Predictability and Dynamics of Extreme Weather Events Over the Indian Subcontinent using Ensemble Sensitivity Analysis in EnKF Data Assimilation System

A thesis submitted
in partial fulfillment for the Degree of

Doctor of Philosophy

by

Babitha George

Department of Earth and Space Sciences
Indian Institute of Space Science and Technology
Thiruvananthapuram – 695547, India
April 2023

सत्यमेव जयते
भारतीय अंतरिक्ष विज्ञान एवं प्रौद्योगिकी संस्थान
(वि.अ.आयोग अधिनियम 1956 की धारा-3 के अधीन मानित विश्वविद्यालय घोषित)
भारत सरकार, अंतरिक्ष विभाग, वलियमला पोस्ट, तिरुवनंतपुरम 695 547 भारत
INDIAN INSTITUTE OF SPACE SCIENCE AND TECHNOLOGY
(A Deemed to be University u/s 3 of the UGC Act, 1956)
Government of India, Department of Space
Valiamala P. O, Thiruvananthapuram 695 547 India
www.iist.ac.in

CERTIFICATE

This is to certify that the thesis entitled **Predictability and Dynamics of Extreme Weather Events Over the Indian Subcontinent using Ensemble Sensitivity Analysis in EnKF Data Assimilation System submitted** by **Babitha George**, to the Indian Institute of Space Science and Technology, Thiruvananthapuram, in partial fulfilment for the award of the degree of **Doctor of Philosophy**, is a *bona fide* record of research work carried out by her under my supervision. The contents of this thesis, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

Dr. Govindan Kutty
Supervisor & Associate Professor
Department of Earth & Space Sciences

डॉ. गोविंदन कुट्टी एम / Dr. Govindan Kutty M
सह आचार्य / Associate Professor
पृथ्वी एवं अंतरिक्ष विज्ञान विभाग
Department of Earth and Space Sciences
भारतीय अंतरिक्ष विज्ञान एवं प्रौद्योगिकी संस्थान
Indian Institute of Space Science and Technology
अंतरिक्ष विभाग,भारत सरकार/Dept.of Space, Govt.of India
तिरुवनंतपुरम/Thiruvananthapuram - 695 547

Dr. Rama Rao Nidamanuri
Professor & Head
Department of Earth & Space Sciences

विभागाध्यक्ष / Head of the Department
पृथ्वी एवं अंतरिक्ष विज्ञान विभाग
Department of Earth and Space Sciences
भारतीय अंतरिक्ष विज्ञान एवं प्रौद्योगिकी संस्थान
Indian Institute of Space Science and Technology
अंतरिक्ष विभाग, भारत सरकार
Dept. of Space, Govt. of India
तिरुवनंतपुरम / Thiruvananthapuram - 695 547

Thiruvananthapuram
April 2023

DECLARATION

I declare that this thesis entitled **Predictability and Dynamics of Extreme Weather Events Over the Indian Subcontinent using Ensemble Sensitivity Analysis in EnKF Data Assimilation System** submitted in partial fulfilment of the degree of **Doctor of Philosophy** is a record of original work carried out by me under the supervision of **Dr. Govindan Kutty**, 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. 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 2023

Babitha George
(SC17D014)

ACKNOWLEDGEMENTS

"Let us be grateful to people who make us happy; they are the charming gardeners who make our souls blossom."

-MARCEL PROUST

Sometimes people come into our lives for a short duration of time. No matter long or short they stay, they leave imprints. I am fortunate enough to have such people who taught me many good things and supported me in different ways to achieve my dreams. I am glad to express my gratitude to all those who have encouraged me in the successful completion of this research work.

First and foremost, I would like to express my profound gratitude to my supervisor Dr. Govindan Kutty, for his consistent support and guidance throughout the research. From the very beginning he has motivated and encouraged me to complete the research on time. I am fortunate to have a great teacher like him with in-depth knowledge and experience in the subject. I was able to approach him anytime and he would clear all my silly doubts with patience. Especially during the time of the COVID pandemic, though I could not physically meet him, he would always respond to my doubts over a call or mail. Working and studying under his guidance was a great honour and privilege.

I am deeply indebted to my doctoral committee member and teacher, Prof. A. Chandrasekar, for enlightening me on the subject, and his constructive suggestions have helped me greatly in shaping my thesis. I am incredibly grateful to the remaining members of the doctoral committee Dr. K. Rajeev, Dr. Arindam Chakraborty and Dr. Rama Rao Nidamanuri, for their valuable comments and constant support during the annual review meetings. I would also like to acknowledge Dr. Anandmayee Tej and Dr. Samir Mandal, former HODs and doctoral committee members for their support and motivation. I also take this opportunity to thank Dr. V.K. Dadhwal, the former Director and Dr. S. Unnikrishnan Nair, the present Director of IIST for providing the necessary resources and lab facilities for carrying out research with ease. I also thank all the faculty members of the ESS department for their support and love over the past five years. Moreover, this endeavor would have been impossible without the financial support from IIST.

I must acknowledge the Director, Indian Institute of Tropical Meteorology, Pune and the Director, National Atmospheric Research Laboratory, Tirupati for giving access to the high-performance computing facilities and making this thesis happen.

I would also like to thank my seniors Dr. Deepak Gopalakrishnan and Dr. Rekha Bharali Gogoi for teaching me the fundamentals of modelling and data assimilation. I thank all my fellow labmates Mr. Mukul, Mr. Vibin, Mr, Rakesh, Mr. Jayakrishnan, Mr. Sam, Mr. Sanjay and Mr. Francis, for the stimulating discussions and for the all the fun we had in all those years. I am also deeply indebted to Mrs. Shalini, Mr. Aswin and Mrs. Saranya for lending all the technical support.

I also thank all my friends at IIST for letting me experience a memorable non-academic life in IIST. To name a few, I thank Dr. Anu, Dr, Aneesha, Dr. Anna, Ms. Geethu and Ms. Sreelekshmi for a wonderful hostel life with outings, movies and cooking experiments. They lend me moral support in stressful situations and made my life in IIST lively.

The completion of this research work would not have been accomplished without the support of my family. I thank my parents for their love and sacrifices for educating and preparing me for my future life. I must also thank my in-laws for their love, care and encouragement. Also, I express my thanks to my sister, brother, brother in-law and sister in-law for boosting me up whenever I am down with stress. Words are not enough to express my gratitude to my husband, Mr. Sebin Pious whose unconditional love and unwavering support gave me strength to complete the thesis without much difficulty.

Babitha George

Abstract

Forecasting extreme weather events using Numerical Weather Prediction (NWP) model is challenging due to the uncertainties associated with the growth of initial errors under chaotic dynamics. Therefore, it is essential to initialize a model with the best estimate of the atmosphere and understand how errors in this initial condition will affect the subsequent forecast. Statistically reliable ensemble predictions from the different realization of initial conditions of the atmosphere are found to be robust in forecasting extreme weather events. The predictability of weather events can be quantified through sensitivity analysis, which essentially indicates how forecast from an NWP system responds to changes in initial conditions. Ensemble sensitivity analysis (ESA) is a linear approach to sensitivity analysis that uses sample statistics to estimate how a scalar forecast function changes with respect to initial conditions. In this research work, ESA is applied to understand the nature and predictability of extreme weather events over the Indian subcontinent by using ensemble analyses and forecasts from an Ensemble Kalman filter (EnKF) data assimilation system. The Weather Research and Forecasting (WRF) is used as the NWP model in this study. In addition, the ensembles from TIGGE European Centre for Medium-Range Weather Forecasts and National Centers for Environmental Prediction Ensemble Prediction Systems have also been employed.

ESA often uses a diagonal approximation to the multivariate regression, leading it to a simple univariate regression, and it is often referred to as univariate ensemble sensitivity analysis. Univariate ESA applied to extreme rainfall event over the Uttarakhand state located in the Western Himalayas indicates that the heavy precipitation is sensitive to the mid-tropospheric trough and moisture fields from the Arabian Sea and the Bay of Bengal. Perturbed initial condition experiments reveal that the initial condition perturbations in the maximum sensitive region can have a large impact on the rainfall. Further, two extreme rainfall events over Kerala in August 2018 (KF18) and 2019 (KF19) are analysed using univariate ESA. In the case of Kerala rainfall in 2018, the results show that the circulations positioned farther east of its mean position over the Western

North Pacific (WNP) are related to stronger precipitation over the response function region. However, in the case of Kerala rainfall in 2019, the moisture-laden low-level flow was more substantial, which favored the development of deep convective clouds and caused extreme rain.

The presence of sampling error can cause the univariate ESA to overestimate the response of a forecast metric to initial conditions. Therefore, univariate ESA is extended to multivariate ESA that utilizes the full covariance matrix. The performance of multivariate ESA over univariate is examined by applying the method to a heavy rainfall event that happened over Chennai in December 2015. The multivariate ESA shows more organized sensitive patterns, unlike univariate sensitivity in which the sensitivity patterns are broadly distributed. Both methods are validated using the perturbed initial condition approach, and it is found that multivariate is more effective in predicting the forecast response closest to the actual model response compared to the univariate ESA.

Further experiments were performed using the multivariate ESA to investigate the general predictability characteristics of tropical cyclones over the Bay of Bengal. Results show that intense, fast-moving, and north-landfalling tropical cyclones exhibit low predictability in its intensity forecast. Intense storms exhibit large initial condition sensitivity than the analysis spread indicating that the low predictability of intense cyclones is likely due to large dynamical perturbation growth. The results of the perturbed initial condition experiment show that the dynamical error growth is faster if the perturbations are smaller in magnitude. It is also found that the error growth associated with moist perturbations is higher for the less predictable tropical cyclones.

Towards the end, we have determined the climatological ensemble sensitivity to identify the optimal locations for deploying the observation network during the Indian summer monsoon. Results show that the precipitation forecasts during the Indian summer monsoon season benefit from the assimilation of observations located over the upstream regions of the forecast metric box.

Table of Contents

4 Multivariate Ensemble Sensitivity Analysis Applied for an Extreme Rainfall Event over the Indian Subcontinent 69

5 Predictability Characteristics of Tropical Cyclones over the Bay of Bengal using Ensemble Sensitivity Analysis 93

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

3DVAR	Three Dimensional Variational
4DVAR	Four Dimensional Variational
DA	Data Assimilation
DART	Data Assimilation Research Testbed
DTE	Difference Total Energy
DVWS	Deep Layer Vertical Wind Shear
EAKF	Ensemble Adjustment Kalman Filter
ECMWF	European Centre for Medium Range Weather Forecasts
EKF	Extended Kalman Filter
EnKF	Ensemble Kalman Filter
EPS	Ensemble Prediction System
ESA	Ensemble Sensitivity Analysis
GDAS	Global Data Assimilation System
GFS	Global Forecast System
IMD	Indian Meteorological Department
JTWC	Joint Typhoon Warning Center
KF18	Kerala Rainfall 2018
KF19	Kerala Rainfall 2019
LAF	Lagged Average Forecasting
LPS	Low-Pressure System
MDBZ	Maximum Composite Reflectivity
MSLP	Minimum Sea Level Pressure
NCAR	National Centre for Atmospheric Research
NCEP	National Centers for Environmental Prediction
NCL	NCAR Graphics Command Language
NCMRWF	National Centre for Medium Range Weather Forecasting
NMM	Non-hydrostatic Mesoscale Model
NOAA	National Ocean and Atmospheric Administration

NWP	Numerical Weather Prediction
PDF	Probability Density Function
RMSE	Root Mean Square Error
SA	Sensitivity Analysis
SKEBS	Stochastic Kinetic-Energy Backscatter Scheme
SLAF	Scaled Lagged Average Forecasting
SVWS	Shallow Layer Vertical Wind Shear
TC	Tropical Cyclone
THORPEX	The Observing system Research and Predictability Experiment
TIGGE	THORPEX Interactive Grand Global Ensemble
TRMM	Tropical Rainfall Measuring Mission
WNP	Western North Pacific
WRF	Weather Research and Forecasting

CHAPTER 1

INTRODUCTION

The evolution of the atmosphere, being a nonlinear dynamical process, is inherently unpredictable as the non-linear equations have no analytical solution. Numerical Weather Prediction (NWP) is an initial-value problem, and its ability to determine the future state of the atmosphere depends on the accuracy of the initial state. The uncertainties in the initial condition and model formulations limit the predictability of weather using NWP models. Therefore, quantifying these uncertainties in the forecasts is crucial in understanding the relevant dynamics and the predictability of the weather event. The uncertainties associated with the initial conditions for the forecast of weather events can be quantified using deterministic and probabilistic approaches. This thesis employs a probabilistic approach (ensemble-based techniques) to understand the predictability and dynamics of extreme weather events over the Indian subcontinent.

1.1 Forecast uncertainty and ensemble forecasting

Though the nonlinear differential equations that describe the atmosphere have no analytical solution, they can be solved using numerical methods. During the 1950s, Charney, FjÖrtoft and Neumann used a two dimensional barotropic model to obtain a numerical solution to the barotropic vorticity equation. In the early 1970s, the global circulation models based on a set of nonlinear differential equations emerged (Lynch, 2008). Later, during the 1980s, increase in computer resources and observational networks led to the development of the regional and mesoscale numerical weather prediction models. This further widened the knowledge on the microphysical processes and dynamics of the atmosphere (Anthes and Warner, 1978). In the 1990s, advanced diagnostic techniques emerged for weather forecasting with the development of high-resolution models and coupled models. However, inaccurate representation of the initial state of the atmosphere, imperfect data assimilation systems, and the parameterization of microphysical processes and non-microphysical processes such as turbulence, surface orography and radiation in the models reduces the accuracy of forecast. Therefore, determining a realistic initial state of the atmosphere becomes impossible (Daley, 1991). Lorenz (1963) showed that the predictability of the atmosphere is limited in time and the error in the initial condition grows with forecast period. Accordingly, the nonlinear behaviour of the atmosphere may be treated in a probabilistic way to quantify the forecast uncertainty. Historically, it was Epstein (1969), who introduced the idea of stochastic-dynamic prediction to acknowledge the uncertainty in the meteorological model forecasts. He used a probabilistic continuity equation to integrate the analysis probability density function (PDF) forward in time, however, this method was computationally expensive even for low-order models. Later, Leith (1974) suggested using multiple forecasts generated from slightly different but equally probable initial atmospheric states to characterize the uncertainty of the prediction. He used an analytical turbulence model to show that Monte Carlo forecasting procedure provides better approximation to stochastic dynamic prediction method in the limit of a smaller number of ensemble members. Hoffmann and Kalnay (1983) combined the forecasts initialized at the initial time and previous times to produce initial ensembles and proposed the idea of lagged average forecasting (LAF). The advantage of LAF over Monte Carlo forecasting is that the initial condition perturbation generated using LAF included the information of dynamics and therefore contained “errors of the day”. Furthermore, the method is straightforward and does not require us to compute perturbations separately. However, LAF requires to estimate the weights of the ensemble members based on their “age”. Ebisuzaki and Kalnay (1991) introduced the scaled lagged average forecasting (SLAF) as a variant of LAF to overcome the disadvantages of LAF. In this approach the perturbations are obtained from the forecast error of forecasts initialized at different initial time. The advantage is that the initial perturbations of ensemble members are of similar size. Pairs of initial perturbations with opposite sign are used in SLAF for better ensemble forecasts.

The advantages of ensemble forecasting include providing ensemble mean forecasts which are found to be more accurate than the deterministic forecasts even beyond first few days, providing guidance on the reliability of the forecasts and providing quantitative basis for probabilistic forecasting (Kalnay, 2002). Ensemble Prediction Systems (EPS) varies mostly in the way the initial perturbations are generated. Broadly, EPS can be classified into two categories- the one with random initial perturbations and the one with initial perturbations that contains information of the underlying dynamic flow. Operational weather forecast centres such as European Centre for Medium Range Weather Forecasting (ECMWF) and National Centres for Environmental Prediction (NCEP) uses ensemble forecasts (Toth and Kalnay, 1993, 1997; Tracton and Kalnay, 1993), which are found to be extremely useful for decision making purposes. Different EPS uses different methods for generating ensembles. For example, at ECMWF, the initial perturbations are obtained using the singular vector approach (Molteni and Palmer, 1993) while at NCEP breeding vectors (Toth and Kalnay, 1997) are used for generating initial perturbations for ensembles. Since 2010, the 50 ensemble members at ECMWF are generated by a combination of singular vectors and ensemble data assimilation (EDA) methods (Isaksen *et al.*, 2010). Forecast ensembles can also be generated by varying model parameterization schemes (Andersson, 1998; Houtekamer and Mitchell, 1998; Stensrud *et al.*, 1999) or stochastic physics schemes in each model forecast (Buizza, Miller and Palmer, 1999). Another strategy is to combine the model integrations from multiple NWP model initialized from perturbed sets of initial conditions (Hou, Kalnay and Droegemeier, 2001; Palmer *et al.*, 2012). Finally, operational forecasts from various centres are considered as initial conditions to produce ensemble forecasts (Wobus and Kalnay, 1995; Krishnamurti *et al.*, 1999; Evans *et al.*, 2000). Ensemble forecasts also paved way for developments in targeted observing systems which identifies regions where additional observations would be useful to reduce forecast errors. An ideal ensemble system should enable identification of biases in the system, account for sources of errors, and carry initial condition perturbations that are consistent with the "errors of the day". Several of such properties can be achieved by designing an ensemble system that employs ensemble-based data assimilation. The ensembles from such a system may follow equal likelihood, where the forecast from all the ensemble members have equal probability of representing the true state of the atmosphere.

1.2 Ensemble data assimilation

Data assimilation (DA) is a powerful method with wide range of applications in the field of atmosphere, ocean and land surface. It combines observational information with the dynamical equations governing the atmosphere to obtain the state of the atmosphere that would better describe the system than a model or observation alone would do (Daley, 1991).

Evensen (1994) proposed the ensemble-based data assimilation based on a Monte Carlo approximation of the traditional Kalman Filter. In ensemble-based data assimilation method, the flow-dependent background error covariances are estimated from an ensemble of forecasts to adjust the background forecast optimally to the newly available observations. As a result, unlike data assimilation techniques that assumes a time invariant background error, the ensemble-based data assimilation methods produce much more accurate analyses and forecasts.

Generally, the ensemble filters are classified into two: stochastic and deterministic. In both classes the ensemble analyses are integrated forward using non-linear forecast models and the key difference between the two classes depends on whether or not random perturbations are added during the update step. Ensemble Kalman Filter (EnKF) is a stochastic ensemble data assimilation algorithm that updates each background ensemble member by introducing random noise to observations at every assimilation cycle. In deterministic filters the observations are not perturbed randomly and it is assumed that an optimal gain exists so that the analysis error covariance can be determined from a modified Kalman gain matrix. The commonly known deterministic filters are the ensemble square root filter (EnSRF, Whitaker and Hamill, (2002)), the ensemble adjustment Kalman Filter (EAKF, Anderson, (2001)) and the ensemble transform Kalman Filter (ETKF, Bishop, Etherton and Majumdar, (2001)). The EAKF data assimilation method is easily available from the Data Assimilation Research Testbed (DART, Anderson *et al.*, (2009)) and is one of the most widely used EnKF algorithm (Singh, Mitta and Upadhyaya, 2015; Hill, Weiss and Ancell, 2016; Kerr, Stensrud and Wang, 2017).

Several studies have applied the ensemble data assimilation systems to various weather systems and found them to be more efficient than other DA systems. Hamill and Snyder (2000) used a hybrid EnKF-three-dimensional variational (3DVAR) analysis scheme to study the contributions of the static and ensemble-based background covariances. They found that giving more weight to ensemble-based covariances improved the initial conditions, particularly in data void regions. Whitaker *et al.* (2004) used surface pressure observations and showed that the ensemble-based data assimilation systems are computationally efficient and more accurate than other assimilation schemes when the observation network is sparse. Studies by Whitaker *et al.*, (2008), Tong and Xue (2008), and Buehner *et al.*, (2010) showed that the ensemble-based data assimilation systems are effective in operational weather prediction models. Whitaker *et al.*, (2008) compared the ensemble data assimilation system implemented in the NCEP Global Forecast System (GFS) with the NCEP Global Data Assimilation System (GDAS). The ensemble data assimilation analyses were found to improve the skill of 24-h forecasts over the Southern Hemisphere extratropical regions. Dowell *et al.*, (2011) used EnKF to produce realistic convective-scale analyses by assimilating ground-based radar reflectivity observations. Thus, EnKF has been applied for various weather system at the global (Houtekamer *et al.*, 2005; Szunyogh *et al.*, 2008; Miyoshi, Sato and Kadowaki, 2010) and the convective scales (Zhang, Snyder and Sun, 2004; Aksoy, Dowell and Snyder, 2010; Houtekamer and Zhang, 2016).

Dirren, Torn and Hakim (2007) used the Weather Research and Forecasting (WRF) model to conduct experiments with the EnKF data assimilation system on a limited area domain over the Pacific Northwest region of the United States. They found reduction in the ensemble-mean analysis error and ensemble spread in both observed and unobserved variables with the assimilation of single type of observation using EnKF data assimilation system. Torn and Hakim (2008b) studied the performance of a pseudo-operational limited-area EnKF system over a 2-year period based on the WRF model. They used the University of Washington EnKF system and mentioned that the gridded datasets from the ensemble system may be useful for understanding the predictability and dynamics of weather systems where flow-dependent analysis errors are important. Torn and Hakim (2008a) used the analyses ensembles from the EnKF system to demonstrate the application of ensemble sensitivity analysis in determining the sensitivity of ensemble forecasts to observations.

Thus, EnKF uses an ensemble of short-term forecasts to approximate the forecast-error statistics and performs parallel data assimilation cycles on these short-term ensemble forecasts to produce an ensemble of analyses (Hamill, 2006). The analyses thus produced are consistent with the analysis-error statistics and can be used as optimal initial conditions for ensemble forecasting. Because of these advantages, EnKF have achieved considerable attention for generating the best estimate of the atmospheric state given the background and the observations. In this thesis, the DART-EAKF data assimilation system is used to produce optimal initial conditions to obtain the relationship between the ensemble forecasts and the initial conditions.

1.3 Ensemble Sensitivity Analysis

Sensitivity analysis (SA) is a reliable method that is employed to quantify the uncertainties imposed by chaos in the forecast of weather events. It identifies the regions where small perturbations in the initial conditions can produce significant impact on the forecast. Therefore, sensitivity regions are geographical locations where initial perturbations will grow rapidly producing uncertainties in the forecasts. Sensitivity analysis is applied to identify potential locations for targeting observations (Snyder, 1996) and to understand the dynamics of forecast errors (Fujii *et al.*, 2008). Earlier studies used adjoint (Rabier *et al.*, 1996; Langland, Shapiro and Gelaro, 2002) or singular vector methods (Gelaro *et al.*, 1998; Palmer *et al.*, 1998; Buizza and Montani, 1999; Palmer and Zanna, 2013) to compute the sensitivity analysis, however these methods require tangent-linear or adjoint models which requires large volume of computational resources.

Alternatively, Torn and Hakim, (2008a) proposed that sensitivity analysis can also be obtained from ensemble-based methods, referred as ensemble sensitivity analysis (ESA). Ensemble sensitivity analysis uses sample statistics to obtain a linear relationship between the initial conditions and the forecast metrics. It was Anderson, (2001) who used ensembles to construct an adaptive observing system to obtain the connection between model states at earlier times and present time. Later, Ancell and Hakim, (2007) made a comparative study between adjoint sensitivity analysis and ESA for a wintertime flow pattern over North America. Their study suggests that compared to adjoint sensitivity, ESA provides accurate estimates of forecast response due to changes in initial conditions. They further showed that ensemble sensitivity is connected to adjoint sensitivity by the analysis error covariance matrix. Torn and Hakim (2009) formally applied ESA to an extratropical cyclone over the Western Pacific. They found a 50% reduction in forecast error associated with the minimum sea level pressure of the storm with the introduction of perturbations to initial conditions based on ESA. Torn, (2010) used ESA to evaluate the impact of initial condition errors on the amplitude and position forecasts of African easterly waves. Zack *et al.*, (2010) applied ESA to the short-range forecasts of wind speed in the Tehachapi Pass of California to identify the locations and variables that affects the accuracy of wind forecasts. Each of these studies highlight the importance of ESA in revealing the atmospheric flow features that are relevant to the predictability of the specific atmospheric event under consideration. ESA has been not only applied to large-scale and mesoscale events (Li *et al.*, 2014; Hill, Weiss and Ancell, 2016), but also to several convective events (Hanley *et al.*, 2013; Bednarczyk and Ancell, 2015; Berman *et al.*, 2017). Thus, ESA reveals the coherent flow features related to the dynamically sensitive locations. It also overcomes the limitations imposed by purely linear methods in obtaining the flow features. Though ESA has been widely applied to reveal the flow features over extratropical regions, its application to tropical regions is limited. Therefore, in this thesis ESA has been applied to extreme weather events that happened over the Indian subcontinent.

1.4 Motivation and thesis objectives

Over the past few decades, the occurrence of extreme weather events has increased drastically all over the world. People across the globe experience the impacts of climate change in the form of extreme weather events. Extreme events occur abruptly in the present and are highly visible, in contrast to the long-term climate change trends which are distant, abstract, complicated and gradual. Weather extremes such as extreme rainfalls, tropical cyclones, heat waves etc are becoming more intense and frequent in many regions of the globe.

According to the most recent ranking by Climate Risk Index, India was ranked the fifth country that is most vulnerable to weather and climate extremes

(Mohanty, 2021). During 2017-2018, extreme climate events has resulted in the displacement of over seven million people in India, Nepal and Bangladesh (Goswami *et al.*, 2006). This suggests that the frequency and amplitude of extreme weather events have increased rapidly in India. For example, in the year 2019, about 73 spells of heatwave were reported and the excess rainfall during the second half of Indian summer monsoon led to floods in the central parts of India. Moreover, the north Indian Ocean witnessed the development of 11 tropical cyclones against a normal of 2 (Ray *et al.*, 2021). At the end of the year, during December there were extreme cold wave spells towards northern India. Studies have recorded an 18% increase of annual extreme weather events from 2007 to 2016 compared to the records from 1997 to 2006 (Ray, Arora and Srivastav, 2019). Some of the recent disastrous extreme rainfall events include the heavy rainfall over Uttarakhand in 2013, Srinagar in 2014, Chennai in 2015, Gujarat in 2017, and Kerala in 2018 and 2019. These heavy rainfall events result in landslides, floods and crop damage with major impacts on the economy, society and the environment. Predicting extreme weather events using the NWP models is still fraught with uncertainties and therefore, statistically reliable ensemble predictions from different realization of initial conditions are found to be robust in forecasting heavy rainfall events (Zhang, Snyder and Rotunno, 2003; Schumacher *et al.*, 2013; Mittermaier and Csima, 2017; Klasa *et al.*, 2018).

Timely monitoring and accurate forecasts of extreme events such as heavy rainfall, cloudburst, tropical cyclones and thunderstorms are necessary to respond effectively to such events. Hence it is important to advance our understanding of the key processes that lead to such events and quantify their predictability characteristics. Several studies have used ensemble analysis and forecasts in investigating extreme weather systems (Yu and Meng, 2016). Torn and Hakim (2008) showed that ESA can be employed as a tool to understand the dynamics and predictability associated with severe weather events. Moreover, there are hardly any studies available in the literature on the application of ESA to extreme weather events over the Indian Subcontinent.

This study aims to address the predictability aspects of extreme weather events over the Indian subcontinent by applying ESA in EnKF data assimilation system. The main objectives of the study are as follows:

- To understand the predictability and dynamics of extreme rainfall events over the Indian subcontinent by applying univariate ensemble sensitivity analysis.
- Extending univariate to multivariate ensemble sensitivity that incorporates the contribution from the full covariance matrix in its calculations and quantify the performance of multivariate ensemble sensitivity
- Quantifying the general predictability characteristics of landfalling tropical cyclones over the Bay of Bengal using multivariate ensemble sensitivity analysis.
- Applying ensemble sensitivity analysis for identifying the optimal observation locations for targeted data assimilation during Indian summer monsoon.

CHAPTER 2

DATA AND METHODS

This chapter provides a detailed account of the data and methods used for the different experiments in the thesis. A brief description of data assimilation and the NWP model used in the thesis are also provided.

2.1 Data Assimilation

Numerical weather prediction (NWP) is an initial value problem, wherein the initial conditions should be as accurate as possible to produce better forecasts. Data assimilation (DA) techniques can be used to provide the best estimate of the initial state of the atmosphere to the NWP model.

Basic concept

The forecasts of a future state of the atmosphere using the fundamental equations governing the evolution of the atmosphere are in line with the laws of physics. However, the fundamental dynamical equations being nonlinear, forecasts are sensitive to the initial conditions. Hence, even with perfect model assumptions, accurate predictions are not possible if the initial states are inaccurate. Optimally combining the initial state with observations may improve its accuracy. Data assimilation is the process of combining an imperfect model with noisy observations to produce the best estimate of the state of the system.

Early techniques of data assimilation involved empirical methods such as Cressman analysis and Newtonian nudging. Both methods are now not used because of their inherent issues in the formulation. The least-square approach attempts to solve the data assimilation problem in a statistically rigorous way by minimizing the difference between the analysis and the truth. Optimal interpolation (OI) data assimilation method is an extension of the least square approximation that has been used by operational forecasters from the 1970s to the 80s. Variational data assimilation methods such as 3DVAR and 4DVAR estimate analysis by minimizing a scalar function called the cost function, which is the weighted distance between the background and the observation. Kalman filter data assimilation approach is a least square problem that finds the best state by estimating optimal weight. For high dimensional systems, an ensemble Kalman filter (EnKF) data assimilation system can be employed that uses sample statistics to estimate the background error covariances.

2.1.1 Ensemble Kalman Filter

Ensemble Kalman filter determines the probability density function (PDF) accurately for the current state of the atmosphere, given all current and past observations. Although computationally prohibitive, Bayesian statistical estimation theory can be used to understand the concepts from the Kalman filter to the ensemble Kalman filter. Bayesian data assimilation consists of two steps: assuming that the pdf of the current state of the atmosphere is available, the first step includes assimilation of recent observations, and the second step includes the propagation of the pdf forward in time until new observations are available. Therefore, given all the current and previous observations, the PDF of the current state of the atmosphere can be expressed as the product of the probability distribution for the current observations and the background. Mathematically, Bayes' rule is given by the equation,

$$P(\boldsymbol{x}_t^t \mid \psi_t) \propto P(\boldsymbol{y}_t \mid \boldsymbol{x}_t^t)\, P(\boldsymbol{x}_t^t \mid \psi_{t-1}) \tag{2.1}$$

Here, $\boldsymbol{x}_t^t$ represents the *n*-dimensional true state of the model at time t and ψ_t indicates a collection of observations. The vector ψ_t consists of observations $\boldsymbol{y}_t$ at the most recent time and all the past observations. Given all the past observations up to the time $t-1$, the background or the prior PDF is obtained by integrating the analysis PDF forward in time from the analysis at $t-1$ by using the continuity equation. However, the Bayesian data assimilation becomes complicated and computationally expensive for model states of the $O(10^7)$ (Hamill, 2006).

The Extended Kalman filter (EKF) is an approximation to Bayesian statistical estimation theory which assumes the error growth to be linear and error distributions to be normal. The two components of the EKF include an update step and a forecast step. In the update step, the estimates of state and forecast uncertainty are adjusted to the new measurements. In the forecast step, the updated state and analysis uncertainty estimate is integrated forward in time until the next set of measurements are available.

Let us consider the background and analysis as $\boldsymbol{x}_t^b$ and $\boldsymbol{x}_t^a$, respectively and the measurements or observations are contained in $\boldsymbol{y}$. Further, $\boldsymbol{P}_t^b$, $\boldsymbol{P}_t^a$, and $\boldsymbol{R}$ are used to indicate the respective covariances for background, analysis, and observations with the assumption that background and observation error distributions are Gaussian. The analysis equation is a linear combination of the observations and the background to minimize the variance in the analysis and is given by,

$$\boldsymbol{x}_t^a = \boldsymbol{x}_t^b + \boldsymbol{K}\left(\boldsymbol{y}_t - \boldsymbol{\mathcal{H}}\left(\boldsymbol{x}_t^b\right)\right) \tag{2.2}$$

Here, the forward observational operator $\boldsymbol{\mathcal{H}}$ maps background fields to the observation space. Let the matrix $\boldsymbol{H}$ denote the linear observation operator with $\boldsymbol{H} = \frac{\partial \boldsymbol{\mathcal{H}}}{\partial \boldsymbol{x}}$. The Kalman gain matrix $\boldsymbol{K}$ is given by

$$\boldsymbol{K} = \boldsymbol{P}_t^b \boldsymbol{H}^T \left(\boldsymbol{H}\boldsymbol{P}_t^b \boldsymbol{H}^T + \boldsymbol{R}\right)^{-1} \tag{2.3}$$

Thus, Kalman gain is a function of the background error covariance matrix, the observation error covariance matrix, and the observation matrix. The Kalman gain $\boldsymbol{K}$ is the weight given to the new observation information to correct the background at the relevant model grid points. In the case of a single observation, $\boldsymbol{R}$ in the above equation becomes scalar. Consequently, the variance $\boldsymbol{H}\boldsymbol{P}_t^b \boldsymbol{H}^T$ reduces to a scalar and the covariance between each model state variable and the model estimate of the observation becomes $\boldsymbol{P}_t^b \boldsymbol{H}^T = \left(\boldsymbol{x}_t^b - \overline{\boldsymbol{x}_t^b}\right)\left(\boldsymbol{H}\boldsymbol{x}_t^b - \overline{\boldsymbol{H}\boldsymbol{x}_t^b}\right)^T$. Under such conditions, $\boldsymbol{K}$ becomes the linear regression coefficient between the model state variable and observation with the independent variable being the new observation information and the dependent variable being the analysis.

