a small box around the location of the maximum rainfall, estimated for each 6h duration (0-6h to 42-48h accumulation). For Case 1, all the model runs have reproduced all the major intense rainfall episodes fairly well (refer Fig.

5.18 a). During 0-6h, 6-12h, and particularly for 42-48h, all the model runs have showed significant overestimation in rainfall, with respect to the TRMM rainfall estimate for the corresponding duration. The most intense downpour occurred during 18-24h, 24-30h, and 30-36h periods and all the experiments captured the intense rainfall activity, however, the uv-BE experiment is found to reproduce the magnitude of rainfall also relatively well. In Case 2 (Fig. 5.18 b), there were three very intense downpour episodes during the 48h period and all the model runs successfully reproduced two most intense rainfall episodes (during 0-6h and 18-24h). However, the model failed to accurately capture the intensity of the heavy rainfall that occurred during 6-12h period and two moderate rainfall episodes during 12-18h and 30-36h duration. For Case 2 also, among the three experiments, the uv-BE experiment is showing relatively better skill in accurately reproducing the intensity and the timely evolution of the rainfall events as compared with the other two experiments. The rainfall simulation for Case 3 by all the model runs are almost comparable and no significant differences are noted (Fig. 5.18 c). All the three experiments have failed to simulate the strongest downpour episode during 12-18h period. However, three moderate intensity rainfall episodes have been reproduced relatively well by the model. It is noted that, all the three experiments have simulated significant amount of spurious rainfall during 36-48h period, particularly the $\psi\chi$-BE experiment show excessive rainfall during the above period.

Overall, the above analysis shows that the model has a reasonable skill in accurately reproducing the evolution of the rainfall intensity, and the above skill is relatively higher with the uv-BE experiment, particularly in simulating the magnitude of rainfall accurately. The results from Fig. 5.17 however suggested that, when the evolution of rainfall intensity over the entire domain is considered, the $\psi\chi$-MBE experiment shows relatively better skill in simulating the location of the rainfall event (in terms of higher ETS) and the time at which the major rainfall episodes occur. Thus, it can be inferred that the $\psi\chi$-MBE formulation, by accounting for the multivariate nature of humidity variable along with the additional balance relations that connect the temperature and surface pressure with velocity potential, is having a significant positive impact on the simulation of heavy rainfall episodes when compared with the $\psi\chi$-BE formulation, which treats humidity as a univariate variable. Nevertheless, despite the fact that the uv-BE experiment is showing relatively lower skill in reproducing the spatial distribution of the rainfall, the same gives improved time evolution of rainfall intensities associated with the heavy rainfall events.

5.9 A comparative study with 3DVar DA system

All the numerical experiments performed for each of the three heavy rainfall cases in this chapter utilized the 4DVar DA system. As compared with the 3DVar DA method, the model dynamics (via the tangent linear/adjoint model) is directly involved in the 4DVar minimization process and the above results in an implicit flow dependence and evolution multivariate correlations among the analysis variables [72]. On the other hand, the 3DVar DA system employs a completely static background error statistics and hence does not account for any kind of flow dependence. Thus, it would be interesting to investigate the impact of the three different BEC formulations (corresponding to $\psi\chi$-BE, $\psi\chi$-MBE, and uv-BE) by employing the 3DVar DA system and to examine how different is the above impact from the impact due to the 4DVar DA system.

Three numerical experiments were conducted for the Uttarakhand heavy rainfall event (Case 1) utilizing the 3DVar DA system by employing three different BEC formulations. As in the case of the 4DVar experiments, five cyclic assimilations were performed utilizing same set of observations (surface and upper-air conventional observations and satellite based wind observations) during the period 00Z 15 to 00Z 16 June 2013. A 24h free forecast was initiated from 00Z 16 to 00Z 17 June 2013 for each of the three BE experiments.

Fig. 5.19 shows the analysis increment in u-wind, v-wind, air temperature, and water vapor mixing ratio at 850 hPa due to the 3DVar assimilation experiments corresponding to the $\psi\chi$-BE (left panel), $\psi\chi$-MBE (middle panel), and uv-BE (right panel) runs. The above results are obtained after the first assimilation cycle, performed at 00Z 15 June 2013. The 3DVar experiments are expected to provide homogeneous and isotropic analysis increments without any flow-dependent structure. There are no marked differences between the increments in the wind fields from the $\psi\chi$-BE (Fig. 5.19 a, d) experiment and the $\psi\chi$-MBE (Fig. 5.19 b, e). The increments in the temperature field due to the $\psi\chi$-MBE (Fig. 5.19 h) is observed to be relatively stronger as compared to the same due to the $\psi\chi$-BE (Fig. 5.19 g). Interestingly, significant adjustments are seen in the humidity analysis field when the fully multivariate BEC ($\psi\chi$-MBE) is utilized for the assimilation. In the $\psi\chi$-BE experiment, the humidity analysis will be influenced by the assimilation of humidity observations only since the moisture control variable is being treated as an univariate variable in the BEC ($\psi\chi$-BE) formulation. On the other hand, the humidity variable will be impacted by the assimilation of wind or pressure fields via the additional balance relations introduced in the $\psi\chi$-MBE BEC formulation, resulting in wide-spread

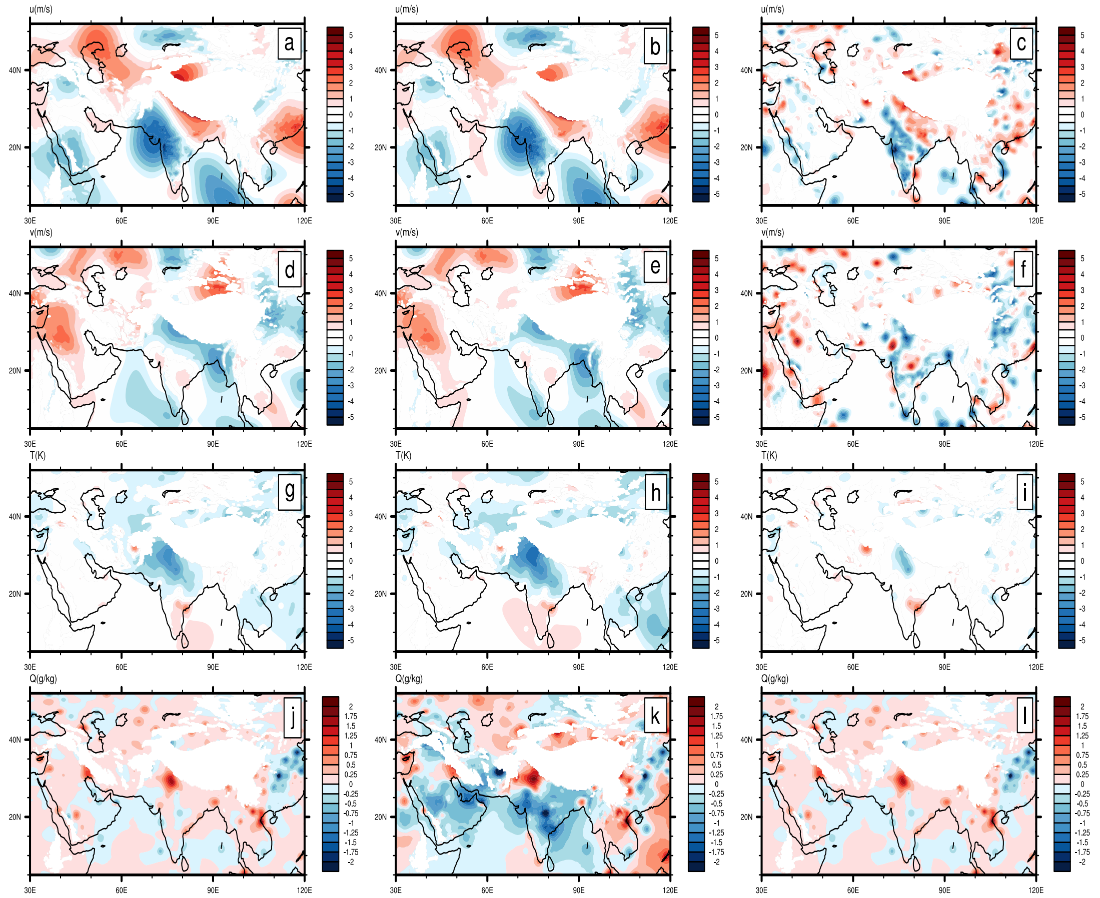

Fig. 5.19: Analysis increment in u-wind (a,b,c), v-wind (d,e,f), temperature (g,h,i), and water vapor mixing ratio (j,k,l) at 850 hPa for $\psi\chi$-BE (a,d,g,j), $\psi\chi$-MBE (b,e,h,k), and uv-BE (c,f,i,l) 3DVar experiments. Fields are valid at 00Z 15 June, 2013.

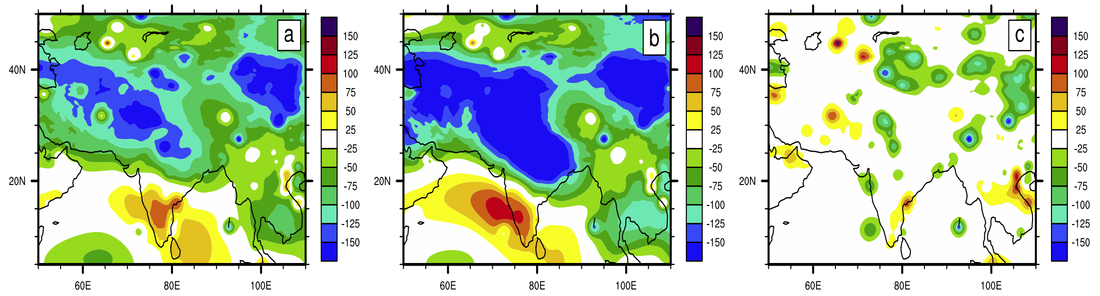

Fig. 5.20: Analysis increment in geopotential height at 200 hPa for $\psi\chi$-BE (a) $\psi\chi$-MBE (b) and uv-BE (c) 3DVar experiments for Case 1.

changes in the moisture field. More importantly, it is to be noted that the humidity field analyzed with the $\psi\chi$-MBE experiment (Fig. 5.19 k) resulted in significant removal of moisture (as revealed from negative analysis increments) at 850 hPa as compared with the humidity field of the $\psi\chi$-BE experiment (Fig. 5.19 j). The most striking observation from Fig. 5.19 is the pattern of the analysis increment corresponding to the uv-BE experiment. In contrast to the wide-spread increments in the case of both the $\psi\chi$-BE and $\psi\chi$-MBE experiments, the analysis increments due to the uv-BE experiment (Fig. 5.19 c, f, i, l) are more sharp and localized, that concentrate mostly over the immediate neighbourhood of the corresponding observations. Most of the regions remain unaffected after the assimilation (as seen from the white shading). It is interesting to note that the analysis increment in the humidity fields from the $\psi\chi$-BE (Fig. 5.19 j) experiment and uv-BE (Fig. 5.19 l) experiment are exactly similar due to the fact that the humidity control variable is univariate in both the BEC formulations.

In the 3DVar DA framework, the analysis increments for the uv-BE experiment are drastically different from the analysis increments in both the $\psi\chi$-based BEC experiments, while, the difference in the analysis increment pattern between the uv-BE experiment and the $\psi\chi$-based BEC experiments with the 4DVar DA system (refer Fig. 5.5) were not as marked as with the 3DVar DA system. The above observation signifies the critical role of the 4DVar DA system in implicitly accounting for the multivariate relationship amongst the analysis variables. In other words, the forward and backward integration of the linear model during the 4DVar minimization process successfully provides for the multivariate correlations among the analysis variables, even when a univariate BEC (uv-BE) is employed for assimilation. In order to compare the increments in geopotential height fields at 200 hPa in the 4DVar experiments for Case 1 (Fig. 5.6a-c), corresponding increments in the height fields due to the 3DVar experiments have also been examined and are shown in Fig. 5.20. As in the case of analysis increments in other fields (u, v, T, and Q), the increments in height field for the uv-BE experiment also differ drastically from the increments in height field for the other two experiments. Most of the regions are unaffected by the 3DVar assimilation when the uv-based BEC is employed. Furthermore, when compared with the analysis increments for Case 1 due to the 4DVar experiments (refer Fig. 5.6 a-c), the increments in height field due to the 3DVar experiment differ considerably, particularly in the magnitude of the increments. The 3DVar results indicate stronger negative analysis increments over most of the regions. Fig. 5.20 also suggests that the observations will affect the analysis field only over an immediate neighbourhood of the location of the observation, when the assimilation is performed utilizing the uv-based univariate BEC.

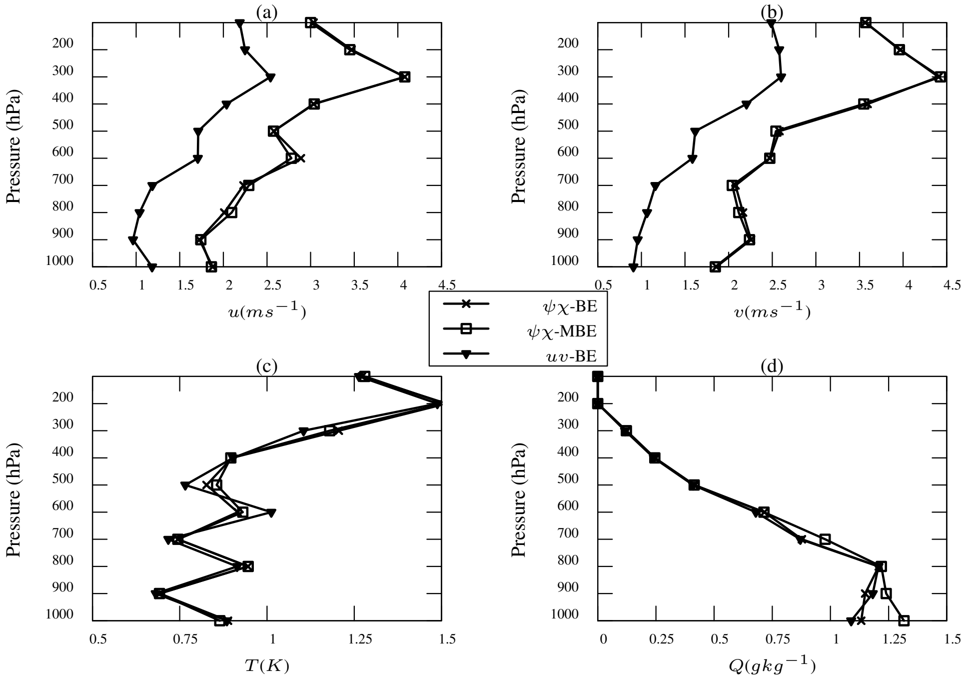

Fig. 5.21: RMS fit to radiosonde observations for u wind (a), v wind (b), temperature (c) and water vapor mixing ratio (d) for all 3DVar BEC experiments.

The RMS-fit error is computed for horizontal wind components, temperature, and the humidity fields with respect to the radiosonde observations, for the 3DVar BEC experiments for Case 1 and are shown in Fig. 5.21. Results from the above figure are consistent with the results from the 4DVar experiments (refer Fig. 5.7) and the analysis fields from the uv-BE experiment, particularly the zonal and meridional wind components are considerably closer to the radiosonde observation. Although the RMS-fit error both the 3DVar and 4DVar experiments follow similar pattern, it is noted the 4DVar analysis fields of respective BEC experiments are relatively closer to the radiosonde observation, when compared with the RMS-fit error for the 3DVar experiments. For instance, the maximum error in the zonal wind component (at 300 hPa) as revealed from the 3DVar experiments are, 2.65 ms^{-1} (uv-BE) and 4.0 ms^{-1} ($\psi\chi$-BE and $\psi\chi$-MBE), whereas, the above errors as revealed from the 4DVar experiments are, 2.5 ms^{-1} (uv-BE) and 3.5 ms^{-1} ($\psi\chi$-BE and $\psi\chi$-MBE). It is worthwhile to note that the reduction in RMS-fit error from the 3DVar experiments to the 4DVar experiments is more apparent in the $\psi\chi$-BE and $\psi\chi$-MBE experiments as compared to the uv-BE experiment. The RMS-fit error in the temperature field or moisture field did not show any significant difference among the three 3DVar BEC experiments.

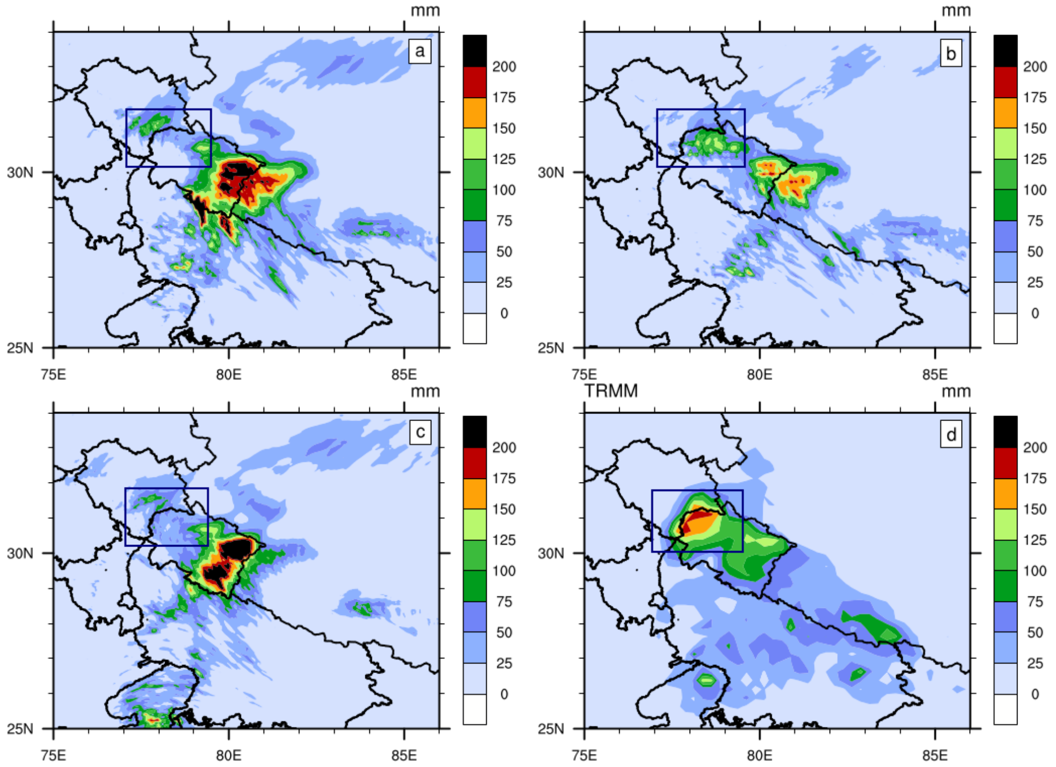

Fig. 5.22: 24h accumulated rainfall for $\psi\chi$-BE (a), $\psi\chi$-MBE (b), uv-BE (c), and TRMM estimate (d) for the 3DVar BEC experiments. Fields are valid at 00Z 17 June 2013.

The 24h accumulated rainfall from the three 3DVar BEC experiments and TRMM estimate, valid at 00Z 17 June 2013 are shown in Fig. 5.22. It is quite evident that none of the experiments have reproduced the maximum rainfall location accurately. Both the $\psi\chi$-BE and uv-BE experiments have simulated excessive rainfall with respect to the TRMM estimate, while, the rainfall simulated by the $\psi\chi$-MBE experiment has successfully suppressed the overestimation. Furthermore, the $\psi\chi$-MBE experiment has reproduced the intense rainfall that occurred over the Kedarnath region (indicated with a rectangular box) to some extent. It can be recalled that (refer Fig. 5.19 k), the 3DVar analysis employing the multivariate BEC ($\psi\chi$-MBE) was found to reduce the moisture content over the Indian region (as seen from the negative analysis increment of magnitude 2 gkg^{-1}). The above can be attributed to reduced overestimation of the rainfall for the $\psi\chi$-MBE experiment as compared with the other two experiments. A comparison of the 24h accumulated rainfall simulated by the 3DVar BEC experiments (Fig. 5.22) and the same by the 4DVar BEC experiments (refer Fig. 5.9) for Case 1 clearly reveal improved skill of the rainfall simulation by the 4DVar experiments. Especially, all the three 4DVar experiments have shown improved skill in reproducing the location of maximum rainfall (near Kedarnath region) as compared with the corresponding skill of the 3DVar experiments. In addition, the overestimation associated with the 3DVar experiments in the rainfall simulation has reduced considerably in the corresponding 4DVar BEC experiments and this signifies the pivotal role of model physics inherent to the 4DVar DA system.

5.10 Application to summer monsoon rainfall over India

The present study has been extended to investigate the impact of different formulations of BEC matrix on the simulation for a fortnight for studying the Indian summer monsoon circulation and rainfall. The study was conducted for 15 sample days of July 2017. As in the case of the heavy rainfall cases, three experiments employing three different BECs ($\psi\chi$-BE, $\psi\chi$-MBE, and uv-BE) were performed in this study also. Conventional upper-air and surface observations, ASCAT ocean surface winds and atmospheric motion vectors were assimilated at 00Z of 15 days starting from 01 July 2017, which was followed by 48h free forecast from each of the analysis fields. Two nested domains with two-way nesting were employed in the study (as shown in Fig. 3.1. The outer domain has 45 km horizontal resolution and the inner domain has 15 km horizontal resolution. Both the domains have 36 terrain following vertical coordinates. The convective processes are parameterized using the Kain-Fritsch scheme, that accounts for deep and shallow convection using mass flux approach. WSM6 scheme was utilized to parameterize the microphysical processes. The boundary layer turbulence is specified following the Yonsei University scheme. The Noah land surface model is employed for the land surface physics. The radiative processes are parameterized following the rapid radiative transfer model and Dudhia scheme for longwave and shortwave radiations respectively.

Fig. 5.23 shows the RMS-fit error in the three analysis fields for zonal wind (a), meridional wind (b), air temperature (c), and water vapor mixing ratio (d) with respect to the radiosonde observations. The RMS-fit error is computed by averaging for all the 15 analysis samples. The results from the above figure are in overall agreement with the results for the three heavy rainfall cases (Fig. 5.7). The uv-BE analysis fields have the lowest RMS-fit error in all the four parameters. The error reduction in the uv-BE analysis field is clearly evident for the wind fields as compared with both temperature and humidity fields. However, there are no marked differences between the errors corresponding to the

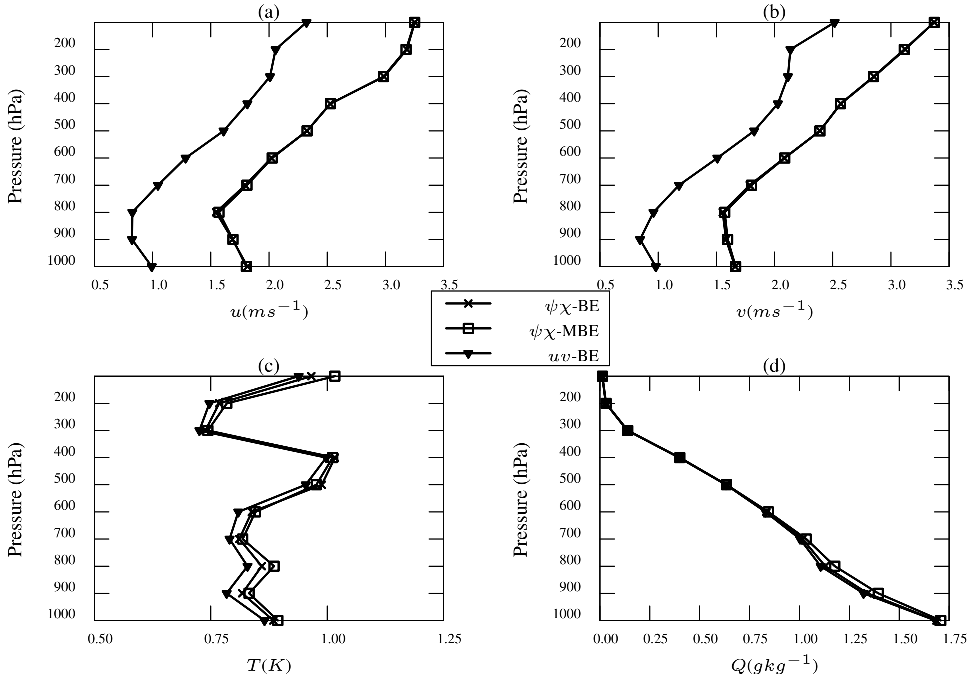

Fig. 5.23: RMS fit to radiosonde observations for u wind (a), v wind (b), temperature (c) and water vapor mixing ratio (d) for all BEC experiments.

$\psi\chi$-BE and $\psi\chi$-MBE analysis fields. For wind fields, both the RMS-fit errors in $\psi\chi$-BE and $\psi\chi$-MBE analysis fields are overlapping each other. For temperature and humidity fields, the $\psi\chi$-BE analysis field is relatively closer to the radiosonde observations as compared with the $\psi\chi$-MBE fields. All the above results are similar and consistent with the results for the three heavy rainfall experiments and hence it can be broadly concluded that the uv-BE analysis fields are invariably closer to the observed fields (radiosonde observations).

The rainfall forecast from all the three experiments were validated against the TRMM rainfall estimates by computing the quantitative skill scores for 24h and 48h accumulated rainfall. The time series of ETS, bias score, FAR, and POD calculated for 24h accumulated rainfall are shown in Fig. 5.24-5.27 respectively. In each of the above figures, upper panel shows the respective skill score for 20 mm rainfall threshold and the lower panel shows the skill score for 40 mm rainfall threshold. It is noted that all the three experiments have almost similar skill in reproducing the rainfall distribution with respect to the TRMM rainfall estimate. A closer examination reveals that, on an average the $\psi\chi$-BE and $\psi\chi$-MBE runs have improved skill scores for most of the times. It is noted that the $\psi\chi$-MBE run is

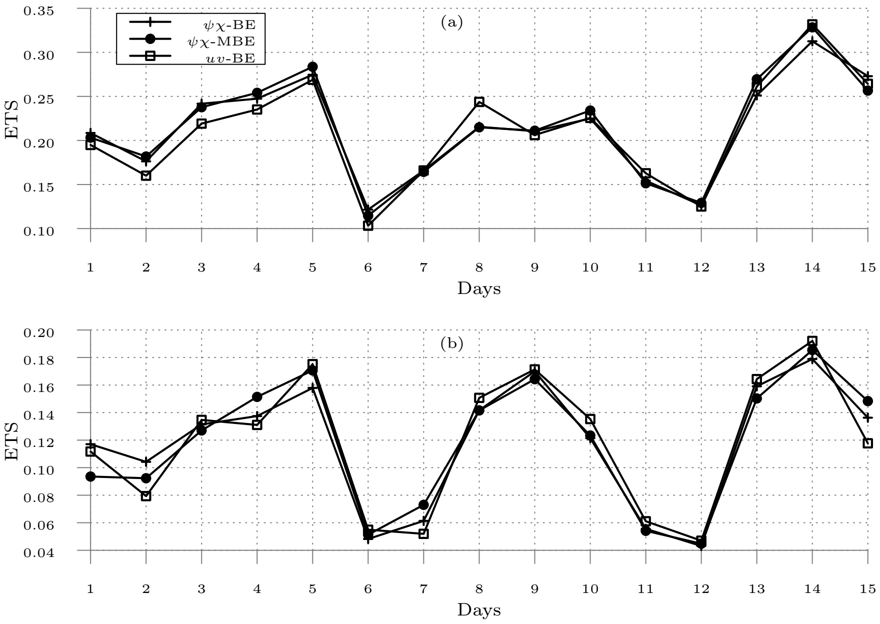

Fig. 5.24: Time series of ETS based on 24h accumulated rainfall for 20 mm (a) and 40 mm (b) rainfall thresholds - for 01-15 July 2017.

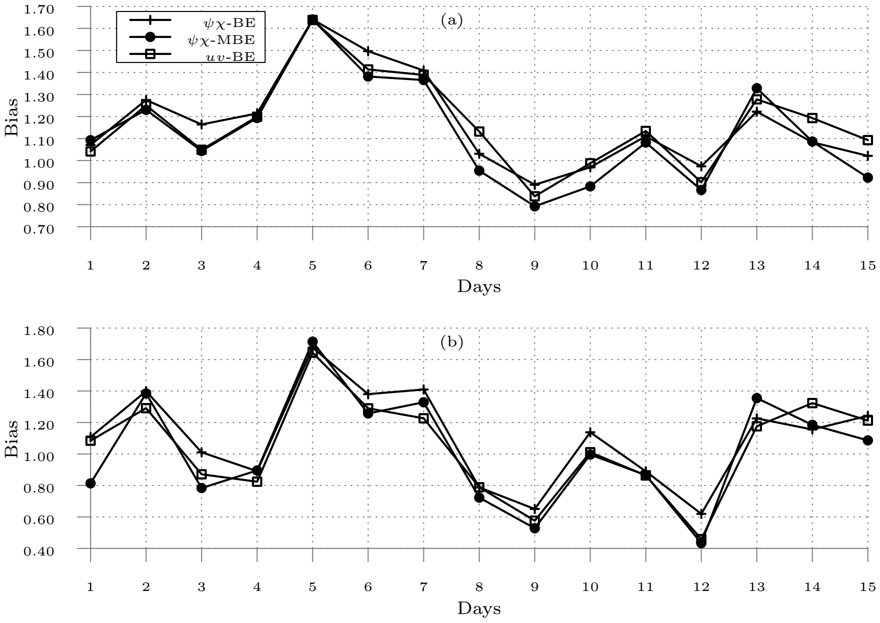

Fig. 5.25: Similar to Fig. 5.24 but for bias score

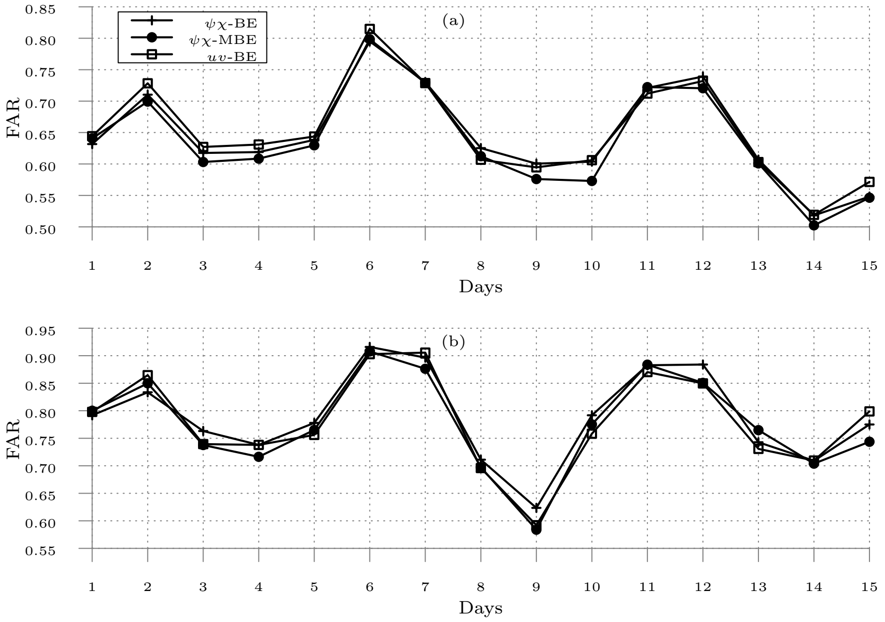

Fig. 5.26: Similar to Fig. 5.24 but for FAR.

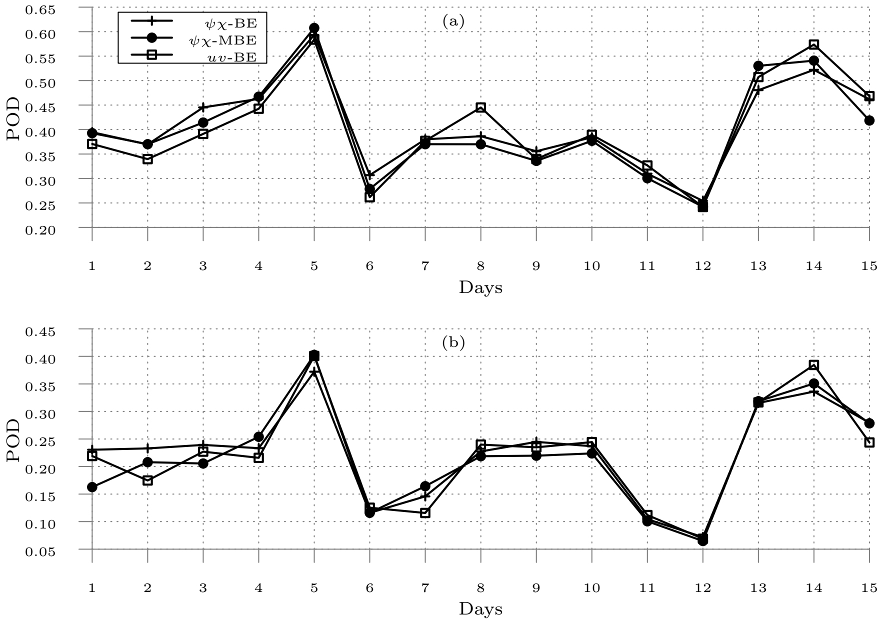

Fig. 5.27: Similar to Fig. 5.24 but for POD.

producing lower false alarms for both of the rainfall thresholds. On the other hand, the uv-BE run is characterized with higher false alarms. Also, the $\psi\chi$-MBE run shows better POD for most of the days, as compared with the other two experiments.

A more general picture can be obtained by analyzing the average skill scores for all the 15 sample days. Fig. 5.28 and Fig. 5.29 depict the 15-day average skill scores based 24h accumulated rainfall and 48h accumulated rainfall respectively. It can be seen that for 20 mm and 40 mm thresholds, the $\psi\chi$-MBE experiment is performing better, whereas, the $\psi\chi$-BE experiment is better reproducing the higher intensity rainfall (80 mm). More importantly, while the $\psi\chi$-BE and the uv-BE experiments tend to overestimate the rainfall, the $\psi\chi$-MBE is consistently associated with bias scores closer to unity. The average FAR and POD scores also suggest that the $\psi\chi$-MBE is able to reproduce the lower and medium intensity rainfall distribution more accurately than the other two experiments.

In the case of 48h accumulated rainfall (Fig. 5.29), the impact of the three different BEC formulations are not much dominant as seen for 24h accumulated rainfall. There are no marked differences among the three model runs in terms of the skill scores. Nevertheless, the $\psi\chi$-BE is associated with better skill in capturing the higher intensity rainfall, as in the case of 24h accumulated rainfall. In addition to that, the bias score suggests that the overestimation is on a lower side for the $\psi\chi$-MBE, as noted in the 24h rainfall forecast.

It is pertinent to note that the skill of the $\psi\chi$-MBE experiment in predicting the high-intensity rainfall (80 mm) has improved for 48h rainfall forecast when compared with the 24h rainfall forecast, as seen from improved ETS and FAR scores. The results of the study hence suggests that the $\psi\chi$-MBE experiment is able to provide better rainfall forecasts at lower and medium rainfall thresholds and, furthermore, the $\psi\chi$-MBE formulation has an additional advantage that, it can successfully suppress the overestimation in rainfall (as noted in the other two experiments).

5.11 Summary

A comprehensive study was carried out to understand the sensitivity of the 4DVar DA system to three different formulations of BECs by employing different sets of control variables. Among the three different BECs employed in the present study, two formulations use stream function and velocity potential as momentum variables ($\psi\chi$-BE and $\psi\chi$-MBE) and the other one utilizes horizontal wind components as momentum variables (uv-BE).

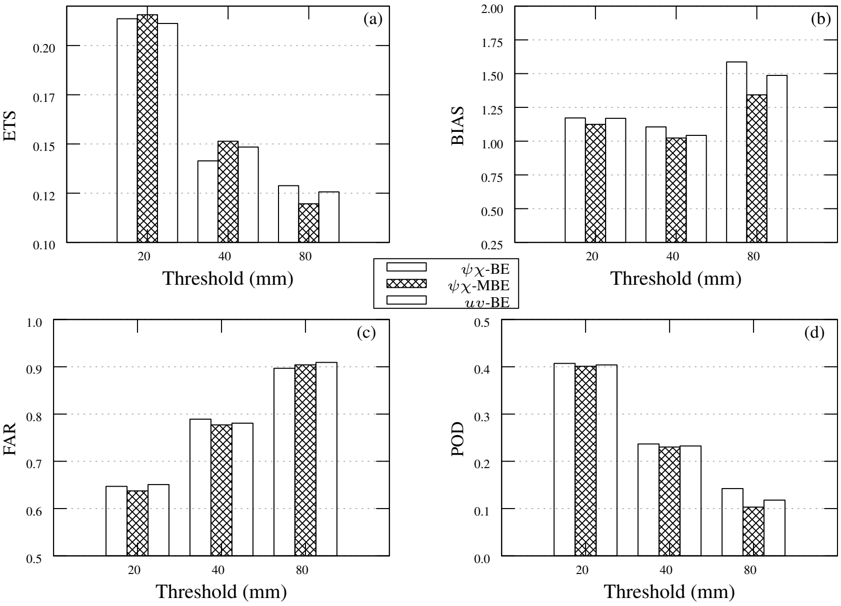

Fig. 5.28: Average quantitative skill scores based on 24h accumulated rainfall for all the three BEC experiments.

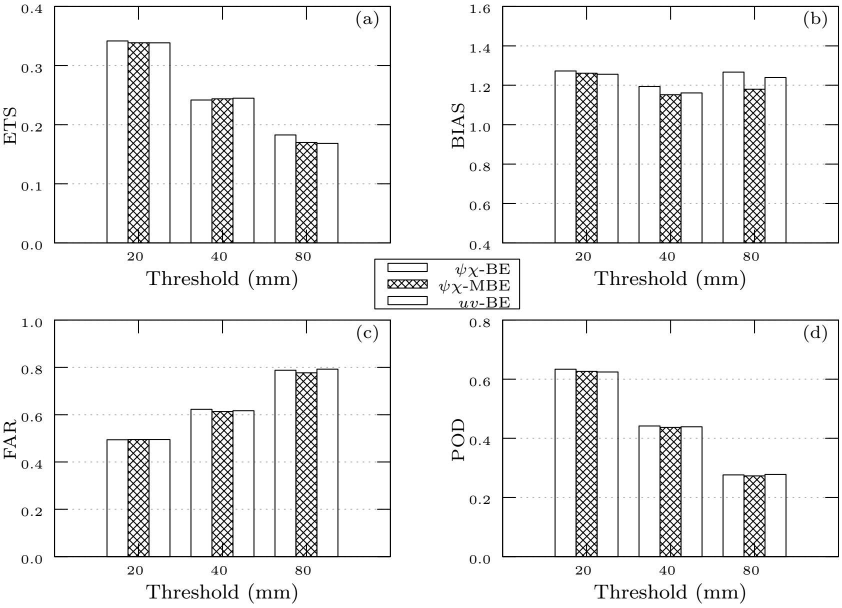

Fig. 5.29: Average quantitative skill scores based on 48h accumulated rainfall for all the three BEC experiments.