Assuming no representation errors in the observation forward operator, the true model state and true observations are related through the equation given by,

$$\boldsymbol{y}_t = \boldsymbol{\mathcal{H}}\left(\boldsymbol{x}_t^t\right) \tag{2.4}$$

However, in the presence of observation or measurement errors ϵ, the actual measurements are defined by the relation

$$\boldsymbol{y} = \boldsymbol{\mathcal{H}}(\boldsymbol{x}_t^t) + \epsilon \tag{2.5}$$

Therefore, the observation error covariance matrix is given by,

$$\begin{aligned}\boldsymbol{R} &= \overline{\epsilon\epsilon^{\top}} \\ &= \overline{\left(\boldsymbol{y} - \boldsymbol{\mathcal{H}}(\boldsymbol{x}_t^t)\right)\left(\boldsymbol{y} - \boldsymbol{\mathcal{H}}(\boldsymbol{x}_t^t)\right)^T} \\ &= \overline{(\boldsymbol{y} - \boldsymbol{y_t})(\boldsymbol{y} - \boldsymbol{y_t})^T}\end{aligned} \tag{2.6}$$

where an overbar indicates the expectation value. Finally, assuming that there is no correlation between the observation and background, the error covariance of the analysis state vector is reduced with respect to the forecast error covariance as

$$\begin{aligned}\boldsymbol{P}_t^a &= \overline{(\boldsymbol{x}_t^a - \boldsymbol{x}_t^t)(\boldsymbol{x}_t^a - \boldsymbol{x}_t^t)^T} \\ &= \overline{\left[\boldsymbol{x}_t^b - \boldsymbol{x}_t^t + \boldsymbol{K}\left(\boldsymbol{y} - \boldsymbol{y_t} - \boldsymbol{\mathcal{H}}(\boldsymbol{x}_t^b) + \boldsymbol{\mathcal{H}}(\boldsymbol{x}_t^t)\right)\right]} \\ &\qquad \overline{\left[\boldsymbol{x}_t^b - \boldsymbol{x}_t^t + \boldsymbol{K}\left(\boldsymbol{y} - \boldsymbol{y_t} - \boldsymbol{\mathcal{H}}(\boldsymbol{x}_t^b) + \boldsymbol{\mathcal{H}}(\boldsymbol{x}_t^t)\right)\right]^T} \\ &= (\boldsymbol{I} - \boldsymbol{KH})\overline{\left(\boldsymbol{x}_t^b - \boldsymbol{x}_t^t\right)\left(\boldsymbol{x}_t^b - \boldsymbol{x}_t^t\right)^T}(\boldsymbol{I} - \boldsymbol{KH})^T + \boldsymbol{K}\overline{(\boldsymbol{y} - \boldsymbol{y_t})(\boldsymbol{y} - \boldsymbol{y_t})^T}\boldsymbol{K}^T \\ &= (\boldsymbol{I} - \boldsymbol{KH})\boldsymbol{P}_t^b(\boldsymbol{I} - \boldsymbol{H}^T\boldsymbol{K}^T) + \boldsymbol{KRK}^T \\ &= \boldsymbol{P}_t^b - \boldsymbol{KHP}_t^b - \boldsymbol{P}_t^b\boldsymbol{H}^T\boldsymbol{K}^T + \boldsymbol{K}\left(\boldsymbol{HP}_t^b\boldsymbol{H}^T + \boldsymbol{R}\right)\boldsymbol{K}^T \\ &= (\boldsymbol{I} - \boldsymbol{KH})\boldsymbol{P}_t^b\end{aligned} \tag{2.7}$$

The background is then updated from the previous analysis time to the next time when the observations are available at $t + 1$. The analysis at time t is integrated forward using a model and is represented as

$$\boldsymbol{x}_{t+1}^{b} = \mathcal{M}(\boldsymbol{x}_t^a) \tag{2.8}$$

where $\mathcal{M}$ is the nonlinear forecast model. Considering $\boldsymbol{M} = \frac{\partial \mathcal{M}}{\partial x}$, the linear tangent model, the background error covariance matrix at time $t+1$ is determined by the equation,

$$\boldsymbol{P}_{t+1}^{b} = \boldsymbol{M}\boldsymbol{P}_t^a\boldsymbol{M}^T + \boldsymbol{Q} \tag{2.9}$$

where $\boldsymbol{Q} = \overline{\eta\eta^T}$ is the model error covariance matrix with η being the stochastic forcing representing model errors. Thus, the formulation favours the propagation of error covariances along with the nonlinear model itself and provides flow-dependent forecast error covariance with the observation information spread optimally during data assimilation.

The above equation is computationally expensive as it estimates the background error covariance which is of the order of $10^7 \times 10^7$. Additionally, in situations where non-linearity is prominent, the EKF may not provide accurate estimates of analysis as the error covariance matrix is propagated in time employing a tangent linear model. Therefore, Evensen (1994) proposed the ensemble Kalman filter (EnKF) which is a Monte Carlo alternative to EKF. In the EnKF, background and forecast error covariances are estimated from a finite set of ensemble members through parallel data assimilation and forecast cycles. A set of ensemble forecasts are used to estimate background error covariances and after that analysis ensembles are produced. The ensemble of analyses is then used to generate an ensemble of short-term forecasts until the next time when the observations are available. Let us assume $\boldsymbol{X}^b$ to be an ensemble of forecasts at time t that randomly sample the model background errors. That is,

$$\boldsymbol{X}^b = \left(\boldsymbol{x}_1^b, \dots, \boldsymbol{x}_m^b\right) \tag{2.10}$$

Here, the subscript represents the ensemble member. With the ensemble mean defined as,

$$\overline{\boldsymbol{x}^b} = \frac{1}{m}\sum_{i=1}^{m} \boldsymbol{x}_i^b \tag{2.11}$$

and perturbation from mean as $\boldsymbol{x}_i'^b = \boldsymbol{x}_i^b - \overline{\boldsymbol{x}^b}$, the flow-dependent background error covariance estimated from a finite number of ensembles is given by

$$\widehat{\boldsymbol{P}}^b = \frac{1}{m-1}\boldsymbol{X}'^b\boldsymbol{X}'^{bT} \tag{2.12}$$

Where $\boldsymbol{X}'^b = \left(\boldsymbol{x}_1'^b, \ldots, \boldsymbol{x}_m'^b\right)$ is a matrix formed from a finite ensemble of perturbations and $\widehat{\boldsymbol{P}}^b$ approach $\boldsymbol{P}^b$ as the number of members approaches infinity.

2.1.2 DART EAKF

The Data Assimilation Research Testbed (DART) is an open-source community facility that delivers software tools for research, development, and education in ensemble data assimilation (Anderson, J. L., T. Hoar, K. Raeder, H. Liu, N. Collins, 2009). Atmospheric scientists, oceanographers, hydrologists, chemists, and geophysicists depend on the DART's ensemble data assimilation algorithms and diagnostic tools for research and operational forecast purposes. The forward operators required to assimilate non-conventional observations such as Global Positioning System (GPS) radio occultation soundings are available in DART data assimilation system. DART also carries codes for applications such as parameter estimation, sensitivity analysis, observing system design and smoothing.

DART assimilates observations using sequential ensemble data assimilation. The ensemble members required for the DART EnKF data assimilation are provided by an ensemble of forecasts obtained randomly from the probability distribution of the model's state. The assimilation process and algorithms used in DART by considering three ensemble members at time t_k is illustrated in Fig. 2.1. The model is integrated forward in time from t_k to t_{k+1}, when the next observations are available. The prior estimates of the observation y are then estimated by applying a forward observation operator h to each model state vector. While the instrument gives the observed value, the instrument error gives observation likelihood. An updated ensemble is then estimated by an ensemble filter which combines the observation, the prior ensemble and the likelihood. DART uses algorithms such as the ensemble Kalman filter (EnKF) and the ensemble adjustment Kalman filter (EAKF) for computing the updated ensemble. EAKF is a deterministic square root filter that has the ability to assimilate observations in the presence of complex nonlinear dynamics. Compared to EnKF, EAKF uses a different algorithm to update the ensemble members using the available observations. In EAKF, the updated ensemble is obtained by the application of a linear operator, **A**, to the prior ensemble or the background. Based on the EnKF Eq. 2.2, the deviation of the ith member from the mean can be written as

$$\boldsymbol{x}'^{a}_{t} = \boldsymbol{x}'^{b}_{t} + \widetilde{\boldsymbol{K}}\left(\boldsymbol{y}'_{t} - \boldsymbol{\mathcal{H}}\left(\boldsymbol{x}'^{b}_{t}\right)\right) \tag{2.13}$$

However, if there is no necessity to perturb the observations, then the above equation simplifies to

$$\boldsymbol{x}'^{a}_{t} = \boldsymbol{x}'^{b}_{t} + \widetilde{\boldsymbol{K}}\boldsymbol{\mathcal{H}}\left(\boldsymbol{x}'^{b}_{t}\right) = \left(\boldsymbol{I} - \widetilde{\boldsymbol{K}}\boldsymbol{\mathcal{H}}\right)\boldsymbol{x}'^{b}_{t} \tag{2.14}$$

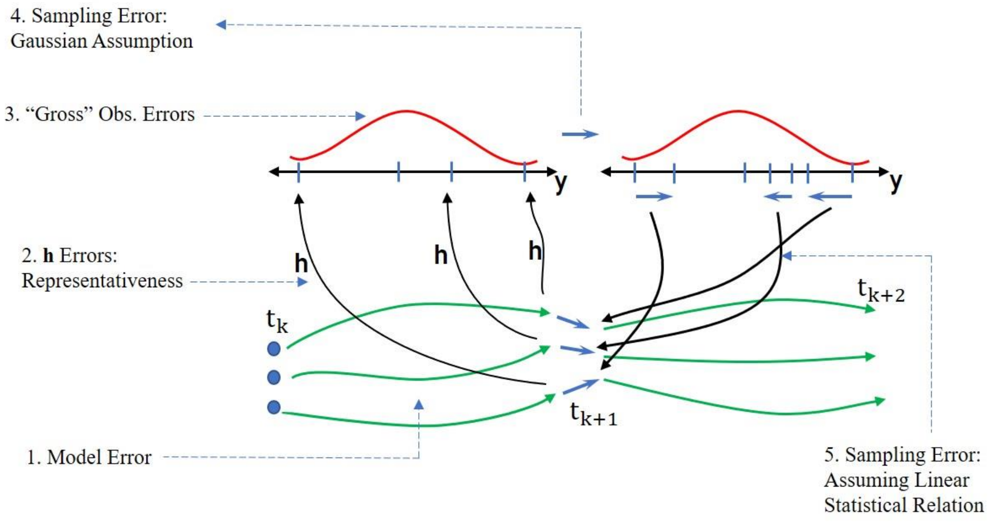

Figure 2. 1. *Idealized assimilation process of an EnKF with errors. The blue asterisk indicates three ensemble members at time* t_k *which are then integrated to time* t_{k+1} *by a forecast model indicated by green vectors. The letter h represents the forward operator and green ticks on the upper left axes indicate observation estimates. The observed value is represented as a red tick while the red curve shows the likelihood. The updated ensemble estimate is represented by blue ticks and increments by blue vectors on the upper right axes. Blue vectors at time* t_{k+1} *are used to represent the updated ensemble estimates.*

where $\widetilde{\boldsymbol{K}}$ is the reduced Kalman gain which is used to update the deviations from the ensemble mean. Therefore, if observations are processed independently, $\boldsymbol{HP}_t^b\boldsymbol{H}^T$ and $\boldsymbol{R}$ become scalars then the reduced Kalman gain is given by,

$$\widetilde{\boldsymbol{K}} = \left(1 + \sqrt{\frac{\boldsymbol{R}}{\boldsymbol{HP}_t^b\boldsymbol{H}^T + \boldsymbol{R}}}\right)^{-1} \boldsymbol{K} \tag{2.15}$$

Here, $\boldsymbol{K}$, $\widetilde{\boldsymbol{K}}$, $\boldsymbol{HP}_t^b$, and $\boldsymbol{P}_t^b\boldsymbol{H}^T$ are all n-dimensional vectors and the quantity in the bracket is a scalar with values between 0 and 1. In the EAKF, the linear operator becomes $\mathbf{A} = \boldsymbol{I} - \widetilde{\boldsymbol{KH}}$ and the updated ensemble has error covariance identical to Eq. 2.7.

2.2 Data Assimilation for Ensemble Generation

Evensen (2003) proposed a method in which ensemble members are generated by adding perturbations in the form of random fields from a specified distribution to an initial best guess of the model states. The ensemble is then integrated over a time interval covering a few characteristic time scales of the dynamical system (Evensen, 2003). This is to make certain that the system is in dynamic balance and achieved appropriate multivariate correlations. Evensen (2004) also proposed an improved sampling scheme that uses randomly generated ensembles, spatially correlated fields, and Singular Value Decomposition (SVD) method to obtain independent perturbations. Above mentioned methods have the disadvantage that they are applied before the assimilation period and therefore, require to spin-up to attain dynamic balance by which time the model equations may alter the prescribed error distribution.

The breeder method and optimal perturbations (singular vectors) are the other two methods for generating ensembles that have been used in operational ensemble forecasting. Both of these methods involve the generation of a set of fastest growing errors. Though these methods can be used for ensemble initialization in certain cases, they are not recommended as a generic approach. The singular vector method generates the fastest growing errors by using an adjoint model, which needs to be developed specifically. Therefore, though the singular vector method has obvious benefits, its use is impractical. On the other hand, the breeder method is simple to apply, however, its ability to estimate forecast error variance with accuracy is questionable.

If the ensemble forecast has a large spread, then the error covariance matrix $\widehat{\boldsymbol{P}}^b$ in Eq. 2.12 will be overestimated and the analysis will tend to overfit the observations. On the contrary, if the ensemble forecast has a smaller spread, then the error covariance matrix $\widehat{\boldsymbol{P}}^b$ will be underestimated and the analysis will tend to underfit the observations. If the forecast error is represented improperly by the ensembles, it will result in a sub-optimal filter. Therefore, while using the EnKF data assimilation technique, it is essential to generate and propagate the ensembles with reasonable variance. The EnKF analysis Eq. 2.2 can be written as a linear combination of the deviations in ensembles and therefore, the independent ensemble deviations make the analysis more efficient. Hence, if the ensembles with linearly independent deviations can be generated then it results in a more efficient assimilation filter.

The nonlinear filtering theory that estimates the probability distribution of a state of the NWP model given a set of observations unifies DA and ensemble generation problem. The continuously cycling EnKF data assimilation will yield flow-dependent initial condition perturbations drawn from an equally likely state. Therefore, the ensembles used in this study conform to the principle of equal likelihood.

2.3 Regional Model WRF

The Weather Research and Forecasting (WRF) model is an atmospheric simulation system and mesoscale non-hydrostatic NWP model designed for operational and research purposes. The WRF model is developed jointly by the National Centre for Atmospheric Research (NCAR), the National Oceanic and Atmospheric Administration (NOAA), the Naval Research Laboratory (NRL), the University of Oklahoma, the Air Force Weather Agency (AFWA) and the Federal Aviation Administration (FAA). The model is suitable for a wide range of applications such as real-time NWP, data assimilation, parameterization schemes, regional climate simulations, atmosphere-ocean coupling, air quality modeling and idealized simulations across scales ranging from convective scales to synoptic scales. The WRF Software Framework (WSF) involves two dynamics solvers: the Advanced Research WRF (ARW) solver and the Nonhydrostatic mesoscale model (NMM) solver which are maintained by NCAR and NCEP respectively. Additionally, the WSF includes physics packages, initialization programs, WRF-Var and WRF-Chem.

The present study uses the WRF-ARW dynamical solver and the corresponding WRF system components are shown in Fig. 2.2.

2.3.1 ARW Equations

The ARW dynamic solver integrates the non-hydrostatic and compressible Euler equations which are presented in the flux form. A terrain-following mass vertical coordinate η is used in the formulation of the ARW equations and it is defined as

$$\eta = \frac{(p_h - p_t)}{\mu} \tag{2.16}$$

where $\mu = p_s - p_t$. Here, p_h refers to the hydrostatic component of the pressure whereas p_s and p_t represent the pressure values along the surface and top boundaries of the model, respectively. The vertical coordinate η varies from a value of 1 at the surface to 0 at the top boundary. Therefore, at any grid point (x, y), the variables are represented in the flux form as $\mathbf{V} = \mu\mathbf{v} = (U, V, W)$, $\Omega = \mu\dot{\eta}$, $\Theta = \mu\theta$

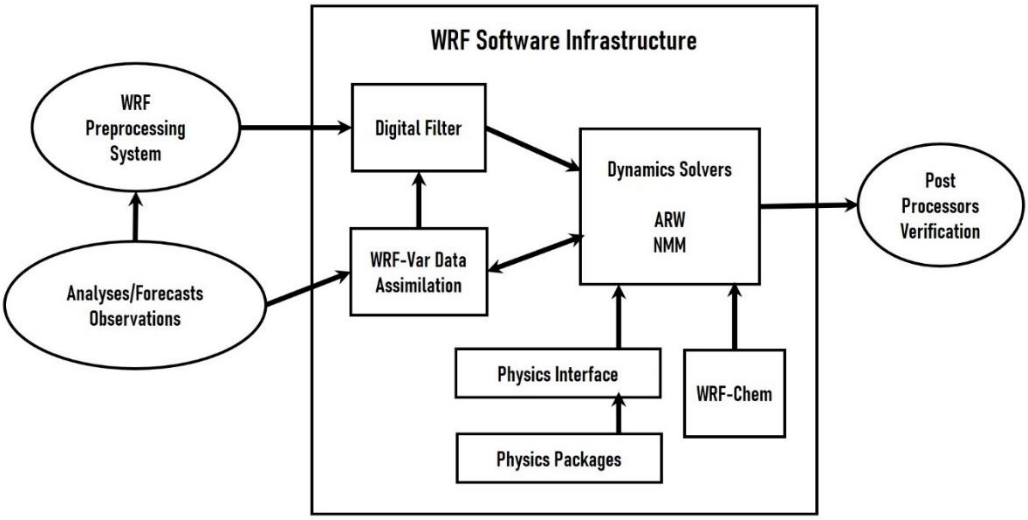

Figure 2. 2. WRF System Components

with $\mu(x,y)$ being the mass per unit area within the vertical column. $\mathbf{v} = (u, v, w)$ indicate the covariant velocities in the two horizontal and vertical directions, respectively whereas $\omega = \dot{\eta}$ represents the contravariant vertical velocity. The equations also contain variables such as the potential temperature (θ), the non-conserved geopotential $\phi = gz$, pressure (p), and the inverse density $\alpha = 1/\rho$. Using these variables, Euler equations in the flux-form can be written as:

$$\partial_t U + (\mathbf{\nabla}.\mathbf{V}u) - \partial_x(p\partial_\eta \phi) + \partial_\eta(p\partial_x \phi) = F_U \tag{2.17}$$

$$\partial_t V + (\mathbf{\nabla}.\mathbf{V}v) - \partial_y(p\partial_\eta \phi) + \partial_\eta(p\partial_y \phi) = F_V \tag{2.18}$$

$$\partial_t W + (\mathbf{\nabla}.\mathbf{V}w) - g(\partial_\eta p - \mu) = F_W \tag{2.19}$$

$$\partial_t \Theta + (\mathbf{\nabla}.\mathbf{V}\theta) = F_\Theta \tag{2.20}$$

$$\partial_t \mu + (\mathbf{\nabla}.\mathbf{V}) = 0 \tag{2.21}$$

$$\partial_t \phi + \mu^{-1}[(\mathbf{V}.\mathbf{\nabla}\phi) - gW] = 0 \tag{2.22}$$

Here, F_U, F_V, F_W, and F_Θ represent the forcing terms due to model physics, turbulent mixing, spherical projections and rotation of the Earth.

Moreover, the diagnostic relation for the inverse density is given by

$$\partial_\eta \phi = -\alpha\mu \tag{2.23}$$

with the equation of state being $p = p_0(R_d\theta/p_0\alpha)^\gamma$, where p_0 is the reference pressure, R_d is the gas constant for dry air and $\gamma = c_p/c_v = 1.4$ is the ratio of specific heat capacities for dry air. The ARW solver supports four projections to the sphere such as Mercator, Lambert conformal, latitude-longitude and Lambert projections. In this thesis, we have used the Lambert projection for all the experiments. Further, the ARW solver uses a third-order Runge-Kutta scheme for the time integration of slow or low-frequency modes and integration over smaller time steps is performed for high-frequency modes. Finally, in the ARW solver, the variables are spatially discretized based on the Arakawa-C grid staggering. The ARW solver also supports horizontal nesting with different options such as two-way nesting, one-way nesting and moving nests. Compared to other models, ARW’s nesting infrastructure can perform nested simulations efficiently on distributed memory parallel computer systems.

The WRF-ARW also offers several categories of physics parameterization schemes to simulate real atmospheric conditions and to perform coarser resolution simulations. They are microphysics, planetary boundary layer, cumulus parameterization, radiation and land-surface model.

2.3.2 Post-processing using NCL

NCL (NCAR Graphics Command Language) is an interpreted programming language with a wide variety of graphic capabilities and many useful functions and procedures. NCL has the advantage that it can read the WRF data directly, make simple plots quickly, and easy to make difference plots. Therefore, this study uses NCL for visualization and analysis purposes.

2.4 Data Used

Data for model initialization: The static data used by the geogrid program in the WPS is downloaded from the WRF page. Since the terrestrial data are time-invariant, these data need to be downloaded only once. The present thesis uses high-resolution geographical dataset in the geogrid program of the WPS. Data source https://www2.mmm.ucar.edu/wrf/users/download/get_sources_wps_geog.html

NCEP-GFS data: After interpolating the time-invariant geographical data, the time-varying meteorological data needs to be interpolated to initialize WRF and this data can be obtained from the NCEP-GFS. The NCEP-GFS uses a global spectral model with spherical harmonic functions and a horizontal resolution of approximately 13km at the equator for 0-10 days. In this study, we have used 6 hourly NCEP-GFS data at $0.25° \times 0.25°$ grided resolution to initialize the WRF model.

Data source: https://rda.ucar.edu/datasets/ds084.3/

Table 2. 1. TIGGE project partners

	Centre	**Country**	**Acronym**	**Ensemble Size**
1	Bureau of Meteorology	Australia	BoM	17
2	China Meteorological Administration	China	CMA	30
3	Centro de Previsao de Tempo e Estudos Climaticos	Brazil	CTPEC	15
4	Deutscher Wetterdienst	Germany	DWD	40
5	Environment and Climate Change Canada	Canada	ECCC	21
6	European Centre for Medium Range Weather Forecasts	Europe	ECMWF	51
7	Indian Meteorological Department	India	IMD	21
8	Japan Meteorological Administration	Japan	JMA	51
9	Korea Meteorological Administration	Korea	KMA	26
10	Meteo-France	France	MF	35
11	National Centres for Environmental Prediction	United States	NCEP	31
12	National Centre for Medium Range Weather Forecasting	India	NCMRWF	12
13	Met Office	United Kingdom	UKMO	18

Surface and upper-air Observations: All the data assimilation experiments in this thesis use NCEP surface and upper-air observations in PREPBUFR format to assimilate observations every 6 hours. The data includes conventional observations such as radiosonde, land surface, marine surface, pilot balloon and aircraft reports from the Global Telecommunication System (GTS).

Data source: https://rda.ucar.edu/datasets/ds337.0/

TIGGE: The Observing system Research and Predictability Experiment (THORPEX) Interactive Grand Global Ensemble (TIGGE) was launched to support the research activities of the THORPEX program by providing operational ensemble forecasts to the international research community. The key objectives of TIGGE are as follows: (1) enhancing international collaboration between universities and operational centers on ensemble prediction, (2) finding new methods to combine ensembles from various sources and correcting for errors such as biases and under/overestimation of spread, (3) understanding the contribution of uncertainties in the model, initial conditions and observations to forecast error, (4) exploring the benefit of ensemble systems to changing uncertainty and (5) development of a future Global Interactive Forecast System (GIFS). The TIGGE dataset consists of ensemble forecasts from 13 global NWP centers (Table 2.1), starting from October 2006 and is available to the research community through the ECMWF and CMA archive portals. Among the thirteen centers, three centers: NCMRWF, IMD and DWD started providing data in recent years. The present study uses ECMWF and NCEP ensemble forecasts at $0.5° \times 0.5°$ grided resolution from TIGGE for the experiments with predictability assessments involving large-scale flow structures.

Data Source: https://apps.ecmwf.int/datasets/data/tigge/levtype=sfc/type=cf/

2.5 Ensemble Sensitivity Analysis

In this section, brief information about ensemble sensitivity analysis (ESA) and its relation to adjoint sensitivity analysis is presented. Sensitive regions are geographical locations where the error growth associated with initial perturbations influences the forecasts within a given region at the verification time. The purpose of sensitivity analysis is to locate the areas for adaptive observations and to understand dynamics. Typically, adjoint or singular-vector methods are used for sensitivity analysis, however, these methods require the tangent-linear or adjoint models which are computationally expensive. In this thesis, an ensemble approach is used to determine how the forecast metrics vary to changes in the initial conditions. Ancell and Hakim (2007) use linear regression and provide a comparison between ensemble and adjoint sensitivity analysis.

Consider a discrete dynamical system, then its evolution with respect to time for the state vector $\boldsymbol{x}$ is described by

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{F}(\boldsymbol{x}) \tag{2.24}$$

where $\boldsymbol{F}(\boldsymbol{x})$ indicates a nonlinear vector function. If $\boldsymbol{F}(\boldsymbol{x})$ can be linearized about a reference trajectory, then the evolution of an initial perturbation is given by

$$\delta\boldsymbol{x}_t = \boldsymbol{R}_{t,t_0}\delta\boldsymbol{x}_0 \tag{2.25}$$

Here, t and t_0 denote the initial and final time, respectively. The resolvent matrix $\boldsymbol{R}_{t,t_0}$ maps the perturbation vector $\delta\boldsymbol{x}_0$ at the initial time into the perturbation vector $\delta\boldsymbol{x}_t$ at forecast time. Let us consider the scalar metric J as the forecast metric, which is a function of the model state at an earlier time t. Then the Taylor expansion can be used to estimate J, for small changes about the control solution

$$J(\boldsymbol{x}_t + \delta\boldsymbol{x}_t) = J(\boldsymbol{x}_t) + \left[\frac{\partial J}{\partial \boldsymbol{x}_t}\right]^T \delta\boldsymbol{x}_t + \cdots, \tag{2.26}$$

Therefore, the change in the forecast metric J is given by

$$\delta J = J(\boldsymbol{x}_t + \delta\boldsymbol{x}_t) - J(\boldsymbol{x}_t) = \left[\frac{\partial J}{\partial \boldsymbol{x}_t}\right]^T \delta\boldsymbol{x}_t \tag{2.27}$$

Substituting for $\delta\boldsymbol{x}_t$ from Eq. 2.25 the above expression becomes

$$\delta J = \left[\frac{\partial J}{\partial \boldsymbol{x}_t}\right]^T \boldsymbol{R}_{t,t_0}\delta\boldsymbol{x}_0 \tag{2.28}$$

Using the algebraic properties of transpose the change in J becomes

$$\delta J = \left[\boldsymbol{R}_{t,t_0}^T \frac{\partial J}{\partial \boldsymbol{x}_t}\right]^T \delta\boldsymbol{x}_0 \tag{2.29}$$

The transpose of the resolvent matrix represents the adjoint of a tangent linear model, that maps the sensitivity gradient at time t backward obtaining the sensitivity gradient with respect to $\boldsymbol{x}_0$ at time t_0.

$$\boldsymbol{R}_{t,t_0}^{T}\frac{\partial J}{\partial \boldsymbol{x}_t}=\frac{\partial J}{\partial \boldsymbol{x}_0} \tag{2.30}$$

Substituting the above equation back into Eq. 2.29 we obtain the change in the forecast metric at time t due to changes in the initial condition as,

$$\delta J=\left[\frac{\partial J}{\partial \boldsymbol{x}_0}\right]^{T}\delta \boldsymbol{x}_0 \tag{2.31}$$

Here $\frac{\partial J}{\partial \boldsymbol{x}_0}$ represents the adjoint sensitivity of the scalar forecast metric to initial conditions and if this quantity is large then it indicates the regions where the errors in the initial conditions grow rapidly during the forecast. However, the adjoint sensitivity is estimated by computing the adjoint of a tangent linear model which is computationally expensive. Therefore, as an alternative, an ensemble of independent samples is used statistically to obtain how the initial condition affects the forecast metric. To obtain the relationship between adjoint and ensemble sensitivity Eq. 2.31 is right multiplied by $\delta \boldsymbol{x}_0^{T}$ and by calculating the expected value

$$E\left[\delta J\delta \boldsymbol{x}_0^{T}=\left[\frac{\partial J}{\partial \boldsymbol{x}_0}\right]^{T}\delta \boldsymbol{x}_0\delta \boldsymbol{x}_0^{T}\right] \tag{2.32}$$

Since $\left[\frac{\partial J}{\partial \boldsymbol{x}_0}\right]^{T}$ is a deterministic quantity applying to the control trajectory (Ancell and Hakim, 2007), the above equation can be written as,

$$cov(\delta J,\delta \boldsymbol{x}_0)=\left[\frac{\partial J}{\partial \boldsymbol{x}_0}\right]^{T}\mathbf{P}^{a} \tag{2.33}$$

where $\mathbf{P}^{a}$ is the initial-time error covariance matrix and cov represents the covariance between the forecast metric and the initial conditions. It is to be noted that δJ and $\delta \boldsymbol{x}_0$ are assumed to have zero mean.

The covariance alone is not enough to obtain the linear regression relationship between the forecast metric and the initial state. Right multiplying the above equation by $\mathbf{P}^{a^{-1}}$ gives the multivariate linear regression that recovers the adjoint sensitivity within sampling error as

$$\left[\frac{\partial J}{\partial \boldsymbol{x}_0}\right]^T = cov(\delta J, \delta \boldsymbol{x}_0)\mathbf{P}^{a^{-1}} \tag{2.34}$$

However, practically it is computationally difficult to invert the matrix $\mathbf{P}^{a^{-1}}$ and is not unique if the number of ensemble members is small. Therefore, a univariate regression approach is used to obtain ensemble sensitivity by regressing to all the initial degrees of freedom

$$\left[\frac{\partial J_e}{\partial \boldsymbol{x}_0}\right]^T = cov(J, \delta \boldsymbol{x}_0)\mathbf{D}^{-1} \tag{2.35}$$

where **D** is a diagonal matrix with the error variance at the initial time and $\frac{\partial J_e}{\partial \boldsymbol{x}_0}$ is the ensemble sensitivity vector. The above equation can also be represented as

$$\frac{\partial J_e}{\partial x} = \frac{cov(\boldsymbol{J}, \boldsymbol{x})}{var(\boldsymbol{x})} \tag{2.36}$$

Here, $\boldsymbol{x}$ and $\boldsymbol{J}$ represents the $1 \times K$ ensemble estimates of the state variable and forecast metric, respectively calculated from K ensemble members. If the diagonal elements are dominant over the off-diagonal elements then the difference between the two regression approaches can be considered to be small. Using Eqs. 2.34 and 2.35, the relationship between adjoint and ensemble sensitivity analysis is obtained as

$$\frac{\partial J_e}{\partial x} = \mathbf{D}^{-1}\mathbf{P}^{a}\frac{\partial J}{\partial \boldsymbol{x}_0} \tag{2.37}$$

The advantage of using ensemble sensitivity is that the matrix **D** is trivial to calculate and hence is computationally more efficient than adjoint sensitivity. If the right-hand side of Eq. 2.36 is multiplied by the ensemble standard deviation, then the ensemble sensitivity can be expressed in the units of the forecast metric. Since ensemble sensitivity is estimated from a small set of ensemble members compared to the number of state variables, it is affected by sampling errors. The issues of sampling error in univariate ensemble sensitivity are addressed by testing for statistical significance similar to Torn and Hakim (2008a). According to their study, a state variable can produce a statistically significant change in the forecast metric if

$$\left|\frac{\partial J}{\partial x}\right| > \delta_s$$

where δ_s indicates the confidence interval on the regression coefficient (Wilks, 2007) which is 99% in this work. If the above equation is satisfied, the null hypothesis of no relationship between the forecast metric and the state variable can be rejected with 99% confidence.

2.5.1 Initial Condition Perturbation

The hypothesis based on ensemble sensitivity analysis is validated quantitatively by applying perturbations to the initial conditions in the most sensitive regions. The perturbed initial conditions thus obtained are then integrated forward by using a non-linear model. The resulting forecast metric is then compared to the change in the forecast metric predicted by ensemble sensitivities. In this work, perturbations are applied only to the initial conditions of one ensemble member. Therefore, the control forecast is considered as the forecast of the ensemble member whose forecast metric values are closest to the ensemble mean. The perturbed ith state variable $\mathbf{x}_{\mathrm{i}}^{\mathrm{p}}$ when a change of α is applied to the state variable in the most sensitive region is given by

$$\mathbf{x}_{\mathrm{i}}^{\mathrm{p}} = \mathbf{x}_{\mathrm{i}}^{\mathrm{a}} + \frac{\partial x_{\mathrm{i}}^{\mathrm{a}}}{\partial x_{\mathrm{s}}} \alpha \tag{2.38}$$

where

$$\frac{\partial x_i^a}{\partial x_s} = \frac{cov(\boldsymbol{x}_i^a, \boldsymbol{x}_s)}{var(\boldsymbol{x}_s)} \tag{2.39}$$

Here, $\boldsymbol{x}_s$ and $\boldsymbol{x}_i^a$ are $1 \times K$ ensemble estimates of the perturbation state variable and ith control analysis state variable, respectively. Thus, this experiment is similar to assimilating a hypothetical observation within the most sensitive grid point where the innovation is α and the observation error covariance is assumed to be zero.

CHAPTER 3

UNIVARIATE ENSEMBLE SENSITIVITY ANALYSIS APPLIED TO EXTREME RAINFALL EVENTS OVER THE INDIAN SUBCONTINENT

3.1 Introduction

During the past few decades, the frequency of occurrence as well as the intensity of extreme rainfall events over the Indian subcontinent have been on the rise (Rajeevan, Bhate and Jaswal, 2008; Roxy *et al.*, 2017). Recent heavy rainfall episodes over the Indian subcontinent such as the rainfall in Mumbai during July 2005 (Sahany, Venugopal and Nanjundiah, 2010), in Uttarakhand during June 2013 (Vellore *et al.*, 2016; Houze *et al.*, 2017), in Chennai during December 2015 (Srinivas *et al.*, 2018; Nikumbh *et al.*, 2021) and in Kerala during 2018 (Mishra *et al.*, 2018; Viswanadhapalli *et al.*, 2019) and 2019 have resulted in huge damage to property and loss of life. Providing accurate predictions of such extreme events in advance can significantly reduce such losses. However, the prediction of extreme rainfall events using NWP models is challenging due to the uncertainties associated with the initial conditions and the models. Statistically reliable ensemble predictions from the different realization of initial conditions of the atmosphere are found to be robust in forecasting heavy rainfall events. Additionally, the forecast ensembles can be employed for understanding the predictability and dynamics of extreme weather events using sensitivity analysis.

Extreme precipitation events over and near the Himalayas are often associated with large-scale synoptic conditions such as the southward intrusion of the upper-level westerly trough and northward propagating moisture-laden low-level circulation (Vellore *et al.*, 2016). The heavy rainfall that occurred over the Uttarakhand state in June 2013 is one such event with strong synoptic forcing that caused massive destruction to life and properties.

*The content of Chapter 3 is taken from "**George, B.**, Kutty, G., 2021. Ensemble sensitivity analysis of an extreme rainfall event over the Himalayas in June 2013. Dynamics of Atmospheres and Ocean. 93 and "**George, B.**, Kutty, G., 2022. Sensitivity analysis applied to two extreme rainfall events over Kerala using TIGGE ensembles. Meteorology and Atmospheric Physics 134, 1–14.*

Although the key factors responsible for the Uttarakhand storm have been recognized in the literature, the uncertainties in the storm producing factors, their relative importance in the development of the storm, and the predictability aspects have not been investigated. In the context of climate change, it has been reported that the unprecedented extreme precipitation episodes and flooding events are on the rise over the Himalayas during the past few decades (Priya *et al.*, 2017). Hence, it is important to advance our understanding of the key processes that lead to such events and quantify its predictability characteristics. Fig. 3.1 shows the spatial distribution of accumulated precipitation from the Tropical Rainfall Measuring

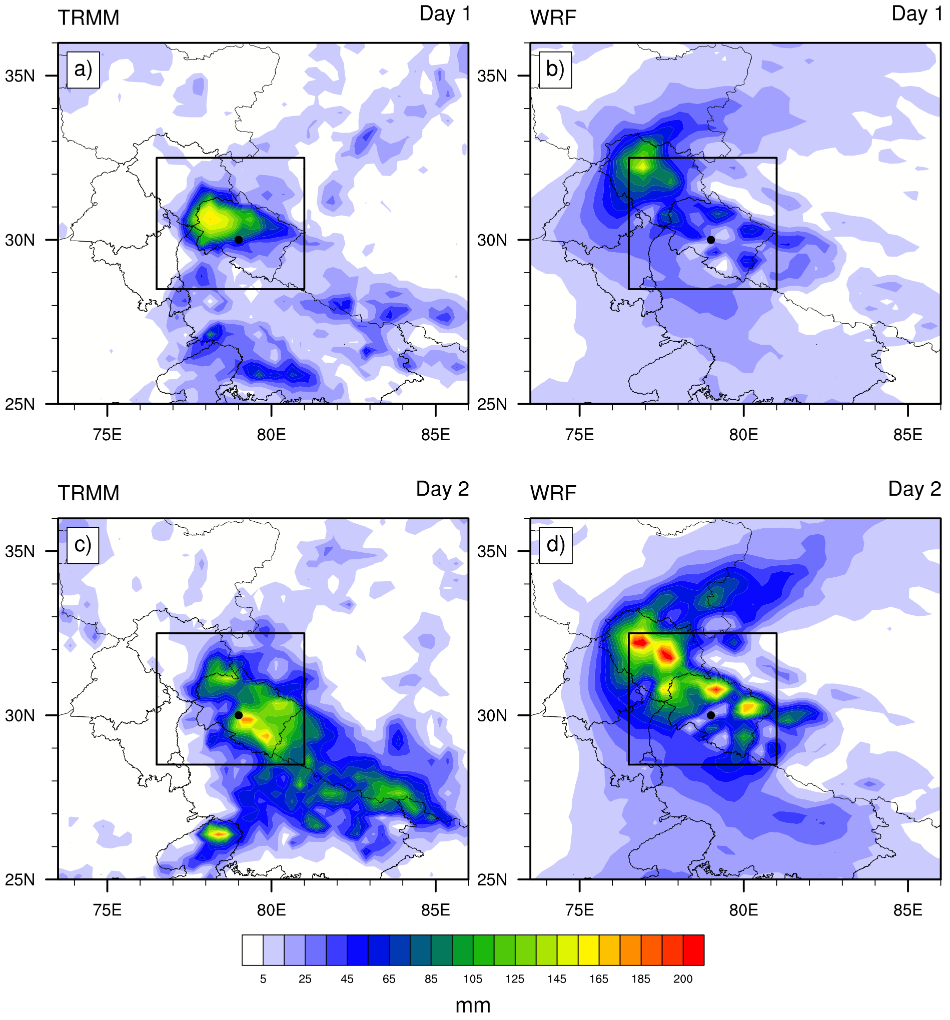

Figure 3. 1. Rainfall distribution of TRMM and ensemble mean on the two days of the rainfall event (D1 and D2). Shading denotes the 24 h accumulated precipitation (a) TRMM and (b) WRF forecast for D1, while (c) TRMM and (d) WRF are for D2 precipitation. The black box denotes the location of the response function. The black dot represents the location of Uttarakhand.

Mission (TRMM) satellite and ensemble mean forecast using the WRF model valid at 1200UTC of 16 and 17, June 2013 (hereafter D1 and D2). As can be seen from Fig. 3.1, the location and intensity errors associated with the forecasted precipitation are substantial and the location of precipitation maxima in the model forecast is shifted northward as compared to the TRMM observations, for both the days. Therefore, one objective of the present study is to identify the uncertainties in the factors that lead to errors in the location and intensity of precipitation maxima using an ensemble framework. Further, the study will also address the predictability aspects of the event in synoptic as well as in convective scales by applying ESA. Another objective of the present chapter is to identify the key synoptic features that lead to uncertainties in the rainfall forecasts of the heavy rainfall events happened over Kerala in 2018 and 2019 by employing the ESA method.

Kerala, the southernmost Indian state, has experienced heavy rainfall and torrential flood episode in August 2018 (hereafter KF18), which produced extensive damages to human life and properties that worth around 3.8 billion US dollars. A similar event is repeated in the very next year in August 2019 (hereafter KF19) causing massive landslides and floods that took lives and left many homeless. The synoptic features that are in common for the two events include the formation of a depression over the Bay of Bengal, stronger LLJ (Low Level Jet) over the Arabian Sea, an off-shore trough over the west coast of the Indian subcontinent, and circulations over the Western North Pacific (WNP). From June 1 to August 19 of the year 2018, Kerala received extensive rainfall of about 2347 mm in contrast to an expected climatological precipitation value of 1650 mm (Padma, 2018). Typically, during southwest monsoon the strong westerly winds from the Arabian Sea brings huge amount of moisture towards the west coast of India and heavy downpour occurs on the windward side of the Western Ghats barrier. If that is the case it needs to be questioned what made the rainfall in the years 2018 and 2019 unique. Numerous studies have worked on the dynamics of KF18, however studies on KF19 are limited. Baisya and Pattnaik (2019) have shown that the synoptic-scale oscillations that favored the development of extreme rainfall over Kerala include the development of an anti-cyclonic circulation towards the northwest of India, a monsoon depression over the Bay of Bengal, and a Phase-6 Madden-Julian

Oscillation (MJO). Mohandas *et al.* (2020) found that the prevalence of a conveyor belt-like flow in the presence of remotely aligned intense tropical circulations is responsible for bringing additional moisture to the Kerala coast, which has resulted in the KF18 event. Viswanadhapalli *et al.* (2019) addressed the KF18 event using a very high-resolution mesoscale model that indicated the presence of convective instability, and strong wind shear has favored deep convection during the event. Kumar *et al.* (2020) found that in addition to the low-pressure system over the Bay of Bengal and the off-shore trough over the west coast, the inflow of dry air from the Middle East region is responsible for setting up an environment that is conducive for the heavy rainfall event. Ashrit *et al.* (2020) evaluated the forecast performance of high-resolution operational NCMRWF Unified Model (NCUM) simulation of KF18 rainfall. They found that at shorter lead times deterministic model forecasts are accurate while the ensemble forecasts are useful at longer lead times. Hunt and Menon (2020) provided a climate-change perspective to the KF18 event using high-resolution models. Several studies addressed the role of reservoirs and dams on the flood associated with the heavy rainfall in 2018 (Mishra and Shah, 2018; Mishra *et al.*, 2018; Sudheer *et al.*, 2019). Investigations are performed to-(i) identify the atmospheric flow features that are important to the predictability of the two heavy rainfall events over Kerala using ESA, (ii) determine the possible similarities and differences in the dynamics of the two events in an ensemble framework, and (iii) examine the importance of remotely aligned circulations in the formation of KF18 and KF19 heavy rainfall events using the ensemble-based approach.

This chapter uses ensemble-based approach to understand the predictability and dynamics of three major heavy rainfall events over the Indian subcontinent, in general. More specifically, the application of univariate ensemble sensitivity analysis to understand the atmospheric flow features that are important to the predictability of extreme rainfall events are discussed in this chapter.

3.2 Case Overview

3.2.1 Uttarakhand rainfall 2013

The state of Uttarakhand is located at the foothills of the Himalayas in the Indian subcontinent. This region comes under the influence of monsoonal as well as large-scale extratropical circulation and hence vulnerable to the intense precipitation episodes. In the year 2013, the Indian Summer Monsoon advanced rapidly towards northern India and covered the entire country by June 16, after its onset over Kerala (Kumar and Krishnamurti, 2016). The Uttarakhand and its adjoining areas experienced a major rainstorm event from 14 to 17 June 2013 that resulted in massive destruction of life and properties due to flood and landslides associated with the rainfall. The India Meteorological Department (IMD) reported an excess anomaly in rainfall over the Uttarakhand that amounts to 847% in the rain rates over the week ending on 19 June 2013. Further, the accumulated 24-hourly rain rates depict an excess of 375% on 17 July 2013 over the Uttarakhand (Dube *et al.*, 2014). Severe convection and associated precipitation are reported a few days before the main event, which has moistened the soil along the Himalayan escarpment and made the region vulnerable to landslides (Houze *et al.*, 2017).

The synoptic conditions that led to the massive rainfall event from 14 – 17 June 2013 are associated with the presence of a mid-level trough in the westerlies and its anomalous extension towards the south. The trough further merged with a westward-moving low-pressure system that induced stronger winds in the south-westerly direction. This flow pattern has supplied sufficient moisture from the Arabian Sea and the Bay of Bengal to the Uttarakhand region (Houze *et al.*, 2017). In addition to the constant advection of moisture from the Arabian Sea and the Bay of Bengal, a stream of dry air aloft in the middle and upper troposphere made the environment over the Uttarakhand conducive for stormy weather (Krishnamurti *et al.*, 2017). A study by Ranalkar *et al.* (2016) also mentions about the favourable synoptic conditions and the role of orography in modulating the rainfall over the Uttarakhand.

3.2.2 Kerala Rainfall 2018

Kerala, the southwestern state of India, is bounded by the Western Ghats Mountain range to the east and the Arabian Sea towards the west. Around 47% of the Kerala state is occupied by the Western Ghats that extends to a height of about 2 km approximately (Fig. 3.2). The highest peak (Anaimudi Peak) of 2695 m in the Western Ghats is located in Kerala and therefore the orography of Western Ghats plays a vital role in the rainfall over the region. The state is influenced by two rainy seasons, the southwest monsoon and northeast monsoon. During southwest monsoon (July-September), the Western Ghats acts as a barrier to the moisture-laden south-westerly monsoon winds from the Arabian Sea. These strong westerly winds along with rain-bearing clouds are forced to ascend at the Ghats and while doing so heavy downpours occur on the windward side. In August 2018, Kerala

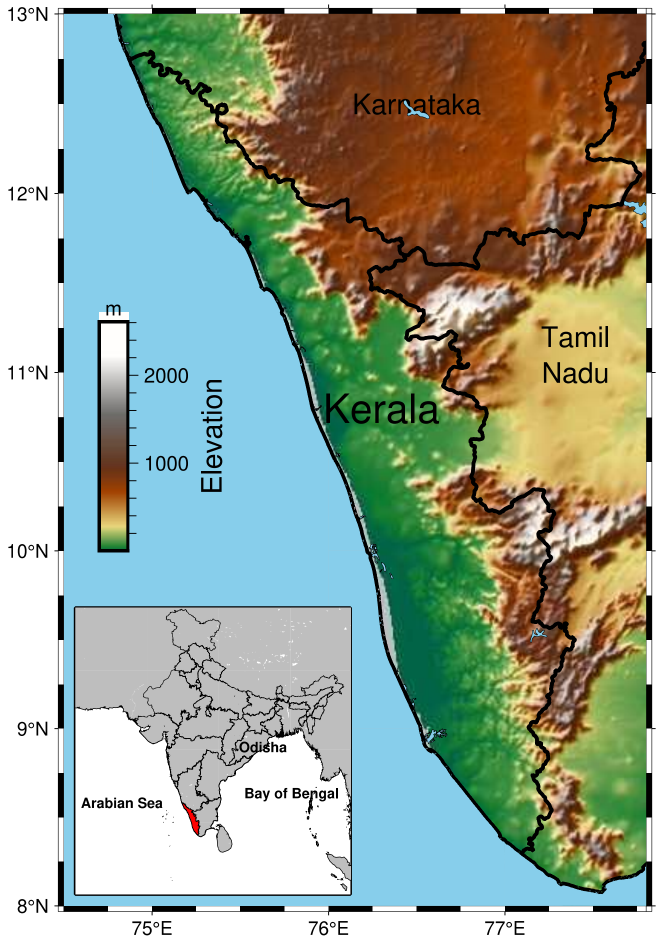

Figure 3. 2. The topographical elevation of the state Kerala. The location of Kerala is highlighted in red in the figure at the bottom left corner.

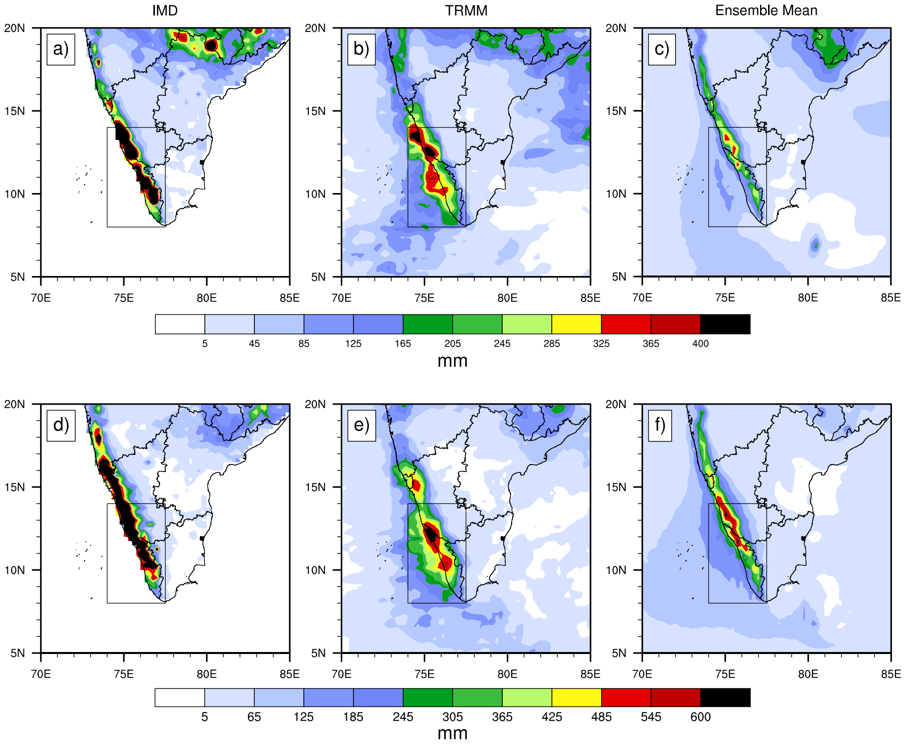

Figure 3.3. The distribution of 5-day accumulated rainfall for August 2018 (top panel) and 2019 (bottom panel). Figures a) and d) represents IMD observation, b) and e) corresponds to TRMM, and c) and f) are the ensemble mean precipitation. The inner solid box represents the area used for rainfall average.

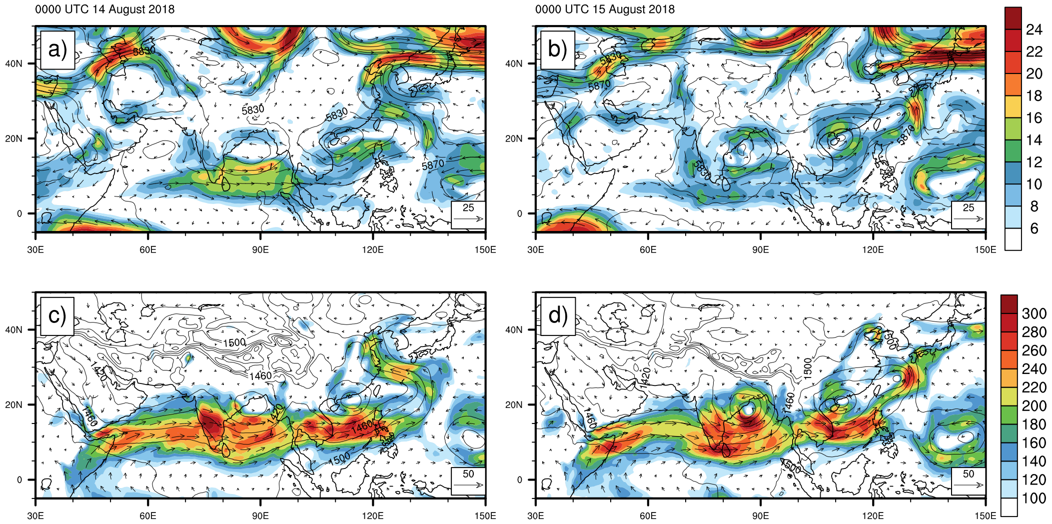

Figure 3.4. The synoptic environment of KF18 using ERA-Interim. The 500-hPa geopotential heights (contoured every 20 gpm), 500-hPa wind (shading) and wind vectors at (a) 0000 UTC 14 August 2018 and (b) 0000 UTC 15 August 2018; 850-hPa geopotential height, moisture flux ($gkg^{-1}ms^{-1}$) and wind vectors at (c) 0000 UTC 14 August 2018 and (d) 0000 UTC 15 August 2018.

received exceptionally high rainfall in two spells centered over 8 – 10 August and 15 – 17 August 2018. The second spell is more severe than the first one, and the incessant rainfall forced the authorities to open the shutters of nearly 33 reservoirs (Mohandas *et al.*, 2020). This has led to devastating floods over many parts of Kerala, and nearly 433 people lost their lives and 5.4 million people are affected (Ashrit *et al.*, 2020). The distribution of 5-day accumulated rainfall from 0000UTC 12 August to 0000UTC 17 August for IMD and TRMM observations and ensemble mean are shown in Fig. 3.3. It can be seen that the amount of rainfall is underestimated in the global operational ensemble forecasts from ECMWF, which can be attributed to the coarser resolution model forecasts.

The synoptic conditions that prevailed during this event are explained in detail in Mohandas *et al.* (2020). The study has shown that the coupled interactions of the intense low-pressure area over the Bay of Bengal and a set of cyclonic circulations over the Western North Pacific (WNP) resulted in an intense low-level wind convergence. Figure 3.4 shows the geopotential height and wind at two levels; *viz* 850 hPa and 500 hPa, during 14 August and 15 August 2018 from the ERA-Interim analysis field, which indicates the existence of a low-pressure system (LPS) near the Head Bay and cyclonic circulations over WNP. At 850 hPa a strong low-level jet is seen extending from the Arabian Sea to WNP (Fig. 3.4c-d). In addition to that, the presence of an off-shore trough can be seen that enhances the wind convergence along the west coast of Kerala.

3.2.3 Kerala Rainfall 2019

Though the impact of the KF18 event is reported to be more severe than KF19, the intensity of rainfall is much stronger in the KF19 event. Since the cumulative rainfall received is lower in the preceding monsoon months of August 2019 as compared to August 2018, the impact of KF19 is not as substantial as KF18. After the onset of ISM in the year 2019, three spells occurred over Kerala, one each during July, August and September. Among these three spells the most intense rainfall happened between the 6th and 11th of August. On 8th August 2019, the Kerala state received above 150 mm rainfall on an average causing flooding over different parts of the state. The percentage departure of daily rainfall over Kerala on 8th August is around 998% of normal (Vijaykumar *et al.*, 2021). Figure 3.5

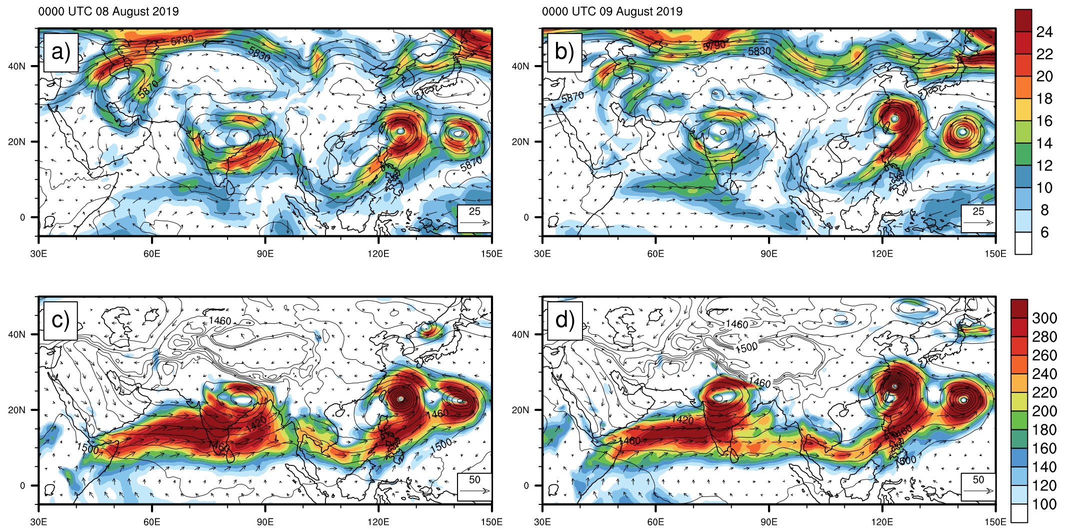

Figure 3. 5. The synoptic environment of KF19 using ERA-Interim. The 500-hPa geopotential heights (contoured every 20 gpm), 500-hPa wind (shading) and wind vectors at (a) 0000 UTC 8 August 2019 and (b) 0000 UTC 9 August 2019; 850-hPa geopotential height, moisture flux ($gkg^{-1}ms^{-1}$) and wind vectors at (c) 0000 UTC 8 August 2019 and (d) 0000 UTC 9 August 2019.

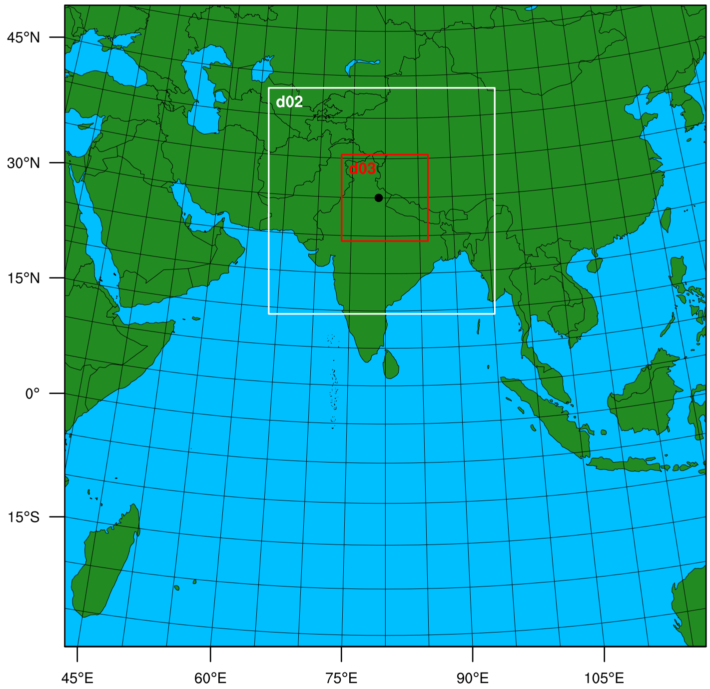

Figure 3. 6. WRF Model domain. The outer (inner) box represents the extent of the 27- (3-) km domain for the nested forecasts. The black dot represents the location of Uttarakhand.

depicts the synoptic environment of the KF19 event from the ERA-Interim reanalysis**.**

Broadly speaking, the key synoptic conditions of KF19 is similar to that in KF18 with depression over the Bay of Bengal, an offshore trough, and circulations over WNP. At 0000 UTC 8 August 2019, the depression is seen over the state Odisha (marked in Fig. 3.2) that moved west-northwestward towards central India in 24 hours. The closed contours indicate that the depression formed over the Bay of Bengal during KF19 is more intense than the KF18 depression. The circulations over WNP are more intense and can be seen located farther east in the WNP compared to the circulations of KF18. Furthermore, ERA reanalysis reveals that the intensity of LLJ is higher during KF19 as compared to that in KF18 (Fig. 3.5c, d). The perpendicular flow of this LLJ to the Western Ghats Mountain barrier enhances orographic rainfall over Kerala.

3.3 Model and Data Assimilation System

For the extreme rainfall over Uttarakhand in 2013 numerical experiments are performed using the ARW-WRF model (Skamarock WC, 2008) of version 3.8.1. The parameterization schemes used in this study are as follows: WRF-Single Moment five-class for microphysics (Hong, Dudhia and Chen, 2004), Kain-Fritsch for cumulus (Kain, 2004), Dudhia for shortwave radiation (Dudhia, 1989), Yonsei University (YSU) for the boundary layer (Hong, Noh and Dudhia, 2006), Rapid Radiative Transfer Model (RRTM) for long wave radiation (Mlawer *et al.*, 1997), and Noah as the land surface model (Chen and Dudhia, 2001). The initial and lateral boundary conditions are generated from the NCEP-GFS data available at 0.5^0 x 0.5^0 horizontal resolution.

The random perturbations for initializing 80 member ensembles are drawn from the background error covariance with the "cv3" option in the WRFVAR system (Barker *et al.*, 2012), and initial ensembles are integrated for 24 h before the first data assimilation cycle. The DART (Anderson *et al.*, 2009) EAKF (Anderson, 2001) system is employed to assimilate observations. Covariance localization and inflation are applied to maintain sufficient ensemble spread and to avoid spurious correlations. Gaspari and Cohn (1999) localization function is used to control the effect of observations with a half-width radius of approximately 950 km. Initial inflation is applied by multiplying the perturbation of 1.02 to inflate the standard deviations of ensembles using the adaptive inflation scheme of Anderson (2009). Further, to investigate the convective-scale characteristics of the Uttarakhand event, a triple-nested forecast is initialized from the analysis of the last data assimilation cycle using a two-way nested approach. The domain is configured with a horizontal grid spacing of 27 km, 9 km, and 3 km (Fig. 3.6). Each domain has 31 levels in the vertical and a model top of 50-hPa. It is to be noted that the cumulus parameterization is turned off at 3 km.

For the extreme rainfall cases over Kerala in 2018 and 2019, the ensemble forecasts are obtained from the ECMWF Observing System Research and Predictability Experiment (THORPEX) Interactive Grand Global Ensemble (TIGGE) data archive. We used global ensemble forecasts for this study, because there were synoptic conditions (extending from WNP to the Arabian Sea) involved in the extreme rainfall over Kerala which could not be studied using a regional model. This study also uses the rainfall observations provided by India Meteorological Department (IMD) (Pai *et al.*, 2014) and TRMM. The 144-h ensemble forecasts are initialized at 0000 UTC 11 August and 0000 UTC 05 August for the years 2018 and 2019, respectively.

3.4 Experimental Design

For the extreme rainfall over Uttarakhand on June 2013, the ensemble members are initialized at 0000 UTC 13 June 2013 by adding random perturbation drawn from a given distribution (Barker *et al.*, 2012). Further, a 24 h hour model forecast is launched from each ensemble member to the first analysis time at 0000 UTC 14 June 2013. Assimilation is performed from 0000 UTC 14 June 2013 to 1200 UTC 15 June 2013, at a 6 h interval on the outer 27 km domain. Subsequently, a 48 h ensemble forecast is launched from the 80 member ensemble analysis. Conventional observation from various sources such as surface synoptic, upper air radiosonde, ship, buoy, aircraft platforms, and satellite-tracked wind observations from NCEP Global Data Assimilation System (GDAS) are used for assimilation. The study uses 24-h accumulated precipitation valid at 1200 UTC of 16 and 17 June

2013 as the forecast metric. The response function is computed for each ensemble member by finding the spatial average over a box of dimension 4° × 4.5° over the Uttarakhand as shown in Fig. 3.1. In this study, the sensitivity of forecast metric with respect to geopotential heights, precipitable water, and horizontal wind components will be explored.

For the extreme rainfall events over Kerala in 2018 and 2019, ESA is used to establish a linear relationship between 5-day area-averaged precipitation and various atmospheric variables of interest at each grid point and forecast hour through ensemble statistics. The forecast metric used is the accumulated precipitation between 24 h to 144 h forecast for both KF18 and KF19, averaged in a box over Kerala as shown in Fig. 3.3. Here the ensemble members are assumed to be equally likely beyond the 24 h forecast.

3.5 Results and Discussions

3.5.1 Uttarakhand rainfall 2013

This section investigates the synoptic and convective scale features associated with the Uttarakhand heavy rainfall using the ESA approach.

3.5.1.1 Features of ensemble sensitivity in synoptic scale

The 24-h accumulated D1 precipitation forecast valid at 1200 UTC 16 June 2013 is used as the response function and the sensitivity with respect to various model variables is evaluated using Eq. 2.36. Fig. 3.7 illustrates the sensitivity of the D1 precipitation to the geopotential heights at 500hPa, which shows broad negative values over the upstream regions of storm location during the initial hours of the forecast. Further, the north-eastward progression of the sensitive region can be observed with an increase in lead time. The area of strongest negative sensitivity is found over the western region of the response function box with a maximum value of approximately 5 mm at 6-h forecast (Fig. 3.7a). It is to be noted that the magnitude of forecast response function increases (decreases) as the state variable increases over a region of positive (negative) sensitivity. This indicates that precipitation in the box is increased by 5 mm per each standard deviation fall in the geopotential height over the western region of the box. An east-west oriented weak dipole develops over the north of the domain starting from 12-h forecast, which is

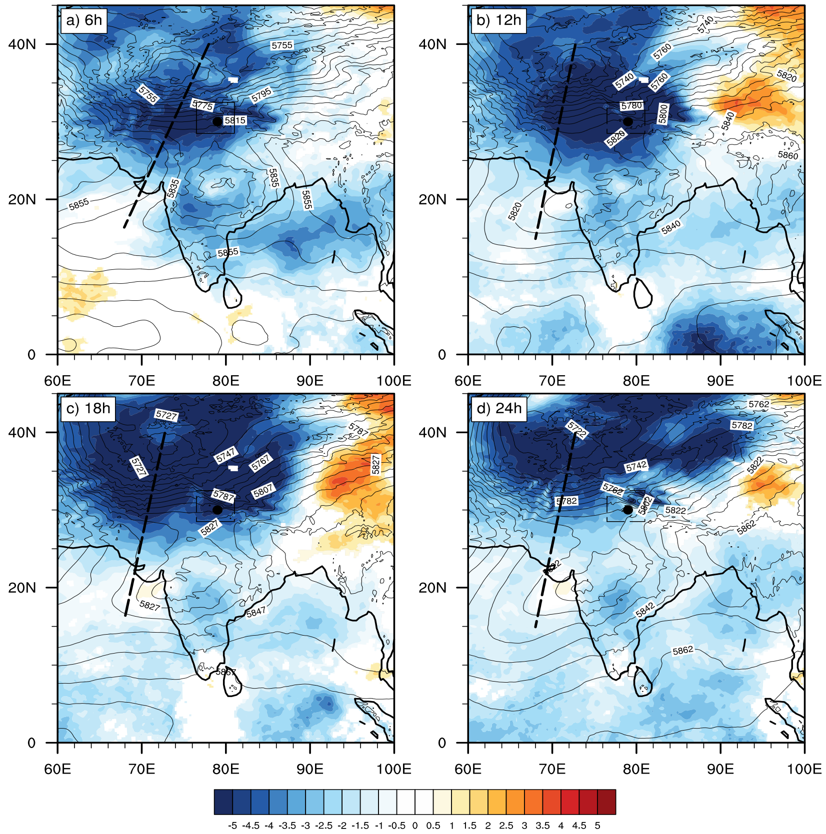

Figure 3.7. Sensitivity (shading) of area-averaged precipitation in D1 to 500hPa geopotential heights at forecast hours (a) 1800 UTC 15 June 2013, (b) 0000 UTC 16 June 2013, (c) 0600 UTC 16 June 2013 and (d) 1200 UTC 16 June 2013. Contours are geopotential heights (every 10m) from the ensemble mean. The black box represents the response function region. The location of Uttarakhand is represented by a black dot. The dashed line represents the trough axis.

indicative of the stronger geopotential height gradient over that region. The occurrence of this dipole feature can be attributed to the position of trough and ridge in the westerlies, and the observed sensitivity pattern implies that any variations in its magnitude or position would lead to variation in the response function. Although the trough is seen extended to far south of the Uttarakhand region, the location of maximum sensitivity is observed over the shortwave trough to the north of the forecast response function location, whose axis is highlighted by the dashed line in Fig. 3.7. The forecast sensitivity of precipitation to the column integrated precipitable water shows large positive values over the box (Fig. 3.8) while an east-west oriented dipole pattern emerges after 6-h forecast. The strong positive

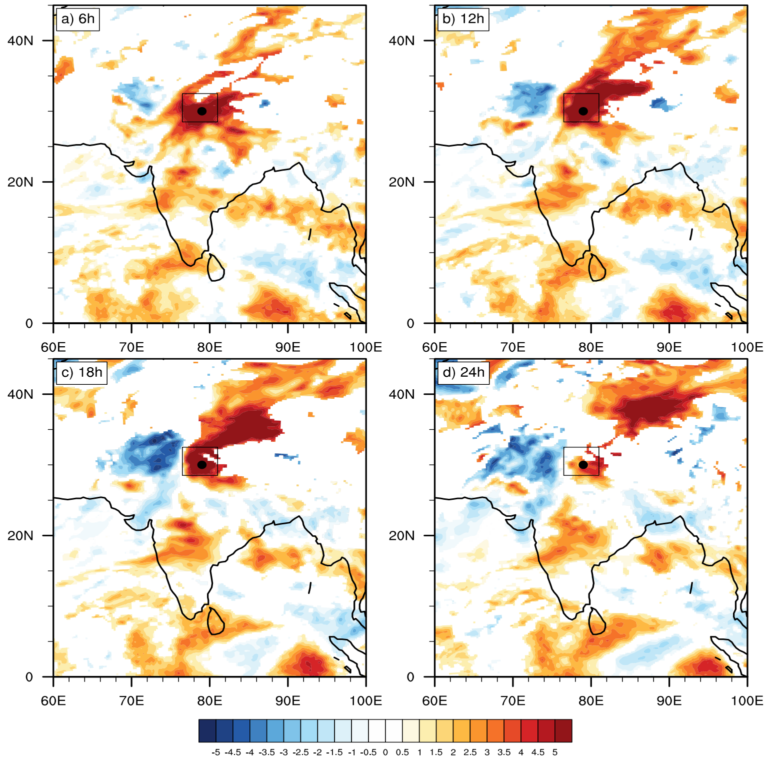

Figure 3.8. Sensitivity (shading) of area-averaged precipitation in D1 to precipitable water at forecast hours (a) 1800 UTC 15 June 2013, (b) 0000 UTC 16 June 2013, (c) 0600 UTC 16 June 2013 and (d) 1200 UTC 16 June 2013. The black box represents the response function region. The location of Uttarakhand is represented by a black dot.