The BEC formulations in $\psi\chi$-BE and $\psi\chi$-MBE differs mainly in the way they treat the moisture control variable: the humidity variable in $\psi\chi$-BE is univariate, while in $\psi\chi$-MBE, the humidity variable is multivariate in nature. The BEC employed in the uv-BE experiment is completely univariate.

The most striking result from the single observation experiment conducted by employing the three different BECs was the drastic difference in the spatial extent of the information spread due to the 4DVar assimilation of single 'u' observation. When both the $\psi\chi$-BE and $\psi\chi$-MBE experiments resulted in significantly wider impact on the analysis field with smooth increments, the impact due to the uv-BE experiment was sharp and highly localized around the neighbourhood of the observation. Real DA experiments were performed with emphasis on three heavy rainfall events that occurred over the Himalayan region, India. The analysis fields indicated that the fields analyzed with the uv-BE experiment is significantly closer to the radiosonde observations as compared with the fields analyzed with both the $\psi\chi$-BE and $\psi\chi$-MBE experiments. The rainfall forecast (accumulated for 48h) verification with respect to the TRMM estimates revealed that the $\psi\chi$-MBE experiment has relatively better skill in reproducing the rainfall distribution, in terms of the statistical skill scores such as ETS, bias, POD, and FAR. Although all the model runs showed a reasonable skill in reproducing the evolution of major rainfall episodes in time, the uv-BE experiment is found to provide for relatively better simulation of time evolution of rainfall intensities. The study was also extended for a 15-days period during the Indian summer monsoon of July 2017, to investigate the impact of employing the above three BECs on the simulation of rainfall over Indian region during the summer monsoon period. The results obtained from the above study were also in agreement with the study on heavy rainfall events, and the $\psi\chi$-MBE experiment was found to have marginally better skill in reproducing the 24h and 48h accumulated rainfall.

Further, a comparative study was performed by employing the 3DVar DA system in the simulation of one heavy rainfall event (Case 1) and examined how the impact of employing different BEC formulations differ in a 3DVar DA system as compared with the results of similar experiments with the 4DVar DA system. The results from the 3DVar experiments also indicated that the uv-BE analysis fields, particularly the wind fields are considerably closer to the radiosonde observations, which is in agreement with the results from 4DVar experiments as well. It is noted that the errors in the 3DVar analysis fields are relatively higher than that of the 4DVar analysis fields. The 24h rainfall forecast revealed that, none of the model runs could accurately reproduce the location of maximum rainfall, also the results indicated significant overestimation in the rainfall simulation with the $\psi\chi$-BE and uv-BE (3DVar) experiments. Nevertheless, the $\psi\chi$-MBE experiment could reproduce moderately intense rainfall over the location of maximum observed rainfall and was found to have relatively better rainfall forecast skill among the three 3DVar experiments.

The results from this detailed study suggest that incorporating humidity as a multivariate variable in the BEC formulation ($\psi\chi$-MBE experiment) does have a definite positive impact on the rainfall forecast. Particularly, the $\psi\chi$-MBE experiment can successfully suppress the excessive overestimation and increased false predictions in rainfall forecast, associated with the $\psi\chi$-BE experiment. Although the uv-BE analysis fields are considerably closer to the radiosonde observations, the same is not able to reproduce the spatial distribution and location of maximum rainfall accurately, as compared with the $\psi\chi$-MBE experiment. The results from single observation experiment as well as the 3DVar experiments revealed that the uv-BE experiments result in very sharp and highly localized analysis increments, in contrast to smoother and wide-spread increments in the case of $\psi\chi$-BE and $\psi\chi$-MBE experiments. The sharp analysis increments in the uv-BE experiment may result in initial imbalances among the variables in the analysis field. In addition, much of the impact due to the uv-BE experiment is concentrated on an immediate neighbourhood of the location of the observation. Hence, spurious signals may be impacting over observation-void regions, which may propagate in space and time as the forecast starts. The above arguments are the possible reasons for the relatively lower skill in reproducing the rainfall pattern by the uv-BE experiment.

Nevertheless, results from the present study are indeed encouraging, and it is expected that the uv-BE experiment may perform better in the case of smaller-scale applications with the assimilation of high-resolution observations such as from radar, where the observation-void regions will be relatively lower.

Chapter 6

Impact of Scatsat-1 Wind Assimilation on the Simulation of Summer Monsoon Rainfall over India

The present chapter examines the impact of assimilating ocean surface winds from Scatsat-1, an Indian scatterometer on the simulation of summer monsoon rainfall over India during the month of July, 2017. The study also aims to compare the performance of impact of assimilating Scatsat-1 winds viz-a-viz the impact of assimilating Advanced Scatterometer (ASCAT) winds.

6.1 Introduction

Remote sensing satellite instruments form an integral part of the current observational system. They provide effective coverage over regions where the conventional observations are unavailable - typically over the oceans. Scatterometers provide near-surface wind speed and wind direction over the oceans on a global scale [149, 150, 151]. Accurate measurement of ocean surface wind speed and direction are extremely important due to their vital role in the ocean-atmosphere interaction. The assimilation of wind information derived from scatterometers are known to to have a positive impact on the simulation of various weather systems including tropical cyclones [127, 152, 67, 149, 153, 150, 154, 151]. Accurate near-surface winds over the ocean help in successfully simulating the turbulent processes together with the energy and moisture transport from the ocean to the atmosphere, all of the above contributing to improved numerical weather forecasts.

India launched its new scatterometer satellite, Scatsat-1 on 26 September 2016 as a continual mission of Oceansat-2. The scatterometer on board Scatsat-1 has a similar design to the earlier Oceansat-2. Studies have documented that assimilating surface winds from Oceansat-2 in NWP models have yielded a clear positive impact on short-range forecasts over India [151, 67, 154]. However, since the successful launch of Scatsat-1 in September 2016, no detailed study has been reported in the literature on the assimilation of Scatsat-1 wind observations over the Indian region. Hence, the present study aims to investigate the impact of assimilating Scaatsat-1 winds on the simulation of Indian summer monsoon rainfall for the entire month of July, 2017. The study will also compare the results of the above impact of Scatsat-1 wind against the impact of assimilating Advanced Scatterometer (ASCAT) winds in simulating the Indian summer monsoon rainfall for the same period.

6.2 Indian summer monsoon 2017 - An overview

The onset of southwest Indian monsoon (SWIM) occurred over Kerala on 30 May 2017, two days ahead of the normal onset date and the SWIM covered the entire country by 19 July 2017. The country received 98% rainfall of the long period average (LPA) during the SWIM of 2017. The rainfall distribution during the monsoon period however, was not uniform - June and July months received 104% and 102% rainfall of LPA while the rainfall recorded in August and September were 87% and 88% of the LPA. Even though most of the central Indian region received normal rainfall (81% to 119% of LPA) during the month July 2017, the rainfall recorded over most of the southern parts of the country was significantly lower (>38% deficiency) during July 2017. The first week of July 2017 and the first two weeks of August 2017 were mostly characterized with break like situation. A total of 14 depressions/low pressure systems formed over the region, out of which 6 had formed in July 2017. By 27 September 2017, dry weather started to prevail over northwestern parts of India and subsequently led to the withdrawal of the summer monsoon from the Indian subcontinent [155].

6.3 Model configuration

The WRF model (v3.8.1) was configured with two nested domains with two way nesting enabled (as seen in 3.1). The outer domain had a horizontal resolution of 45 km and the inner domain was configured with 15 km horizontal resolution. Both the domains have 36 terrain-following vertical levels. The Kain-Fritsch scheme was employed for representing the convective processes, which accounts for both the deep and shallow convection using the mass flux approach. The microphysical processes follow the WRF single moment six-class scheme, that includes graupel phase and related processes. The boundary layer turbulence was specified following the Yonsei University scheme. The Noah land surface model provides for the land surface physics. The radiative processes are parameterized utilizing the rapid radiative transfer model and Dudhia scheme for longwave and shortwave radiations respectively.

6.4 Experimental design

Three numerical experiments were conducted in the present study employing the 4DVar DA method (WRFDA v3.8.1): (i) a CTRL run with the assimilation of conventional surface and upper-air observations and AMVs (without any scatterometer observation), (ii) a ASCAT run with the assimilation of ASCAT winds in combination with all the observations in CTRL experiment, and (iii) a SCATSAT run with the assimilation of Scatsat-1 winds in combination with all the observation in CTRL experiment. All the experiments were performed for 30 sample days during the month of July 2017. 4DVar assimilation of the above-mentioned observations were performed for the respective experiments only at 00Z of every day (one-time assimilation after a spin-up run for 6h), starting from 01 July 2017 to 30 July 2017. GFS 24-forecast fields were used as initial and boundary conditions. All the 30 analysis fields were subjected to short-range forecast lasting 48h in free forecast mode, without additional assimilation of observations.

Table 6.1: Typical number of observations used for assimilation at 00Z 10 July 2017.

Observation type	Number of observations used
Synoptic	3178
Soundings	81
Metar report	1874
Pilot	51
AMVs	38000
ASCAT winds	29668
Scatsat-1 winds	41240

6.5 Various scatterometer observations used in the study

6.5.1 Scatsat-1

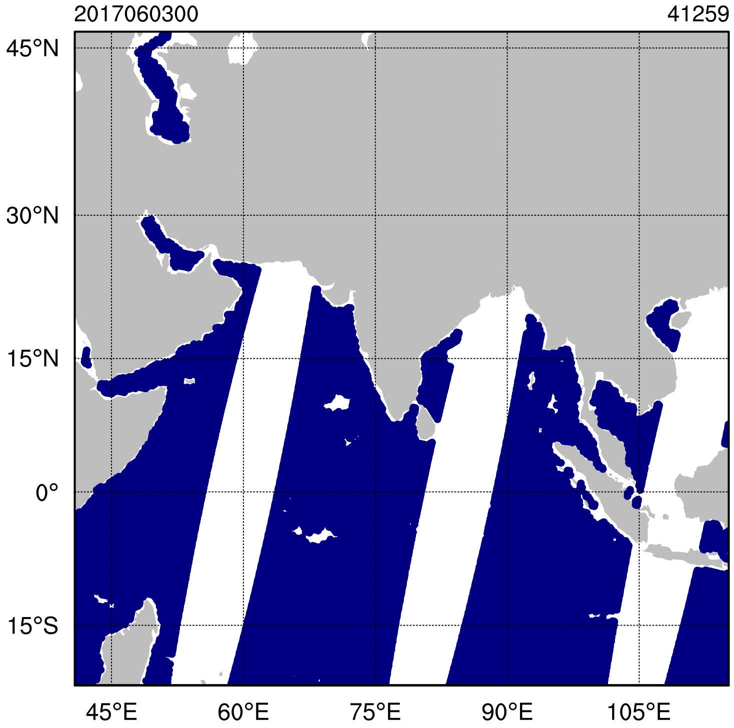

Fig. 6.1: Distribution of Scatsat-1 observation over the study area available for assimilation at 00Z 10 July 2017.

The Scatterometer satellite-1 (Scatsat-1) is a polar, sun synchronous satellite, launched by Indian Space Research Organization (ISRO) on 26 September 2016. It is a continuity mission to the earlier Oceansat-2, the latter carried a Ku-band scatterometer, ocean color monitor and a radio occultation sounder. The scatterometer on board Scatsat-1 is a Ku-band, pencil beam scatterometer that operates at 13.515 GHz. Situated at an altitude of 720 km, Scatsat-1 provides wind information (both speed and direction) at a horizontal resolution of 25 x 25 km at ground. The inner beam of the swath has a coverage of 1400 km and the outer beam has a coverage of 1840 km. The wind speed is sampled at an accuracy of 1.8 ms^{-1} and the accuracy for wind direction is 20°. Scatsat-1 has a revisit time of two days. Fig. 6.1 shows the typical distribution of the wind observation from Scatsat-1 over the experiment domain at 00Z, 10 July 2017.

6.5.2 Advanced Scatterometer (ASCAT)

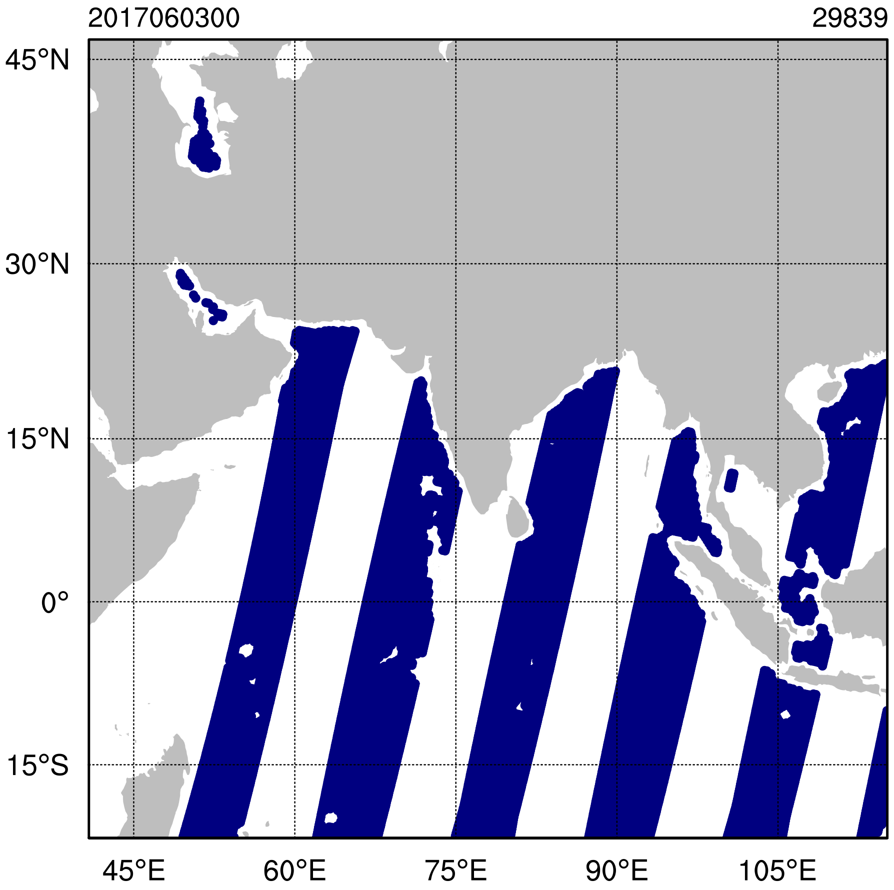

Fig. 6.2: Distribution of ASCAT observation over the study area available for assimilation at 00Z 10 July 2017.

The advanced scatterometer (ASCAT) is an active scatterometer on board Metop satellite of EUMETSAT, and is operational since May 2007. It is positioned at an altitude of 837 km in a polar orbit and consists of a C-band radiometer which operates at 5.2 GHz. The radiometer has three fan-beam antenna that are oriented at 45°, 90°, and 135° with reference to the track of the satellite. ASCAT carries two swaths of 500 km width, each having 25 km horizontal resolution. The wind observations from ASCAT are considered to be less sensitive to rain since it operates at C-band, as compared to Scatsat-1, that operates at Ku-band instruments. Fig. 6.2 shows the typical distribution of the wind observation from ASCAT over the experiment domain at 00Z 10 July 2017.

6.6 Results and discussion

6.6.1 Impact on the analysis fields

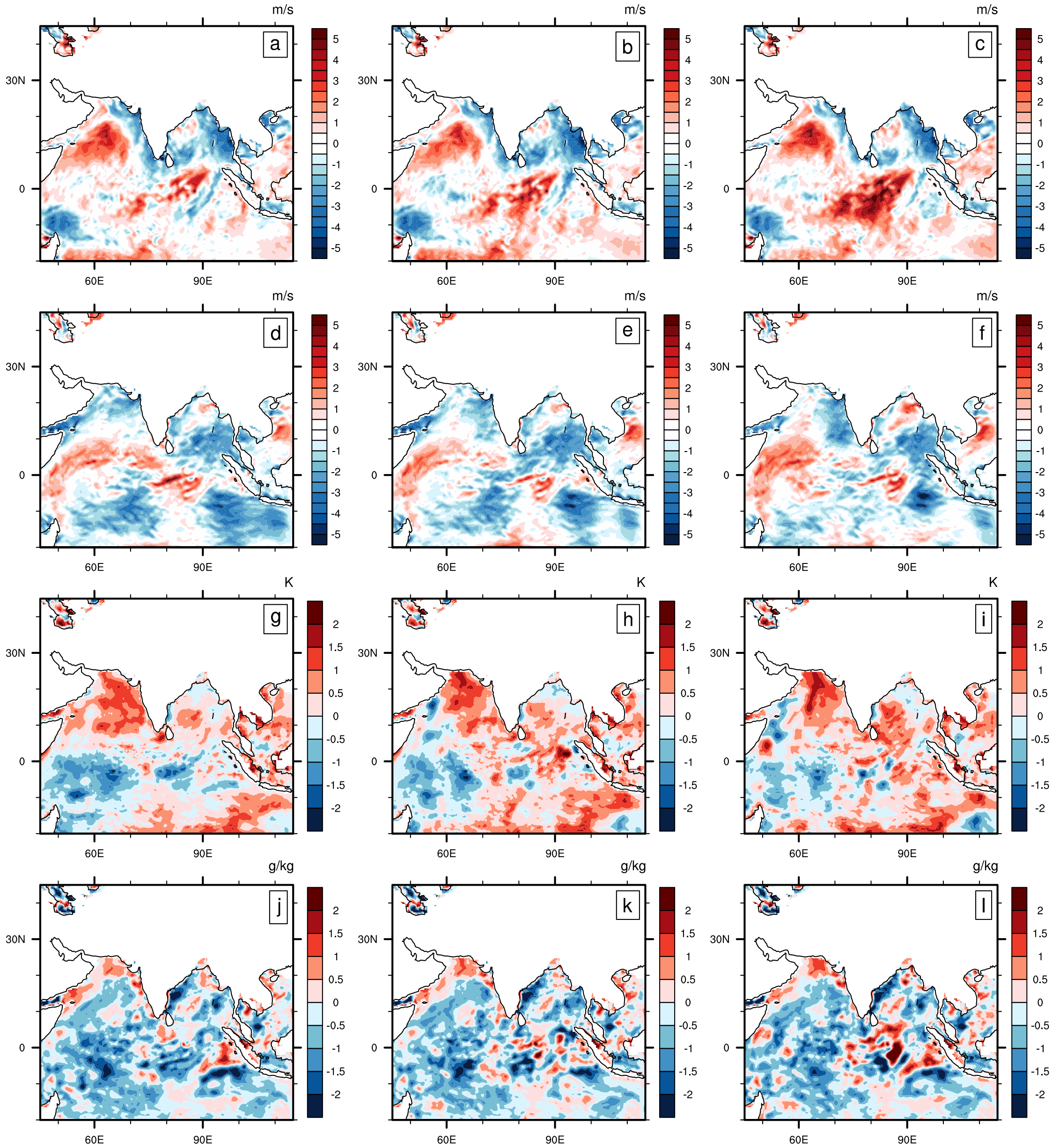

Fig. 6.3: Analysis increment in u (a,b,c), v (d,e,f), T (g,h,i), and water vapor mixing ratio (j,k,l) for CTRL run (a,d,g,j), ASACT run (b,e,h,j), and SCATSAT run (c,f,i,l).