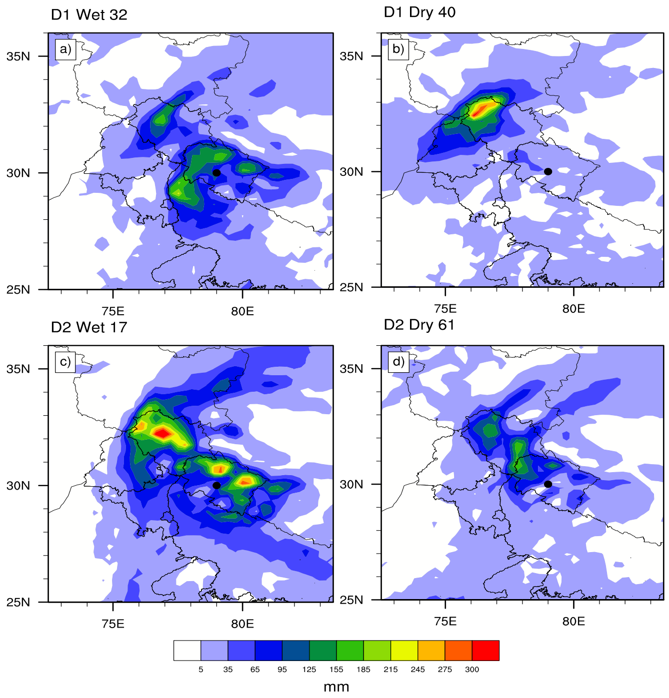

Figure 3.9. Spatial distribution of Rainfall from WET and DRY ensemble for D1 and D2. Shading (every 30mm) denotes 24 h accumulated precipitation of (a) WET and (b) DRY for D1; (c) WET and (d) DRY for D2.

sensitivity suggests that the increase in column integrated precipitable water increases the precipitation over the box while the region of negative sensitivity over the western edge of the box indicates the existence of a dry condition due to the southward descending branch of the westerlies. Joseph *et al.* (2015) confirms the existence of a southward intruding branch in the westerlies that brings cold dry air from the polar region to the Indian subcontinent during the Uttarakhand rainfall event. The positive sensitivity area over the central regions of the Indian subcontinent can be seen extending towards the Arabian Sea and the Bay of Bengal, which is indicative of the source region of moisture for the precipitation event. To understand the ESA results further, the flow structures of ensemble members with different forecast performance are evaluated. For this purpose, an ensemble member with maximum and minimum response function is selected from the set of 80 ensemble members (Bednarczyk and Ancell, 2015). The maximum and minimum response function occurred for member 32 (WET) and member 40 (DRY) with precipitation magnitude of 60 mm/day and 17 mm/day, respectively. Further, the quantitative precipitation skill scores of these two ensembles are evaluated with

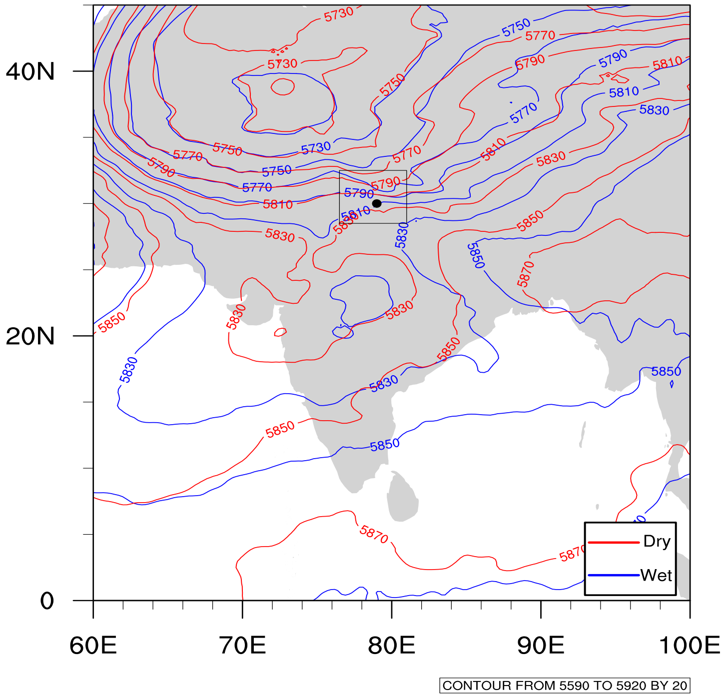

Figure 3.10. The 500-hPa geopotential height contours for WET and DRY ensembles (contoured every 20m) for 24-h forecast. The blue (red) colour represents the geopotential height of WET (DRY) ensemble member.

respect to the TRMM observations using Equitable Threat Score (ETS). The ETS value of 1 assumes the forecast skill to be perfect and a value of 0 or less represents no skill for precipitation forecast. The ETS for ensemble members 32 and 40 are 0.21 and -0.05, respectively. It is found that the precipitation maxima of DRY ensemble is located far north of the Uttarakhand region and depicts stronger precipitation as compared to WET ensemble member (Fig. 3.9). A comparison of 500 hPa geopotential height of DRY and WET ensemble members indicate that the position of the trough in DRY over the most sensitive region is shifted north of WET ensemble member (Fig. 3.10). We hypothesize that the northward shift in the

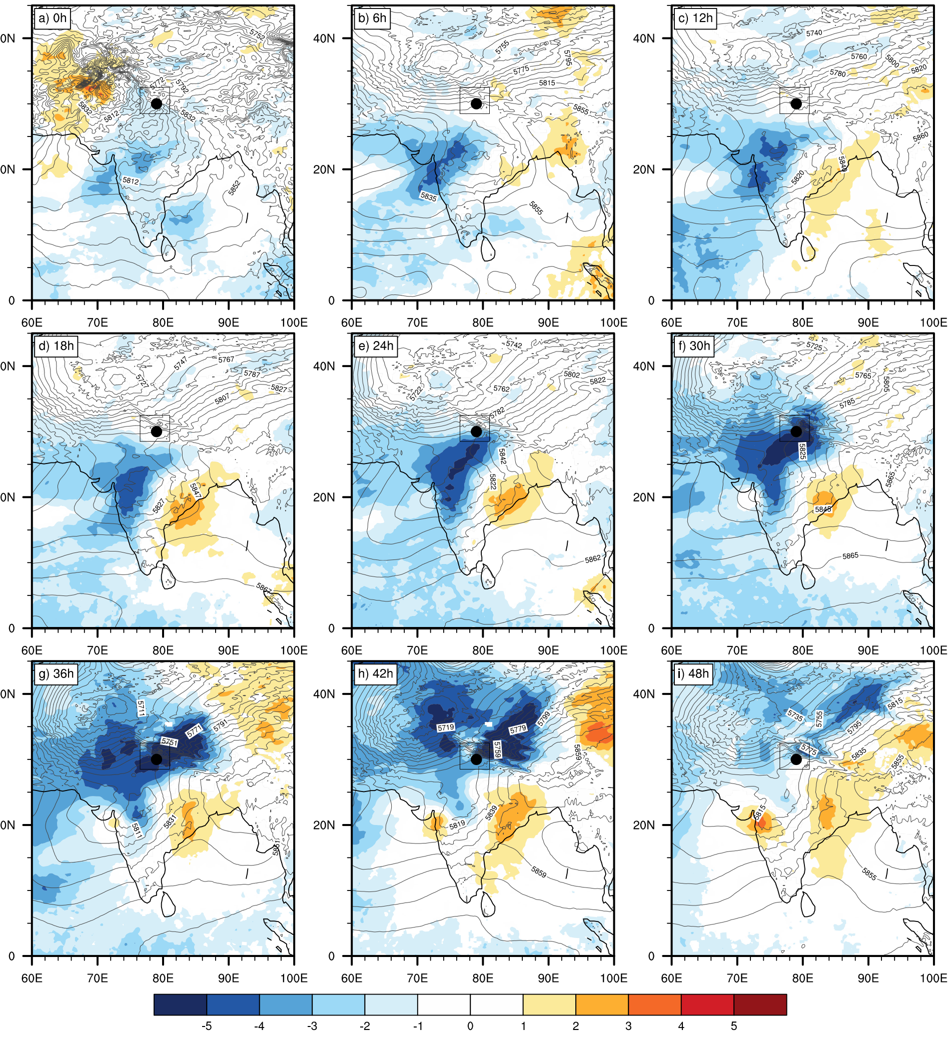

Figure 3.11. Sensitivity (shading) of area-averaged precipitation in D2 to 500hPa geopotential heights at (a) 1200 UTC 15 June 2013, (b) 1800 UTC 15 June 2013, (c) 0000 UTC 16 June 2013, (d) 0600 UTC 16 June 2013, (e) 1200 UTC 16 June 2013, (f) 1800 UTC 16 June 2013, (g) 0000 UTC 17 June 2013, (h) 0600 UTC 17 June 2013 and (i) 1200 UTC 17 June 2013. Contours are ensemble-mean geopotential heights (every 10m).

trough position over the sensitive region has resulted in the displacement of precipitation maxima to the north of the Uttarakhand region in DRY ensemble member.

Fig. 3.11 shows the ensemble sensitivity of 24-h accumulated precipitation forecast valid at 1200 UTC 17 June 2013 to 500 hPa geopotential heights at various forecast hours. In contrast to the results obtained for D1, the prominent negative sensitive regions are observed far southwest of Uttarakhand at 6-h forecast, which then progresses north-eastward during the later hours of the forecast. Further, an east-west oriented sensitivity dipole feature is observed starting from 6-h forecast (Fig. 3.11b– h). This is indicative of the sensitivity of D2 precipitation to the southward extension of the mid-latitude trough, which is shown to have intensified

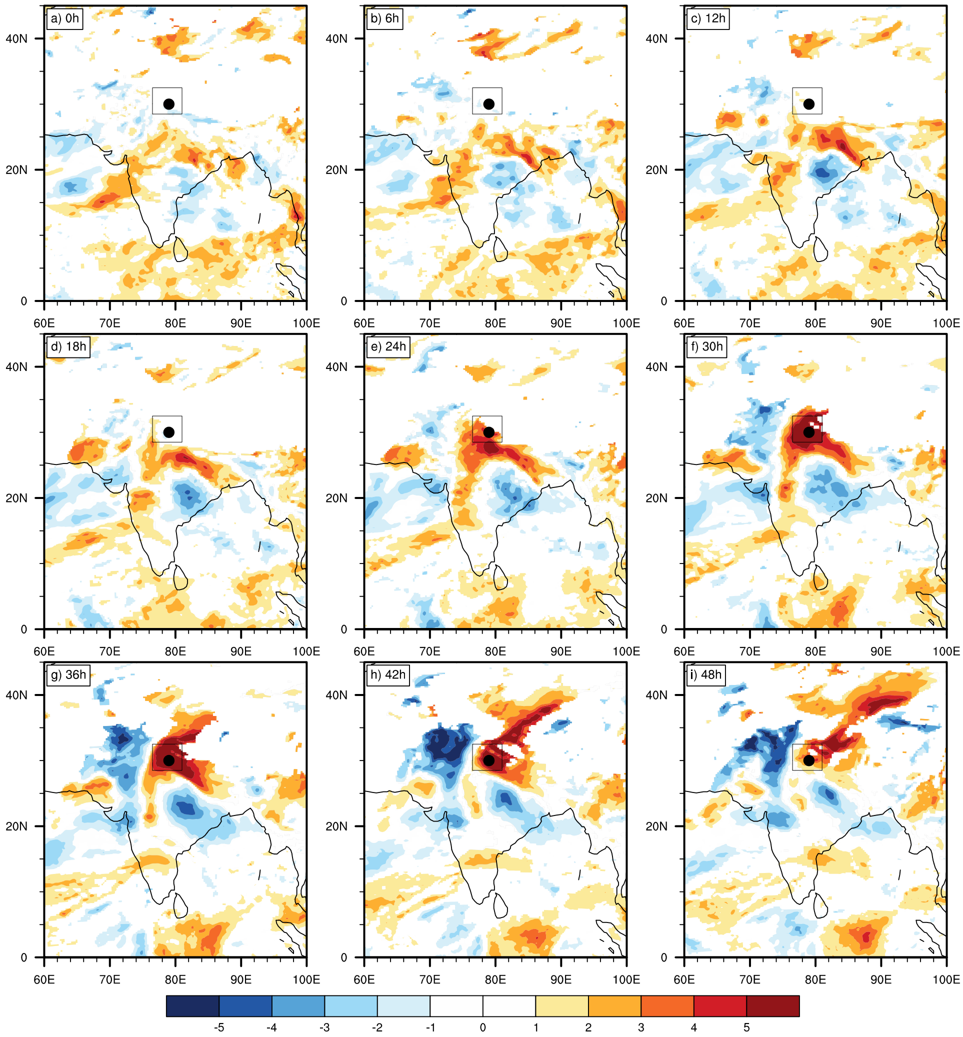

Figure 3.12. Same as in Figure 3.11, but for precipitable water

the flow of moist air towards the Uttarakhand region (Ranalkar *et al.*, 2016). The sensitivity features confirm that the D2 precipitation over the Uttarakhand is significantly influenced by the reduced heights to the south of the storm location, unlike D1 precipitation. The presence of positive sensitivity from 36 to 48-h forecast (Fig. 3.11g – i) over the far east of the Uttarakhand illustrates that enhanced ridging over the Tibetan Plateau increases the precipitation event. Fig. 3.12 illustrates the sensitivity of D2 precipitation to the column integrated precipitable water. The sensitive regions with positive values can be seen extending from the Uttarakhand region towards the Arabian Sea. Earlier studies such as Chevuturi and

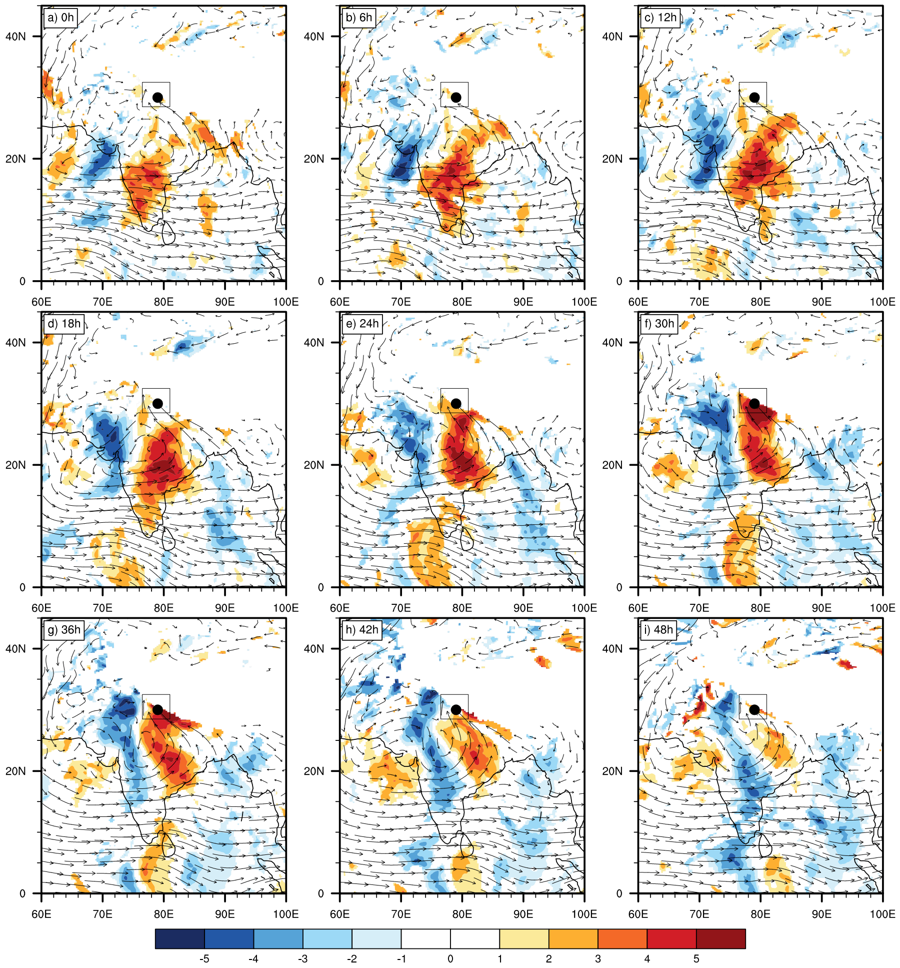

Figure 3.13. Sensitivity (shading) of area-averaged precipitation in D2 to 700 hPa meridional wind at (a) 1200 UTC 15 June 2013, (b) 1800 UTC 15 June 2013, (c) 0000 UTC 16 June 2013, (d) 0600 UTC 16 June 2013, (e) 1200 UTC 16 June 2013, (f) 1800 UTC 16 June 2013, (g) 0000 UTC 17 June 2013, (h) 0600 UTC 17 June 2013 and (i) 1200 UTC 17 June 2013. Vectors represent ensemble mean 700 hPa wind (vectors 20m/s). The black box represents the response function region. The location of Uttarakhand is represented by a black dot.

Dimri (2016) illustrated the role of stronger moisture-laden south-westerly wind from the Arabian Sea in the extreme precipitation over the Uttarakhand. Further, the sensitive region is also seen extending towards the Head Bay, which is highlighting the significance of moisture from the Bay of Bengal in D2 precipitation. Additionally, a sensitivity dipole is observed over the west of the response function box, similar to that in D1 precipitation at later forecast hours, however, with an enhanced magnitude of sensitivity. The sensitivity of D2 precipitation to 700 hPa meridional wind also shows a dipole near the forecast response function location that indicates the presence of the southward extending trough (Fig. 3.13). Further, a WET (member 17) and DRY (member 61) ensemble corresponding to maximum and minimum D2 precipitation, respectively, are selected. The area-averaged precipitation is 85 mm/day and 41 mm/day with an Equitable Threshold (ETS) of 0.2 and 0.01 for the WET and DRY ensemble, respectively. The comparison of 500-hPa geopotential height in WET and DRY ensemble member depicts that the trough in the former is extended further south of the latter in the most sensitive regions (Fig. 3.14). The result confirms that the deepening of the southward protruding shortwave trough and its northeast-

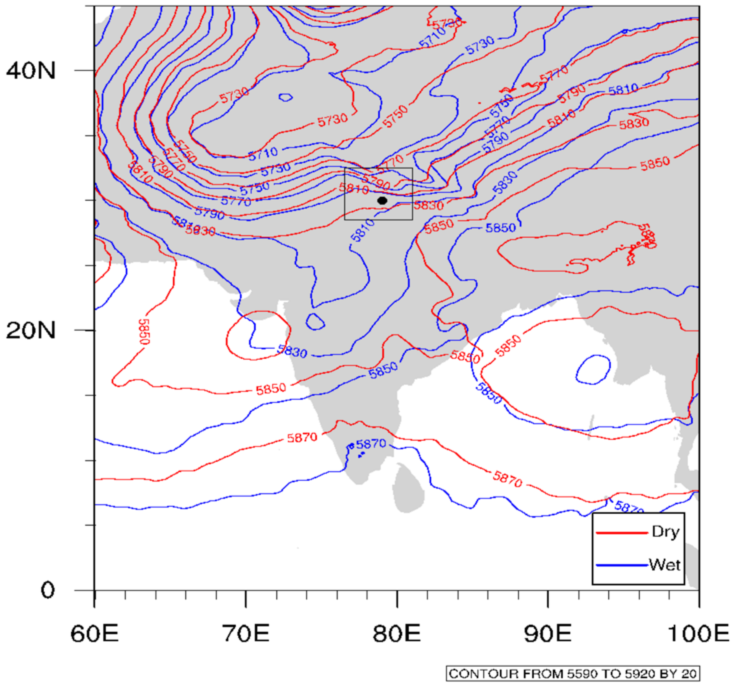

Figure 3.14. Same as in Figure 3.10, but for 42-h forecast

southwest orientation, as indicated by the sensitivity patterns, enhances the D2 precipitation over the Uttarakhand region.

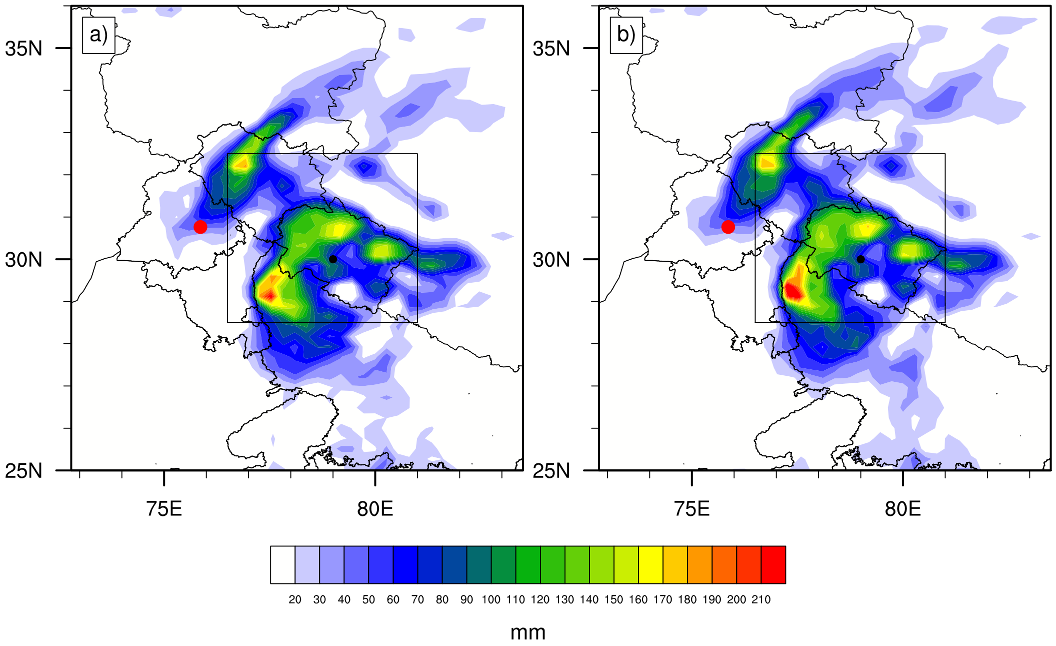

Figure 3.15. The D1 precipitation for forecast from (a) Control and (b) perturbed initial condition for perturbation amplitude $\alpha = -10m$. The red solid circle is the grid point where perturbation is applied.

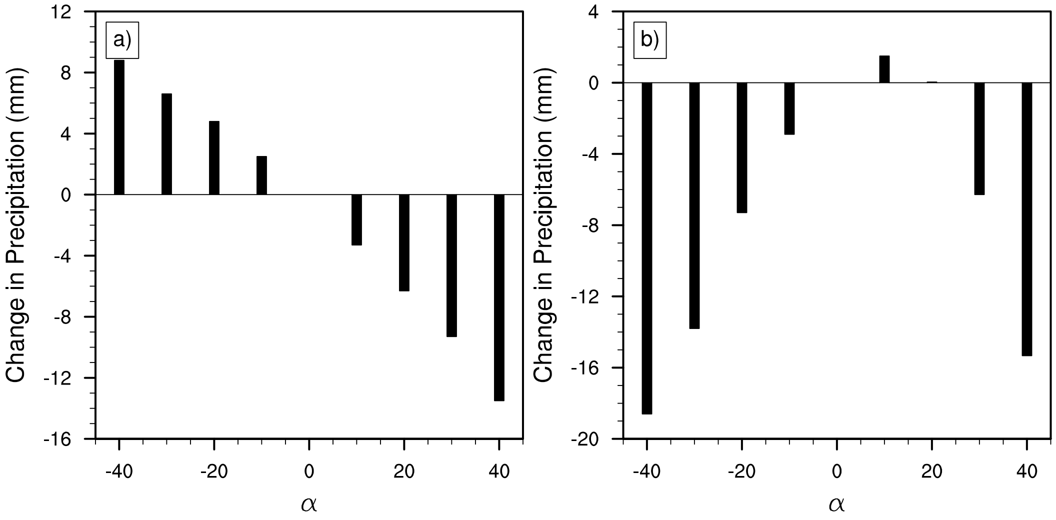

Figure 3.16. The change in precipitation in the perturbed ensemble forecast against the baseline (control) forecast for a range of perturbation amplitudes for (a) D1 and (b) D2**.**

3.5.1.2 Perturbed Initial condition experiment

Results from the previous section show that the position of the shortwave trough has a significant impact in the location and intensity of the forecasted precipitation. To test this idea quantitatively, perturbations are applied to the analysis ensembles in the most sensitive regions and integrated the model forward from the perturbed initial conditions.

The perturbation with an amplitude of $\alpha = -10m$ is applied to the grid point that lies in the most sensitive region (Fig. 3.15) of D1 precipitation. The forecast from the perturbed initial condition has shown an increase of 2.5 mm in precipitation in the response function box when the geopotential height at 500 hPa in the most sensitive region is reduced by 10 m (standard deviation). The initial condition is perturbed with a range of values of α for both D1 and D2 precipitation and the predicted change in precipitation is shown in Fig. 3.16. For D1, the increase (decrease) in the geopotential height at the sensitive grid point decreases (increases) the precipitation in the box which depicts that the magnitude of precipitation is inversely proportional to α. However, the amplitude of perturbation is not correlated linearly with D2 precipitation. For D2, the increase in precipitation is observed to the northwest of the Uttarakhand with decrease in α and hence the area-averaged precipitation inside the box shows a decrease (Fig. 3.17). This indicates that perturbation in the sensitive region may produce displacement in the precipitation location, in addition to its variations for intensity. As the forecast length increases the perturbation may grow nonlinearly and we hypothesize this as a possible reason for the observed variations in D1 and D2 results.

3.5.1.3 Features of Ensemble Sensitivity in Convective Scale

The organization and evolution of convection are expected to be well represented in convection-permitting resolution ensembles. Houze *et al.* (2017) observed that the precipitation event over the Uttarakhand is primarily stratiform, with convective cells embedded to it occasionally. To understand the finer scale thermodynamic process of the Uttarakhand precipitation event, ESA is performed on the ensembles at a horizontal resolution of 3 km. We consider the maximum composite reflectivity (MDBZ) and maximum vertical wind (MaxW) in a vertical column as the forecast

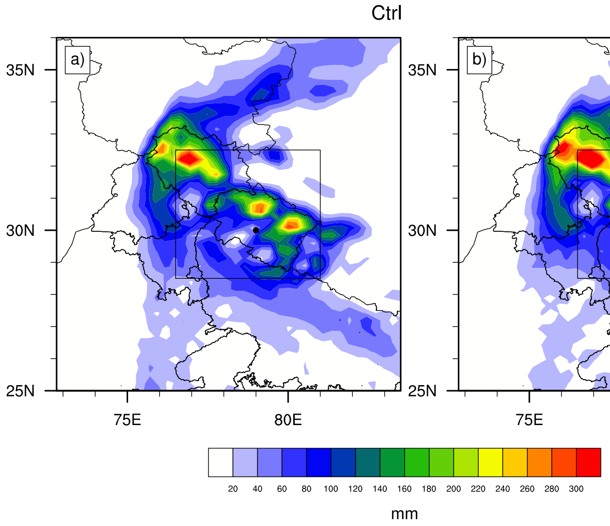

Figure 3.17. Spatial distribution of D2 precipitation forecast from (a) control and (b) perturbed ensemble. The perturbation amplitude is $\alpha = -10m$.

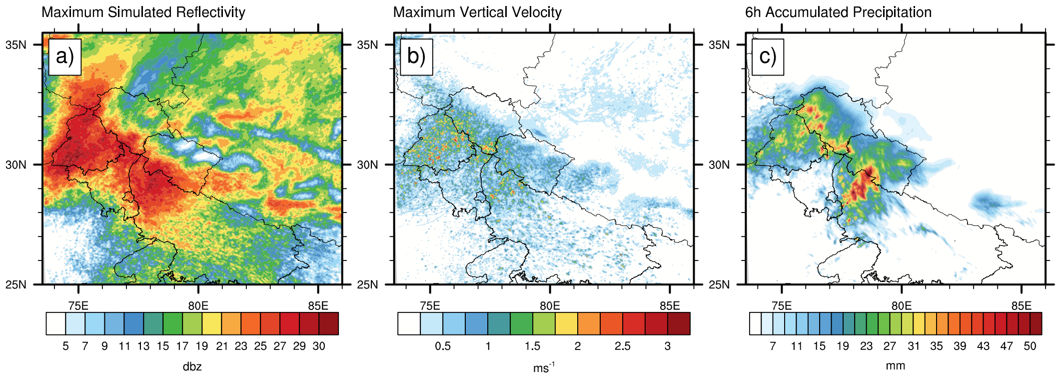

Figure 3.18. Ensemble spread of (a) MDBZ, (b) MaxW, and (c) 6-h accumulated precipitation at forecast hour 18.

metric here. As can be seen from Fig. 3.18, the spread of MaxW is maximized to the west of the Uttarakhand region, and its southeast extent is much smaller as compared to the spread of MDBZ. To understand how the sensitivity varies with the location of the response function, ESA is performed at two regions; one centred over the Uttarakhand region and the other one over the northwest of the storm location.

Fig. 3.19a-c illustrates the sensitivity of MDBZ to the temperature at 850 hPa. The reflectivity forecast is negatively sensitive to the temperature at 850 hPa at all the forecast hours, which indicates that the increase in reflectivity over the box is associated with cooler temperature.

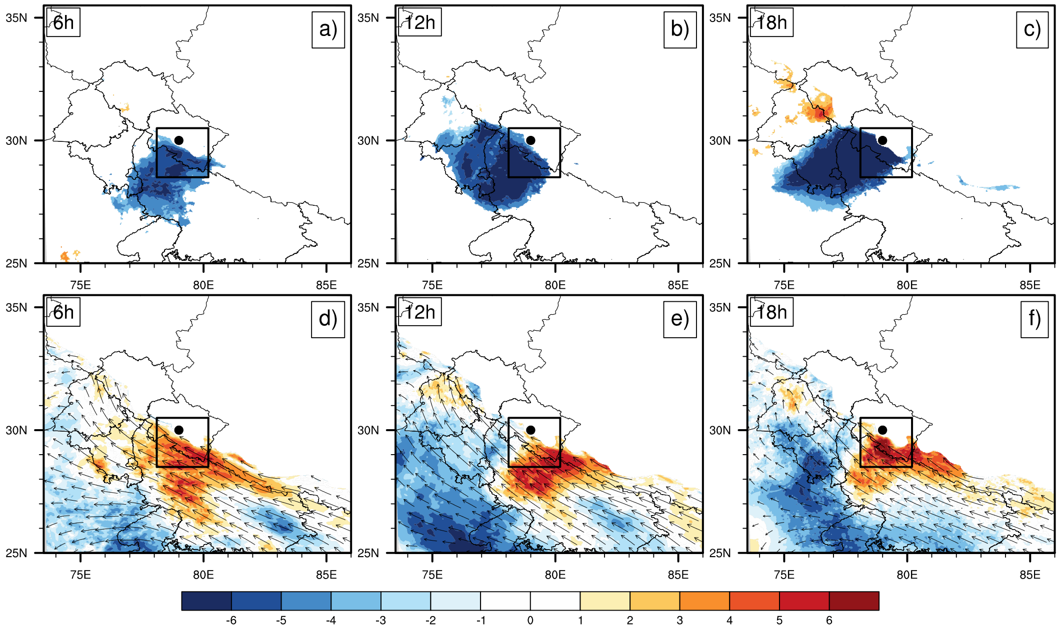

Figure 3.19. Sensitivity (shading) of MDBZ at 18-h to temperature (top) and moisture flux (bottom) at 850-hPa in the convective scale for a box over Uttarakhand at (a) 1800 UTC 15 June 2013, (b) 0000 UTC 16 June 2013 and (c) 0600 UTC 16 June 2013. The arrows indicate the ensemble mean wind vectors. The black dot represents the location of Uttarakhand.

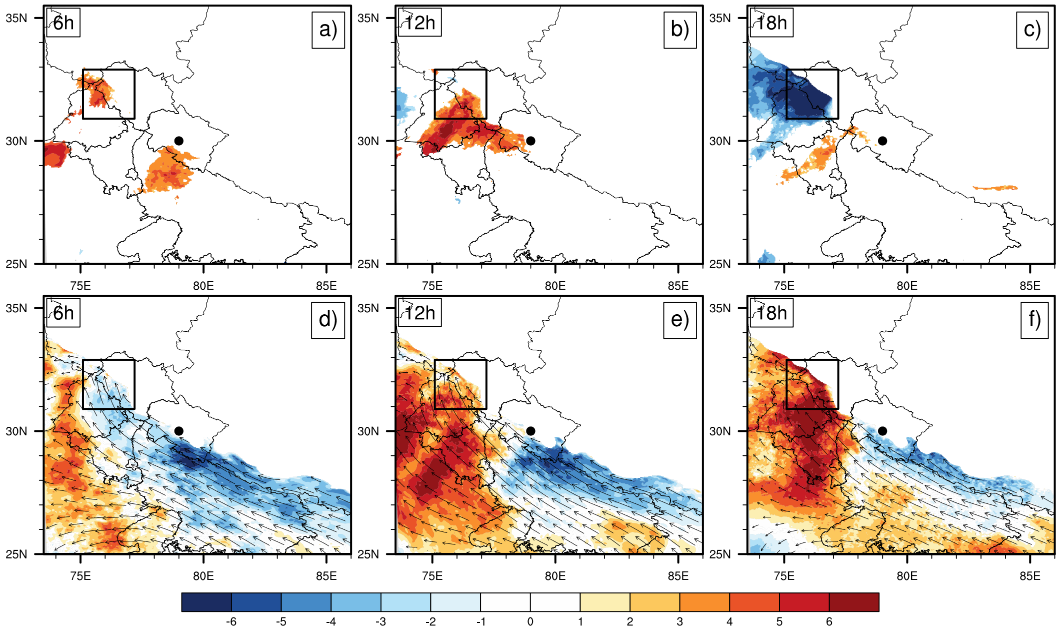

Figure 3.20. Similar to Figure 3.19 but for a box over northwest of Uttarakhand.