The impact of assimilating three different sets of observations corresponding to three experiments are examined by analyzing the analysis increment in the horizontal wind components, temperature and humidity fields at 1000 hPa and are shown in Fig. 6.3. In general, there are no marked differences in the pattern of increments among the three experiments. However, the intensity of impact in all the four fields does vary in all the three experiments. In the zonal component of velocity (Fig. 6.3 a, b, c), positive increments are seen over the central Arabian Sea while, stronger negative increments are observed over the eastern Bay of Bengal in all the three experiments. It is interesting to note that the positive increments in u-wind are relatively higher in ASCAT and considerably more for SCATSAT experiments. The above indicates that the SCATSAT experiment is simulating stronger u-wind fields than the CTRL and ASCAT experiments. Increments in the meridional wind fields (Fig. 6.3 d, e, f) suggest that, the extent of negative increments are more over most of the oceanic regions. The negative increments (for v wind) in the SCATSAT experiment are notably lower over west-central Indian Ocean as compared to the other two experiments. This also indicates that the SCATSAT experiment simulates more stronger wind field (v-wind). The temperature field also shows notable differences among the three experiments. As seen in the case of wind fields, the temperature simulated by the SCATSAT experiment is also on the higher side when compared with the CTRL run and ASCAT run. The large extent of positive increments can be observed over the northern Arabian Sea and central Bay of Bengal for the SCATSAT experiment. Correspondingly, the extent of negative increments in temperature (over western Indian Ocean) are lower for the SCATSAT experiment. Similarly, the magnitude of humidity field simulated with the SCATSAT experiment is slightly higher, as seen from the marginally higher positive increments over central and eastern Indian Ocean. Hence, the analysis increments suggest that the fields analyzed with the assimilation of Scatsat-1 winds (SCATSAT experiment) are in general characterized with stronger winds, higher temperature and humidity fields.

The most striking observation from Fig. 6.3 is that, the CTRL experiment, in which no scatterometer observations are assimilated also produced increments patterns similar to ASCAT experiment and SCATSAT experiment, where scatterometer observations are assimilated. The above suggests that, the general pattern of lower level wind field is getting successfully captured even without the assimilation of ocean surface winds from the scatterometer. In order to validate the above inference, the distribution of atmospheric motion vectors assimilated in CTRL experiment (which happens to be the same in the other two experiments as well) is examined further. Fig. 6.4 (a) shows the distribution of AMVs at lower levels, i.e, 1000-900 hPa while Fig. 6.4 (b) depicts the total available

AMVs over the troposphere (1000-100 hPa) over the domain. The above-mentionedfigure suggests that, there are a very large number of wind observations (AMVs) over the lower troposphere over the Indian Ocean region. Furthermore, there is an extensive and dense coverage of AMVs (Fig. 6.4 b) over the Indian Ocean region and Indian subcontinent (typically ~40000 AMVs were assimilated at 00Z of everyday). The AMVs available in the lower troposphere would account for the overall behaviour of winds over the oceanic region as revealed from the increments corresponding to the location of lower level AMVs assimilated. Hence, the CTRL experiment is also expected to perform reasonably well as compared with the ASCAT experiment and SCATSAT experiment, despite the fact that the scatterometer observations are not assimilated in the CTRL experiment.

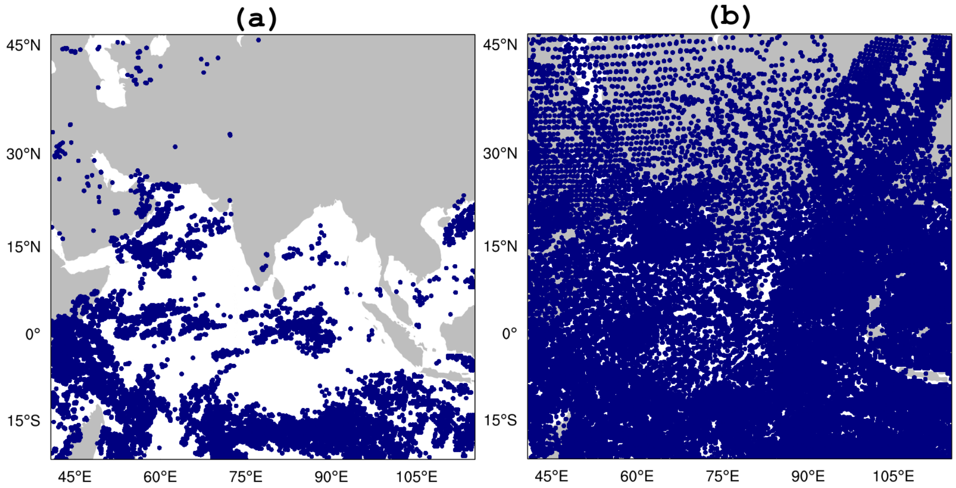

Fig. 6.4: Typical distribution of AMVs over the experiment domain: AMVs available at lower levels (1000-900 hPa) (a) and the total available AMVs (b).

The validation of the analysis fields with the radiosonde observations is examined by estimating the RMS-fit error in the zonal wind, meridional wind, air temperature and water vapor mixing ratio and are shown in Fig. 6.5. It is seen that, there are no discernible differences in the wind fields analyzed by the three experiments. A marginal difference in temperature field is seen and, the RMS-fit error for the SCATSAT temperature field is found to be slightly higher. There are, however, no notable differences among the humidity fields simulated by the three experiments. The three experiments did not show any marked differences in the winds fields, probably due to the very large number of AMVs assimilated in all the three experiments, that predominantly account for most of the wind observations over the domain.

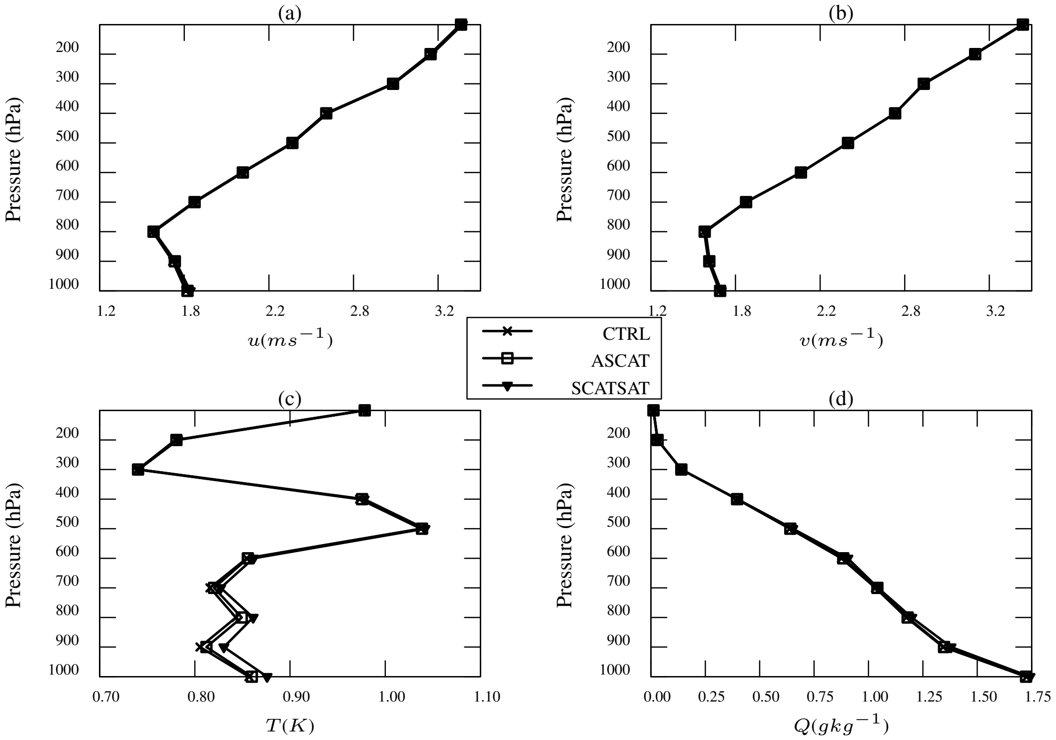

Fig. 6.5: RMS fit to radiosonde observations for u wind (a), v wind (b), temperature (c) and water vapor mixing ratio (d) for CTRL, ASCAT, and SCATSAT runs.

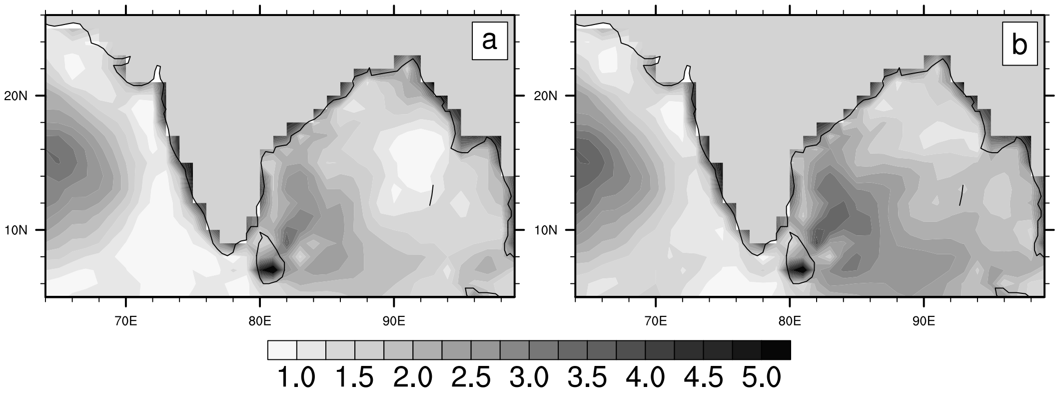

Fig. 6.6: RMSE in wind speed for ASCAT run (a) and SCATSAT run (b) with respect to the wind speed from OAFlux data.

Furthermore, the RMSE in surface wind speeds from the analysis fields have been estimated with respect to the wind speed obtained from Objectively Analyzed wind speed data from Woods Hole Oceanographic Institution. Fig. 6.6 depicts the RMSE in wind speed for ASCAT run (a) and SCATSAT run (b) for the entire month. It can be noted that there are no significant differences in the RMSE of wind speed among the ASCAT and SCATSAT experiments, and both follow a similar pattern. Nevertheless, the RMSE for SCATSAT simulated surface wind speed is slightly higher as compared with that of ASCAT simulated surface wind speed, particularly over the Bay of Bengal region. In general, over the Arabian Sea region over which the low level jet (LLJ) manifests during the Indian summer monsoon period, the RMSE in wind speed is on a higher side for both the experiments.

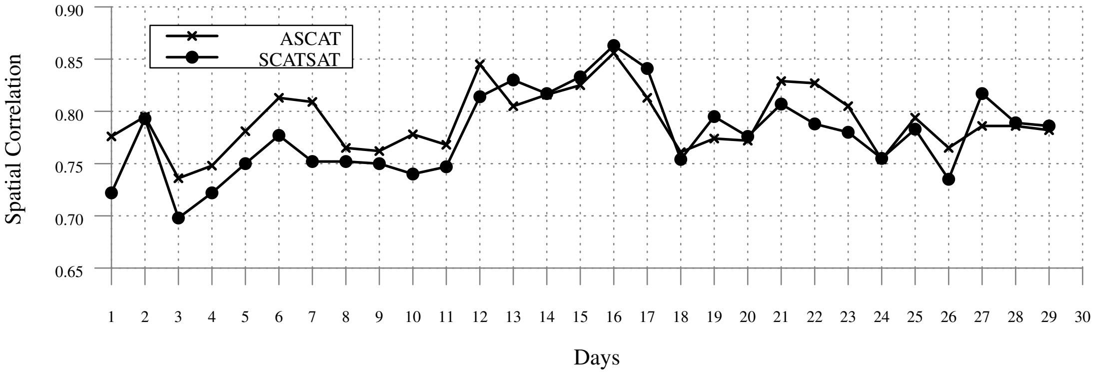

Fig. 6.7: Time series of spatial correlation in wind speed for ASCAT run and SCATSAT run with respect to the wind speed from OA Flux data over the entire second domain.

The spatial correlation between the surface wind speed simulated by the ASCAT and SCATSAT experiments (analysis fields) with the wind speed from the Objectively Analyzed (OA) flux data is examined and the time evolution of the same for the month of July, 2017 is shown in Fig. 6.7 over the entire second domain. The above figure suggests that both the experiments have simulated the surface wind speed reasonably well and on most of the days the correlation of ASCAT and SCATSAT analyzed surface wind speed with the OA flux data is above 75%. As revealed from Fig. 6.6, the ASCAT experiment is simulating the surface wind speed slightly more accurately and for most of the days, the correlation between the ASCAT wind and OA flux wind is slightly higher when compared with the SCATSAT wind, notably for the first 10 days. After first 10-12 days, the spatial correlation results from both the ASCAT experiment and SCATSAT experiment are somewhat similar and comparable. The first week of July was characterized with a break condition in the summer monsoon. The break monsoon is associated with the southeastward extension of the LLJ over the eastern Arabian Sea with no LLJ over the peninsular India and strength of LLJ is also found to be higher during the break periods [156, 157, 158]. The initial disagreement with the SCATSAT experiment for the first 10 days may be due to the presence of above normal wind speed during the first week of July.

6.6.2 Forecast verification

The short-range model forecasts (24h and 48h) corresponding to the three experiments were compared with the ERA-interim reanalysis and are shown in Fig. 6.8 - Fig. 6.9. Fig. 6.8 shows the 24h and 48h mean sea level pressure (MSLP) forecast fields from CTRL, ASCAT, and SCATSAT experiments, together with MSLP from ERA-interim reanalysis that are valid at 00Z 11 July 2017 (a-d) and 00Z 12 July 2017 (e-h). Similarly, geopotential height at 500 hPa (Fig. 6.9), temperature at 850 hPa (Fig. 6.10), and water vapor mixing ratio at 850 hPa (Fig. 6.11) are also shown. It is evident that, all the model runs have captured the position and intensity of monsoon trough quite well, when compared with the MSLP from ERA-interim reanalysis. However, the low pressure systems along the foothills of Himalayas simulated in all the experiments are relatively stronger. It is important to note that, there are no marked differences among the MSLP fields simulated by the three experiments. The geopotential field height values simulated by the model runs (Fig. 6.9) are in general higher as compared with the ERA-interim reanalysis fields. Nevertheless, all the experiments have reproduced the spatial pattern of the height fields and the position of low pressure system reasonably well. In the model simulated height fields, most of the oceanic region are characterized by higher height field values with notables differences seen over the Bay of Bengal region. The ERA-interim reanalysis reveals that lower height fields prevail over the northern Bay of Bengal, including the head-Bay. Marginal differences can be observed in SCATSAT simulated 24h and 48h height fields, particularly, the lower height fields extend over a larger region near the west coast of India in the 48h forecast fields (Fig. 6.9 g). The temperature fields (Fig. 6.10) indicate that the model has reproduced the seasonal heat low over Pakistan/northwest India quite well. The general pattern of model simulated temperature distribution is also in good agreement with that of ERA-interim reanalysis fields. As seen for the MSLP and height fields, no marked differences in the temperature fields among the three experiments are observed. In Fig. 6.11, the ERA-interim fields indicate that the moisture distribution is more concentrated along the foothills of Himalayas, the above being co-located with monsoon trough. It is pertinent to note that the 24h forecast from the SCATSAT experiment simulated relatively drier air mass over eastern India, as compared to CTRL and ASCAT experiments. Apart

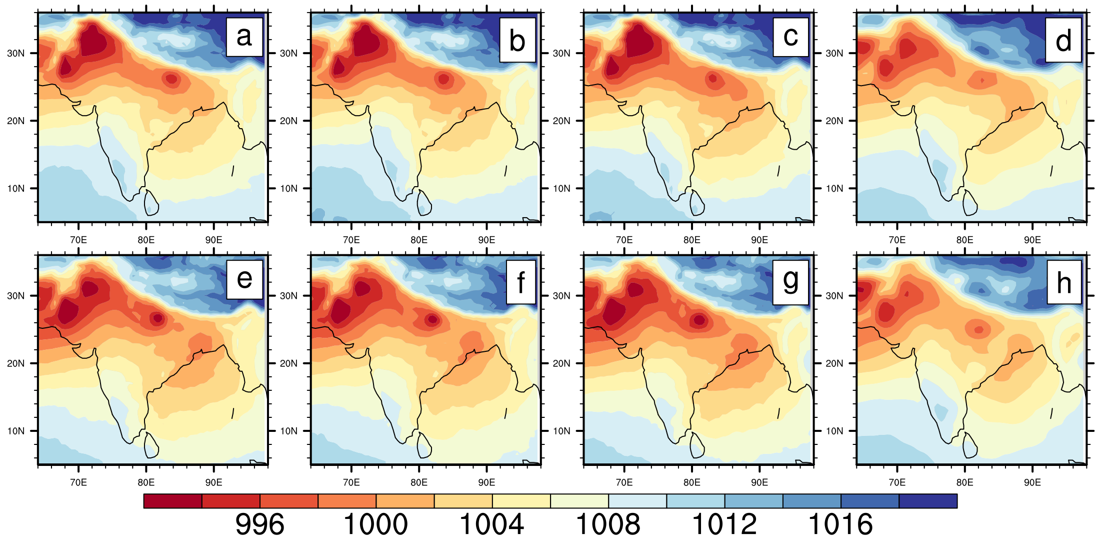

Fig. 6.8: 24h forecasts (a, b, c) and 48h forecasts (e, f, g) of MSLP for CTRL (a, e), ASCAT (b, f) and SCATSAT (c, g) run. Corresponding MSLP fields from ERA-interim reanalysis (d, h) are also shown. Model forecasts were initialized at 00Z 10 July 2017.

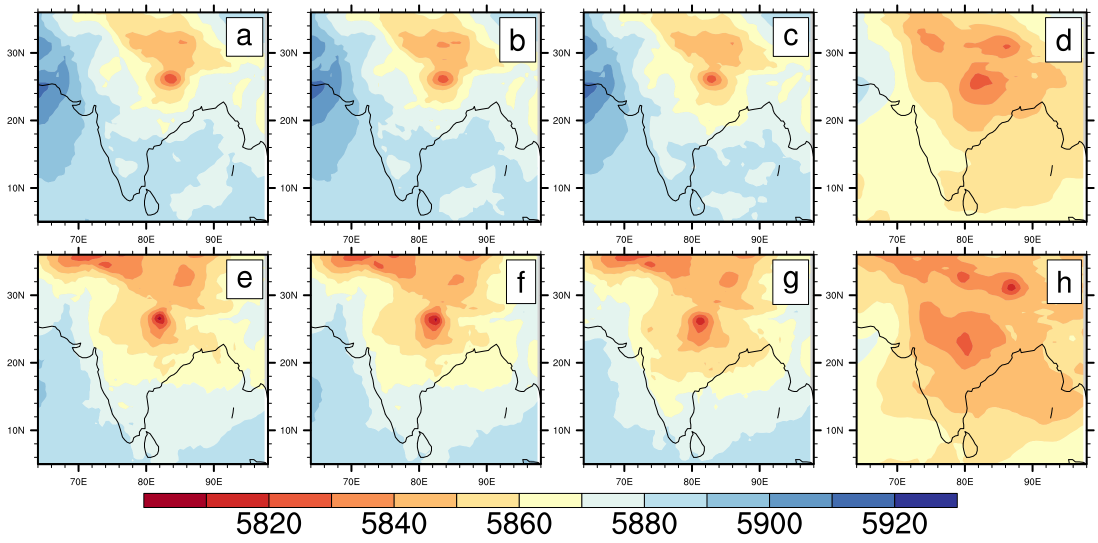

Fig. 6.9: Similar to Fig. 6.8 but for geopotential height fields at 500 hPa.

from the above feature, all the three experiments show similar pattern of moisture distribution in the 24h and 48h forecasts, and are overall in good agreement with the corresponding ERA-interim fields.

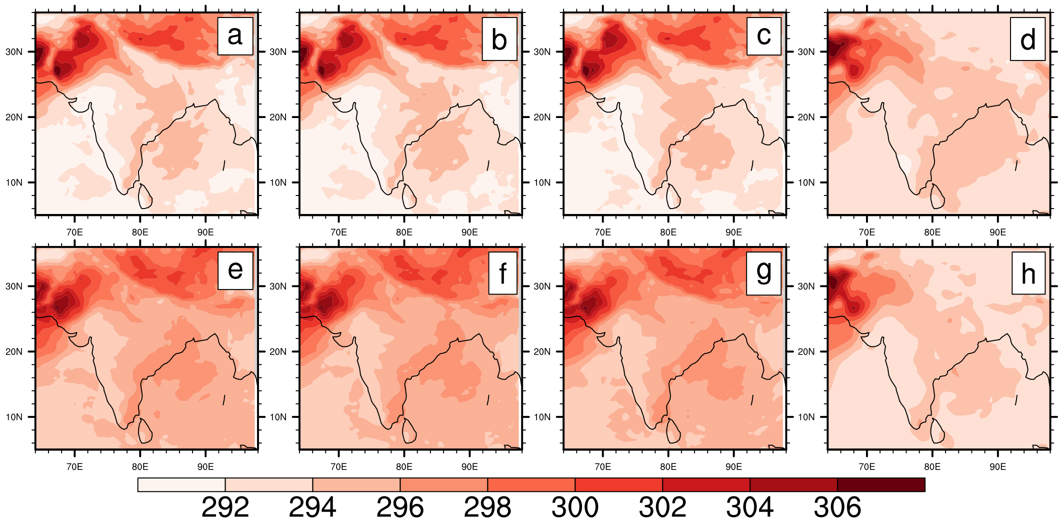

Fig. 6.10: Similar to Fig. 6.8 but for temperature fields at 850 hPa.

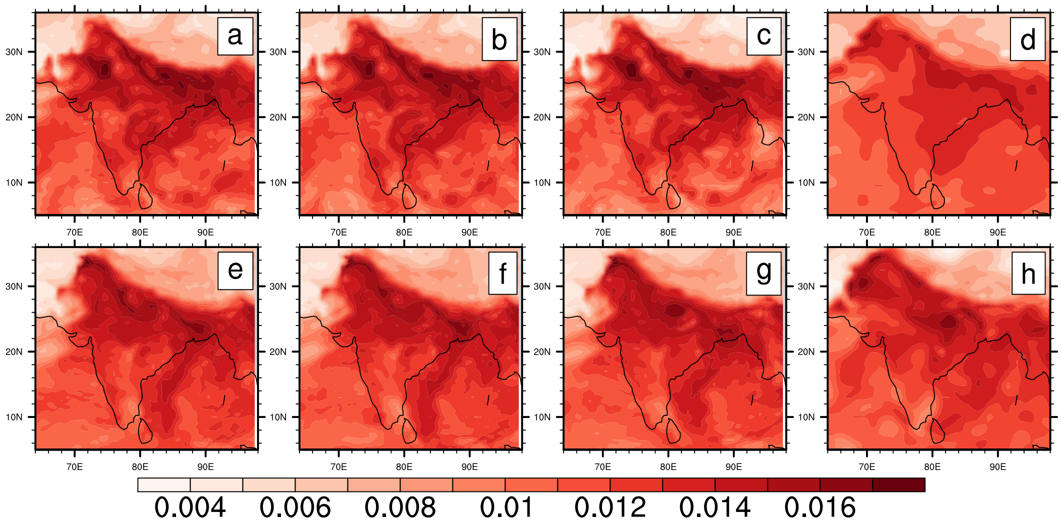

Fig. 6.11: Similar to Fig. 6.8 but for water vapor mixing ratio fields at 850 hPa.

The 24h and 48h accumulated rainfall forecast by the three model runs have been verified quantitatively by computing the skill scores ETS, bias, FAR, and POD. The skill scores are calculated with respect to the TRMM rainfall estimates (3B42). The average skill scores (averaged over 30 samples) for 24h and 48h accumulated rainfall are shown in Fig. 6.12 and Fig. 6.13 respectively. It is noted that, the ASCAT and SCATSAT experiments have similar skills in reproducing the 24h accumulated rainfall up to 30 mm rainfall threshold. The bias score for 24h rainfall forecast is found to be consistently lower for ASCAT run and considerably higher for SCATSAT run for the higher rainfall thresholds. The CTRL run has slightly more number of false alarms as compared to the other two experiments, for most of the thresholds, whereas the ASCAT run has the least number of false alarms consistently for all the rainfall thresholds. The SCATSAT experiment shows marginal improvement in reproducing the rainfall distribution in terms of slightly higher POD values, especially for lower rainfall thresholds.

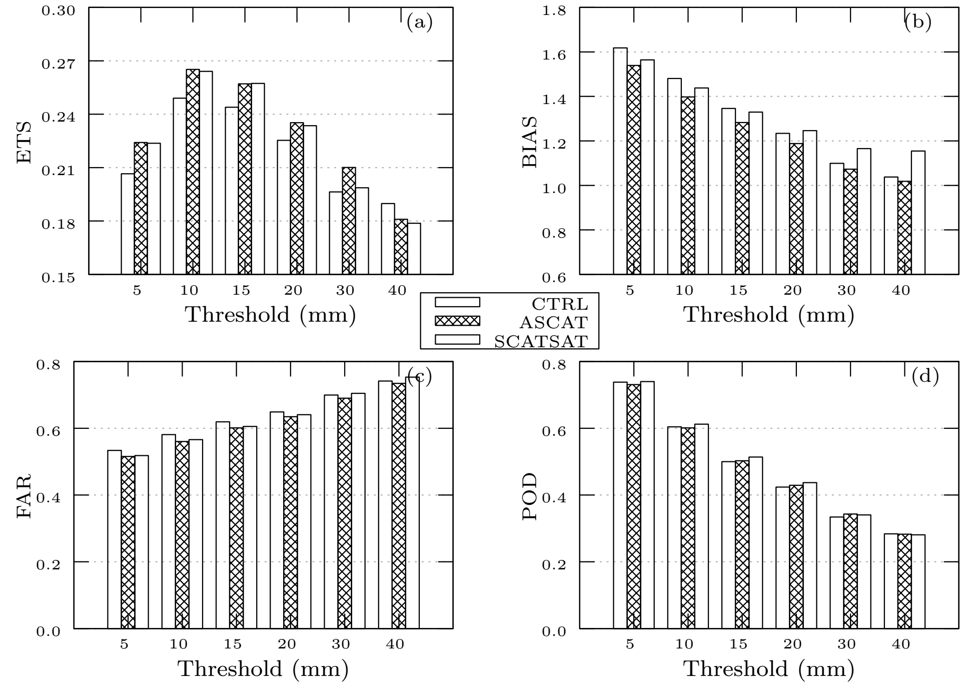

Fig. 6.12: Quantitative skill scores for 24h accumulated rainfall for all the three experiments. Skill scores are averaged for 30 samples.

The results for 48h rainfall forecast also indicate that, both the ASCAT and SCATSAT experiments have better skill in reproducing the low and the medium level of rainfall distribution, while the skill scores are comparable for the higher rainfall thresholds for all the three experiments. Consistent with the results for 24h rainfall forecasts, the ASCAT run shows less overestimation as compared with the other two experiments. When compared with the bias score of the CTRL run, the SCATSAT experiment shows slightly lower overestimation, particularly in lower and medium rainfall thresholds. Similarly, the CTRL experiment has relatively more number of false alarms for all the rainfall thresholds, when compared with the other two experiments. Further, among the ASCAT and SCATSAT

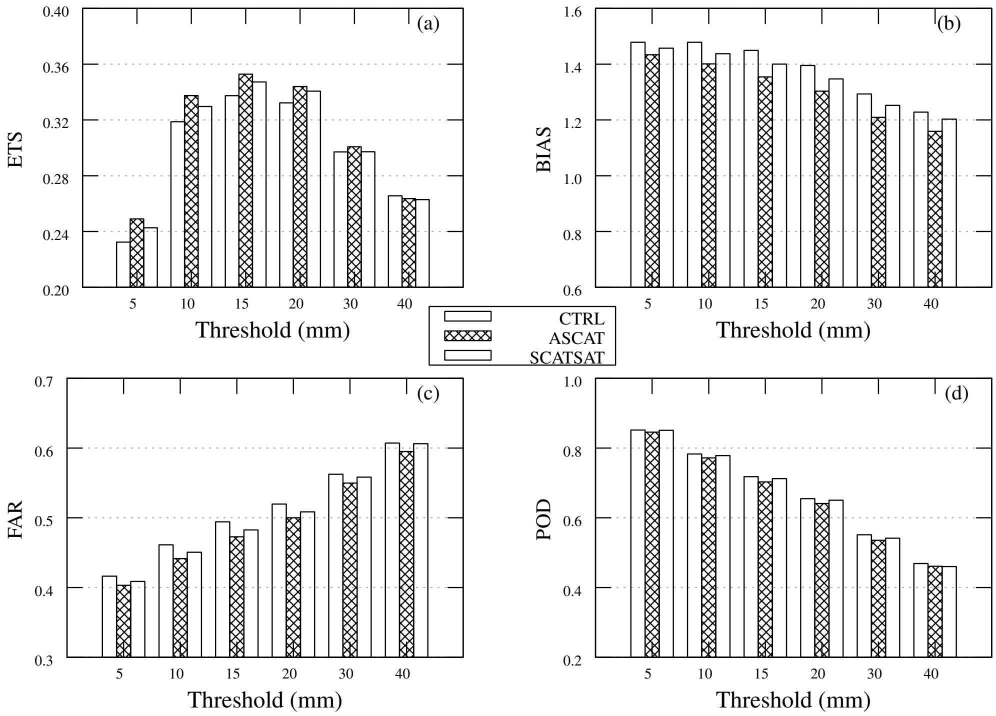

Fig. 6.13: Similar to Fig. 6.12, but for 48h accumulated rainfall.

experiments, the false alarms are relatively lower for the ASCAT run. As noticed from the results for 24h rainfall forecast, the SCATSAT run has slightly higher probability of detection in capturing the rainfall distribution, as revealed from higher POD values especially for the lower rainfall thresholds. In general, the positive impact of assimilating scatterometer winds are more evident on the 24h rainfall forecasts rather than on the 48h rainfall forecasts. Moreover, for higher rainfall thresholds, the impact of assimilating scatterometer winds are less apparent.

Fig. 6.14 shows the 24h accumulated rainfall valid at 00Z 25 July 2017 corresponding to CTRL run (a), ASCAT run (b), SCATSAT run (c), and TRMM estimate (d). All the model runs have reproduced the intense rainfall in the vicinity of the northern Bay of Bengal, near to the head bay. However, there are marked differences in the location of maximum rainfall in each of the experiments. The rainfall band along the eastern Bay of Bengal is not reproduced in any the three model runs. The SCATSAT run has simulated a region of intense rainfall over the Bay of Bengal, which is not observed in the other two model runs and in the TRMM.

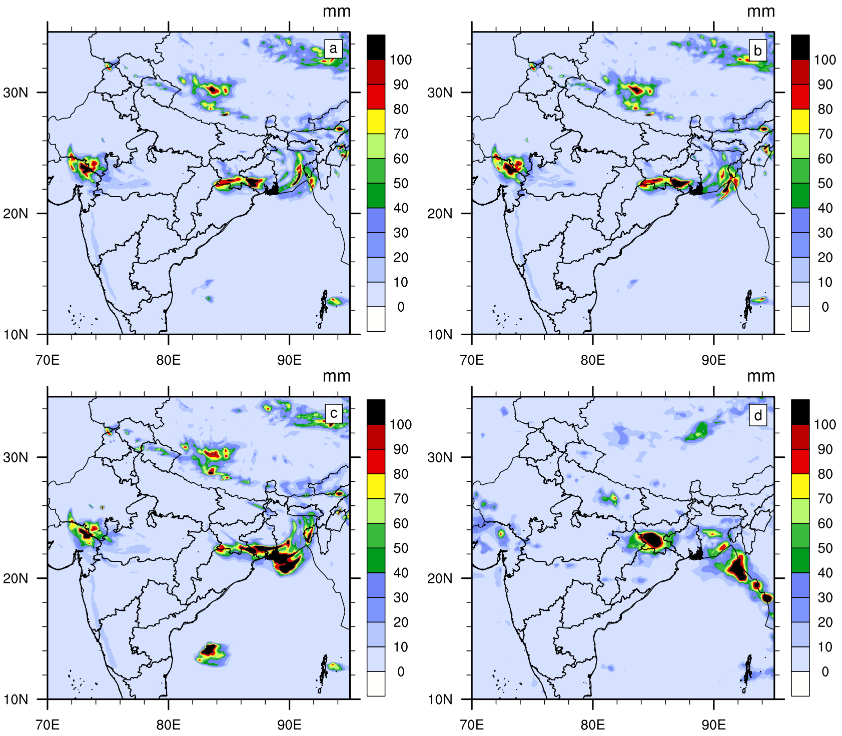

Fig. 6.14: 24h accumulated rainfall valid at 00Z 25 July 2017 corresponding to CTRL run (a), ASCAT run (b), SCATSAT run (c), and TRMM estimate (d).

Table 6.2: Relative impact of ASCAT and Scatsat-1 winds on 24h rainfall forecast.

Threshold (mm)	No: of days when ASCAT/SCATSAT has					
	higher ETS		higher POD		lower FAR	
	ASCAT	SCATSAT	ASCAT	SCATSAT	ASCAT	SCATSAT
5	14	15	7	22	15	14
10	16	13	10	19	19	10
15	16	13	15	14	18	11
20	18	11	13	14	18	11
30	18	11	14	15	20	9
40	20	9	14	15	21	8

Table 6.2 highlights the number of days in which SCATSAT and ASCAT experiment have better skill in reproducing the 24h rainfall. Similarly, Table 6.3 summarizes the comparative analysis among the skill of ASCAT and SCATSAT experiments in simulating 48h accumulated rainfall. It is noted that, the SCATSAT run has higher POD value on most of the sample days. Particularly, for about 20 days in the month of July, the SCATSAT run shows higher POD for 48h rainfall forecast, as compared to the ASCAT run for 40

mm rainfall threshold. At the same time, the SCATSAT run has more number of false alarms during the month of July. It is interesting to note that, for 24h rainfall forecast, the skill (in terms of ETS) of the SCATSAT experiment is decreasing with increase of rainfall threshold, a feature that is not observed for the 48h rainfall forecast.

Table 6.3: Relative impact of ASCAT and Scatsat-1 winds on 48h rainfall forecast.

Threshold (mm)	No: of days when ASCAT/SCATSAT has					
	higher ETS		higher POD		lower FAR	
	ASCAT	SCATSAT	ASCAT	SCATSAT	ASCAT	SCATSAT
5	19	9	11	17	20	8
10	17	11	11	17	20	8
15	17	11	11	17	20	8
20	17	11	10	18	20	8
30	16	12	9	19	18	10
40	17	11	8	20	18	10

6.7 Summary

The impact of assimilating the ocean surface winds observed by the Indian scatterometer on board Scatsat-1 has been investigated on the simulation of Indian summer monsoon and its associated rainfall for the month of July 2017. The skill of the model in simulating the Indian southwest monsoonal system with the assimilation of Scatsat-1 winds has been further compared viz-a-viz with a simulation that assimilated ASCAT winds.

Both ASCAT and SCATSAT analyzed surface wind speed are found to have reasonable correlation ($\sim$75%) with respect to the wind speed obtained from OA flux wind speed. It is observed that both ASCAT and SCATSAT experiments have positive impact on the simulation of Indian summer monsoon and associated rainfall, especially in simulating low and medium intensity rainfall amounts. The results of the quantitative verification of 24h and 48h rainfall forecasts indicate that, among the SCATSAT and ASCAT experiments, the ASCAT experiment has slightly higher skill scores. It is pertinent to note that the SCATSAT experiment provides for improved skill in capturing the rainfall distribution, in terms of POD values especially at lower rainfall thresholds. Studies have reported that the C-band radiometers are less sensitive to rain, while the Ku-band radiometers are more susceptible to contamination from rainfall [159, 160]. This could be a possible reason for improved performance of the model with the assimilation of ASCAT winds as compared to the assimilation of Scatsat-1 winds. Also, the CTRL experiment (in which no scatterometer observations are included), is also found to have similar patterns of analysis increments and reproduces the general overall features of the monsoonal system. Especially, the CTRL run also provides for reasonable rainfall forecasts as compared to the other two experiments, notably at higher rainfall thresholds. The assimilation of dense coverage of wind observations over the domain (predominantly AMVs) contribute to successful reproduction of the general features of the tropospheric circulation and hence contribute to reasonable performance of the model (in terms of the simulation of MSLP, geopotential height, temperature and moisture fields) even without the assimilation of ocean surface winds, observed from the scatterometers.

Chapter 7

Conclusions

"In all chaos there is a cosmos, in all disorder a secret order"

– Carl Gustav Jung

The variational techniques are being utilized for data assimilation purpose by global NWP centers since 1990s. Sustained efforts were pursued to further advance the DA capabilities through developing more sophisticated DA procedures. With the advent of remote sensing observational systems and more powerful computational resources, the effective utilization of the available observations became possible through appropriate assimilation methods, which finally contributed to improved initial conditions for the NWP models. The present thesis explores on a few aspects of the variational DA methods with emphasis on the weather systems over India, through regional DA studies using the state-of-the-art WRF model.

Knowing fully well that the studies which examine the relative performance of the 4DVar DA system over the 3DVar DA system in the simulation of severe weather systems like tropical cyclones over the Indian region are extremely limited, a preliminary study was carried out (in Chapter 3) to analyze the skill of the WRF model in simulating two TCs, cyclone Thane and cyclone Hudhud with regional 3DVar and 4DVar initialization. The results from the study revealed considerably improved representation of the vortex features of the tropical cyclone in the 4DVar analysis fields (at 00h forecast). The forecast fields also indicated notable and marked improvements in the simulation of cyclone track and intensity for the 4DVar run, especially for the first 24h of forecast lead time. Furthermore, precipitation forecasts from the 4DVar run also showed clear positive impact in the case of TC Thane. A marked degradation in the track and rainfall forecast of TC Hudhud with the 4DVar run have been attributed to the manifestation of anomalous northeasterlies in the upper troposphere.