The convection requires warmer air to provide buoyant air parcels, and hence, it can be inferred that the chances of active convection near the location of response function are limited. However, the sensitivity of MDBZ with respect to 850-hPa moisture flux shows strikingly different features as compared to temperature sensitivity (Fig. 3.19d-f). The moisture flux indicates positive values over and near the response function location while the southeast regions of the box show negative sensitivity. So, when linked together, the negative and positive sensitivities of temperature and moisture depicts the presence of cool and moist air near the response function location. Hence, it may be inferred that the stronger bands of precipitation features observed near the Uttarakhand are less likely due to the convective processes, and the cooler temperature may be associated with a cloud shadowing effect. Cloud shadows results in alternate light and dark episodes with sudden changes occurring in the presence of low-level forced cumulus clouds (Kivalov and Fitzjarrald, 2018). However, when the forecast response function box is relocated to the north of its original position, positive sensitivity region is seen inside the box during 6-h and 12-h forecast (Fig. 3.20a-c). This indicates that stronger reflectivity is associated with warmer air temperature, which implies chances of stronger convection associated with the precipitation to the north of the Uttarakhand region.

3.5.2 Kerala rainfall 2018

3.5.2.1 Ensemble sensitivity analysis

As indicated in section 3.2.2, the mid-tropospheric conditions play a crucial role in the development of the KF18 event. Therefore, it is of interest to calculate the sensitivity between the precipitation forecasts and the mid-tropospheric features. Henceforth, ESA uses the spatially averaged 5-day accumulated rainfall as the response function over the region shown in Fig. 3.3. It is to be noted that the magnitude of the forecast response function (which is, precipitation in this case) increases as the state variable increases (decreases) over a region of positive (negative) sensitivity.

The sensitivity of response function to the 500 hPa geopotential averaged over the 5 days shows strong negative sensitivity over the Arabian Sea, which

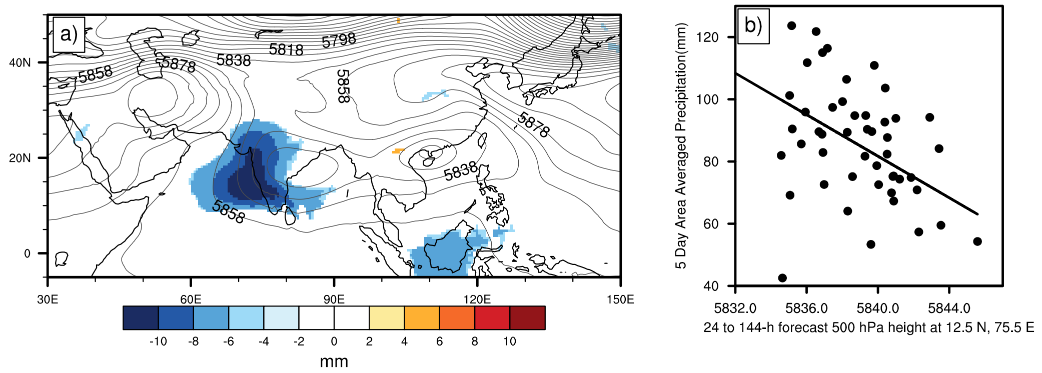

Figure 3.21. (a) Sensitivity of 5-day accumulated precipitation with time-averaged 24-h to 144-h forecast 500 hPa geopotential height (shaded) and (b) scatterplot of the 50 ensemble members (indicated by filled circles), with forecast 24-h to 144-h 500 hPa geopotential height at the point of maximum precipitation along the abscissa and forecast metric along the ordinate for KF18. The linear least-squares fit line is also shown.

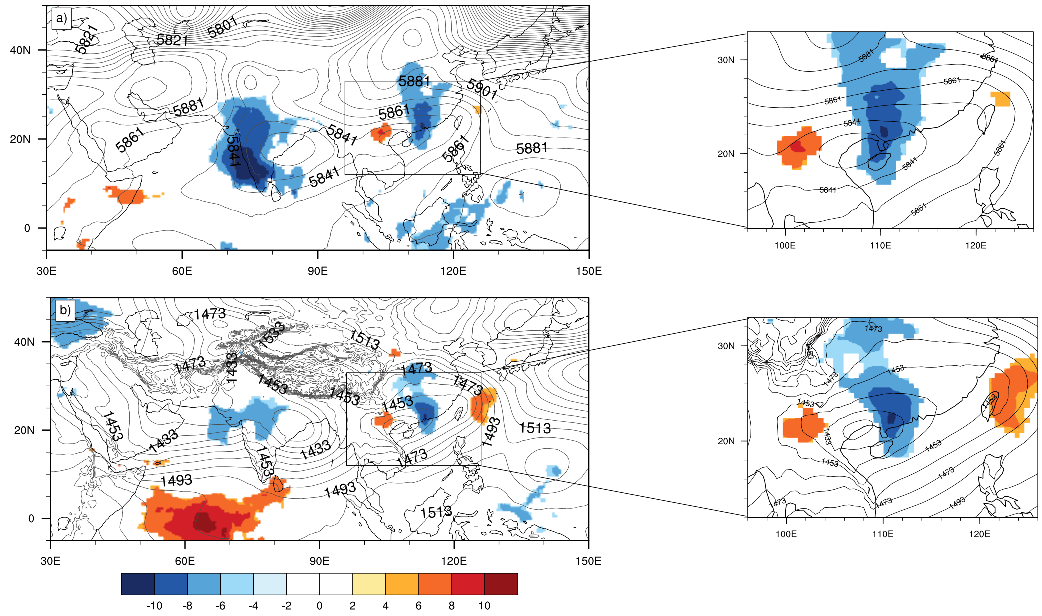

Figure 3.22. Sensitivity (shaded) of the 5-day accumulated precipitation averaged over the box in Fig. 3.3 to a) 500 hPa and b) 850 hPa geopotential heights at 0000 UTC 15 August 2018. Contours are the ensemble mean geopotential heights.

indicates that lower geopotential heights in this region increase the precipitation over the response region (Fig. 3.21a). To be more specific, the ensemble members with lower heights over the highly sensitive region produce more rainfall over Kerala. Quantitatively, for every one standard deviation fall in heights over the sensitive region, the precipitation in the box increases by 10 mm. It is evident from Fig. 3.21b that the lower the geopotential height, the higher the precipitation will be in the box. Furthermore, the sensitivity at 500 hPa level shows the presence of a dipole pattern near the circulation over the WNP at 0000 UTC August 15 (Fig. 3.22a). Bednarczyk and Ancell (2015) showed that the existence of such dipole-like patterns in ensemble sensitivity is often associated with positional shifts of the synoptic systems. The gradient of sensitivity patterns demonstrates the existence of a gradient in geopotential and stronger wind circulations over the WNP, in connection with the KF18 event. To be more precise, the ensemble sensitivity structure indicates that greater height fall needs to occur over the WNP to increase the response function, and any shift in the location of these features may affect the precipitation patterns over Kerala. We found that the dipole exhibits a westward shift in its position with respect to time, which is indicative of the westward movement of circulations over the WNP. From Fig. 3.4 it can be observed that the alignment of circulations over South China sea and Western North Pacific and the low-pressure system over the Bay of Bengal are responsible for sustaining the intense westerly flow during the KF18 rainfall period. Therefore, it is possible that the series of tropical systems over the Bay of Bengal, SCS and WNP maintained an intense zonal alignment such that the momentum transported by this chain of tropical system kept the low-level flow active for a couple of days. Such an interaction might have further intensified the transport of additional amount of moisture to Kerala, triggering orographic lifting at the Western Ghats and enhanced rainfall over the windward regions. This confirms the findings of Mohandas *et al.* (2020) that the presence of remotely aligned tropical circulations extending from WNP to Kerala has influenced the KF18 event. The sensitivity of the forecast metric to 850 hPa geopotential heights is shown in Fig. 3.22b. The negative sensitive region over the southwest coast of India at 500 hPa is absent at 850 hPa indicating the importance of mid-tropospheric conditions in KF18. To probe whether the positional shift in the sensitivity dipole has any impact on the KF18 event, ESA is applied by considering the longitudinal position of WNP circulation as the forecast metric.

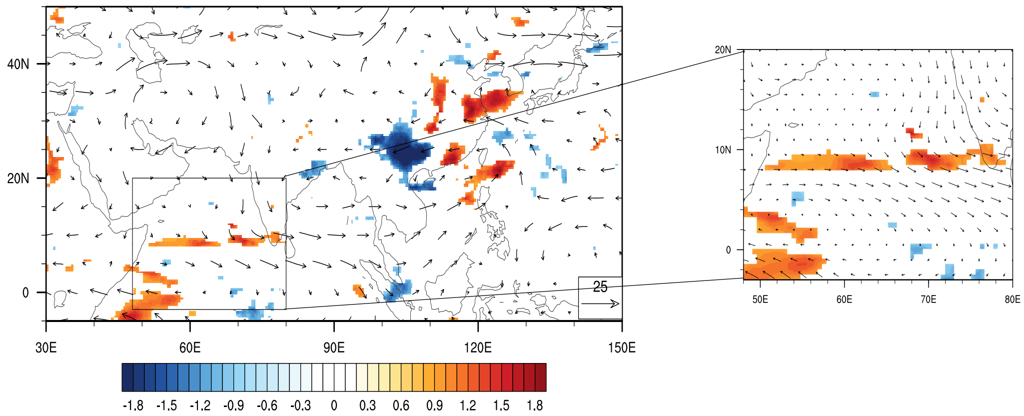

Figure 3.23. Sensitivity of longitude of the circulation over WNP to 500 hPa wind at 1200 UTC 14 August 2018. The black vectors are the ensemble mean wind vectors at the corresponding forecast hour.

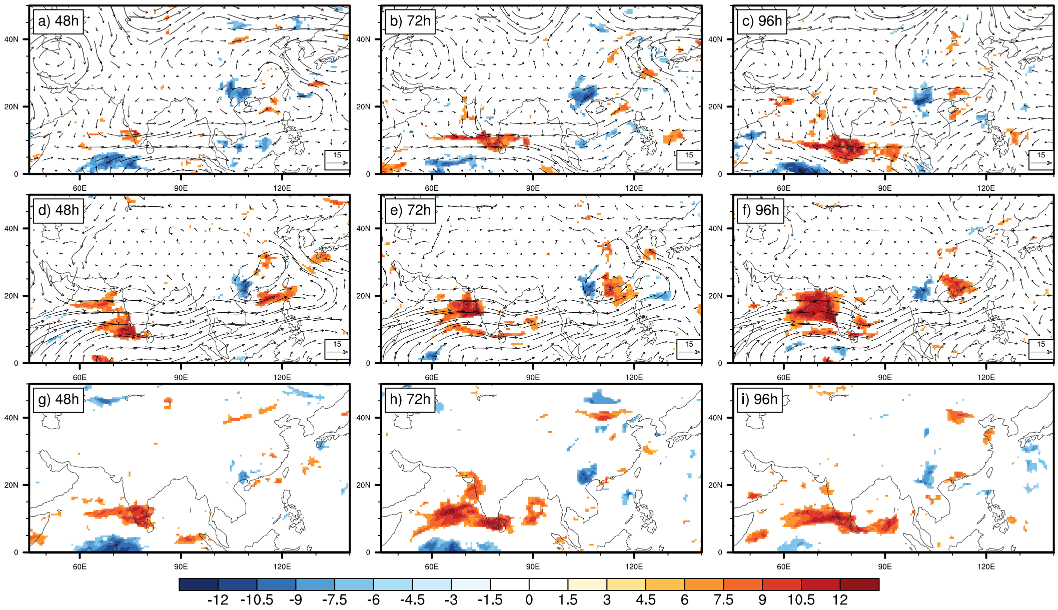

Figure 3.24. The sensitivity of the 5-day accumulated precipitation averaged over the box in Fig. 3.3 to 500 hPa moisture flux (shading) at (a) 0000 UTC 13 August 2018, (b) 0000 UTC 14 August 2018, and (c) 0000 UTC 15 August; (d)-(f) similar to (a)-(c) but for 850 hPa moisture flux and (g)-(i) for Total Colum Water (TCW). The black vectors represent the ensemble mean wind at the appropriate pressure level and time.

The longitudinal position, which is used as a proxy for the positional shift of circulations, is estimated for each ensemble member by computing the position of maximum relative vorticity at 850 hPa in a specified box (from $15°N$ to $30°N$ and $100°E$ to $120°E$) over WNP. The longitude thus obtained is then used as a forecast metric in the equation of ESA, and the sensitivity to 500 hPa wind is calculated. Figure 3.23 depicts that the position of circulation over WNP shows positive sensitivity to 500 hPa wind over the Arabian Sea and the Kerala coast. Positive sensitivity suggests that when the circulation is shifted to the east (*i.e.* as the longitude value increases), the mid-tropospheric wind over the Arabian Sea increases. As indicated before, the mid-tropospheric feature has significantly influenced the rainfall, which in turn confirms the role of WNP circulations in KF18. Further, the moisture flux over 500 hPa level also shows substantial positive sensitivity values near south-western peninsular India (Fig. 3.24a-c). Further it should be noticed that the sensitivity to 850 hPa moisture flux (Fig. 3.24 e and f) shows positive sensitivity values upstream of the response region at forecast hours 72 and 96. Looking backward at forecast hour 48 (Fig. 3.24d) positive sensitivity is noticed near Kerala coast. This indicates that the rainfall forecasts are sensitive to the low-level moisture reaching the Kerala coast at an earlier forecast hour, and if there is an increment in this feature, it will increase the amount of rainfall over Kerala. But at later hours the positive sensitivity is shifted to north of the response region possibly because of the anomalous alignment between the LPS over the Bay of Bengal and the circulations over the WNP. Additionally, the sensitivity of total column water (TCW) shows a large positive area over the southern tip of peninsular India, extending from the Arabian Sea to the Bay of Bengal (Fig. 3.24d-f). The positive sensitivity to moisture variables shows the significance of moisture supply from the Arabian Sea in the precipitation over the response region. Furthermore, the moisture gets embedded in the synoptic flow that extends to the Bay of Bengal region.

3.5.2.2 Ensemble analysis

The synoptic-scale processes that are favorable and detrimental for the event are further analyzed by classifying the ensembles into a good and bad set based on the quantitative precipitation forecast (QPF) skill, which is estimated using Equitable

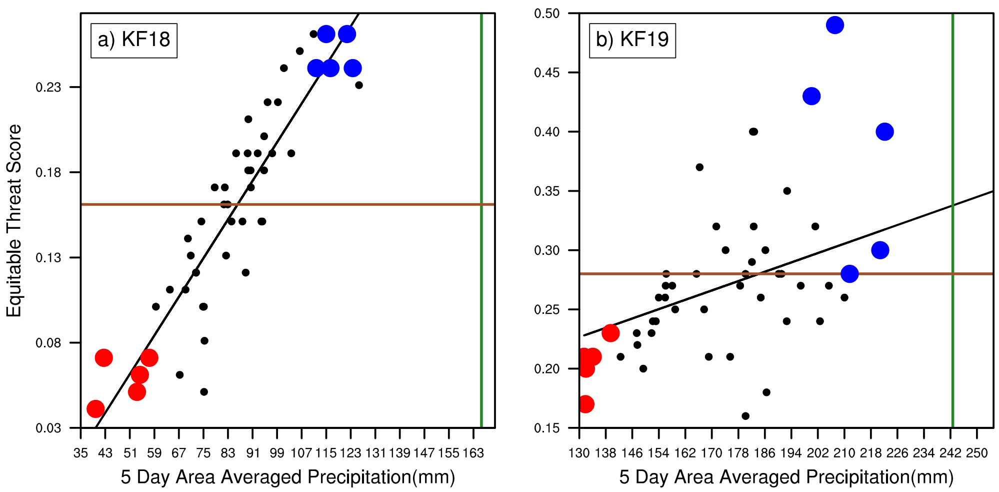

Figure 3.25. The 5-day accumulated area-averaged precipitation in (a) KF18 and (b) KF19 verses ETS scatterplot for 150 mm and 200 mm threshold, respectively. The ETS is calculated over the black box shown in Fig.3.3. GOOD (BAD) members are indicated as blue (red) filled circles and other members are represented as black filled circles. The green line represents the observed forecast metric and the brown line indicates ensemble mean ETS. The best fit line is represented by the black line.

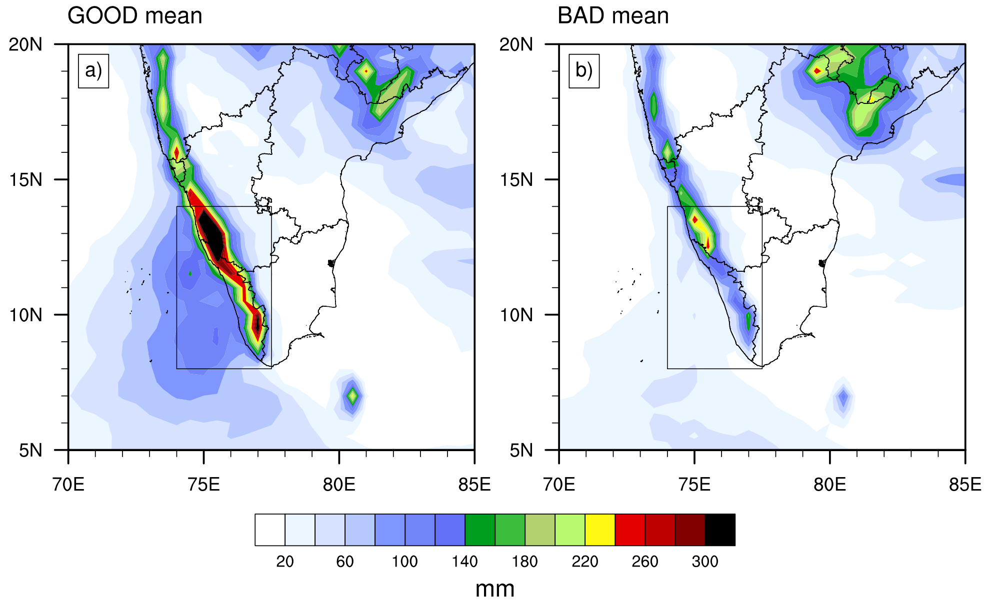

Figure 3.26. The distribution of 5-day accumulated rainfall in KF18 for (a) GOOD and (b) BAD.

Threat Score (ETS). The ETS is computed with respect to TRMM rainfall observations at a threshold of 150 mm. Figure 3.25a shows ETS estimated for individual ensemble members against the area-averaged precipitation over the response function box. Based on the ETS score and the 5-day area-averaged rainfall, the ensembles 1, 11, 31, 37, 38 are classified as GOOD, and 12, 17, 26, 35, 43 as BAD members. The composite of precipitation in GOOD ensembles shows that precipitation patterns of KF18 events extend all over the Kerala coast while in BAD members, the distribution of rainfall is not very extensive as in GOOD and the heavier precipitation is concentrated mostly towards the north (Fig. 3.26).

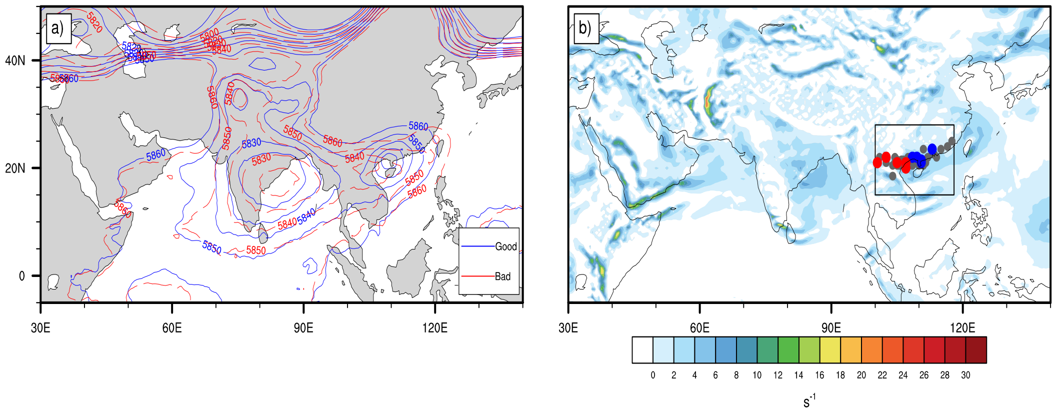

Figure 3.27. (a) Comparison between the GOOD (blue contours, every 10 gpm) and BAD (red contours, every 10 gpm) members for 500-hPa geopotential heights; (b) the maximum relative vorticity location at 850 hPa in each ensemble members at 0000 UTC 15 August 2018; blue (red) filled circles indicate good (bad) members and grey filled circles represents all the other members. Shading indicates relative vorticity at 850 hPa.

GOOD members depict a deeper trough over Kerala as compared to BAD members, which can be indicated as one of the potential reasons for the enhanced rainfall during the KF18 event (Fig. 3.27a). Another notable feature is the shift in the position of the circulation over the WNP in the GOOD and BAD members. The GOOD members show an eastward shift of WNP circulations from their mean position as compared to BAD members (Fig. 3.27b). The circulation that is shifted east might be responsible for maintaining the zonal alignment with the LPS over the Bay of Bengal. However, in doing so this feature pulled moisture from the Bay of Bengal and simultaneously amplified the monsoon circulation over the southern Indian region. It is worth noting that the results are consistent with that of ensemble sensitivity patterns obtained in the previous section. Further, it is found from Fig. 3.28a and 3.28b that the moisture flux over the mid-troposphere is substantially higher over the south of the Indian subcontinent in GOOD.

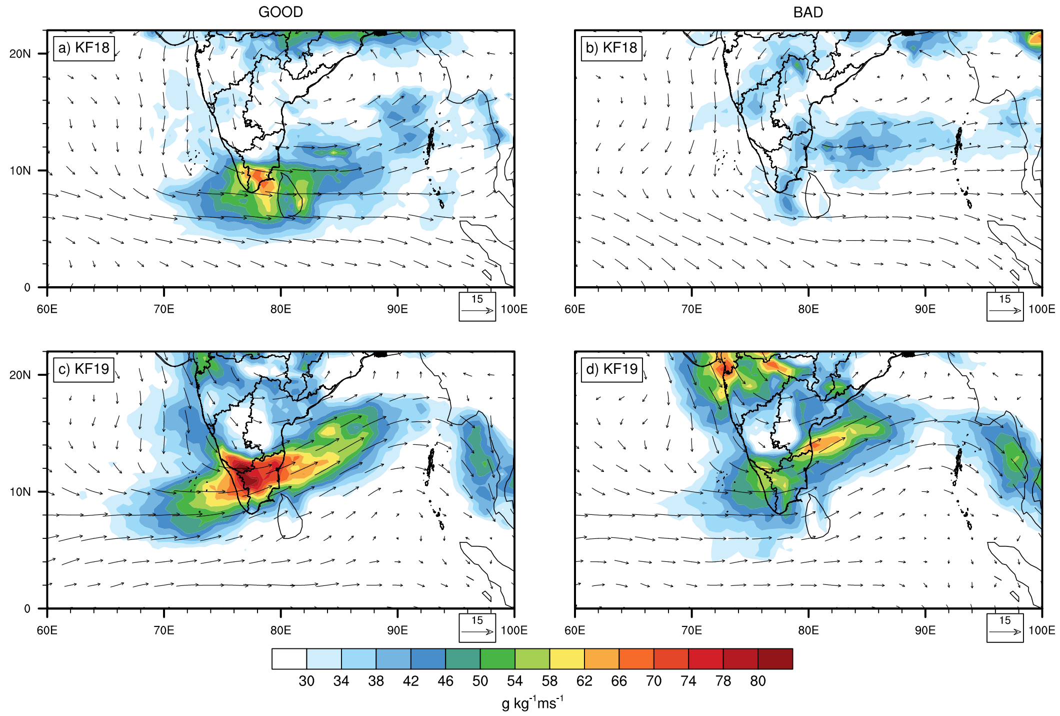

Figure 3.28. Comparison between the composite moisture flux at 500-hPa of (a), (c) GOOD and (b), (d) BAD at 0000 UTC 15 August 2018 for (a), (b) KF18 and (c), (d) KF19.

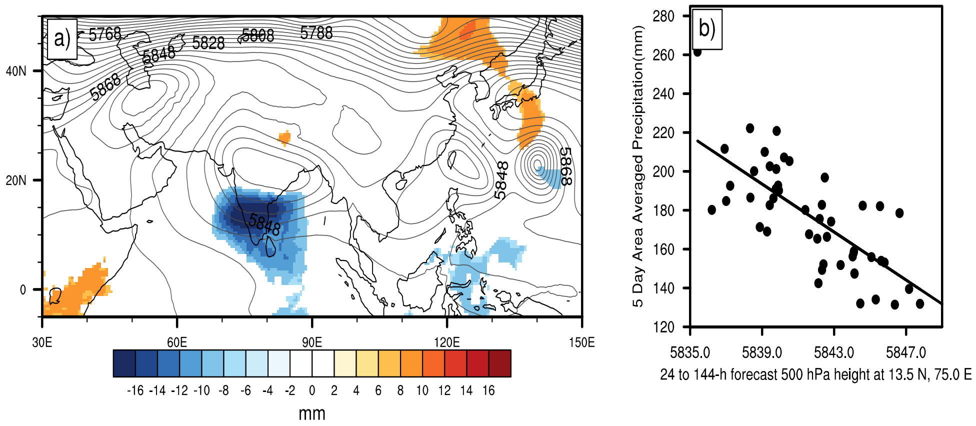

Figure 3.29. (a) The sensitivity of 5-day accumulated precipitation with time-averaged 24-h to 144-h forecast 500 hPa geopotential height (shaded) and (b) scatterplot of the 50 ensemble members (indicated by filled circles), with forecast 24-h to 144-h 500 hPa geopotential height at the point of maximum precipitation along the abscissa and forecast metric along the ordinate for KF19. The linear least-squares fit line is also shown.

3.5.3 Kerala Rainfall 2019

3.5.3.1 Ensemble sensitivity analysis

Similar to KF18, ESA is performed for KF19 events and the results are presented in this section. Figure 3.29a shows the broad regions of negative sensitivity for temporally averaged 500 hPa geopotential heights to the 5-day accumulated precipitation forecast in the response region. As compared to the KF18 event, the sensitive region in the KF19 event is seen concentrated over peninsular India. Moreover, the maximum sensitivity values indicate that for every one standard deviation fall in geopotential height, the response function increases by 16 mm, which is about 60% stronger than the KF18 event. One of the potential reasons for enhanced precipitation over the response region can be attributed to the higher sensitivity of precipitation to the midlevel trough and to the greater geographical proximity. Figure 3.29b depicts the relationship between temporally averaged 500 hPa heights at the maximum precipitation location $13.5°N$ and $75°E$, and the 5-day area-averaged precipitation. It is evident from the figure that the lower geopotential height enhances the response function over the box. The absence of sensitivity dipole suggests that the circulations over WNP may not have a significant impact on KF19 (Fig. 3.30).

The sensitivity to 500 hPa moisture flux shows a significant positive sensitive region over the Arabian Sea and peninsular India at 0000 UTC 9 August 2019 (Fig. 3.31c), which depicts the significance of mid-tropospheric moisture flux in KF19. However, Figure 3.31f indicates that rather than the 500 hPa moisture flux the forecast metric is sensitive to the low-level (850 hPa) moisture flux. The LLJ play a major role in the summer monsoon rainfall over the Indian subcontinent. The low-level westerly winds impinging the Western Ghats increase water vapor transportation and low-level convergence, which promotes orographic precipitation over Kerala. Therefore, any change in this low-level flow influences the rainfall over the Western Ghats. The ensemble sensitivity analysis shows that for one standard deviation change in the 850 hPa moisture flux the predicted change in the forecast metric is more than 14 mm. Such an understanding on change in low-level westerly moisture flux convergence is essential for predicting deep convection and the torrential rain over Kerala. The larger spatial extent of positive sensitivity to

TCW over the western coast of peninsular India indicates that the supply of moisture from the Arabian Sea is important for sustaining the precipitation for KF19 (Fig. 3.31f). Moreover, the magnitude of positive sensitivity for TCW is higher in KF19 than in KF18.

3.5.3.2 Ensemble analysis

The ensemble forecast of the KF19 event is examined by comparing the forecast performances of different ensemble members. Similar to the KF18 event, the ensemble members are classified as GOOD (14, 15, 30, 34, and 44 ensemble members) and BAD (2, 13, 20, 42, and 49) set based on the values of ETS and forecast metric. The analysis of 5-day area-averaged precipitation reveals that most of the ensemble members underpredicted the rainfall (Fig. 3.25b). The shallow slope of the regression line shows that the correlation between 5-day area-averaged precipitation and ETS scores is not strong. Similar to August 2018, the ensemble members are characterized by differences in geopotential heights, wind, and moisture. The comparison of 500 hPa geopotential heights between the GOOD and BAD members shows that the heights in GOOD are deeper than that in BAD. In contrast to KF18, the circulations over WNP do not show any significant difference in the position or intensity among the GOOD and BAD ensembles of KF19. From Fig. 3.28c, d it is clear that the forecasts of GOOD ensembles depict enhanced mid-tropospheric moisture flux compared to BAD.

3.6 Summary

The present chapter focuses on the catastrophic events that happened over Kerala in the years 2018 and 2019, and the torrential flood episode over Uttarakhand in 2013. This study uses univariate ensemble sensitivity analysis to understand the atmospheric flow features that are important to the predictability of these three heavy rainfall events that happened over the Indian subcontinent.

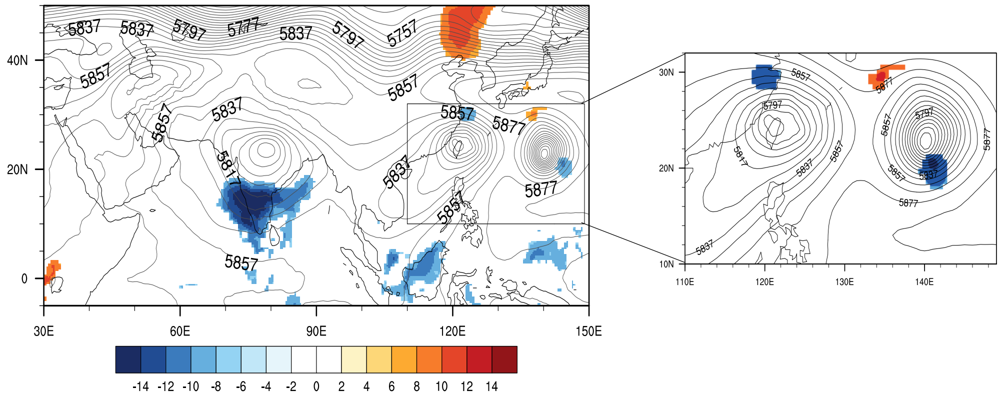

Figure 3.30. The sensitivity of the 5-day accumulated precipitation averaged over the box in Fig. 1 to the 500 hPa geopotential heights (shading, every 2 mm) at 0000 UTC 9 August 2019. The black contours (every 10 m) are the ensemble mean geopotential heights at the appropriate time.

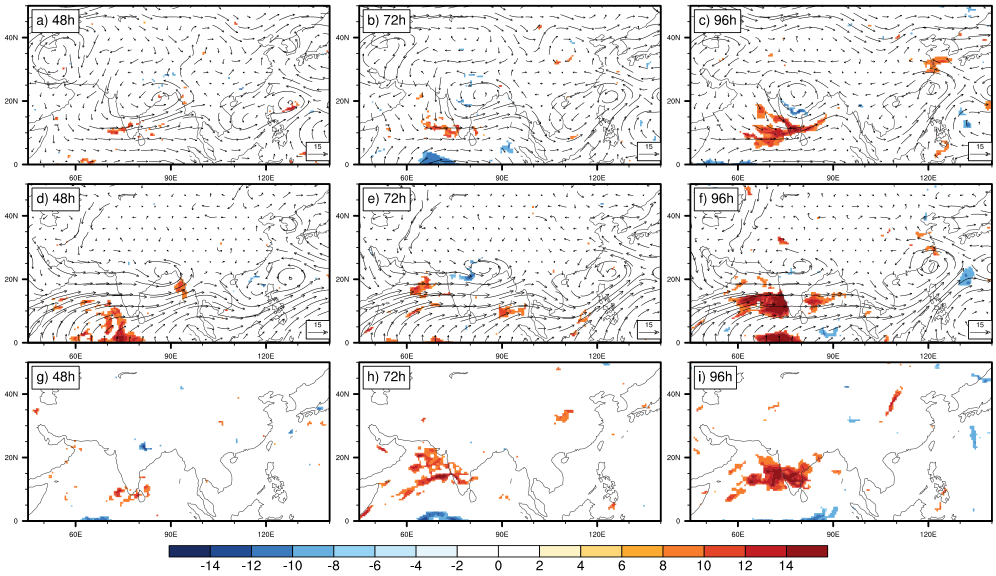

Figure 3.31. The sensitivity of the 5-day accumulated precipitation averaged over the box in Fig. 1 to 500 hPa moisture flux (shading) at (a) 0000 UTC 7 August 2019, (b) 0000 UTC 8 August 2019 and (c) 0000 UTC 9 August 2019; (d)-(f) similar to (a)-(c) but for 850 hPa moisture flux and (g)-(i) for Total Colum Water (TCW). The black vectors represent the ensemble mean wind at the appropriate pressure level and time.