Goaded by the encouraging results from the previous chapter (Chapter 3), a more systematic study was undertaken in Chapter 4 to quantify the improved performance of the 4DVar DA system over the 3DVar DA system in the simulation of four TCs over the Indian region by generating a large number of analysis/forecast samples. Apart from the conventional surface and upper-air observations and satellite based wind observations, radiance information from different satellites were also assimilated in the study. The analysis increment from the 4DVar experiments showed clear flow-depended structure, owing to the implicit evolution of model errors. The results of the study indicated that the 4DVar analysis fields (at 0h forecast) led to significant improvement ($\sim$50%) in reproducing the intensity and location of the storm, over the corresponding 3DVar analysis fields. The model physics involved in the tangent linear/adjoint model integration during the 4DVar minimization plays an important role in modulating the moisture flux in the analysis fields, which in turn contributed to improved rainfall simulation associated with the TCs. On an average, the 4DVar runs showed an improvement of 33% in simulating the position of the storm and 17% improvement in simulating the intensity of the storm over the corresponding 3DVar runs at 48h forecast lead time. Interestingly, the 4DVar run successfully simulated the rapid intensification phase of TC Phailin, during the period 00Z 10 to 00Z 11 October 2013, whereas, the corresponding 3DVar run could not simulate the rapid intensification process realistically. The above inference suggests that the 3DVar DA system maybe inadequate for accurate simulation of rapidly evolving weather systems. Moreover, it can also be inferred that the asynoptic observations are more appropriately utilized with the 4DVar DA method, which allows for the observations to be assimilated at the exact time of observation.

In general, the error statistics of the model (the BEC matrix) has a critical role in modulating the impact of the observations on the analysis field. In the variational methods, particularly in the 3DVar procedure, the model errors are expressed as a climatological mean and hence do not provide for the time evolution of the model errors. However, in the 4DVar approach the model integration involved in the 4DVar minimization procedure implicitly provides for the evolution of model errors. Since the 4DVar studies which have explored the impact of employing different BEC formulations on the analysis and forecast fields are relatively less, Chapter 5 was devoted to address the above problem with emphasis on heavy rainfall events over the Himalayan region, India. Three different BEC formulations, which employ three different sets of control variables were considered in the study. Among the three BEC formulations utilized in Chapter 5, two formulations utilize stream function and velocity potential as momentum variables ($\psi\chi$-BE and $\psi\chi$-MBE) and the other one uses horizontal wind components as momentum variables (uv-BE). In the $\psi\chi$-BE experiment, the humidity variable is univariate and in the $\psi\chi$-MBE experiment, the humidity variable is multivariate in nature. On the other hand, all the control variables in the uv-BE experiment are univariate in nature. The results from the single observation 4DVar assimilation experiment suggest that both the $\psi\chi$-BE and $\psi\chi$-MBE experiments provide for wider impact on the analysis fields, whereas, the impact is highly localized with the uv-BE experiment. The real 4DVar DA experiments indicated that the uv-BE analysis fields are consistently closer to the radiosonde observations. The 48h rainfall forecasts from the $\psi\chi$-MBE experiment, on an average, reproduces the observed rainfall distribution, more accurately than the other two experiments. Nevertheless, the uv-BE experiment has better skill in reproducing the time evolution of the rainfall intensity. When the above study was extended to a case study of the Indian summer monsoon period, the results obtained with the latter are in agreement with the results for the heavy rainfall cases. The uv-BE analysis fields are considerably closer to the radiosonde observed fields and the $\psi\chi$-MBE experiment results showed marginal improvement in rainfall forecasts as compared with the other two experiments. In general, the results from different BEC experiments suggest that incorporating humidity as a multivariate variable in the BEC formulation ($\psi\chi$-MBE experiment) provides for a clear positive impact on the rainfall forecasts. In effect, the $\psi\chi$-MBE experiment is found to suppress the overestimation and and associated false predictions in rainfall forecast seen in the $\psi\chi$-BE experiment. A comparative study with 3DVar experiments suggests that, even for the univariate BEC employed in the 4DVar DA system, the multivariate correlations among the analysis variables would evolve implicitly in the 4DVar experiments.

Finally, the impact of assimilating ocean surface winds from the scatterometer on board Scatsat-1 on the simulation of Indian summer monsoon and associated rainfall during the month of July, 2017 is investigated in Chapter 6. Further, the above impact is then compared with the impact of assimilating ocean surface winds from ASCAT also. Both the SCATSAT and ASCAT experiments are found to reproduce the surface winds with reasonable accuracy with respect to the surface wind speed data obtained from OA flux project(>75% correlation). The results from the study showed that the assimilation of winds from Scatsat-1 do have clear positive impact on the simulation of 24h and 48h accumulated rainfall, especially for lower and medium rainfall intensities. Further, the comparative study with the ASCAT experiments indicated that, both ASCAT and SCATSAT experiments have similar skill in reproducing the 24h rainfall distribution, however, the ASCAT experiment showed relatively better skill in reproducing the 48h rainfall distribution. It has been reported in literature that the observations from a C-band scatterometer are more reliable during rainfall when compared with the observations from a Ku-band scatterometer, owing to the fact that the Ku-band radiometers are more susceptible to contamination from rain. Hence, the above inference may provide the possible reason for improved skill of the ASCAT experiments in rainfall simulation.

The major findings of the present study are highlighted below.

- The WRF-4DVar DA system contributed to significant improvements in reproducing the initial vortex structure of all the four tropical cyclone cases that are investigated in this study.

- For tropical cyclone forecasts, the 4DVar run showed an improvement of 17-50% in intensity simulation (MSLP) and an improvement of 22-57% in simulating the TC track, at different forecast lead times.

- The moist-physics scheme involved in the WRF-4DVar contributed positively to the analysis fields through modulating the humidity fields, which in turn yielded improved rainfall forecasts.

- The 4DVar DA system implicitly provides for the multivariate correlations among the analysis variables, even when a completely univariate BEC is utilized in the assimilation.

- The analysis fields as obtained from the uv-BE experiment are consistently closer to the radiosonde observed fields.

- Representation of moisture variable in the multivariate form in BEC formulation ($\psi\chi$-MBE) have a significant impact in suppressing the overestimation and associated false predictions in the rainfall forecast, seen in the $\psi\chi$-BE experiment.

- The uv-BE experiment provides for relatively accurate simulation of time evolution of the rainfall intensities.

- Assimilating ocean surface winds from Scatsat-1 yielded clear positive impact on the 24h and 48h rainfall forecasts, notably at lower and medium rainfall thresholds.
- For 24h rainfall forecasts, assimilation of Scatsat-1 winds have positive impact similar to the assimilation of ASCAT winds. However, the ASCAT run showed better impact for 48h rainfall forecast, especially for higher thresholds as compared to the SCATSAT run.

Future scope

The present study does not provide for experiments utilizing any of the Kalman filter based assimilation methods. The assimilation procedures based on the hybrid methods which blend the advantages of both variational and Kalman filter method together, have remarkably good skills as compared to the 4DVar DA method. The comparative analysis of the results of the present study with the results from Kalman filter based methods and/or from the hybrid methods could have provided interesting results, particularly in the context of tropical weather systems. Hence, an interesting and important extension of the experiments in the present study would be application of Kalman filter based and/or hybrid DA methods (like 4DEnVar), which may have greater positive impact on the accurate simulation of rapidly evolving tropical weather systems. Another interesting aspect would be to investigate the impact of the uv based BEC with the assimilation of high-resolution observations (eg: radar), since the uv-based BEC formulation is known to have better impact over smaller region having high-resolution observations. Results from the present study also reiterate that accurate prediction of the location and the magnitude of highest rainfall associated with extreme rainfall events are still very challenging. Hence, observation targeting data assimilation studies may be carried out employing adjoint sensitivity analysis for severe weather systems over India.

Bibliography

[1] S. G. Philander, *Our affair with El Niño: How we transformed an enchanting Peruvian current into a global climate hazard.* Princeton University Press, 2006.

[2] J. G. Haltiner, *Numerical Weather Prediction.* John Wiley & Sons, Inc., New York, 1971.

[3] L. F. Richardson, *Weather Prediction by Numerical Process.* Cambridge University Press, 1922.

[4] P. Lynch, *The Emergence of Numerical Weather Prediction: Richardson's Dream.* Cambridge University Press, 2006.

[5] J. G. Charney, "Dynamic Forecasting by Numerical Process," in *Compendium of Meteorology.* Springer, 1951, pp. 470–482.

[6] N. A. Phillips, "The general circulation of the atmosphere: A numerical experiment," *Quarterly Journal of the Royal Meteorological Society*, vol. 82, no. 352, pp. 123–164, 1956.

[7] A. Chandrasekar, *Basics of Atmospheric Science.* PHI Learning Pvt. Ltd., 2010.

[8] P. D. Thompson, "A history of numerical weather prediction in the United States," *Bulletin of the American Meteorological Society*, vol. 64, no. 7, pp. 755–769, 1983.

[9] E. Kalnay, *Atmospheric Modeling, Data Assimilation and Predictability.* Cambridge university press, 2003.

[10] R. Panofsky, "Objective weather-map analysis," *Journal of Meteorology*, vol. 6, no. 6, pp. 386–392, 1949.

[11] B. Gilchrist and G. P. Cressman, “An experiment in objective analysis,” *Tellus*, vol. 6, no. 4, pp. 309–318, 1954.

[12] R. Daley, *Atmospheric Data Analysis*. Cambridge university press, 1993.

[13] O. Talagrand, “Assimilation of observations, an introduction,” *Journal of the Meteorological Society of Japan. Ser. II*, vol. 75, no. 1B, pp. 191–209, 1997.

[14] L. S. Gandin, “Objective analysis of meteorological fields,” *Israel program for scientific translations*, vol. 242, 1963.

[15] D. F. Parrish and J. C. Derber, “The National Meteorological Center’s spectral statistical-interpolation analysis system,” *Monthly Weather Review*, vol. 120, no. 8, pp. 1747–1763, 1992.

[16] J. M. Lewis and J. C. Derber, “The use of adjoint equations to solve a variational adjustment problem with advective constraints,” *Tellus A*, vol. 37, no. 4, pp. 309–322, 1985.

[17] A. C. Lorenc and F. Rawlins, “Why does 4D-Var beat 3D-Var?” *Quarterly Journal of the Royal Meteorological Society*, vol. 131, no. 613, pp. 3247–3257, 2005.

[18] G. Evensen, “Using the extended Kalman filter with a multilayer quasi-geostrophic ocean model,” *Journal of Geophysical Research: Oceans*, vol. 97, no. C11, pp. 17 905–17 924, 1992.

[19] ——, “Inverse methods and data assimilation in nonlinear ocean models,” *Physica D: Nonlinear Phenomena*, vol. 77, no. 1-3, pp. 108–129, 1994.

[20] F. Rabier, “Overview of global data assimilation developments in numerical weather-prediction centres,” *Quarterly Journal of the Royal Meteorological Society*, vol. 131, no. 613, pp. 3215–3233, 2005.

[21] F.-X. L. Dimet and O. Talagrand, “Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects,” *Tellus A*, vol. 38, no. 2, pp. 97–110, 1986.

[22] J. C. Derber, “A variational continuous assimilation technique,” *Monthly Weather Review*, vol. 117, no. 11, pp. 2437–2446, 1989.

[23] P. Courtier, E. Andersson, W. Heckley, D. Vasiljevic, M. Hamrud, A. Hollingsworth, F. Rabier, M. Fisher, and J. Pailleux, “The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation,” *Quarterly Journal of the Royal Meteorological Society*, vol. 124, no. 550, pp. 1783–1807, 1998.

[24] F. Rabier, A. McNally, E. Andersson, P. Courtier, P. Unden, J. Eyre, A. Hollingsworth, and F. Bouttier, “The ECMWF implementation of three-dimensional variational assimilation (3D-Var). II: Structure functions,” *Quarterly Journal of the Royal Meteorological Society*, vol. 124, no. 550, pp. 1809–1829, 1998.

[25] E. Andersson, J. Haseler, P. Undén, P. Courtier, G. Kelly, D. Vasiljevic, C. Brankovic, C. Gaffard, A. Hollingsworth, C. Jakob *et al.*, “The ecmwf implementation of three-dimensional variational assimilation (3D-Var). III: Experimental results,” *Quarterly Journal of the Royal Meteorological Society*, vol. 124, no. 550, pp. 1831–1860, 1998.

[26] P. Gauthier, C. Charette, L. Fillion, P. Koclas, and S. Laroche, “Implementation of a 3D variational data assimilation system at the Canadian meteorological centre. Part I: The global analysis,” *Atmosphere-Ocean*, vol. 37, no. 2, pp. 103–156, 1999.

[27] A. Lorenc, S. Ballard, R. Bell, N. Ingleby, P. Andrews, D. Barker, J. Bray, A. Clayton, T. Dalby, D. Li *et al.*, “The Met. Office global three-dimensional variational data assimilation scheme,” *Quarterly Journal of the Royal Meteorological Society*, vol. 126, no. 570, pp. 2991–3012, 2000.

[28] J.-N. Thepaut and P. Courtier, “Four-dimensional variational data assimilation using the adjoint of a multilevel primitive-equation model,” *Quarterly Journal of the Royal Meteorological Society*, vol. 117, no. 502, pp. 1225–1254, 1991.

[29] F. Rabier and P. Courtier, “Four-dimensional assimilation in the presence of baroclinic instability,” *Quarterly Journal of the Royal Meteorological Society*, vol. 118, no. 506, pp. 649–672, 1992.

[30] I. M. Navon, X. Zou, J. Derber, and J. Sela, “Variational data assimilation with an adiabatic version of the NMC spectral model,” *Monthly Weather Review*, vol. 120, no. 7, pp. 1433–1446, 1992.

[31] P. Courtier, J.-N. Thépaut, and A. Hollingsworth, “A strategy for operational implementation of 4D-Var, using an incremental approach,” *Quarterly Journal of the Royal Meteorological Society*, vol. 120, no. 519, pp. 1367–1387, 1994.

[32] F. Rabier, H. Järvinen, E. Klinker, J.-F. Mahfouf, and A. Simmons, “The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics,” *Quarterly Journal of the Royal Meteorological Society*, vol. 126, no. 564, pp. 1143–1170, 2000.

[33] J.-F. Mahfouf and F. Rabier, “The ECMWF operational implementation of four-dimensional variational assimilation. II: Experimental results with improved physics,” *Quarterly Journal of the Royal Meteorological Society*, vol. 126, no. 564, pp. 1171–1190, 2000.

[34] M. Janisková, J.-N. Thépaut, and J.-F. Geleyn, “Simplified and regular physical parameterizations for incremental four-dimensional variational assimilation,” *Monthly Weather Weview*, vol. 127, no. 1, pp. 26–45, 1999.

[35] P. Gauthier and J.-N. Thepaut, “Impact of the digital filter as a weak constraint in the preoperational 4DVar assimilation system of Météo-France,” *Monthly Weather Review*, vol. 129, no. 8, pp. 2089–2102, 2001.

[36] G. Desroziers, G. Hello, and J.-N. Thépaut, “A 4D-Var re-analysis of FASTEX,” *Quarterly Journal of the Royal Meteorological Society*, vol. 129, no. 589, pp. 1301–1315, 2003.

[37] M. Zupanski, “Regional four-dimensional variational data assimilation in a quasi-operational forecasting environment,” *Monthly Weather Review*, vol. 121, no. 8, pp. 2396–2408, 1993.

[38] N. Gustafsson and X.-Y. HUANG, “Sensitivity experiments with the spectral HIRLAM and its adjoint,” *Tellus A*, vol. 48, no. 4, pp. 501–517, 1996.

[39] A. C. Lorenc, “Development of an operational variational assimilation scheme,” *Journal of the Meteorological Society of Japan. Ser. II*, vol. 75, no. 1B, pp. 339–346, 1997.

[40] N. Gustafsson, P. Lonnberg, and J. Pailleux, “Data assimilation for high-resolution limited-area models,” *Journal of the Meteorological Society of Japan. Ser. II*, vol. 75, no. 1B, pp. 367–382, 1997.

[41] X.-Y. Huang, N. Gustafsson, and E. Källén, "Using an adjoint model to improve an optimum interpolation-based data-assimilation system," *Tellus A*, vol. 49, no. 2, pp. 161–176, 1997.

[42] N. Gustafsson, L. Berre, S. Hörnquist, X.-Y. Huang, M. Lindskog, B. Navascues, K. Mogensen, and S. Thorsteinsson, "Three-dimensional variational data assimilation for a limited area model. Part I: General formulation and the background error constraint," *Tellus A*, vol. 53, no. 4, pp. 425–446, 2001.

[43] M. Lindskog, N. Gustafsson, B. Navascues, K. S. Mogensen, X.-Y. Huang, X. Yang, U. Andrae, L. Berre, S. Thorsteinsson, and J. Rantakokko, "Three-dimensional variational data assimilation for a limited area model: Part II: Observation handling and assimilation experiments," *Tellus A*, vol. 53, no. 4, pp. 447–468, 2001.

[44] N. Gustafsson, "Status and performance of HIRLAM 4D-Var," *HIRLAM Newsletter*, vol. 51, no. 8, 2006.

[45] M. Stengel, P. Undén, M. Lindskog, P. Dahlgren, N. Gustafsson, and R. Bennartz, "Assimilation of SEVIRI infrared radiances with HIRLAM 4D-Var," *Quarterly Journal of the Royal Meteorological Society*, vol. 135, no. 645, pp. 2100–2109, 2009.

[46] J. Dudhia, "A nonhydrostatic version of the Penn State–NCAR mesoscale model: Validation tests and simulation of an Atlantic cyclone and cold front," *Monthly Weather Review*, vol. 121, no. 5, pp. 1493–1513, 1993.

[47] D. M. Barker, W. Huang, Y.-R. Guo, A. Bourgeois, and Q. Xiao, "A three-dimensional variational data assimilation system for MM5: Implementation and initial results," *Monthly Weather Review*, vol. 132, no. 4, pp. 897–914, 2004.

[48] W. C. Skamarock, J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, M. G. Duda, W. Huang, Xiang-Yu Wang, and J. G. Powers, "A description of the advanced research WRF version 3," NCAR Technical Note, NCAR/TN-475+STR, Tech. Rep., 2008.

[49] X.-Y. Huang, Q. Xiao, D. M. Barker, X. Zhang, J. Michalakes, W. Huang, T. Henderson, J. Bray, Y. Chen, Z. Ma *et al.*, "Four-dimensional variational data assimilation for WRF: Formulation and preliminary results," *Monthly Weather Review*, vol. 137, no. 1, pp. 299–314, 2009.

[50] X. Zhang, X.-Y. Huang, and N. Pan, "Development of the upgraded tangent linear and adjoint of the weather research and forecasting (WRF) model," *Journal of Atmospheric and Oceanic Technology*, vol. 30, no. 6, pp. 1180–1188, 2013.