For the extreme rainfall over Uttarakhand during 14-17 June 2013, the ensemble forecasts initialized from an EnKF DA system using the WRF model is used to understand the dynamics and predictability. ESA is employed in synoptic and convective scale ensembles to identify the multi-scale aspect of the weather event. Further, predictability aspects of the heavy rainfall event are explored by applying perturbations to the analysis ensembles in the most sensitive regions and analysing its impact on the forecast metric.

The precipitation forecast is sensitive to the mid-tropospheric trough and moisture fields for both days, in general. The D1 precipitation shows negative sensitivity to the geopotential height at 500 hPa over the upstream regions of the storm location, which then progresses eastward with time. The midlatitude westerly trough is seen extended to far south of the storm location, however, the maximum sensitivity is observed north of the Uttarakhand region for D1 precipitation. This shows that the southward intrusion of the trough has not significantly influenced the D1 precipitation. Further analysis revealed that the positional shift of the trough over the sensitive region has resulted in the displacement of precipitation bands. In contrast to the results obtained for D1, sensitive region to D2 precipitation is observed over the southward intruded branch of the trough, which then progresses north-eastward along with the trough. A sensitivity dipole to geopotential height is observed for both D1 and D2 precipitation. However, the location of the dipole varies significantly in both days. The sensitivity of total precipitable water to the D2 precipitation is seen extending towards Head Bay, which highlights the significance of moisture from the Bay of Bengal in the D2 precipitation. Perturbed initial condition experiments are performed to quantitatively ascertain the initial condition sensitivity to the precipitation episode. On reducing the geopotential height at 500 hPa by 10 m at the most sensitive grid point, the D1 precipitation in the response region increases by 2.5 mm and thus difference in forecast metric between the control and predicted ensemble members compare well with sensitivity values (5 mm). These results show that the initial condition perturbations in the maximum sensitive region can have a large impact on the intensity and location of precipitation forecast. Perturbations introduced in the initial conditions (IC) over the most sensitive region over the west of the storm location indicate significant variations in the forecast location of precipitation. IC perturbed experiments show that the perturbation amplitude is correlated linearly with predicted change in precipitation, which becomes nonlinear as the forecast length increases. To determine the finer scale thermodynamic process of the Uttarakhand precipitation event, ESA is applied on the convection-permitting ensembles. The sensitivity of MDBZ forecast to temperature reveals that D1 precipitation is mostly non-convective. However, when the location of the response function is moved north-westward, the sensitivity patterns show signs of convective precipitation.

Next univariate ensemble sensitivity analysis is also applied to understand the atmospheric flow features that are important to the predictability of two heavy rainfall events formed over the southern peninsular region of the Indian subcontinent. The ensembles required for ESA are obtained from the ECMWF TIGGE data archive. One of the important characteristic features of the Indian summer monsoon is the strong low-level westerly winds over the central Arabian Sea and the peninsular India. The rainfall activity over Indian subcontinent is strongly dependent on the moisture supplied to this region by the low-level winds.

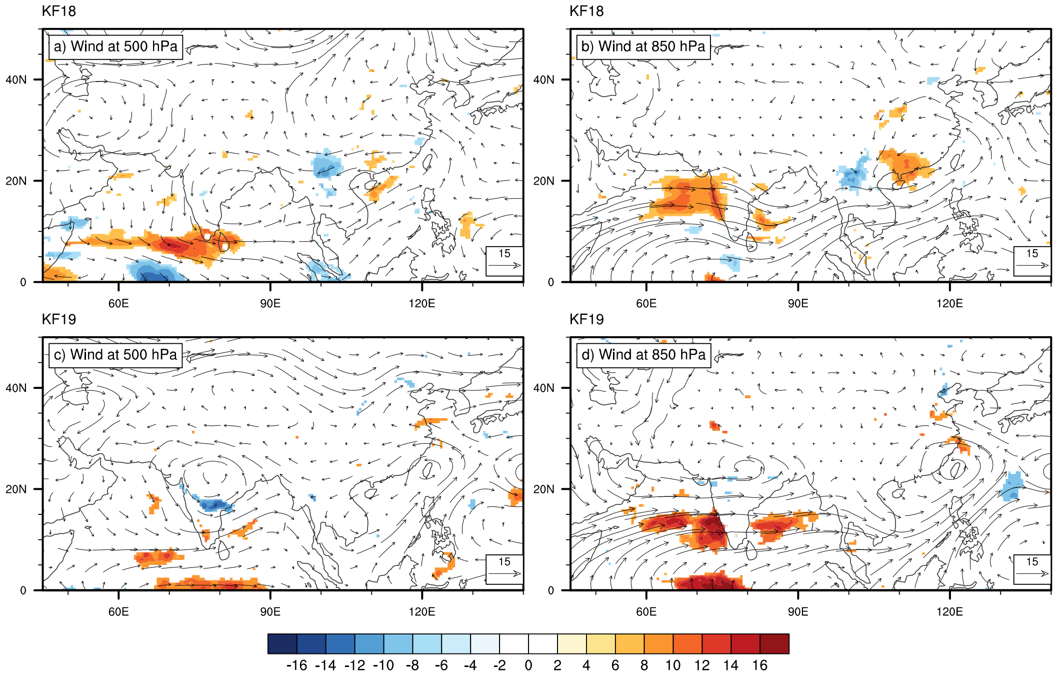

Figure 3.32. The sensitivity (shading) of 5day area-averaged accumulated precipitation to (a) 500 hPa wind and (b) 850 hPa wind for KF18; (c) and (d) similar to (a) and (b) but for KF19. The vectors represent ensemble mean wind.

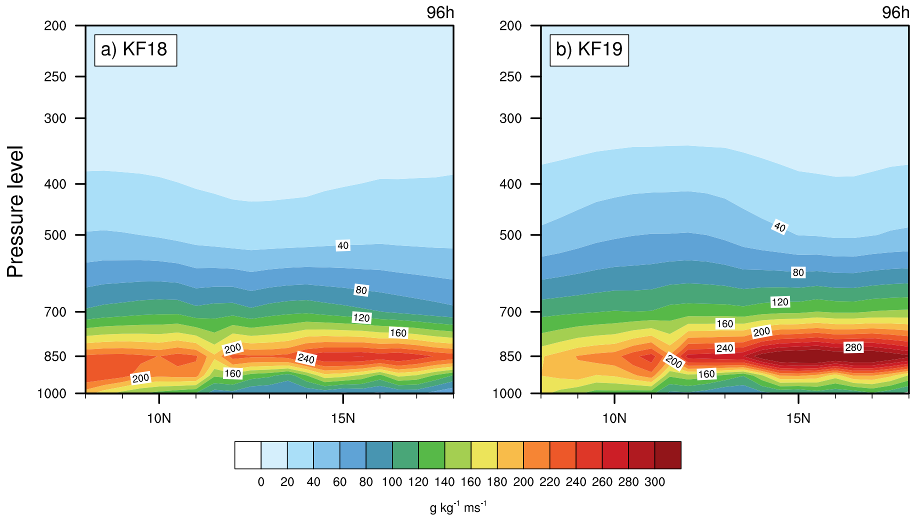

Figure 3.33. Vertical distribution of moisture flux averaged between the longitude (74.5°E – 77.0°E) at (a) 0000 UTC 15 August 2018 and (b) 0000 UTC 9 August 2019.

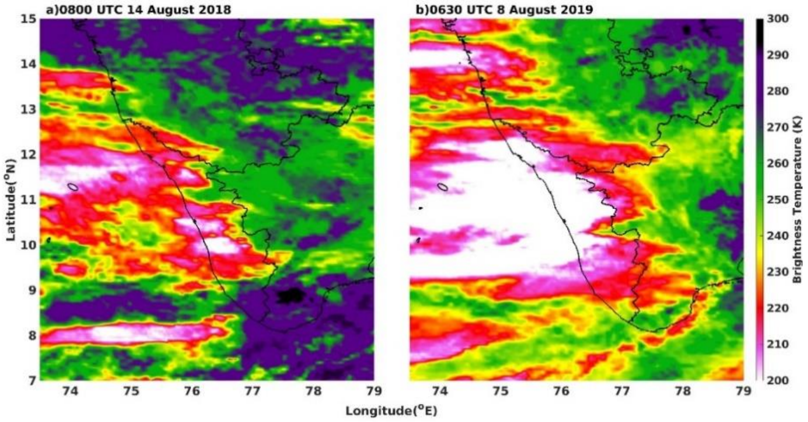

Figure 3.34. INSAT-3D satellite imagery for brightness temperature at 0800 UTC 14 August 2018 and 0630 UTC 8 August 2019.

Ultimately, any change in the low-level moisture flux will have a direct impact on the rainfall over the peninsular regions of India. The sensitivity patterns shown in Figs. 3.32b and 3.32d clearly indicates the importance of low-level moisture in the extreme rainfall events over Kerala on 2018 and 2019. Quantitatively, a one unit change in the low-level moisture flux might increase the KF18 and KF19 rainfall by 10mm and 16mm, respectively. However, it is understood that the mid-tropospheric flow has a significant role in the formation of KF18 than the KF19 event, which is evident from the prominent positive sensitivity over the Arabian Sea (Fig. 3.32a). On the other hand, flow conditions at the lower levels of the atmosphere are more important for the KF19 event, and the influx of moisture at 850 hPa is stronger in KF19 than in KF18 (Fig. 3.33). The synoptic-scale flow patterns at lower levels in the KF19 event are intense and persistent toward the Western Ghats for a sustained period. Analysis of the INSAT-3D (Indian National Satellite System) imagery during KF18 and KF19 events depicts that the KF19 event is dominated by deep convective clouds (Fig. 3.34).

The spatial extent of negative sensitivity over the Arabian Sea and Peninsular India at 500 hPa level for KF18 and KF19 events, indicates the significance of lower geopotential heights to the precipitation over the response region. To be more specific, the ensemble members with lower heights over the highly sensitive region produce more rainfall over Kerala. Sensitivity analysis indicates that greater height fall needs to occur over the WNP to increase the KF18 precipitation, and any shift in the location of these features may affect the precipitation patterns over Kerala. Additionally, the results indicate that the circulations positioned farther east of its mean position are related to stronger precipitation over the response function region. To be more specific, the LPS over the Bay of Bengal and the circulations over the WNP maintained an alignment such that the momentum transported by this system triggered additional supply of moisture towards Kerala coast. Any shift in this alignment might produce an impact on the direction of westerly winds towards Kerala. As can be seen from Figs. 3.28a, b the direction of mid-tropospheric winds is normal over Kerala in GOOD ensembles whose circulations over WNP show a shift towards east. While the alignment of circulations over WNP and LPS over the Bay of Bengal influences the

KF18, the absence of sensitivity dipole suggests that the circulations over WNP have not impacted the KF19 event. The mid-tropospheric flow over the response region has shown stronger positive sensitivity for the KF18 event while the flow conditions over the lower levels depict higher sensitivity values for the KF19 event. The moisture-laden low-level flow is more substantial in the case of KF19, which has favored the development of deep convective clouds. Thus, it is found that nevertheless, the synoptic conditions of KF18 and KF19 look similar, the underlying dynamics governing the extremely heavy rainfall events are different.

CHAPTER 4

MULTIVARIATE ENSEMBLE SENSITIVITY ANALYSIS APPLIED TO AN EXTREME RAINFALL OVER THE INDIAN SUBCONTINENT

4.1 Introduction

The ensemble sensitivity analysis proposed by Torn and Hakim (2008) accounts for the contribution of diagonal elements of the covariance matrix. The sensitivity analysis obtained by regressing the forecast variable independently on each analysis variable is known as univariate ensemble sensitivity (Torn and Hakim, 2008a). In the previous chapter, we have used the univariate ensemble sensitivity analysis to understand the dynamics of three extreme rainfall events over the Indian subcontinent. However, the univariate ensemble sensitivity may overestimate the forecast responses due to sampling errors.

In this chapter, a multivariate ensemble sensitivity is introduced that accounts for the collective contributions from all state variables across all nearby grid points simultaneously that is expected to ameliorate the sampling error in the analysis system (Hacker and Lei, 2015). Further, Hacker and Lei (2015) used a two-scale Lorenz model to investigate the impact of diagonal approximation and possible ways to reduce the sampling error. They found that in the presence of fast scales (large-scale synoptic activity), model errors, and fewer observations, multivariate ensemble sensitivity has superior skill in predicting forecast responses compared to univariate ensemble sensitivity. Limpert and Houston (2018) demonstrated storm-scale ensemble sensitivity analysis and the application of multivariate regression to a supercell thunderstorm for targeting observations. Ren et al. (2019) applied multivariate ensemble sensitivity analysis to a tropical cyclone over the western North Pacific

*The content of Chapter 4 is taken from "**George, B.**, Kutty, G., 2022. Multivariate ensemble sensitivity analysis applied for an extreme rainfall over Indian subcontinent. Atmospheric Research 277. 1-14.*

Ocean. Using perturbed initial condition experiments they showed that the predicted response using multivariate ensemble sensitivity is more accurate than univariate ensemble sensitivity analysis.

This chapter examines the various aspects of the multivariate sensitivity analysis in a nonlinear, quasi-operational forecast model. The formulation has been tested in meso- and convective scales in the presence of model errors to understand the impact of diagonal approximation for ensemble sensitivity estimates in the context of an extreme rainfall event that devastated a south-eastern state of the Indian subcontinent, Chennai, on December 2015. This is done by incorporating the contribution from the full covariance matrix in the ensemble sensitivity calculations.

4.2 Case Overview

On 1st December 2015, unprecedented rainfall led to flooding in the Chennai city of the Indian subcontinent. The 24-h precipitation started from 1st December, 0830 Local Time (LT) varied between 77 mm to 494 mm as reported by the city commercial centre, and the recorded citywide average precipitation is estimated to be 286 mm (Herring *et al.*, 2016). The event caused flash flooding leading to widespread damages to life and property. A low-pressure system (LPS) formed over the southern Bay of Bengal on 27th November 2015 made landfall over the Tamil Nadu coast on 1st December 2015, which was responsible for bringing convective systems over Chennai. Chakraborty (2016) has shown that the upper-tropospheric anticyclones that were present over the Bay of Bengal and the Arabian Sea have guided the convective systems to Chennai city. A study by Phadtare (2018) showed that the orographic blocking by the Eastern Ghats obstructed the propagation of the low-pressure system from moving inland and caused the convective systems to remain stationary over the coast. Studies have shown that (Reshmi Mohan *et al.*, 2018; Singh *et al.*, 2018; Srinivas *et al.*, 2018) use of a high-resolution mesoscale model and better model parameterization schemes provides an improved forecast of the Chennai extreme rainfall. Krishnamurthy et al. (2018) has shown that the surge of moisture from the warm Bay of Bengal induced by the formation of low-pressure systems triggered the precipitation over Chennai on 1st December 2015.

4.3 Model and Data Assimilation System

The numerical weather prediction model employed in this chapter is the Advanced Research Weather Research and Forecasting version 3.8.1. The boundary and initial conditions on a 0.5° × 0.5° global latitude-longitude grid are obtained from NCEP-GFS for the WRF model simulation. Random perturbations are added to the initial and lateral boundary conditions to generate 90 ensemble members (Barker *et al.*, 2004). A horizontal grid spacing of 27 km with 36 vertical levels is chosen for the coarser-resolution parent domain with model top at 50 hPa. The two inner nested domains are configured using the one-way nesting strategy with a horizontal grid spacing of 9 km and 3 km. All domains utilized Yonsei University (Hong, Noh and Dudhia, 2006) scheme for the boundary layer, Noah land surface (Chen and Dudhia, 2001), the Rapid Radiative Transfer Model (RRTM) for longwave (Mlawer *et al.*, 1997), shortwave scheme (Dudhia, 1989), and Thompson microphysical (Thompson, Rasmussen and Manning, 2004) schemes. The Kain-Fritsch (Kain, 2004) cumulus parameterization scheme is adopted for 27 km and 9 km domains, and the 3 km inner domain is run without cumulus parametrization.

Observations are assimilated using an EAKF (Anderson, 2001) available in DART (Anderson et al., 2009) for updating the ensemble perturbations. About 8390 observations are being assimilated over the model domain after processing through the DART quality control (QC) field. Higher values of DART QC field are considered poorer and lower values are regarded as better. Nearly 76.33% observations fail DART QC check and about 635 observations are available in the vicinity of the low-pressure system over the Bay of Bengal. The assimilated observations are as follows: radiosonde winds, temperature and moisture; aircraft winds and temperature; satellite-derived winds; surface winds, temperature, moisture and pressure; and marine temperature, winds and moisture. These observations are assimilated in every 6 hours interval on the outer domain for three days, starting from 0000 UTC 28 November 2015 to 0000 UTC 1 December 2015. Subsequently, 72-h, 48-h and 24-h ensemble forecasts are initialised from the analysis ensemble members on 0000 UTC 29 November 2015, 0000 UTC 30 November 2015 and 0000 UTC 1 December 2015, respectively. During the cycling phase, spatially and temporally varying adaptive inflation (Anderson, 2007; Anderson *et al.*, 2009) and covariance localization (Houtekamer and Mitchell, 1998; Hamill, Whitaker and Snyder, 2001) are applied to maintain the ensemble spread. The Gaspari-Cohn localization function (Gaspari and Cohn, 1999) with horizontal half-widths of 765 km is used as covariance localization and no vertical localization has been used in this chapter.

Table 4. 1. Mathematical notation and dimensions

K	Ensemble size
N	Model state dimension
J	Scalar forecast metric
$\boldsymbol{J}$	$K \times 1$ forecast response vector
$\boldsymbol{x}$	$N \times 1$ analysis state vector
$\boldsymbol{x}_n^e$	$1 \times K$ nth ensemble member state vector
$\boldsymbol{\beta}$	$N \times 1$ vector of regression coefficients
$\boldsymbol{X}$	$N \times K$ matrix with ensemble of state vector perturbations
σ_n	Standard deviation of nth state vector
$\boldsymbol{\rho}$	$N \times N$ matrix of the covariance localization function
$\boldsymbol{x}^a$	$1 \times K$ analysis state vector
$\boldsymbol{x}_s$	$1 \times K$ perturbing state vector
$\boldsymbol{x}^p$	$1 \times K$ vector of perturbed ensemble state variable
α	Perturbation amplitude

4.4 Multivariate Ensemble Sensitivity Analysis

Torn and Hakim (2008) introduced univariate ensemble sensitivity as the slope obtained from the linear regression between a scalar forecast metric $\boldsymbol{J}$ and a state variable $\boldsymbol{x}$. All the mathematical notations used in this section are provided in Table 4.1. The linear relationship between $\boldsymbol{J}$ and $\boldsymbol{x}$, for a slope b and error ε, is given by the linear regression model,

$$\boldsymbol{J} = b\boldsymbol{x} + \varepsilon \tag{4.1}$$

For an ensemble of size K, let the forecast response vector be represented by $\boldsymbol{J}$ and the state vector as $\boldsymbol{x}$. $\boldsymbol{J}$ and $\boldsymbol{x}$ has the dimensions $K \times 1$ and $N \times 1$, respectively, where N indicates the dimension of the state vector. The ensemble perturbations corresponding to the *n*th state variable is the $1 \times K$ vector, $\boldsymbol{x}_n^e$. The ensemble sensitivity of a scalar forecast metric to changes in a state variable is often estimated by solving the ordinary least squares equation and is expressed as

$$b_n = \frac{\partial J}{\partial x_n} = \frac{cov(\boldsymbol{J}, \boldsymbol{x}_n^e)}{var(\boldsymbol{x}_n^e)}, \quad n = 1,2, \dots, N \tag{4.2}$$

The above equation gives the change in $\boldsymbol{x}$ needed to produce a given change in $\boldsymbol{J}$, and therefore, provides the linear statistical estimate of ensemble sensitivity. The sensitivities thus obtained may be used to estimate the forecast response by assimilating a hypothetical observation. This is performed by multiplying the sensitivity with an expected analysis increment from a new or hypothetical observation in a data assimilation system. Consider assimilating a hypothetical observation that is of the same type as the *n*th state variable. Let σ_n be the corresponding ensemble spread that is equal to the analysis increment. The change in forecast response obtained by introducing the new observation is,

$$\delta J_n = \sigma_n \times \frac{\partial \boldsymbol{J}}{\partial \boldsymbol{x}_n} \tag{4.3}$$

It is understood that the covariance matrix of $\boldsymbol{x}$ is approximated as a matrix with diagonal elements alone to estimate the univariate ensemble sensitivity. Hacker and Lei (2015) and Ren et al. (2019) found that the univariate ensemble sensitivity can overestimate the forecast response to an assimilated observation. Hacker and Lei (2015) proposed the multivariate ensemble sensitivity analysis that incorporates the full covariance matrix by using multivariate regression to estimate the ensemble sensitivity. Using sample statistics, the multivariate linear prediction equation is given by,

$$J = b_1 x_1 + b_2 x_2 + \dots + b_n x_n + \varepsilon \tag{4.4}$$

Solving the above equation yields the regression coefficient $\boldsymbol{\beta}$, a $N \times 1$ vector, that gives a linear statistical estimate of multivariate sensitivity,

$$\boldsymbol{\beta} = \frac{\partial J}{\partial \boldsymbol{x}} \tag{4.5}$$

The multivariate regression retains the full covariance matrix when computing the multivariate regression in the subspace spanned by the ensembles and the regression coefficient is given by

$$\boldsymbol{\beta} = \frac{\partial J}{\partial \boldsymbol{x}} = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-\mathbf{1}}\boldsymbol{J} \tag{4.6}$$

Here, $\boldsymbol{X}$ is a $N \times K$ matrix, with ensemble perturbations in each row. Further, this multivariate ensemble sensitivity is multiplied by the analysis increment $\delta\boldsymbol{x}$ of the *n*th state variable to obtain the change in the forecast response after assimilating a hypothetical observation of the same type as the state variable.

$$\delta J_n = \left(\frac{\partial J}{\partial \boldsymbol{x}}\right)^T \delta\boldsymbol{x} \tag{4.7}$$

The analysis increment $\delta\boldsymbol{x}$ is a $N \times 1$ vector that consists of elements δx given as,

$$\delta x_i = \begin{cases} \sigma_n, & i = n, \\ \sigma_n \times \dfrac{cov(x_i, x_n)}{var(x_i)}, & i = 1, \ldots, n-1, n+1, \ldots, N \end{cases} \tag{4.8}$$

Thus, elements δx contains σ_n of *n*th state variable and correlations with the other state variables. However, if the off-diagonal elements of the covariance matrix of x are ignored, then the δJ_n in Eq. (4.7) becomes equivalent to that in Eq. (4.3).

Localization function is applied to multivariate ensemble sensitivity analysis using fifth-order piecewise polynomial with spatial correlation called the Gaspari-Cohn localization function (Gaspari and Cohn, 1999). The localization is introduced as Schur or Hadamard product as $\boldsymbol{\rho} \circ (\delta\boldsymbol{x})$, where "∘" represents the element-by-element product. Accordingly, with covariance localization the multivariate ensemble sensitivity analysis can be represented as follows

$$\delta J_n = \left(\frac{\partial J}{\partial \boldsymbol{x}}\right)^T \boldsymbol{\rho} \circ (\delta\boldsymbol{x}) \tag{4.9}$$

The 24-h accumulated precipitation averaged over a box in a chosen area is used as a forecast metric for the ensemble sensitivity analysis. The selected model state variables include the geopotential height, water vapour mixing ratio, potential temperature, sea level pressure and wind. Firstly, the sensitivity patterns obtained via the univariate and multivariate ensemble sensitivity analysis are identified and compared. The localization scale used in the multivariate ensemble sensitivity is equal to 1600 km in horizontal which is approximately the same as the scale used in the EnKF assimilation cycle. Next, three additional experiments are carried to explore the impact of lead time, localization and model errors on the precipitation forecasts.

4.5 Results and Discussion

4.5.1 Dynamics of rainfall event using wet and dry ensemble members

A better understanding of the dynamics of the weather event may be obtained from differencing subsets of ensemble members based on the intensity of the precipitation. A set of 5 ensemble members with maximum and minimum precipitation over the forecast response region is categorized into wet and dry ensembles, respectively. Figure 4.1 presents the difference in the synoptic features associated with the wet and dry ensemble members. The geopotential height of the subsets differs considerably over the Bay of Bengal at 850 hPa level (Fig. 4.1a), which indicates that the occurrence of a deeper low-level region over the Bay of Bengal favoured the extreme rainfall over the response region. Moreover, the vertical extent of the convective system seems higher in the wet members as compared to dry members. The difference in 300 hPa geopotential heights between the wet and dry members show two significant positive regions towards the east and west of the Indian subcontinent (Fig. 4.1b), which confirms the role of anticyclones in the upper troposphere over the west and east of the Indian subcontinent in the precipitation over Chennai as shown by Chakraborty (2016). Figure 4.1c represents the sea level pressure contours for wet and dry members and closed circulation for blue contours indicates the presence of a low-pressure for wet members over the Bay of Bengal near the response region. Figure 4.1d indicates that the precipitation location in the wet members is shifted to the north of the response region as compared to the dry members.

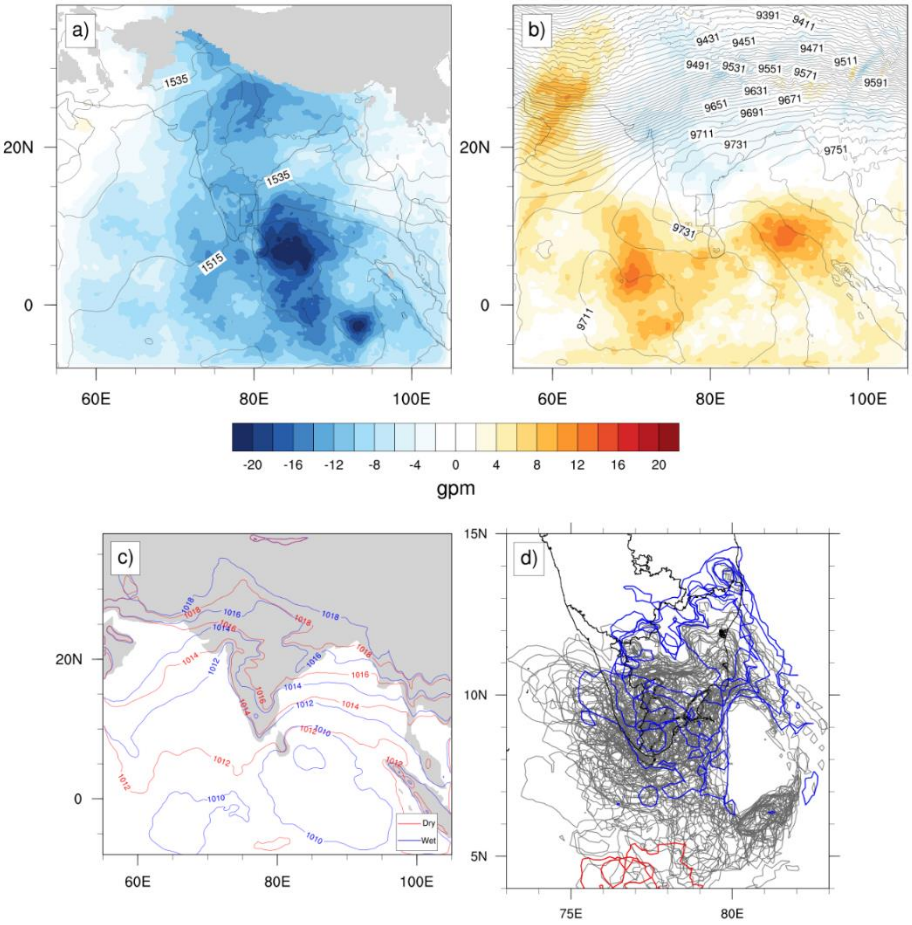

Figure 4.1. Difference between the wet and dry ensemble members on analysis (a) 850 hPa geopotential heights and (b) 300 hPa geopotential height. Contours of analysis (c) sea level pressure for wet (blue) and dry (red) ensemble members. (d) Spaghetti plot showing the predicted 100 mm rainfall contour valid at 0000 UTC 2 December 2015 from the 90 ensemble members. The ensemble-mean 100 mm contour is shown in thick black line. The blue contours represent the wet ensemble members and the red contours represent the dry ensemble members. Black contours in (a) and (b) are the ensemble mean geopotential heights for the respective pressure levels.

4.5.2 Univariate and multivariate ensemble sensitivity

The ensemble sensitivity from univariate and multivariate computed using the 90-member ensemble system are compared to understand the effects of assuming off-diagonal elements as zero in the ESA approach. It is to be noted that the 24-h accumulated precipitation forecast valid at the 72-h lead time is used as the forecast metric for sensitivity analysis in the following sections. As the event considered in this chapter is an extreme precipitation event, rainfall is used as the forecast metric. Figure 4.2 represents the sensitivity of precipitation forecast to the mean sea level pressure in the initial condition. The regions of negative (positive) sensitivity

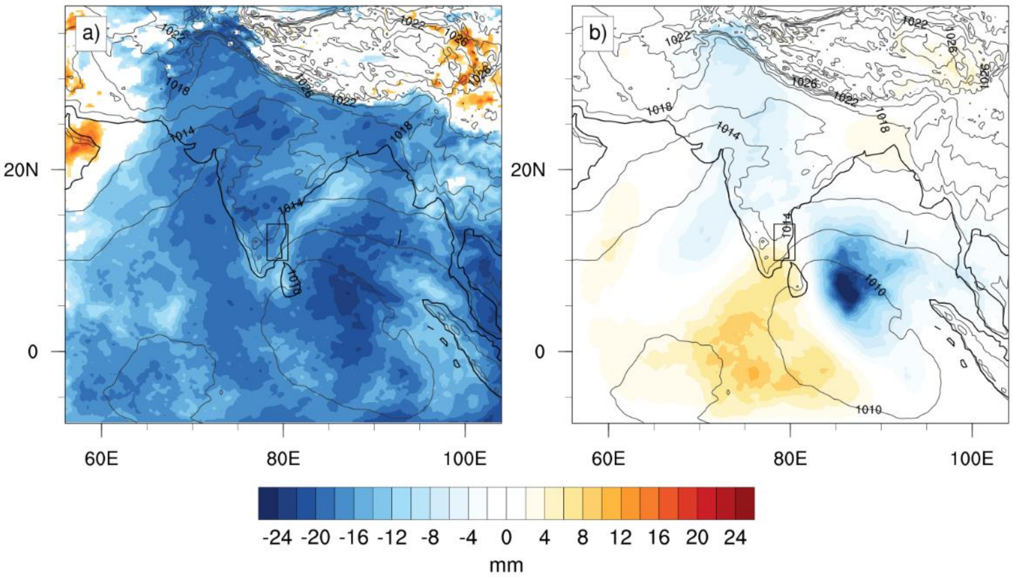

Figure 4.2. Sensitivity of 24-h accumulated area-averaged precipitation valid at 72-h lead time to analysis sea level pressure for (a) univariate and (b) multivariate. Contours (every 2 hPa) are the ensemble mean sea level pressure. The black box represents the response region.

suggest that the forecast metric will be higher if the initial sea level pressure in the sensitive region is lower (higher). More specifically, if the sea level pressure in the negative sensitive region shown in Fig. 4.2a changes by one standard deviation, then the precipitation forecast within the box will change by 22 mm. Both univariate and multivariate ensemble sensitivity analysis estimates a stronger negative sensitive region over the Bay of Bengal indicating that the sea level pressure in the sensitive region plays a significant role in the predictability of rainfall over the response region. However, univariate ensemble sensitivity calculations indicate broader regions of negative sensitivity that span throughout the domain, while multivariate depicts more organized patterns of sensitivity, notably as a dipole over the south of the Indian Ocean region (Fig. 4.2b). Such organized sensitivity patterns reveal the key synoptic-scale factors that influences the forecast metric. Additionally, these sensitive regions identified by the ensemble sensitivity analysis can be used as the potential locations for targeting additional observations to reduce the forecast error. A closer comparison between the results in the previous section reveals that the region of maximum sensitivity values is similar to the regions where maximum difference among the subsets is observed. However, the positive region of the dipole is not observed in Fig. 4.2a.

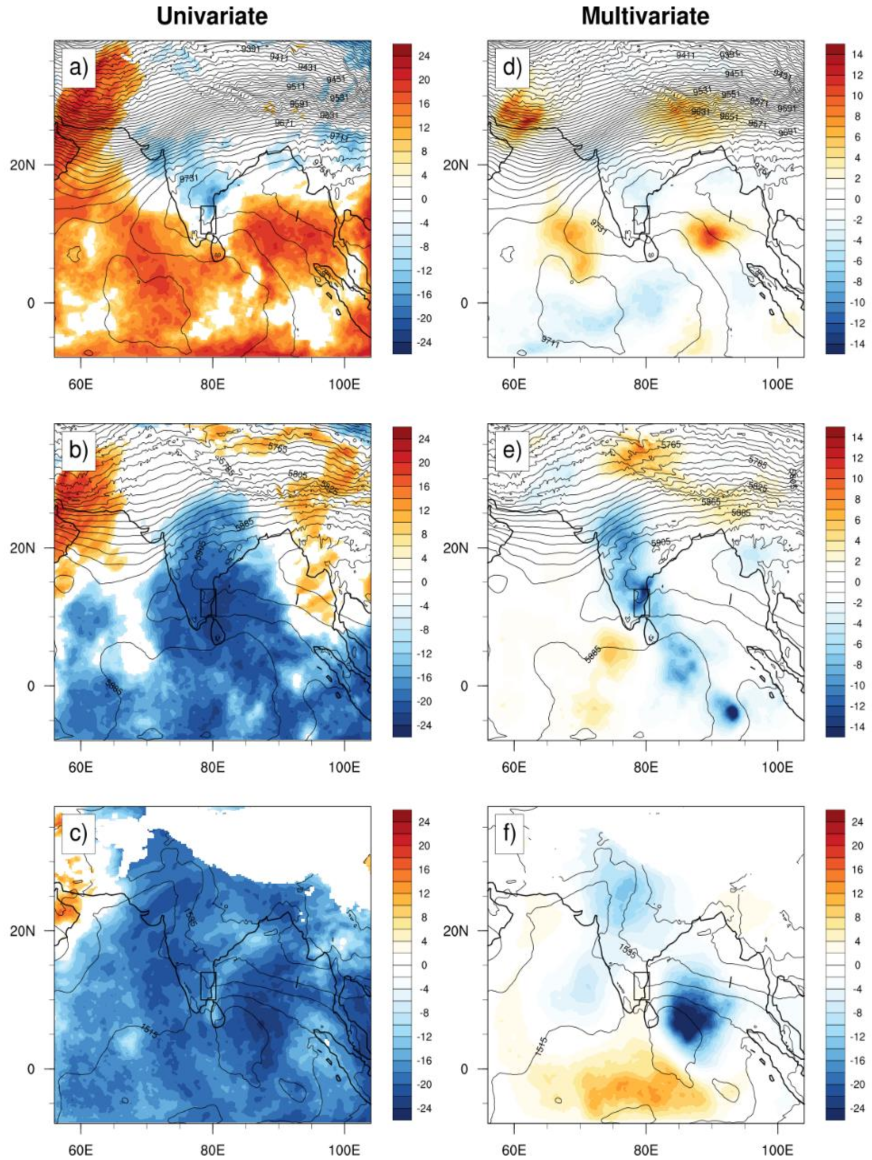

Figure 4.3. Shading (mm) indicates (a-c) the univariate ensemble sensitivity and (d-f) the multivariate ensemble sensitivity of 24-h accumulated area-averaged precipitation valid at 72-h lead time to analysis geopotential heights at 850 hPa (bottom), 500 hPa (middle) and 300 hPa (top). Contours (every 10 gpm) are geopotenial heights from ensemble mean.