[51] X. Zhang, X.-Y. Huang, J. Liu, J. Poterjoy, Y. Weng, F. Zhang, and H. Wang, "Development of an efficient regional four-dimensional variational data assimilation system for WRF," *Journal of Atmospheric and Oceanic Technology*, vol. 31, no. 12, pp. 2777–2794, 2014.

[52] Y. K. Sasaki and J. S. Goerss, "Satellite data assimilation using NASA Data Systems Test 6 observations," *Monthly Weather Review*, vol. 110, no. 11, pp. 1635–1644, 1982.

[53] J. Eyre, G. Kelly, A. McNally, E. Andersson, and A. Persson, "Assimilation of TOVS radiance information through one-dimensional variational analysis," *Quarterly Journal of the Royal Meteorological Society*, vol. 119, no. 514, pp. 1427–1463, 1993.

[54] J. C. Derber and W.-S. Wu, "The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system," *Monthly Weather Review*, vol. 126, no. 8, pp. 2287–2299, 1998.

[55] S.-H. Chen, F. Vandenberghe, G. W. Petty, and J. F. Bresch, "Application of SSM/I satellite data to a hurricane simulation," *Quarterly Journal of the Royal Meteorological Society*, vol. 130, no. 598, pp. 801–825, 2004.

[56] X. Fan and J. S. Tilley, "Dynamic assimilation of MODIS-retrieved humidity profiles within a regional model for high-latitude forecast applications," *Monthly Weather Review*, vol. 133, no. 12, pp. 3450–3480, 2005.

[57] S. Sandeep, A. Chandrasekar, and D. Singh, "The impact of assimilation of AMSU data for the prediction of a tropical cyclone over India using a mesoscale model," *International Journal of Remote Sensing*, vol. 27, no. 20, pp. 4621–4653, 2006.

[58] R. Singh, P. Pal, C. Kishtawal, and P. Joshi, "The impact of variational assimilation of SSM/I and QuikSCAT satellite observations on the numerical simulation of Indian Ocean tropical cyclones," *Weather and Forecasting*, vol. 23, no. 3, pp. 460–476, 2008.

[59] S.-H. Chen, Z. Zhao, J. S. Haase, A. Chen, and F. Vandenberghe, “A study of the characteristics and assimilation of retrieved MODIS total precipitable water data in severe weather simulations,” *Monthly Weather Review*, vol. 136, no. 9, pp. 3608–3628, 2008.

[60] V. Xavier, A. Chandrasekar, R. Singh, and B. Simon, “The impact of assimilation of MODIS data for the prediction of a tropical low-pressure system over India using a mesoscale model,” *International Journal of Remote Sensing*, vol. 27, no. 20, pp. 4655–4676, 2006.

[61] V. Xavier, A. Chandrasekar, H. Rahman, D. Niyogi, and K. Alapaty, “The effect of satellite and conventional meteorological data assimilation on the mesoscale modeling of monsoon depressions over India,” *Meteorology and Atmospheric Physics*, vol. 101, no. 1-2, pp. 65–92, 2008.

[62] M. Govindankutty, A. Chandrasekar, A. Bohra, J. P. George, and M. Dasgupta, “The impact of assimilation of MODIS observations using WRF-VAR for the prediction of a monsoon depression during September 2006,” *Open Atmospheric Science Journal*, vol. 2, pp. 68–78, 2008.

[63] M. Govindankutty and A. Chandrasekar, “Impact of assimilation of ATOVS temperature and humidity and SSM/I total precipitable water on the simulation of a monsoon depression,” *Natural Hazards*, vol. 59, no. 3, pp. 1647–1669, 2011.

[64] A. Routray, U. Mohanty, D. Niyogi, S. Rizvi, and K. K. Osuri, “Simulation of heavy rainfall events over Indian monsoon region using WRF-3DVAR data assimilation system,” *Meteorology and Atmospheric Physics*, vol. 106, no. 1-2, pp. 107–125, 2010.

[65] A. Routray, S. Kar, P. Mali, and K. Sowjanya, “Simulation of monsoon depressions using WRF-VAR: Impact of different background error statistics and lateral boundary conditions,” *Monthly Weather Review*, vol. 142, no. 10, pp. 3586–3613, 2014.

[66] R. Singh, C. Kishtawal, P. Pal, and P. Joshi, “Assimilation of the multisatellite data into the WRF model for track and intensity simulation of the Indian Ocean tropical cyclones,” *Meteorology and Atmospheric Physics*, vol. 111, no. 3-4, pp. 103–119, 2011.

[67] P. Kumar, K. H. Kumar, and P. K. Pal, “Impact of Oceansat-2 scatterometer winds and TMI observations on Phet cyclone simulation,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 6, pp. 3774–3779, 2013.

[68] M. Greeshma, C. Srinivas, V. Yesubabu, C. Naidu, R. Baskaran, and B. Venkatraman, “Impact of local data assimilation on tropical cyclone predictions over the Bay of Bengal using the ARW model.” *Annales Geophysicae*, vol. 33, no. 7, 2015.

[69] M. Dhanya, D. Gopalakrishnan, A. Chandrasekar, S. K. Singh, and V. Prasad, “The impact of assimilating MeghaTropiques SAPHIR radiances in the simulation of tropical cyclones over the Bay of Bengal using the WRF model,” *International Journal of Remote Sensing*, vol. 37, no. 13, pp. 3086–3103, 2016.

[70] A. Routray, U. Mohanty, K. K. Osuri, S. Kar, and D. Niyogi, “Impact of satellite radiance data on simulations of Bay of Bengal tropical cyclones using the WRF-3DVAR modeling system,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 4, pp. 2285–2303, 2016.

[71] P. Kumar, C. Kishtawal, and P. Pal, “Impact of satellite rainfall assimilation on weather research and forecasting model predictions over the Indian region,” *Journal of Geophysical Research: Atmospheres*, vol. 119, no. 5, pp. 2017–2031, 2014.

[72] J. Ban, Z. Liu, X. Zhang, X.-Y. Huang, and H. Wang, “Precipitation data assimilation in WRFDA 4D-Var: Implementation and application to convection-permitting forecasts over united states,” *Tellus A: Dynamic Meteorology and Oceanography*, vol. 69, no. 1, p. 1368310, 2017.

[73] R. N. Bannister, “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*, vol. 134, no. 637, pp. 1951–1970, 2008.

[74] V. Rakesh and P. Goswami, “Impact of background error statistics on 3D-Var assimilation: Case study over the Indian region,” *Meteorology and Atmospheric Physics*, vol. 112, no. 1-2, pp. 63–79, 2011.

[75] ——, “Impact of background error statistics on forecasting of tropical cyclones over the north Indian ocean,” *Journal of Geophysical Research: Atmospheres*, vol. 116, no. D20, 2011.

[76] Y. Xie and A. E. MacDonald, “Selection of momentum variables for a three-dimensional variational analysis,” *Pure and Applied Geophysics*, vol. 169, no. 3, pp. 335–351, 2012.

[77] Y. Chen, S. R. Rizvi, X.-Y. Huang, J. Min, and X. Zhang, “Balance characteristics of multivariate background error covariances and their impact on analyses and forecasts in tropical and Arctic regions,” *Meteorology and Atmospheric Physics*, vol. 121, no. 1-2, pp. 79–98, 2013.

[78] M. Dhanya and A. Chandrasekar, “Impact of variational assimilation using multivariate background error covariances on the simulation of monsoon depressions over India.” *Annales Geophysicae (09927689)*, vol. 34, no. 2, 2016.

[79] X. Li, M. Zeng, Y. Wang, W. Wang, H. Wu, and H. Mei, “Evaluation of two momentum control variable schemes and their impact on the variational assimilation of radarwind data: Case study of a squall line,” *Advances in Atmospheric Sciences*, vol. 33, no. 10, pp. 1143–1157, 2016.

[80] J. Sun, H. Wang, W. Tong, Y. Zhang, C.-Y. Lin, and D. Xu, “Comparison of the impacts of momentum control variables on high-resolution variational data assimilation and precipitation forecasting,” *Monthly Weather Review*, vol. 144, no. 1, pp. 149–169, 2016.

[81] R. Laprise, “The euler equations of motion with hydrostatic pressure as an independent variable,” *Monthly weather review*, vol. 120, no. 1, pp. 197–207, 1992.

[82] L. J. Wicker and W. C. Skamarock, “Time-splitting methods for elastic models using forward time schemes,” *Monthly Weather Review*, vol. 130, no. 8, pp. 2088–2097, 2002.

[83] A. Arakawa and V. R. Lamb, “Computational design of the basic dynamical processes of the UCLA general circulation model,” *General circulation models of the atmosphere*, vol. 17, pp. 173–265, 1977.

[84] J. B. Klemp, W. C. Skamarock, and J. Dudhia, “Conservative split-explicit time integration methods for the compressible nonhydrostatic equations,” *Monthly Weather Review*, vol. 135, no. 8, pp. 2897–2913, 2007.

[85] J. Dudhia, “Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model,” *Journal of Atmospheric Sciences*, vol. 46, no. 20, pp. 3077–3107, 1989.

[86] A. A. Lacis and J. Hansen, “A parameterization for the absorption of solar radiation in the earth’s atmosphere,” *Journal of Atmospheric Sciences*, vol. 31, no. 1, pp. 118–133, 1974.

[87] E. J. Mlawer, S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, “Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave,” *Journal of Geophysical Research: Atmospheres*, vol. 102, no. D14, pp. 16 663–16 682, 1997.

[88] F. Chen and J. Dudhia, “Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and sensitivity,” *Monthly Weather Review*, vol. 129, no. 4, pp. 569–585, 2001.

[89] T. G. Smirnova, J. M. Brown, and S. G. Benjamin, “Performance of different soil model configurations in simulating ground surface temperature and surface fluxes,” *Monthly Weather Review*, vol. 125, no. 8, pp. 1870–1884, 1997.

[90] T. G. Smirnova, J. M. Brown, S. G. Benjamin, and D. Kim, “Parameterization of cold-season processes in the MAPS land-surface scheme,” *Journal of Geophysical Research: Atmospheres*, vol. 105, no. D3, pp. 4077–4086, 2000.

[91] S.-Y. Hong and H.-L. Pan, “Nonlocal boundary layer vertical diffusion in a medium-range forecast model,” *Monthly Weather Review*, vol. 124, no. 10, pp. 2322–2339, 1996.

[92] S.-Y. Hong, Y. Noh, and J. Dudhia, “A new vertical diffusion package with an explicit treatment of entrainment processes,” *Monthly weather review*, vol. 134, no. 9, pp. 2318–2341, 2006.

[93] Z. I. Janjić, “The step-mountain coordinate: Physical package,” *Monthly Weather Review*, vol. 118, no. 7, pp. 1429–1443, 1990.

[94] ——, *Nonsingular implementation of the Mellor-Yamada level 2.5 scheme in the NCEP Meso model.* NCEP office note, 2001, no. 437.

[95] H. Morrison, J. Curry, and V. Khvorostyanov, "A new double-moment microphysics parameterization for application in cloud and climate models. Part I: Description," *Journal of the Atmospheric Sciences*, vol. 62, no. 6, pp. 1665–1677, 2005.

[96] D. J. Stensrud, *Parameterization Schemes: Keys to Understanding Numerical Weather Prediction Models.* Cambridge University Press, 2009.

[97] E. Kessler, "On the distribution and continuity of water substance in atmospheric circulations," in *On the distribution and continuity of water substance in atmospheric circulations. Meteorological Monographs.* American Meteorological Society, Boston, MA, 1969, vol. 10, pp. 1–84.

[98] Y.-L. Lin, R. D. Farley, and H. D. Orville, "Bulk parameterization of the snow field in a cloud model," *Journal of Climate and Applied Meteorology*, vol. 22, no. 6, pp. 1065–1092, 1983.

[99] S. Rutledge and P. Hobbs, "The mesoscale and microscale structure and organization of clouds and precipitation in midlatitude cyclones. Part XII: A diagnostic modeling study of precipitation development in narrow cloud-frontal rainbands," *Journal of Atmospheric Sciences*, vol. 41, no. 20, pp. 2949–2972, 1984.

[100] S.-Y. Hong, J. Dudhia, and S.-H. Chen, "A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation," *Monthly Weather Review*, vol. 132, no. 1, pp. 103–120, 2004.

[101] S.-Y. Hong and J.-O. J. Lim, "The WRF single-moment 6-class microphysics scheme (WSM6)," *J. Korean Meteor. Soc*, vol. 42, no. 2, pp. 129–151, 2006.

[102] G. Thompson, R. M. Rasmussen, and K. Manning, "Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis," *Monthly Weather Review*, vol. 132, no. 2, pp. 519–542, 2004.

[103] J. S. Kain, "The Kain–Fritsch convective parameterization: An update," *Journal of Applied Meteorology*, vol. 43, no. 1, pp. 170–181, 2004.

[104] J. S. Kain and J. M. Fritsch, "A one-dimensional entraining/detraining plume model and its application in convective parameterization," *Journal of the Atmospheric Sciences*, vol. 47, no. 23, pp. 2784–2802, 1990.

[105] ——, “Convective parameterization for mesoscale models: The Kain-Fritsch scheme,” in *The representation of cumulus convection in numerical models*. Springer, 1993, pp. 165–170.

[106] Z. I. Janjić, “The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes,” *Monthly Weather Review*, vol. 122, no. 5, pp. 927–945, 1994.

[107] ——, “Comments on "Development and evaluation of a convection scheme for use in climate models",” *Journal of the Atmospheric Sciences*, vol. 57, no. 21, pp. 3686–3686, 2000.

[108] G. A. Grell and D. Dévényi, “A generalized approach to parameterizing convection combining ensemble and data assimilation techniques,” *Geophysical Research Letters*, vol. 29, no. 14, pp. 38–1, 2002.

[109] H.-l. Pan and W.-S. Wu, “Implementing a mass flux convection parameterization package for the NMC medium-range forecast model,” *NMC office note*, vol. 409, no. 40, pp. 20–233.

[110] A. Arakawa and W. H. Schubert, “Interaction of a cumulus cloud ensemble with the large-scale environment, Part I,” *Journal of the Atmospheric Sciences*, vol. 31, no. 3, pp. 674–701, 1974.

[111] G. A. Grell, “Prognostic evaluation of assumptions used by cumulus parameterizations,” *Monthly Weather Review*, vol. 121, no. 3, pp. 764–787, 1993.

[112] A. C. Lorenc, “Analysis methods for numerical weather prediction,” *Quarterly Journal of the Royal Meteorological Society*, vol. 112, no. 474, pp. 1177–1194, 1986.

[113] K. Ide, P. Courtier, M. Ghil, and A. C. Lorenc, “Unified notation for data assimilation: Operational, sequential and variational,” *Journal of the Meteorological Society of Japan. Ser. II*, vol. 75, no. 1B, pp. 181–189, 1997.

[114] P. Courtier and O. Talagrand, “Variational assimilation of meteorological observations with the direct and adjoint shallow-water equations,” *Tellus A: Dynamic Meteorology and Oceanography*, vol. 42, no. 5, pp. 531–549, 1990.

[115] P. Gauthier, M. Tanguay, S. Laroche, S. Pellerin, and J. Morneau, “Extension of 3DVar to 4DVar: Implementation of 4DVar at the meteorological service of Canada,” *Monthly weather review*, vol. 135, no. 6, pp. 2339–2354, 2007.

[116] F. Bouttier and P. Courtier, “Data assimilation concepts and methods March 1999,” *Meteorological training course lecture series. ECMWF*, p. 59, 2002.

[117] Q. Xiao, Y.-H. Kuo, Z. Ma, W. Huang, X.-Y. Huang, X. Zhang, D. M. Barker, J. Michalakes, and J. Dudhia, “Application of an adiabatic WRF adjoint to the investigation of the May 2004 McMurdo, Antarctica, severe wind event,” *Monthly Weather Review*, vol. 136, no. 10, pp. 3696–3713, 2008.

[118] L. Hascoët and V. Pascual, “Tapenade 2.1 user’s guide,” INRIA, Tech. Rep. RT-0300, 2004. [Online]. Available: https://hal.inria.fr/file/index/docid/69880/filename/RT-0300.pdf

[119] I. D. Rutherford, “Data assimilation by statistical interpolation of forecast error fields,” *Journal of the Atmospheric Sciences*, vol. 29, no. 5, pp. 809–815, 1972.

[120] M. Fisher, “Background error covariance modeling,” in *Proceedings of the seminar on recent developments in data assimilation for atmosphere and ocean*. ECMWF, Reading, UK, 2003, pp. 45–63.

[121] R. F. Adler, G. J. Huffman, D. T. Bolvin, S. Curtis, and E. J. Nelkin, “Tropical rainfall distributions determined using TRMM combined with other satellite and rain gauge information,” *Journal of Applied meteorology*, vol. 39, no. 12, pp. 2007–2023, 2000.

[122] Z. S. Haddad, E. A. Smith, C. D. Kummerow, T. Iguchi, M. R. Farrar, S. L. Durden, M. Alves, and W. S. Olson, “The TRMM ’day-1’radar/radiometer combined rain-profiling algorithm,” *Journal of the Meteorological Society of Japan. Ser. II*, vol. 75, no. 4, pp. 799–809, 1997.

[123] L. Leslie, J. LeMarshall, R. Morison, C. Spinoso, R. Purser, N. Pescod, and R. Seecamp, “Improved hurricane track forecasting from the continuous assimilation of high quality satellite wind data,” *Monthly Weather Review*, vol. 126, no. 5, pp. 1248–1258, 1998.

[124] S.-H. Chen, “The impact of assimilating SSM/I and QuikSCAT satellite winds on hurricane Isidore simulations,” *Monthly Weather Review*, vol. 135, no. 2, pp. 549–566, 2007.

[125] R. H. Langland, C. Velden, P. M. Pauley, and H. Berger, “Impact of satellite-derived rapid-scan wind observations on numerical model forecasts of hurricane Katrina,” *Monthly Weather Review*, vol. 137, no. 5, pp. 1615–1622, 2009.

[126] R. Singh, C. Kishtawal, P. Pal, and P. Joshi, “Improved tropical cyclone forecasts over north Indian Ocean with direct assimilation of AMSU-A radiances,” *Meteorology and Atmospheric Physics*, vol. 115, no. 1-2, pp. 15–34, 2012.

[127] S. M. Leidner, L. Isaksen, and R. N. Hoffman, “Impact of NSCAT winds on tropical cyclones in the ECMWF 4DVar assimilation system,” *Monthly Weather Review*, vol. 131, no. 1, pp. 3–26, 2003.

[128] C. Köpken, G. Kelly, and J.-N. Thépaut, “Assimilation of Meteosat radiance data within the 4D-Var system at ECMWF: Assimilation experiments and forecast impact,” *Quarterly Journal of the Royal Meteorological Society*, vol. 130, no. 601, pp. 2277–2292, 2004.

[129] S. Laroche, P. Gauthier, M. Tanguay, S. Pellerin, and J. Morneau, “Impact of the different components of 4DVar on the global forecast system of the meteorological service of Canada,” *Monthly Weather Review*, vol. 135, no. 6, pp. 2355–2364, 2007.

[130] G. J. Haltiner and R. T. Williams, *Numerical Prediction and Dynamic Meteorology*. John Wiley & Sons Inc, 1980.

[131] N. Žagar, N. Gustafsson, and E. Källén, “Variational data assimilation in the tropics: The impact of a background-error constraint,” *Quarterly Journal of the Royal Meteorological Society*, vol. 130, no. 596, pp. 103–125, 2004.