Similar results are obtained for the sensitivity of precipitation forecast to geopotential height at the lower, middle, and upper troposphere when the univariate and multivariate sensitivity is compared (Fig. 4.3). As the extreme rainfall event is closely associated with the low-pressure system formed over the Bay of Bengal, it is a natural choice to consider the geopotential height as the analysis field. Broader regions of positive sensitivity are observed in Fig. 4.3 for univariate sensitivity at upper levels, while middle and lower troposphere regions depict pronounced negative sensitive regions over the south of the computational domain. In multivariate approach, the sensitive regions are located closer to the forecast response function box in the middle and lower levels. It is hypothesized that the observed variations in the univariate and multivariate ensemble sensitivity patterns originate from the sampling error associated with the ensemble statistics. If the error in the predicted change of the metric is proportional to the sampling error variance, it introduces overconfidence in the analysis, which essentially leads to the overestimation of ensemble sensitivity. Multivariate ensemble sensitivity calculation with localization accounts for sampling error and therefore better predicts forecast metric than univariate ensemble sensitivity approach.

The arguments are further verified by performing perturbed initial condition experiment as given in section 2.5.1. To reduce the computational costs, an ensemble member whose forecast metric is close to the ensemble mean is considered for performing perturbed initial condition experiments. The ensemble member thus selected is then perturbed at 100 randomly chosen grid points using the Eqs. 2.38 and 2.39 in Chapter 2:

$$\mathbf{x}_{\mathrm{i}}^{\mathrm{p}} = \mathbf{x}_{\mathrm{i}}^{\mathrm{a}} + \frac{\partial \mathrm{x}_{\mathrm{i}}^{\mathrm{a}}}{\partial \mathrm{x}_{\mathrm{s}}} \alpha \tag{2.38}$$

where

$$\frac{\partial x_i^a}{\partial x_s} = \frac{cov(\boldsymbol{x}_i^a, \boldsymbol{x}_s)}{var(\boldsymbol{x}_s)} \tag{2.39}$$

With $\boldsymbol{x}_s$ and $\boldsymbol{x}_i^a$ being the $1 \times K$ ensemble estimates of the perturbation state variable and ith control analysis state variable, respectively. The actual model response is obtained by integrating the model forward in time from each perturbed initial condition. Figure 4.4 shows the scatter diagram comparing the predicted response to the actual response obtained from the forecasts initialized from 100 perturbed initial conditions. It can be seen that while univariate ensemble sensitivity overestimates the predicted response, the multivariate ensemble sensitivity is closer to the actual response indicating improved results. The root-mean-square error (RMSE) estimated from the univariate and multivariate ensemble sensitivity compared to actual model response are 23 and 12 mm, respectively. Similar results are obtained by perturbing other variables such as wind, water vapour mixing ratio, and potential temperature.

4.5.3 Impact of model errors

Model error is the hardest source of uncertainty to quantify in a weather forecast and no matter how developed the models are or how fast the supercomputers have become the model imperfections will always have an impact on the forecast responses. Hacker and Lei (2015) proposed that the skill in predicting a nonlinear forecast response from the linear assumptions becomes superior when model error information is incorporated in the ensemble sensitivity analysis. To account for the model error and to analyse its impact on sensitivity calculations, a new set of ensemble members with small-amplitude stochastic perturbations using stochastic kinetic-energy backscatter scheme (SKEBS: Berner et al., 2011, 2009; Shutts, 2005) are generated for sensitivity calculations using univariate and multivariate approach. Figure 4.5 indicates that the magnitude of sensitivity has increased for both univariate and multivariate approaches when SKEB ensembles are used. In addition to that, the multivariate sensitivity using SKEB ensembles shows substantial variations in its magnitude and spatial distribution as compared to univariate method. For instance, it can be seen that the positive sensitivity patterns observed over the south of the Indian Ocean in the No SKEB experiment have been found shifted to the Arabian Sea and Indian landmass in the SKEB experiment (Fig. 4.5c and 4.5d). We hypothesize that the use of the SKEB scheme has increased the forecast spread, which might have resulted in the enhanced magnitude of ensemble sensitivity calculations (Fig. 4.6). The results are consistent with the study performed by Bednarczyk and Ancell (2015), which indicates that greater variance in the response function may lead to increase in the magnitude of sensitivity. Another interesting feature is that the bimodal distribution observed in Fig. 4.4 is absent when SKEB scheme is employed. We presume that the bimodal distribution is associated with the under dispersive ensemble system and using SKEB ensembles can treat such problems to a certain extent.

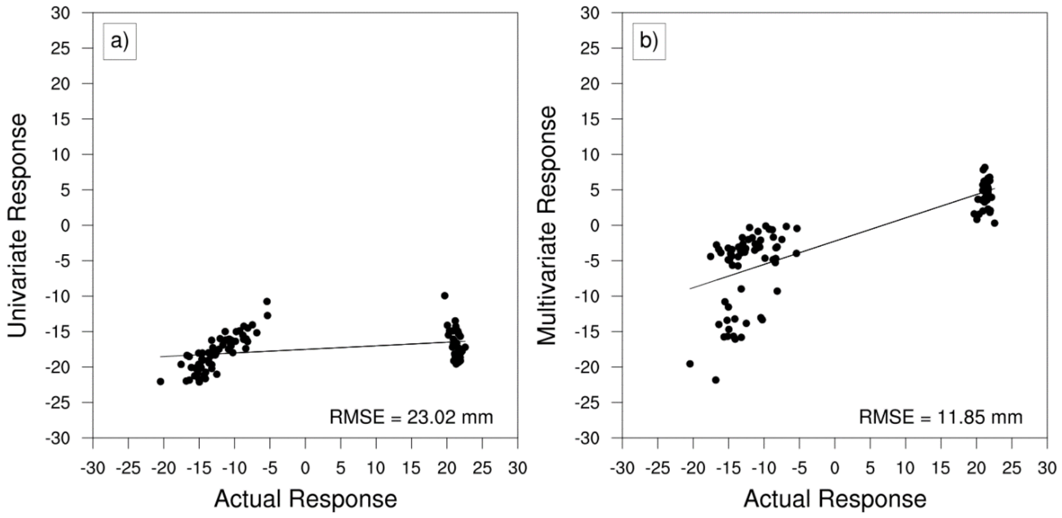

Figure 4.4. Actual response obtained from perturbation versus predicted response from (a) univariate ensemble sensitivity analysis and (b) multivariate ensemble sensitivity analysis for 72-h lead time. The black line indicates the least squares best-fit line.

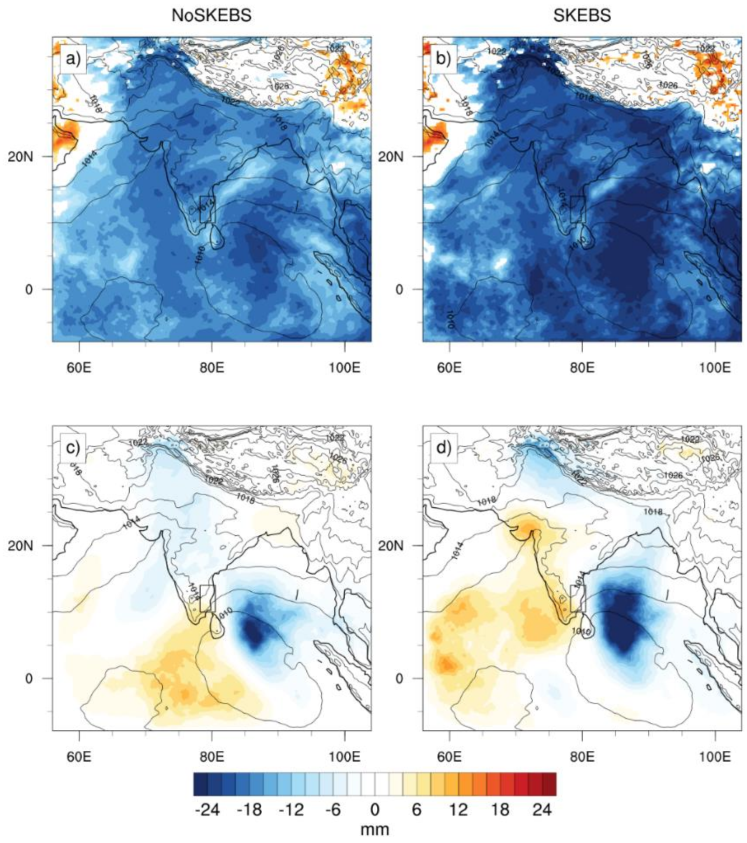

Figure 4.5. Sensitivity of 24-h accumulated area-averaged precipitation valid at 72-h lead time to analysis sea level pressure from univariate (top panel) and multivariate (bottom panel) methods for SKEBS and no SKEBS ensembles. Contours (every 2 hPa) are mean sea level pressure from ensemble mean.

The predicted responses from univariate and multivariate sensitivities are further evaluated using perturbed initial condition experiment. Figure 4.7 shows the scatter diagram of predicted responses from sensitivity analysis versus the actual model response, which indicates that predicted responses using both the methods with SKEB ensembles are lower when compared to the actual response. Compared to univariate, the forecast response estimated by multivariate compares better with the actual response, which is evident from the RMSE values estimated for univariate and multivariate sensitivity.

4.5.4 Impact of localization

As mentioned in section 4.5.2 sampling error associated with the finite (here 80) ensemble members is ameliorated using a localization function in multivariate ensemble sensitivity. To further understand the influence of the specification of localization cut-off radius in the multivariate approach, sensitivity studies with different localization radius in the multivariate method are attempted. A total of six experiments are conducted by increasing and decreasing the cut-off radius by 25%, 50% and 75% of the reference cut-off radius (1600 km).

A significant reduction (increase) in the sensitivity magnitude is observed when the cut-off radius of localization is decreased (increased). The sensitivity decreases almost by half of its magnitude when the cut-off radius is reduced from 1200 km to 800 km (Fig. 4.8a, b). However, on increasing the cut-off radius the sensitivity magnitude increases only gradually (Fig. 4.9). Further, the actual response obtained from the perturbed initial condition experiment is compared with the predicted response by the multivariate ensemble sensitivity. Figure 4.8d-f indicates that the performance of the multivariate approach depends on the optimal choice of localization radius and if insufficient localization is applied, the spurious long-distance correlation contaminates the performance of the multivariate ensemble sensitivity method.

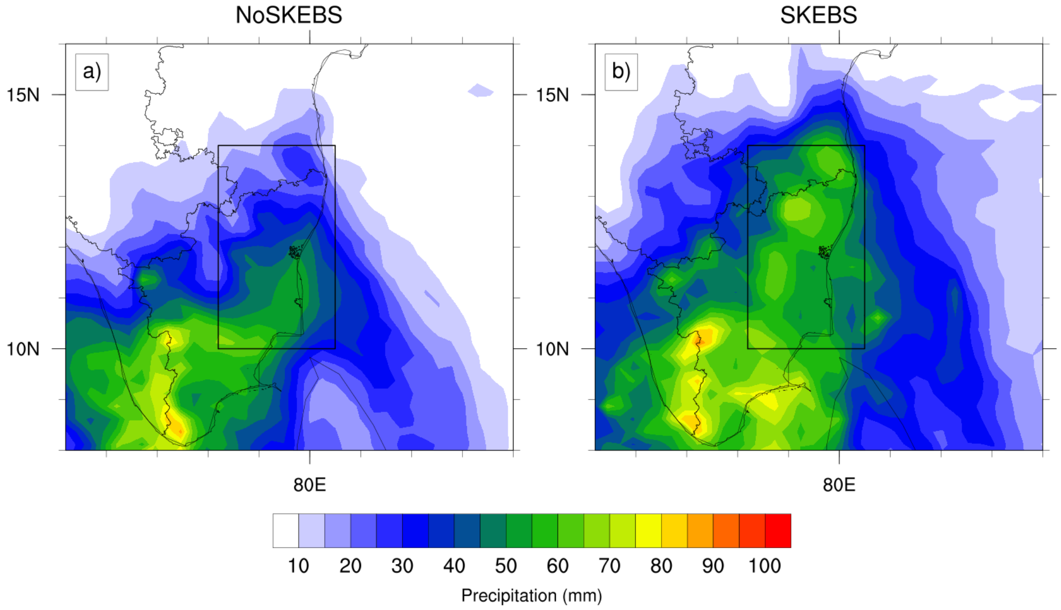

Figure 4.6. Ensemble spread of 72-h accumulated precipitation valid at 0000 UTC 2 December 2015 with (a) NoSKEBS and (b) SKEBS. The black box represents the response region.

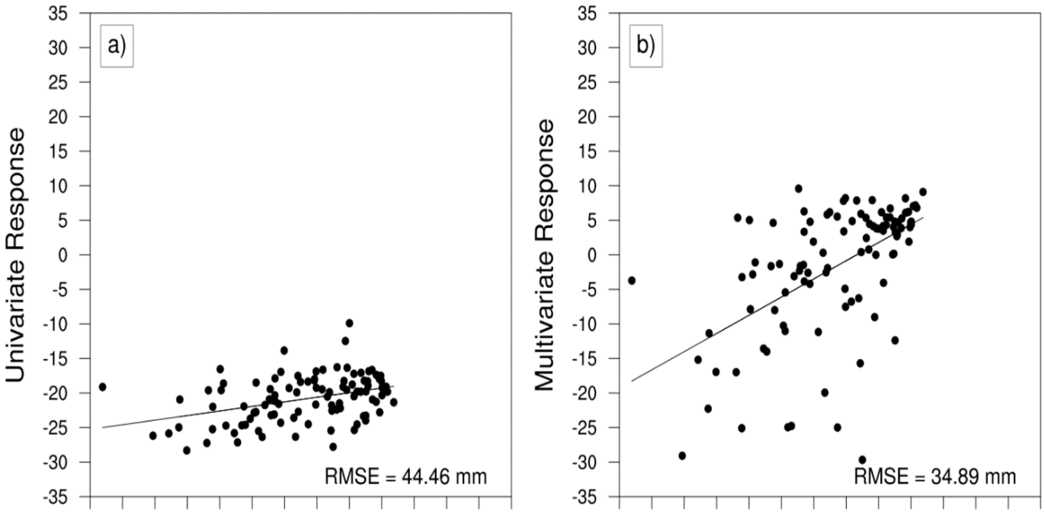

Figure 4.7. Actual response versus predicted response from (a) univariate ensemble sensitivity analysis and (b) multivariate ensemble sensitivity analysis for 72-h ensemble forecasts with SKEBS. The black line indicates the least squares best-fit line.

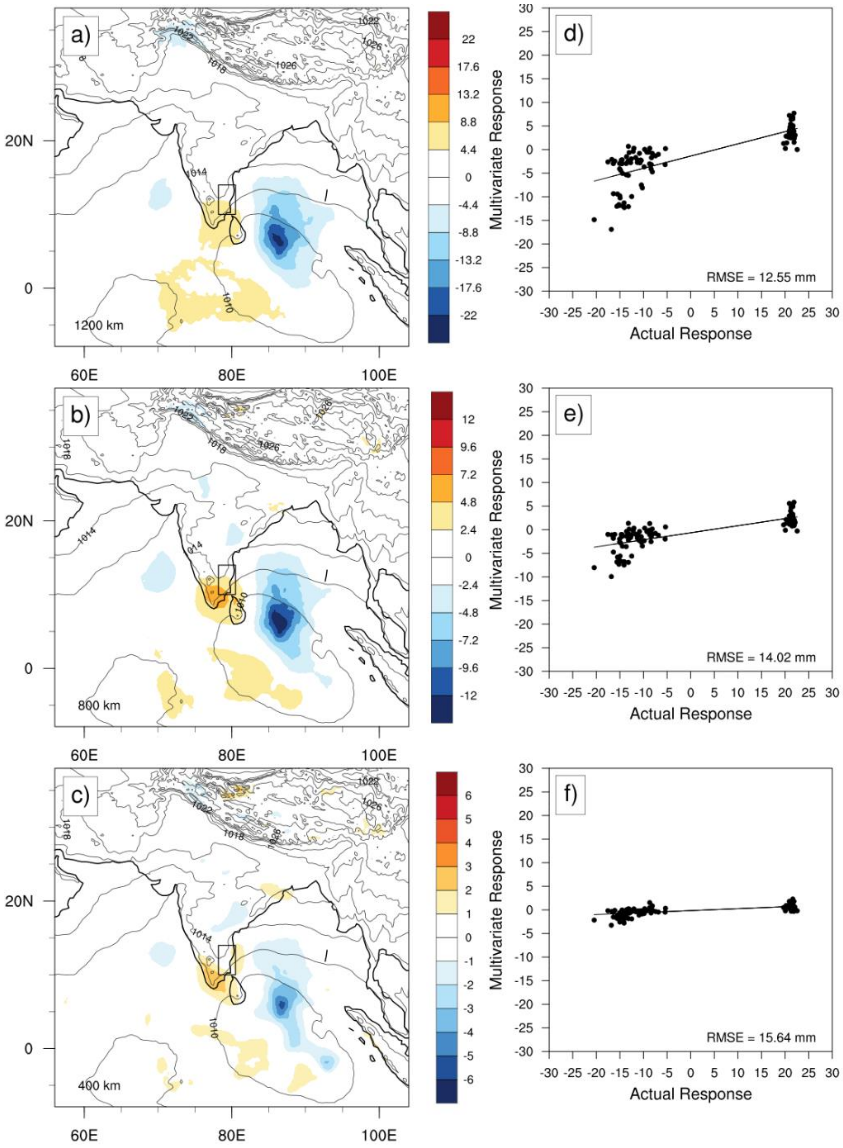

Figure 4.8. Sensitivity (left) of 24-h accumulated area-averaged precipitation valid at 72-h lead time to analysis sea level pressure and perturbation experiment results (right) for different localization cut off radius (a & d)1200 km, (b & e) 800 km and (c & f) 400 km. Contours (every 2 hPa) are mean sea level pressure from ensemble mean. The black box represents the response region.

4.5.5 Impact of lead time

The impact of various forecast lead times on the univariate and multivariate ensemble sensitivity analysis is explored in this section. The univariate and multivariate ensemble sensitivities are estimated at 72-h, 48-h and 24-h lead time with analysis variables initialized at 0000 UTC 29 Nov 2015, 0000 UTC 30 Nov 2015 and 0000 UTC 1 Dec. It should be noted that the forecast metric remains the same and the analysis variables are set to vary in the two sensitivity methods. Figure 4.10 shows the sensitivity obtained for univariate and multivariate with respect to sea level pressure at different lead times. While the broader region of negative sensitivity values is obtained for univariate sensitivity analysis at all times, multivariate ensemble sensitivity depicts more organized and localized negative sensitive regions. The magnitude of multivariate response is quite higher than that of the univariate method during 48-h and 72-h of forecast whereas the magnitude of sensitivity estimated by both the methods for the 24-h forecast is comparable. Furthermore, the multivariate depicts dipole pattern of sensitivity near response region, which clearly indicates the significance of the position of the low-pressure system over the Bay of Bengal for the Chennai rainfall event. The results of the perturbation experiment (Fig. 4.11) show that at all lead times, the predicted response from multivariate ensemble sensitivity analysis compares well with the actual model response. At the 72-h lead time, the RMSE (23.05 mm) from the univariate is nearly twice of the RMSE (11.85 mm) from the multivariate analysis. However, at the 24-h lead time the RMSE from the multivariate sensitivity is slightly higher than that from the univariate method, possibly because of the weak correlations between the forecast response function and perturbation variable. Thus, it is understood that the predicted responses using multivariate ensemble sensitivity are more accurate than univariate, especially at longer lead times when nonlinearity becomes significant.

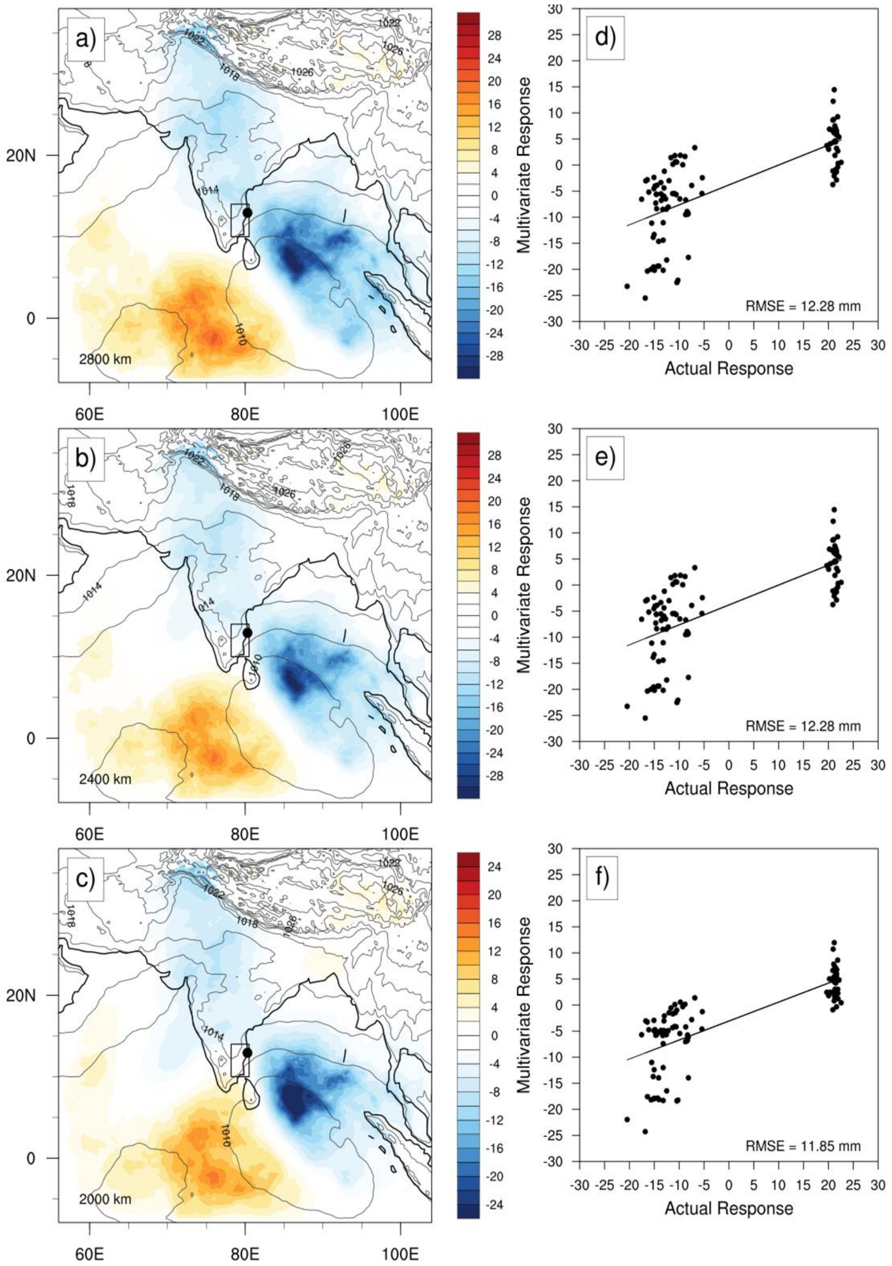

Figure 4.9. Sensitivity (left) of 24-h accumulated area-averaged precipitation valid at 72-h lead time to analysis sea level pressure and perturbation experiment results (right) for different localization distances (a), (b) 2800 km, (c), (d)2400 km, and (g), (h) 2000 km. Contours are the ensemble mean sea level pressure. The black box represents the response region. The black dot represents the location of Chennai.

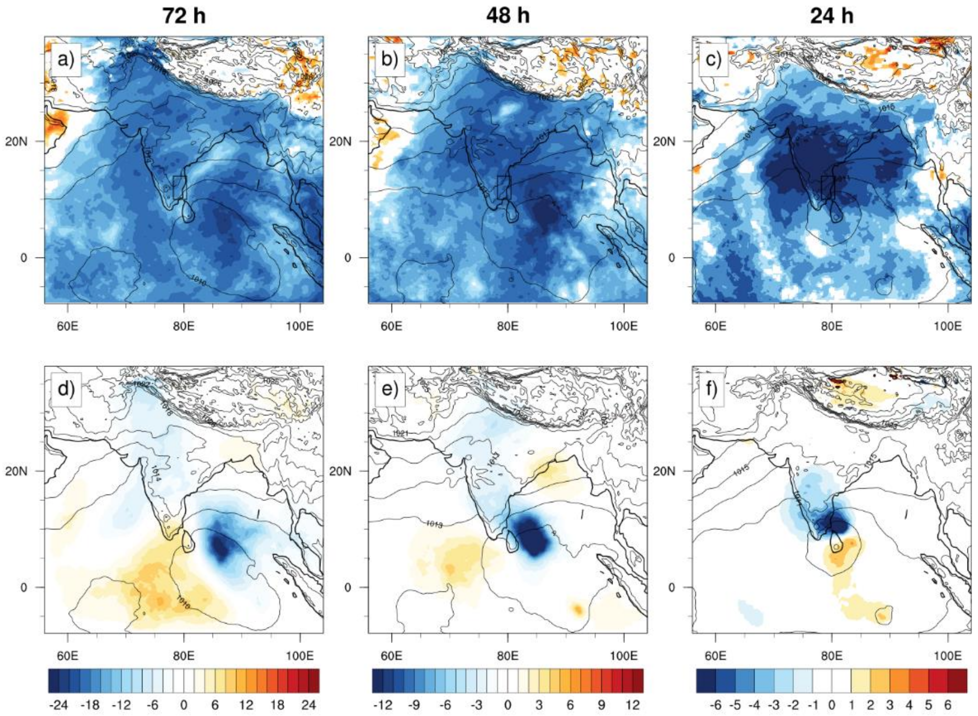

Figure 4.10. Shading (mm) represents the univariate (top) and multivariate (bottom) ensemble sensitivity of 24-h accumulated area-averaged precipitation w.r.t sea level pressure for (a & d) 72-h, (b & e) 48-h and (c & f) 24-h lead times. Contours (every 2 hPa) are mean sea level pressure from ensemble mean.

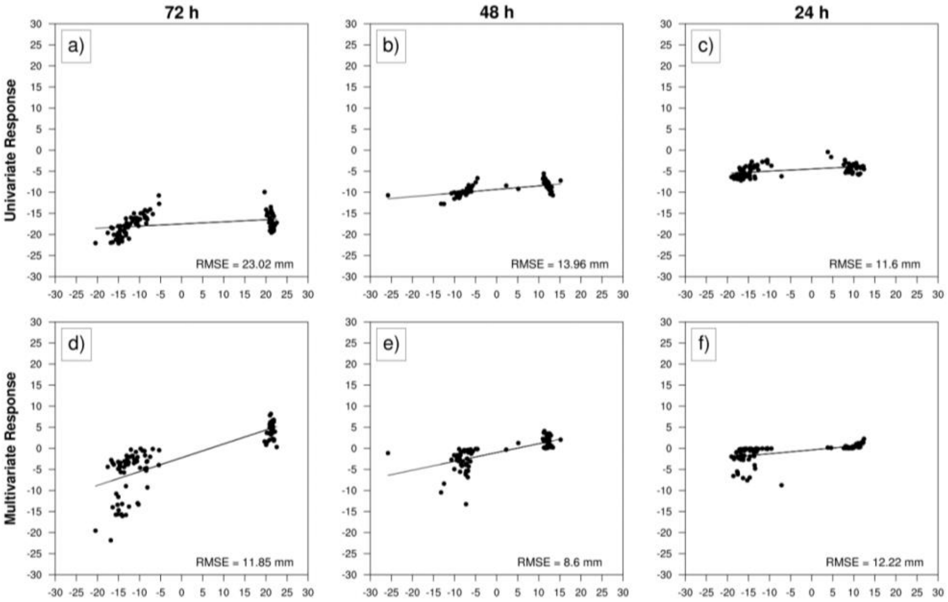

Figure 4.11. Actual response versus predicted response from univariate ensemble sensitivity (top panel) analysis and multivariate ensemble sensitivity analysis (bottom panel) for (a & d) 72-h, (b & e) 48-h and (c & f) 24-h lead times. The black line indicates the least squares best-fit line.

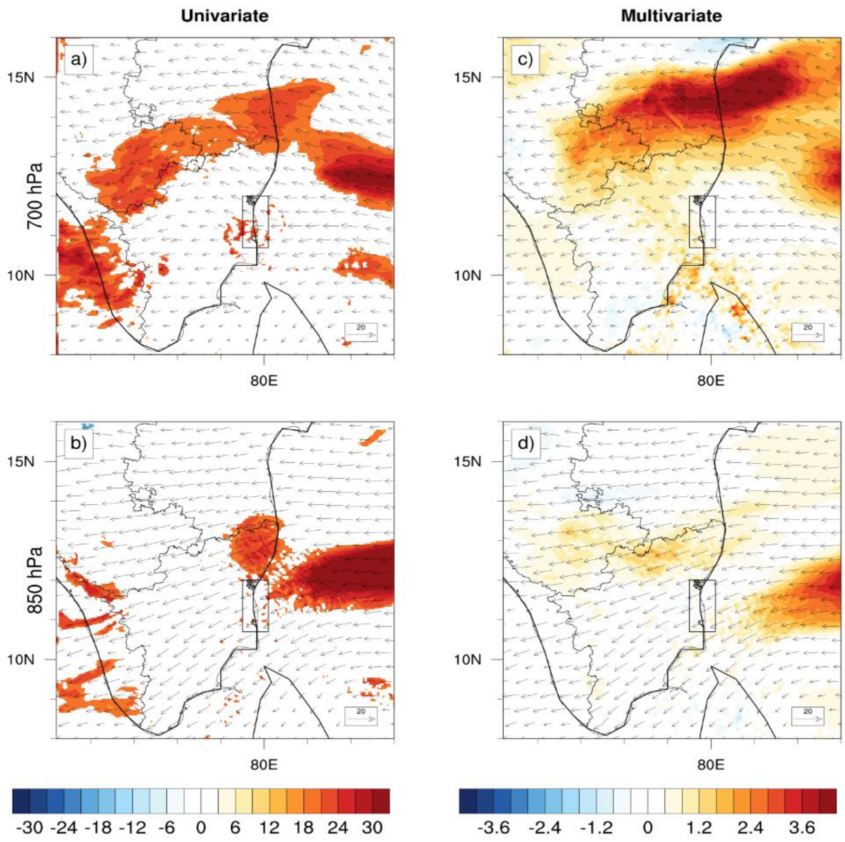

Figure 4.12. Shading (dBZ) represents the sensitivity of area-averaged maximum simulated reflectivity valid at 0000 UTC 2 December 2015 to analysis water vapor mixing ratio at 700 hPa (top panel) and 850 hPa (bottom panel) for univariate (left) and multivariate (right). Vectors are the wind vectors at the appropriate pressure level from ensemble mean. The black box represents the response region.

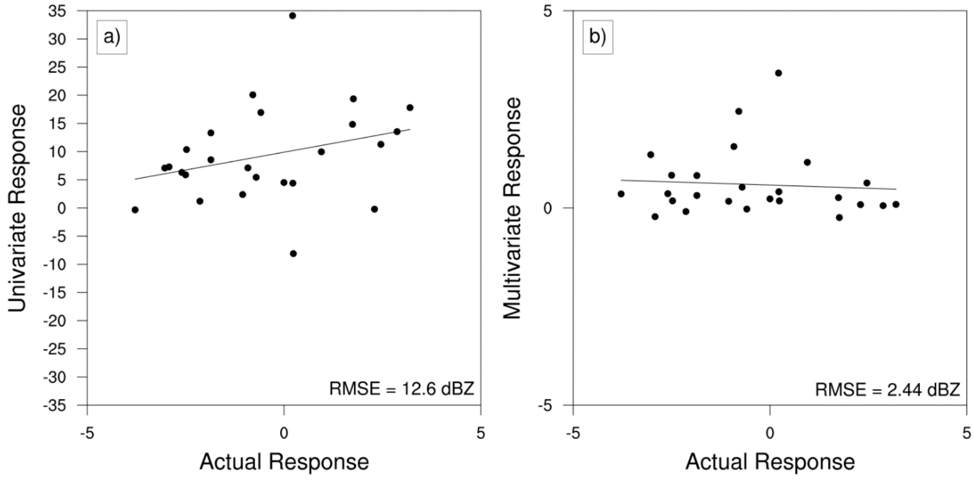

Figure 4.13. Actual response versus predicted response from (a) univariate ensemble sensitivity analysis and (b) multivariate ensemble sensitivity analysis for 24-h convective ensemble forecasts. The black line indicates the least squares best-fit line.

4.5.6 Convection permitting ensembles

The nature of the evolution of initial condition error in convection-permitting forecasts is significantly different from a coarser resolution model forecast. Therefore, the magnitude and spatial distribution of sensitivity patterns may vary substantially in meso- and convective flow scales. The performance of univariate and multivariate ensemble sensitivities in convection scale forecasts has been examined in this section. The univariate and multivariate sensitivity experiments are repeated for the ensemble forecasts and analysis obtained from the 3 km domain. Section 4.3 provides an explanation on obtaining the ensemble forecasts and analysis from the 3 km inner domain.