[132] K. A. Emanuel, “The theory of hurricanes,” *Annual Review of Fluid Mechanics*, vol. 23, no. 1, pp. 179–196, 1991.

[133] J. C. Chan, “Identification of the steering flow for tropical cyclone motion from objectively analyzed wind fields,” *Monthly weather review*, vol. 113, no. 1, pp. 106–116, 1985.

[134] B. Harris and G. Kelly, “A satellite radiance-bias correction scheme for data assimilation,” *Quarterly Journal of the Royal Meteorological Society*, vol. 127, no. 574, pp. 1453–1468, 2001.

[135] R. N. Bannister, “A review of forecast error covariance statistics in atmospheric variational data assimilation. II: Modelling the forecast error covariance statistics,” *Quarterly Journal of the Royal Meteorological Society*, vol. 134, no. 637, pp. 1971–1996, 2008.

[136] F. Rawlins, S. Ballard, K. Bovis, A. Clayton, D. Li, G. Inverarity, A. Lorenc, and T. Payne, “The Met Office global four-dimensional variational data assimilation scheme,” *Quarterly Journal of the Royal Meteorological Society*, vol. 133, no. 623, pp. 347–362, 2007.

[137] M. Zupanski, D. Zupanski, T. Vukicevic, K. Eis, and T. V. Haar, “CIRA/CSU four-dimensional variational data assimilation system,” *Monthly Weather Review*, vol. 133, no. 4, pp. 829–843, 2005.

[138] C. Fischer, T. Montmerle, L. Berre, L. Auger, and S. E. ŞTEFĂNESCU, “An overview of the variational assimilation in the ALADIN/France numerical weather-prediction system,” *Quarterly Journal of the Royal Meteorological Society*, vol. 131, no. 613, pp. 3477–3492, 2005.

[139] J. Sun, D. W. Flicker, and D. K. Lilly, “Recovery of three-dimensional wind and temperature fields from simulated single-Doppler radar data,” *Journal of the Atmospheric Sciences*, vol. 48, no. 6, pp. 876–890, 1991.

[140] J. Sun and N. A. Crook, “Dynamical and microphysical retrieval from Doppler radar observations using a cloud model and its adjoint. Part I: Model development and simulated data experiments,” *Journal of the Atmospheric Sciences*, vol. 54, no. 12, pp. 1642–1661, 1997.

[141] J. Gao, M. Xue, A. Shapiro, and K. K. Droegemeier, “A variational method for the analysis of three-dimensional wind fields from two Doppler radars,” *Monthly Weather Review*, vol. 127, no. 9, pp. 2128–2142, 1999.

[142] W.-S. Wu, R. J. Purser, and D. F. Parrish, “Three-dimensional variational analysis with spatially inhomogeneous covariances,” *Monthly Weather Review*, vol. 130, no. 12, pp. 2905–2916, 2002.

[143] Q.-S. Chen and Y.-H. Kuo, “A consistency condition for wind-field reconstruction in a limited area and a harmonic-cosine series expansion,” *Monthly Weather Review*, vol. 120, no. 11, pp. 2653–2670, 1992.

[144] M. Krysta, S. R. Rizvi, and X.-Y. Huang, “A new formulation of WRFDA analysis control variables,” in *10th annual WRF users’ workshop*. Citeseer, 2009, pp. 23–26.

[145] H. Wang, X.-Y. Huang, J. Sun, D. Xu, S. Fan, J. Zhong, and M. Zhang, “A comparison between the 3/4DVAR and hybrid ensemble-VAR techniques for radar data assimilation,” in *35th Conference on Radar Meteorology*, 2013.

[146] R. K. Vellore, M. L. Kaplan, R. Krishnan, J. M. Lewis, S. Sabade, N. Deshpande, B. B. Singh, R. Madhura, and M. R. Rao, “Monsoon-extratropical circulation interactions in Himalayan extreme rainfall,” *Climate Dynamics*, vol. 46, no. 11-12, pp. 3517–3546, 2016.

[147] D. S. Pai and S. C. Bhan, “Monsoon 2013: A report,” National Climate Centre, India Meteorological Department, Pune, Tech. Rep., 2013.

[148] A. Dimri and A. Chevuturi, “Model sensitivity analysis study for western disturbances over the Himalayas,” *Meteorology and Atmospheric Physics*, vol. 123, no. 3-4, pp. 155–180, 2014.

[149] J. Figa and A. Stoffelen, “On the assimilation of Ku-band scatterometer winds for weather analysis and forecasting,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 4, pp. 1893–1902, 2000.

[150] L. Isaksen and A. Stoffelen, “ERS scatterometer wind data impact on ECMWF’s tropical cyclone forecasts,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 4, pp. 1885–1892, 2000.

[151] R. Singh, P. Kumar, and P. K. Pal, “Assimilation of Oceansat-2-scatterometer-derived surface winds in the weather research and forecasting model,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 4, pp. 1015–1021, 2012.

[152] R. Atlas, R. Hoffman, S. Leidner, J. Sienkiewicz, T.-W. Yu, S. Bloom, E. Brin, J. Ardizzone, J. Terry, D. Bungato *et al.*, “The effects of marine winds from scatterometer data on weather analysis and forecasting,” *Bulletin of the American Meteorological Society*, vol. 82, no. 9, pp. 1965–1990, 2001.

[153] K. B. Katsaros, E. B. Forde, P. Chang, and W. T. Liu, “QuikSCAT’s sea winds facilitates early identification of tropical depressions in 1999 hurricane season,” *Geophysical Research Letters*, vol. 28, no. 6, pp. 1043–1046, 2001.

[154] M. Dhanya and A. Chandrasekar, “Improved rainfall simulation by assimilating Oceansat-2 surface winds using ensemble Kalman filter for a heavy rainfall event over south India,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 12, pp. 7721–7726, 2014.

[155] O. P. Sreejith, D. S. Pai, and M. Mohapatra, “Monsoon 2017: A report,” National Climate Centre, India Meteorological Department, Pune, Tech. Rep., 2018.

[156] K. Raghavan, “Break-monsoon over India,” *Monthly Weather Review*, vol. 101, pp. 33–43, 1973.

[157] R. Krishnan, C. Zhang, and M. Sugi, “Dynamics of breaks in the Indian summer monsoon,” *Journal of the Atmospheric Sciences*, vol. 57, no. 9, pp. 1354–1372, 2000.

[158] S. Gadgil and P. Joseph, “On breaks of the Indian monsoon,” *Journal of Earth System Science*, vol. 112, no. 4, pp. 529–558, 2003.

[159] L. Yu and X. Jin, “Insights on the OAFlux ocean surface vector wind analysis merged from scatterometers and passive microwave radiometers (1987 onward),” *Journal of Geophysical Research: Oceans*, vol. 119, no. 8, pp. 5244–5269, 2014.

[160] J. C. L. Chan and J. D. Kepert, *Global Perspectives on Tropical Cyclones: From Science to Mitigation*. World Scientific, 2010, vol. 4.

[161] C. Srinivas, D. Bhaskar Rao, V. Yesubabu, R. Baskaran, and B. Venkatraman, “Tropical cyclone predictions over the Bay of Bengal using the high-resolution Advanced Research Weather Research and Forecasting (ARW) model,” *Quarterly Journal of the Royal Meteorological Society*, vol. 139, no. 676, pp. 1810–1825, 2013.

[162] V. Yesubabu, C. Srinivas, K. Hariprasad, and R. Baskaran, “A study on the impact of observation assimilation on the numerical simulation of tropical cyclones JAL and THANE using 3DVAR,” *Pure and Applied Geophysics*, vol. 171, no. 8, pp. 2023–2042, 2014.

[163] R. Singh, C. Kishtawal, and P. Pal, “Impact of ATOVS radiance on the analysis and forecasts of a mesoscale model over the Indian region during the 2008 summer monsoon,” *Pure and Applied Geophysics*, vol. 169, no. 3, pp. 425–445, 2012.

[164] R. Singh, S. P. Ojha, C. Kishtawal, and P. Pal, “Quality assessment and assimilation of Megha-Tropiques SAPHIR radiances into WRF assimilation system,” *Journal of Geophysical Research: Atmospheres*, vol. 118, no. 13, pp. 6957–6969, 2013.

[165] R. Singh, S. P. Ojha, C. Kishtawal, P. Pal, and A. K. Kumar, “Impact of the assimilation of insat-3d radiances on short-range weather forecasts,” *Quarterly Journal of the Royal Meteorological Society*, vol. 142, no. 694, pp. 120–131, 2016.

List of Publications

Refereed Journals

1. Deepak Gopalakrishnan and A. Chandrasekar (2018). "On the Improved Predictive Skill of WRF Model with Regional 4DVar Initialization: A Study with North Indian Ocean Tropical Cyclones" **IEEE Transactions on Geoscience and Remote Sensing**, 56,6, 3350-3357.
2. Deepak Gopalakrishnan and A. Chandrasekar (2018). "Improved 4DVar Simulation of Indian Ocean Tropical Cyclones using a Regional Model" **IEEE Transactions on Geoscience and Remote Sensing**, 56,9, 5107-5114.

Manuscript under preparation

1. Deepak Gopalakrishnan and A. Chandrasekar. "How Sensitive is the WRF-4DVar Assimilation System to the Control Variables?: Study on Indian Heavy Rainfall Events Using Limited Area Model" (under review in Journal of Geophysical Research: Atmospheres).
2. Deepak Gopalakrishnan and A. Chandrasekar. "Impact of Assimilating Scatsat-1 Ocean Surface Winds on the Simulation of Indian Summer Monsoon Rainfall Using WRF-4DVar" (under preparation).

Refereed Conferences

1. Deepak Gopalakrishnan; Chandrasekar, A; Bhate, Jyoti; Kesarkar, Amit P., "Impact of assimilation of satellite radiance in the simulation of tropical cyclone Phailin using

WRF model", paper presented at 19th National Space Science Symposium, Vikram Sarabhai Space Centre, Kerala, India (09-12 February 2016).

2. Deepak Gopalakrishnan; Chandrasekar, A, "Improved Simulation of Tropical Cyclone Hudhud with Regional 4DVar Initialization: A Case Study Using WRF Model", paper presented at International Tropical Meteorology Symposium, SAC, India (07-10 November 2017).

3. Deepak Gopalakrishnan; A. Chandrasekar, "Sensitivity of the 4DVar assimilation system to the control variables: A case study on Uttarakhand cloudburst event", paper presented at TROPMET-2018, Banaras Hindu University, Varanasi (24-27 October, 2018).

Spin-off publications

1. Dhanya, M, Gopalakrishnan, D, A. Chandrasekar, S.K. Singh and V.S. Prasad (2016), "The impact of assimilating MeghaTropiques SAPHIR radiances in the simulation of tropical cyclones over the Bay of Bengal using the WRF model", International Journal of Remote Sensing, 37:13, 3086-3103.

2. Dhanya, M., Deepak Gopalakrishnan; Chandrasekar, A; Singh, S K; Prasad, V.S; "Impact of Megha-Tropiques SAPHIR radiance assimilation on the simulation of tropical cyclones over Bay of Bengal", (2016, Conference.Proceedings, SPIE).

Appendix A

WRF Model Configurations

1. WRF and WRFDA namelist options used in Chapter 3

Dynamics	Non-hydrostatic
Center of the domain	13°N and 80°E
Number of domains	2
Number of east-west grids	350 and 520
Number of south-north grids	350 and 451
Horizontal grid spacing	27 km and 9 km
Mode of nesting	Two-way nesting
Map projection	Mercator
Number of vertical coordinates	36
Pressure at the top	50 hPa
Horizontal grid system	Arakawa-C grid
Integration time step	120 seconds (for 27 km domain)

Model physics

Cloud microphysics	WSM 5-class scheme (option 4)
Cumulus convection	Kain-Fritsch scheme (option 1)
Planetary boundary layer	Yonsei University scheme (option 1)
Land surface physics	Noah LSM (option 2)
Shortwave radiation	Dudhia scheme (option 1)
Longwave radiation	RRTM scheme (option 1)
Surface physics	Eta similarity scheme (option 2)

WRFDA-3DVar

Var-4D	false
Observation format	1 (BUFR)
Number of FGAT times	1
Thinning	True
Maximum number of iterations	100
Convergence criterion	0.01
CV option	5
Tuning factors of BEC variables	1.0
Humidity CV option	1
Balance type	3 (geostrophic + cyclostrophic)

WRFDA-4DVar

Var-4D	true
Observation format	1 (BUFR)
Number of FGAT times	1
Thinning	True
Maximum number of iterations	100
Convergence criterion	0.01
CV option	5
Time step for TL integration	120
Tuning factors of BEC variables	1.0
Humidity CV option	1
Balance type	3 (geostrophic + cyclostrophic)

All the WRF physics schemes employed in the numerical experiments performed in Chapter 3 are chosen following the similar studies conducted for the simulation of tropical cyclones over Bay of Bengal and Arabian Sea [161, 162, 68, 69, 70]. In the WRFDA, all the observations are used in the PREPBUFR format. Since background field at only one time (at the analysis time) is employed for the 3DVar assimilation, FGAT option was set to 1. The BEC employed in the study has been generated by considering 60 sample forecasts of 12h and 24h forecast lengths that are valid at same time employing the CV5 option.

2. WRF and WRFDA namelist options used in Chapter 4

Dynamics	Non-hydrostatic
Center of the domain	13°N and 80°E
Number of domains	2
Number of east-west grids	350 and 520
Number of south-north grids	350 and 451
Horizontal grid spacing	27 km and 9 km
Mode of nesting	Two-way nesting
Map projection	Mercator
Number of vertical coordinates	36
Pressure at the top	50 hPa
Horizontal grid system	Arakawa-C grid
Integration time step	120 seconds (for 27 km domain)

Model physics

Cloud microphysics	WSM 5-class scheme (option 4)
Cumulus convection	Kain-Fritsch scheme (option 1)
Planetary boundary layer	Yonsei University scheme (option 1)
Land surface physics	Noah LSM (option 2)
Shortwave radiation	Dudhia scheme (option 1)
Longwave radiation	RRTM scheme (option 1)
Surface physics	Eta similarity scheme (option 2)

WRFDA-3DVar

Var-4D	false
Observation format	1 (BUFR)
Number of FGAT times	1
Thinning	True
Maximum number of iterations	100
Convergence criterion	0.01
CV option	5
Qc-rad	true
RTM-option	2

Only-sea-rad	false
Use-varbc	true
Tuning factors of BEC variables	1.0
Humidity CV option	1
Balance type	3 (geostrophic + cyclostrophic)

WRFDA-4DVar

Var-4D	true
Observation format	1 (BUFR)
Number of FGAT times	1
Thinning	True
Maximum number of iterations	100
Convergence criterion	0.01
CV option	5
Qc-rad	true
Thinning-mesh	120.0 km
RTM-option	2
Only-sea-rad	false
Use-varbc	true
Time step for TL integration	120
Tuning factors of BEC variables	1.0
Humidity CV option	1
Balance type	3 (geostrophic + cyclostrophic)

All the namelist options for the model configuration in Chapter 4 are similar to the options in Chapter 3, except for the changes in WRFDA namelist corresponding to the radiance observations. All the radiance observations are used in BUFR format. Thorough quality check for the radiance observations has been performed before the assimilation, following the procedure discussed in [163]. Thinning for the radiance observations is performed at 120.0 km mesh. The Community Radiative Transfer Model (RTM=2) is employed for simulating the model equivalent brightness temperature. The radiance observations are affected by bias and the variational bias correction method (varbc) within the WRFDA module is applied for removing the bias.

3. WRF and WRFDA namelist options used in Chapter 5

Dynamics	Non-hydrostatic
Center of the domain	30°N and 75°E
Number of domains	3
Number of east-west grids	375, 412, and 460
Number of south-north grids	250, 400, and 409
Horizontal grid spacing	27 km, 9 km and 3 km
Mode of nesting	Two-way nesting
Map projection	Lambert
Number of vertical coordinates	36
Pressure at the top	50 hPa
Horizontal grid system	Arakawa-C grid
Integration time step	120 seconds (for 27 km domain)

Model physics

Cloud microphysics	Eta-Ferrier scheme (option 5)
Cumulus convection	Kain-Fritsch scheme (option 1)
Planetary boundary layer	Yonsei University scheme (option 1)
Land surface physics	Noah LSM (option 2)
Shortwave radiation	Dudhia scheme (option 1)
Longwave radiation	RRTM scheme (option 1)
Surface physics	Eta similarity scheme (option 2)

WRFDA-4DVar

Var-4D	true
Observation format	1 (BUFR)
Number of FGAT times	1
Thinning	True
Maximum number of iterations	100
Convergence criterion	0.01
CV options	5, 6, and 7
Time step for TL integration	120
Tuning factors of BEC variables	1.0

Humidity CV option	1
Balance type	3 (geostrophic + cyclostrophic)

The model physics schemes employed in the numerical simulations performed in Chapter 5 to study the mountainous rainfall events are similar to the schemes in the previous chapters, except the cloud microphysics scheme. Since the detailed study by Dimri and Chevuturi in 2014 [148] on the simulation of western disturbances over Himalayan region has shown that the Eta-Ferrier scheme of microphysics parameterization provides improved rainfall forecasts, the present study also utilized the same parameterization scheme for representing the cloud microphysical processes. The finer domain in the study has a horizontal resolution of 3 km and the cumulus parameterization scheme is not employed in the same to allow for explicit convection. In WRFDA, three CV options (5, 6, and 7) are utilized in this chapter. All other WRFDA namelist options are similar to options employed in Chapter 3 and 4.

4. WRF and WRFDA namelist options used in Chapter 6

Dynamics	Non-hydrostatic
Center of the domain	13°N and 80°E
Number of domains	2
Number of east-west grids	185 and 277
Number of south-north grids	185 and 277
Horizontal grid spacing	45 km and 15 km
Mode of nesting	Two-way nesting
Map projection	Mercator
Number of vertical coordinates	36
Pressure at the top	50 hPa
Horizontal grid system	Arakawa-C grid
Integration time step	180 seconds (for 45 km domain)

Model physics	
Cloud microphysics	WSM 6-class scheme (option 5)
Cumulus convection	Kain-Fritsch scheme (option 1)
Planetary boundary layer	Yonsei University scheme (option 1)

Land surface physics	Noah LSM (option 2)
Shortwave radiation	Dudhia scheme (option 1)
Longwave radiation	RRTM scheme (option 1)
Surface physics	Eta similarity scheme (option 2)

WRFDA-4DVar

Var-4D	true
Observation format	2 (ASCII)
Number of FGAT times	1
Thinning	True
Maximum number of iterations	100
Convergence criterion	0.01
CV options	5
Time step for TL integration	180
Tuning factors of BEC variables	1.0
Humidity CV option	1
Balance type	3 (geostrophic + cyclostrophic)

The model physics option chosen in Chapter 6 are following the month-long studies for Indian summer monsoon in the existing literature [164, 165]. Since the present chapter deals with Indian summer monsoon system which is having a larger horizontal extent, horizontal resolutions of 45 km and 15 km are chosen instead of 27 km and 9 km. Also, as the number of assimilations for a month-long study will be considerably higher, reducing the horizontal resolutions will help to reduce the computational cost significantly. CV5 option is used for modeling the BEC matrix.