The forecast response function used in the ensemble sensitivity analysis for convection scale ensembles is the area-averaged maximum simulated reflectivity valid at 0000 UTC 2 December 2015 and the analysis variables considered are potential temperature and water vapor mixing ratio. Figure 4.12 shows the sensitivity estimates of forecasted reflectivity to analysis moisture at 850 hPa and 700 hPa estimated using univariate and multivariate methods. It can be seen that the magnitude of sensitivity estimated by the univariate is significantly higher than the multivariate method. In convection-permitting ensembles, the accuracy of estimated correlations from the ensemble statistics may be suboptimal due to pronounced nonlinearity and we attribute this to be a potential reason for the observed difference in magnitude of sensitivity of univariate and multivariate sensitivity estimates. The sensitivity patterns estimated by the univariate and multivariate ensemble sensitivity methods show similarity in spatial distribution, in general. To verify the performance of methods, the perturbation initial condition experiment is conducted using the convective ensembles at 25 randomly selected grid points (Fig. 4.13). It is found that the predicted response by multivariate sensitivity falls within the range of actual response with an RMSE of 2.4 dBZ. However, univariate sensitivity predicts a higher forecast response compared to the actual response with RMSE equal to 12.6 dBZ. Thus, the predicted response from multivariate ensemble sensitivity analysis using the convective ensembles provides better estimates than the univariate estimates of sensitivity. The error growth is expected to be greater in convective scales and the multivariate sensitivity calculation with localization is found to be substantially beneficial in finer scales.

4.6 Summary

Univariate ensemble sensitivity has been widely applied to understand the dynamics of forecast errors associated with various weather systems using ensemble statistics. However, the method ignores the contribution of diagonal elements of the covariance matrix in its calculations and it may potentially overestimate the response. On the other hand, the multivariate ensemble sensitivity analysis incorporates the contribution from the full covariance matrix in the sensitivity calculations. This work evaluates the performance of univariate and multivariate ensemble sensitivity analysis methods in the backdrop of an extreme rainfall event over the Tamil Nadu coast in India. The 90-member ensemble analyses and forecasts are generated using data drawn from a cycling DART EnKF data assimilation system and WRF model. Univariate and multivariate ensemble sensitivity is estimated for meso- as well as convection scale ensemble members and the effectiveness of the methods is demonstrated by validating the response against actual response generated from perturbed initial condition experiment.

The analysis performed on the subsets of ensemble members based on the intensity of precipitation confirms the role of anticyclones in the upper troposphere over the west and east of the Indian subcontinent in the precipitation over Chennai. Results from ensemble sensitivity analysis indicate that the univariate ensemble sensitivity calculations for precipitation forecasts depict broader regions of negative sensitivity that extend throughout the domain, while multivariate depicts more organized patterns with a dipole over the south of the Indian Ocean region. Perturbation initial condition experiments suggest that the multivariate ensemble sensitivity is more effective in predicting the forecast response closest to the actual model response compared to the univariate ensemble sensitivity. To account for the model error and to analyse its impact on sensitivity calculations, a new set of ensemble members with small-amplitude stochastic perturbations using stochastic kinetic-energy backscatter scheme are generated for sensitivity calculations using the univariate and multivariate approaches. Compared to univariate, the forecast response estimated by multivariate with SKEBS ensembles compares better with the actual response, which is evident from the RMSE values estimated for univariate and multivariate sensitivity. To account for sampling errors resulting from finite ensembles, covariance localization is used in multivariate ensemble sensitivity. It is found that the performance of the multivariate approach depends on the optimal choice of localization radius and if insufficient localization is applied, the spurious long-distance correlation contaminates the performance of the multivariate ensemble sensitivity method. The impact of various forecast lead times on the univariate and multivariate ensemble sensitivity analysis indicates that responses using multivariate ensemble sensitivity are more accurate than univariate, especially at longer lead times when nonlinearity becomes significant. The performance of univariate and multivariate methods in convection-permitting scale is examined by using the high-resolution ensemble forecasts and it is found that the multivariate sensitivity with localization substantially improves the estimates in finer scales.

The analysis of this chapter suggests that multivariate ensemble sensitivity provides better performance than univariate ensemble sensitivity in the presence of model error, non-linearity, and when the observations are sparse.

CHAPTER 5

PREDICTABILITY CHARACTERISTICS OF TROPICAL CYCLONES OVER THE BAY OF BENGAL USING ENSEMBLE SENSITIVITY ANALYSIS

5.1 Introduction

Tropical cyclones (TCs) are one of the most devastating natural disasters due to their economic and social impacts. During the past few decades, there have been drastic improvements in the TC track forecasts, however, the forecasting of tropical cyclone intensity requires further research (DeMaria *et al.*, 2014). Even though advanced data assimilation methods, numerical models and availability of a wide range of observations reduce the errors in the track forecasts, small-scale moist processes and the chaotic nature of the internal dynamics of a TC make it difficult to predict the intensity with a great amount of accuracy. That is, the inherent chaotic nature of the TC puts limits on intrinsic and practical predictability. Therefore, it is crucial to have an understanding on these predictability limits to improve the TC modeling forecasts. Several studies have shown that the ambient factors of TC, such as the sea surface temperature (Tao and Zhang, 2014) and vertical wind shear (Zhang and Tao, 2013) strongly influence the predictability (Torn, 2016). Many other studies have focused on the role of internal processes, such as initial intensity of vortex (Munsell *et al.*, 2017; Liu *et al.*, 2018; Nystrom *et al.*, 2018), the initial size of vortex (Xu and Wang, 2010; Chen, Cheung and Lee, 2011; Carrasco, Landsea and Lin, 2014; Guo and Tan, 2017) and the evolution of the inner-core (Van Sang, Smith and Montgomery, 2008; Zhang and Sippel, 2009; Rogers, 2010; Liu *et al.*, 2018) on the predictability of TC intensity. Besides the chaotic nature of TC, the practical predictability is also limited by model uncertainty and initial condition uncertainty. The former is related to model physics, air-sea interaction and model resolution, and the latter depends on data assimilation techniques and the assimilated observations.

Ensemble-based methods are widely accepted in the research community and operational centers as the ensemble mean provide less forecast error than the deterministic forecasts (Krishnamurti *et al.*, 1997; Zhang and Krishnamurti, 1999; Mackey and Krishnamurti, 2001). Another benefit is that these methods provide flow-dependent estimates of the background and forecast uncertainty. One important advantage of examining predictability using ensemble-based methods is that we can determine both ensemble sensitivity and analysis/forecast spread (Ancell and Hakim, 2007; Torn and Hakim, 2008a). While the ensemble spread of an optimal ensemble system can be regarded as a proxy for predictability, ensemble sensitivity provides a measure of dynamical error growth of initial-time perturbations (Ancell and Hakim, 2007). The analysis covariances in an ensemble framework can be used to describe the likely initial-time perturbations. In short, if the forecast spread is large, it could be either because of large initial spread and small dynamical growth or because of small initial spread and large dynamical growth. Ancell and McMurdie, (2013) showed that the forecast variance can be reduced by assimilating targeted observations at the potential locations for certain types of storms. That is forecast spread can be reduced by reducing initial-time spread through targeted data assimilation.

Studies have shown that among the set of worldwide TCs that caused human fatalities of 5000 or more in the last 300 years, more than 75% occurred in the Bay of Bengal (Singh, 2007; Attri and Tyagi, 2010; Needham, Keim and Sathiaraj, 2015). The densely populated areas at the rim of the Bay of Bengal and the shallow continental shelf makes it vulnerable to TCs (Jakobsen *et al.*, 2006). As the tropical cyclones that develop in the Bay of Bengal are considered to be lethal globally owing to the number of fatalities, it is essential to understand the predictability characteristics of TCs that develop in this basin. Though several studies have taken place and still going on about various aspects of TCs over the Bay of Bengal, very few have ventured into the predictability study of TCs in this basin (Belanger *et al.*, 2012). In this chapter, the general predictability characteristics of landfalling TCs that formed across the Bay of Bengal from 2000 to 2020 are quantified using the multivariate ensemble sensitivity analysis. The chapter also compares the dynamical perturbation growth with initial-time spread for the least and most predictable cyclones. In addition to that, predictability of TC Sidr has been investigated in detail using ensemble sensitivity analysis and ensemble perturbation experiments.

5.2 Model and Data Assimilation System

The version 3.8.1 of Weather Research and Forecasting (WRF; Skamarock WC, (2008)) model with a horizontal resolution of 27 km and 36 vertical levels together with the DART EAKF is used for generating ensemble analyses and forecasts. The initial and lateral boundary conditions are obtained from the National NCEP FNL (final) data with a 1-degree by 1-degree grids. The 90 ensembles and boundary conditions are generated using the RANDOM CV3 option in the WRF-3DVAR system (Barker *et al.*, 2004). The ensemble perturbations are allowed to grow for a period of 12 hours and the analysis ensembles are generated by assimilating observations every 6-h for half a day. The assimilated observations include: satellite winds, Aircraft Communication Addressing and Reporting System (ACARS) aircraft wind and temperature observations, radiosonde wind, temperature and moisture observations, and surface temperature, wind and altimeter observations. Satellite radiance observations are not assimilated. The ensemble analyses thus generated consist of long-range spurious correlations due to sampling error. Therefore, adaptive inflation and covariance localization (Gaspari and Cohn, 1999) are used to mitigate sampling errors in the analysis step. Background for the next assimilation cycle is provided by integrating the model for 6 hours from each analysis for each cyclone. Finally, an extended forecast (48-h) is initialized before the tropical cyclone makes landfall. Parametrization schemes used in the model are as follows: Yonsei University (YSU) for the planetary boundary layer (Hong, Noh and Dudhia, 2006), Kain-Fritsch for cumulus (Kain, 2004), Noah as the land surface model (Chen and Dudhia, 2001), WRF single-moment 5-class for the microphysics scheme (Hong, Dudhia and Chen, 2004), Rapid Radiative Transfer Model (RRTM) for the long wave radiation (Mlawer *et al.*, 1997), and Dudhia for shortwave radiation (Dudhia, 1989).

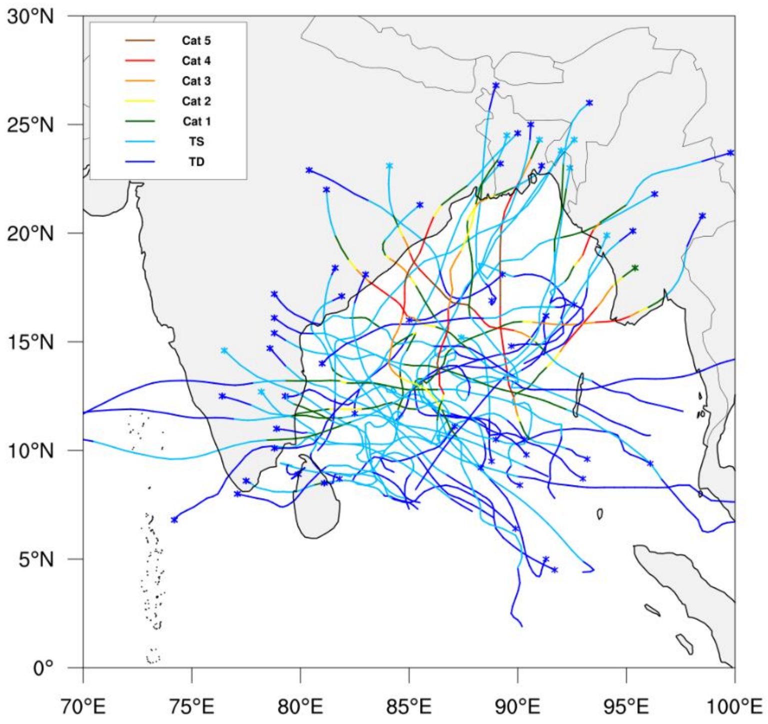

Figure 5.1. Tracks of the tropical cyclones formed over the Bay of Bengal from 2000 to 2020.

5.3 Experimental Design

For this study, 45 cyclones formed over the Bay of Bengal from 2000 to 2020 are selected by keeping the minimum wind speed at 40 knots. The tracks of these cyclones are shown in Fig. 5.1. The forecast metrics used for ensemble sensitivity and spread calculations are the intensity error and track error obtained with respect to the observed intensity and track. The best track analysis from Joint Typhoon Warning Center (JTWC) is used to calculate the model forecast errors. The ensemble spread at the final forecast time (48 h) is used as a proxy for predictability assuming a well-calibrated ensemble system. It is to be noted that larger spread indicates lower predictability and smaller ensemble spread shows higher predictability.

The ensemble sensitivity of the 48-h forecast metric for each cyclone is calculated with respect to initial minimum sea level pressure (MSLP), deep layer vertical wind shear (DVWS) between 850 and 200 hPa, shallow layer vertical wind shear (SVWS) between 850 and 600 hPa, relative humidity at 925hPa (RH925) and relative humidity at 600hPa (RH600). After obtaining the ensemble sensitivity field for each cyclone, the maximum absolute value of the ensemble sensitivity for each variable is found. The maximum absolute value of sensitivity is used as the proxy for the potential dynamical growth of error. For different groups of cyclones such as intense, fast, and north-landfalling, the maximum ensemble spread for each cyclone is averaged together and compared to other groups of cyclones. The same process is repeated for maximum ensemble sensitivity.

All the results in the study are subject to statistical significance using the bootstrap method as follows (Wilks, 2007). Suppose there are two groups with members a and b, respectively. Our focus is to verify the difference in the mean values of the two groups by the bootstrap method. For this a new group with a members and another new group with b members are constructed by random selection of members with replacement from the $(a + b)$-union of the original two groups. Then, the difference between the mean values of the two new groups is computed. This whole process is then repeated 10000 times. Finally, the original difference, which needs to be verified is then considered significant if it achieves the 0.05 or the 95% confidence level.

5.4 Results and Discussions

5.4.1 General predictability characteristics

The tropical cyclones considered in this study are classified into different categories based on their intensity, translational speed and landfalling location to understand the general predictability characteristics. Figure 5.2 shows the forecast spread in MSLP and track for each of the categories. Based on the maximum wind speed TCs are first classified into weak (< 44 knots) and intense cyclones (>44 knots). Intense TCs are found to exhibit large forecast spread in SLP and hence low predictability

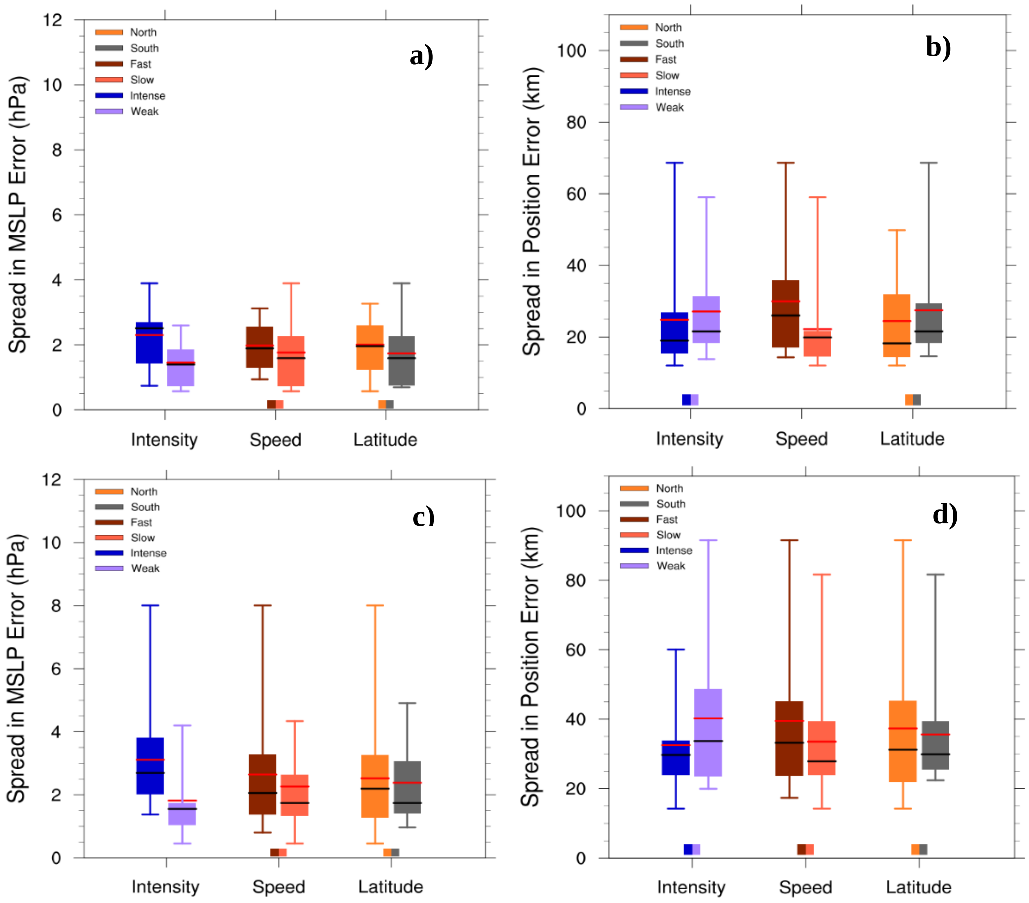

Figure 5.2. Ensemble spread in MSLP error (left panel) and position error (right panel) at 24 h (top panel) and 48 h (bottom panel) for the different categories of cyclones. The top and bottom end of each box represents the upper and lower quartile, respectively. The vertical lines represent the range between the extreme values of each group. The two-colour squares at the bottom indicate significantly different pairs of the respective group. The red horizontal line represents the mean and the black horizontal line indicates mode of each group.

(Fig. 5.2a, c). Similarly, it has been found that intense cyclones show comparatively less forecast spread in track, and therefore these cyclones show high predictability in terms of position as compared to the weak cyclones (Fig. 5.2b, d). Further, from Fig. 5.3a nominal difference is noticed in the analysis SLP spread between the weak and intense TCs. However, the difference in the mean of maximum absolute sensitivity is higher in the case of weak cyclones compared to intense cyclones and the difference is statistically significant. It can be found that the upper quartile is higher in the case of intense cyclones than the weak cyclones for both analysis spread and sensitivity. It means that even though the mean is higher, the number of intense cyclones showing ensemble sensitivity to SLP is higher than the number of weak cyclones. This signifies that the low predictability of intense cyclones is associated with the large intrinsic potential for perturbation growth. Intense cyclones are also found to exhibit high initial condition sensitivity of MSLP to the TC environment and core fields, as seen from Fig. 5.4a. This further suggests that the low predictability of intense cyclones is also linked to the dynamical perturbation growth associated with the environmental and core fields. Figure 5.3d shows the track analysis spread and the position error sensitivity to SLP. It can be seen that the initial condition sensitivity of weak cyclones to SLP is higher than the intense cyclones, which suggests that perturbation growth is a crucial factor in the low predictability of track for weak cyclones.

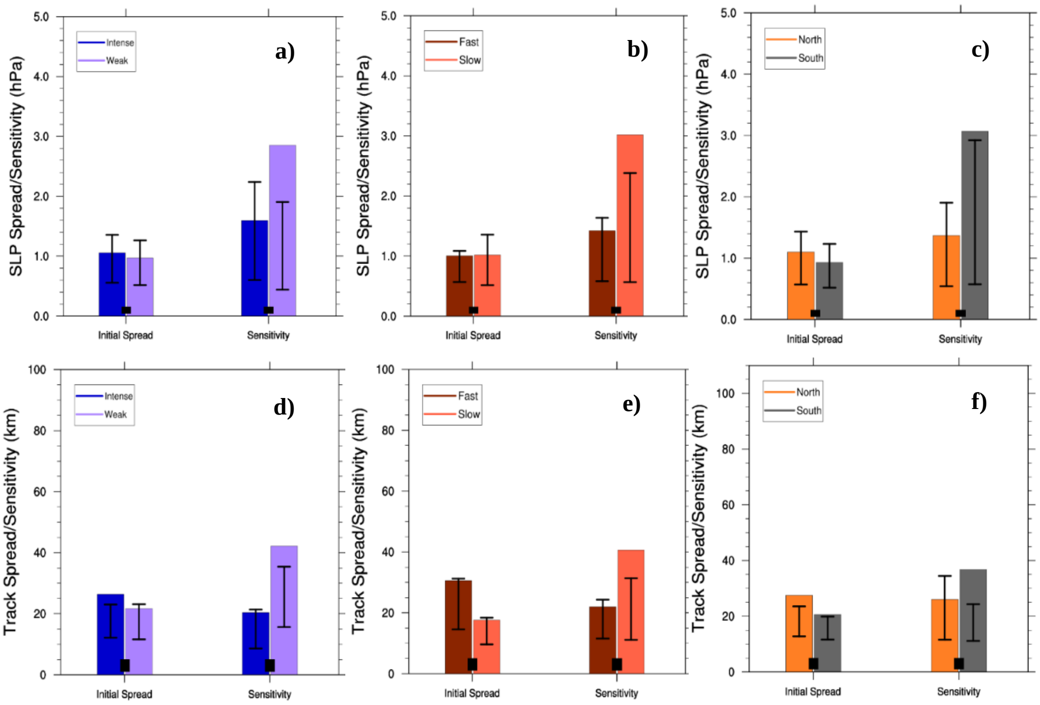

Figure 5.3. Analysis spread in SLP and maximum absolute ensemble sensitivity of MSLP error to SLP (top panel); analysis spread in track and maximum absolute ensemble sensitivity of position error to SLP (bottom panel) for the different categories of cyclone. The bars represent the mean ensemble sensitivity corresponding to each group and the vertical lines are the ranges between the upper and lower quartiles of each group. The small black box at the bottom indicates significantly different pairs.

Again, based on the translational speed the cyclones are further categorized into fast and slow-moving cyclones, where fast-moving cyclones have a translational speed greater than 7.8 knots and slow-moving cyclones have a translational speed less than 7.8 knots. It has been found that fast-moving cyclones exhibit significantly large forecast spread in both SLP and position of the TCs compared to slow-moving storms indicating low predictability in terms of intensity and track (Fig. 5.2). From Fig. 5.3b the analysis spread and initial condition sensitivity to SLP is found to be lower in fast-moving cyclones. However, fast-moving cyclones are also found to show statistically significant sensitivity to the TC core features (Fig. 5.4b). Therefore, core features might have an important role to play in the lower predictability of intensity of fast-moving cyclones. Figure 5.3e shows that the analysis spread of SLP in the fast-moving cyclones is higher than the magnitude of the initial condition sensitivity, which indicates that the low predictability in the track of the fast-moving cyclones is due to the initial spread in the SLP. Also, the position error of faster storms shows higher sensitivity to core features, as seen in Figure 5.4e. Thus, the initial error in the SLP together with the intrinsic error growth in both near surface and midtropospheric humidity at the core

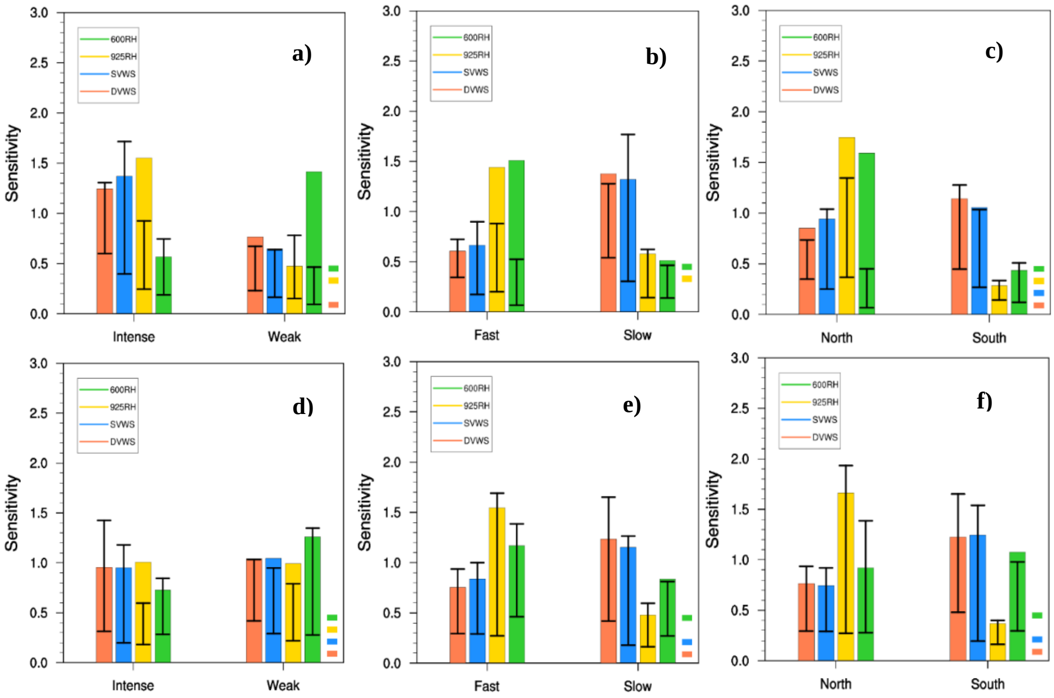

Figure 5.4. The maximum absolute ensemble sensitivity of MSLP error (top panel) and position error (bottom panel) to DVWS, SVWS, 925RH and 600RH for the different categories of cyclone. The bars represent the mean ensemble sensitivity corresponding to each group and the vertical lines are the ranges between the upper and lower quartiles of each group. The small coloured boxes at the right end indicates significantly different pairs.

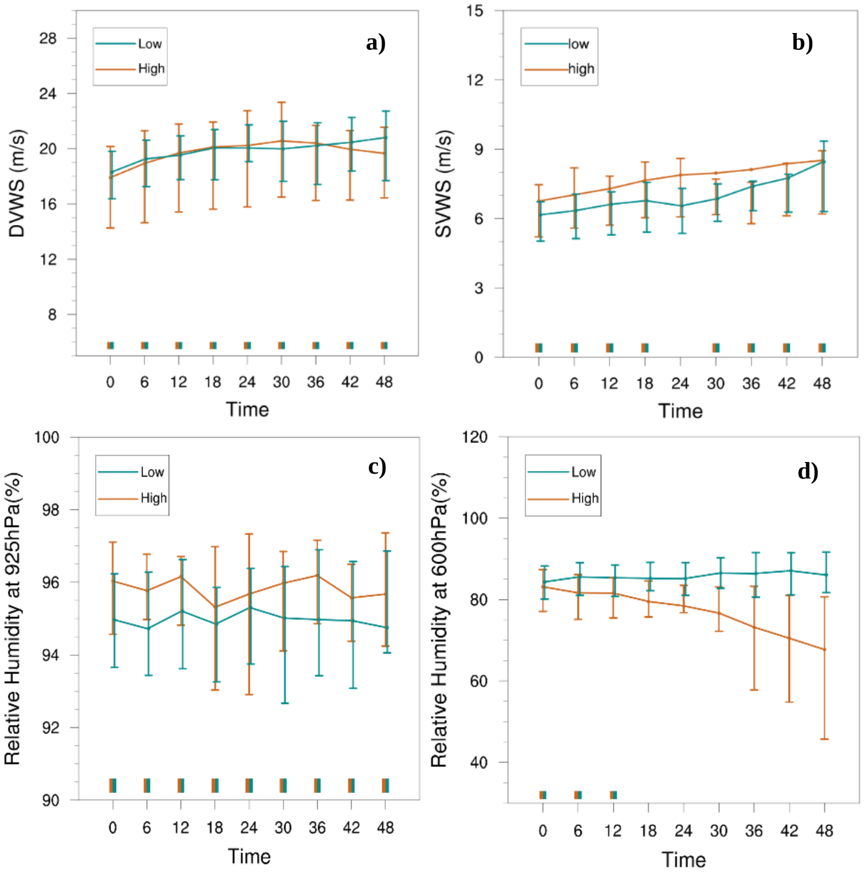

Figure 5.5. Mean (a) DVWS, (b) SVWS, (c) 925RH (d) 600RH at different forecast hours for low and high predictability cyclone classification based on TC intensity error spread at forecast hour 48. The vertical lines are the ranges between the upper and lower quartiles of each group. The two-colour squares at the bottom indicate significantly different pairs at that time.

of faster storms can be considered as a plausible reason for the low track predictability of the fast-moving cyclones.

Another category of TCs is also considered in this study based on the landfalling latitude. The cyclones that make landfall north of 17° latitude are considered north falling cyclones and that make landfall south of 17° are south falling cyclones. The average maximum spread in intensity and position are compared for north and south falling cyclones in Fig. 5.2. Interestingly, north falling cyclones exhibit higher spread in both intensity as well as track compared to south falling cyclones. This suggests that cyclones that make landfall towards the north of the Indian subcontinent are less predictable in terms of intensity and position when compared to the south falling cyclones. Figure 5.3c compares the analysis spread and sensitivity of intensity and position error for north and south falling cyclones. It can be seen that north falling cyclones show higher analysis spread in SLP while the ensemble sensitivity to SLP is higher for the south falling cyclones. Therefore, it can be interpreted that the large forecast spread in the intensity of the north falling cyclones is due to the initial spread in SLP rather than dynamical perturbation growth. Further, the north falling cyclones show higher initial condition sensitivity to the core features, especially to the near surface humidity (Fig. 5.4c). From Fig. 5.3f it is clear that the lower predictability in the track of the north falling cyclones is due to both initial spread and the initial condition sensitivity to SLP compared to the south falling cyclones. The north falling cyclones also show higher ensemble sensitivity to the core fields, which can also be considered a factor that contributes to the low predictability of north falling cyclones (Fig. 5.4f).

5.4.2 Factors affecting the predictability of the tropical cyclones

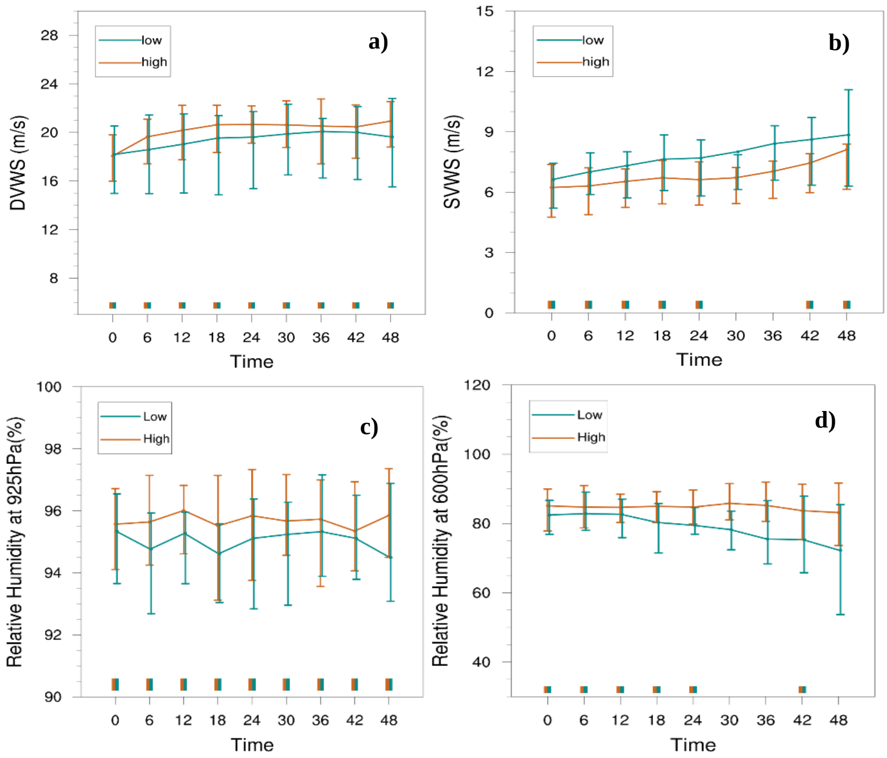

Figure 5.6. Similar to Figure 5.5, but for TC track error spread.

Next, we examine how the different factors of cyclones affect the predictability of tropical cyclone intensity. For this, the cyclones are classified into low predictability cyclones and high predictability cyclones based on the ensemble spread of MSLP error at 48 h. Bootstrap sampling is performed to test the statistical significance of the difference between the mean of the two classifications. Figure 5.5 shows the evolution of DVWS, SVWS, 925RH and 600RH for low predictability and high predictability cyclones as well as the six hourly lower and upper quartiles of each group. There is hardly any difference in the mean value of DVWS between the low and high predictability cyclones (Fig. 5.5a). However, it should be noticed that the range between the upper and lower quartiles is different among the two categories. This means that 50% of cyclones with low predictability have values of DVWS between 18 to 21 m/s at 24 h and 50% of high predictability cyclones have values between 14 to 22 m/s. This suggests that low predictability cyclones have significantly higher values of DVWS than high predictability cyclones. Further, the difference in the mean value of SVWS between low and high predictability cyclones is slightly higher during initial forecast hours with higher SVWS in high predictability cyclones (Fig. 5.5b). From Fig. 5.5c, it is clear that low predictability cyclones exhibit lower values of near surface humidity within the core of the cyclones whereas the midtropospheric humidity (Fig. 5.5d) is high in low predictability cyclones compared to high predictability cyclones. Moreover, it can be noticed that the core features produce a considerable difference between the low and high predictability cyclones rather than the environmental features. The cyclones are again classified into low and high predictability based on the track error spread at forecast hour 48 (Fig. 5.6). It is found that cyclones with low track predictability exhibit low and high values of DVWS and SVWS, respectively. Besides, cyclones with low track predictability include lower near surface humidity and midtropospheric humidity. We have seen how the different parameters evolve in low and high predictability cyclones. Such differences in characteristics between low and high predictability cyclones could be explained either by analysing the error in the initial conditions or the dynamical error growth in the initial perturbations. Figure 5.7a, b shows the analysis spread in SLP and initial condition sensitivity to various atmospheric fields. It can be seen that low intensity predictability cyclones are characterized by less analysis spread in SLP. On the other hand, the intensity of low predictability cyclones shows comparatively high ensemble sensitivity to all the atmospheric fields. The high ensemble sensitivity of low predictability cyclones compared to high predictability cyclones reveals the importance of dynamical perturbation growth in the low predictability of cyclones. Similar analysis of track predictability shows (Fig. 5.7c) a comparatively higher analysis spread for low predictability cyclones. Nevertheless, cyclones with low track predictability show generally large initial condition sensitivity to all fields except SLP (Fig. 7d). Therefore, the initial spread in SLP and dynamical error growth associated with core and environmental fields make the cyclones to be less predictable in track.

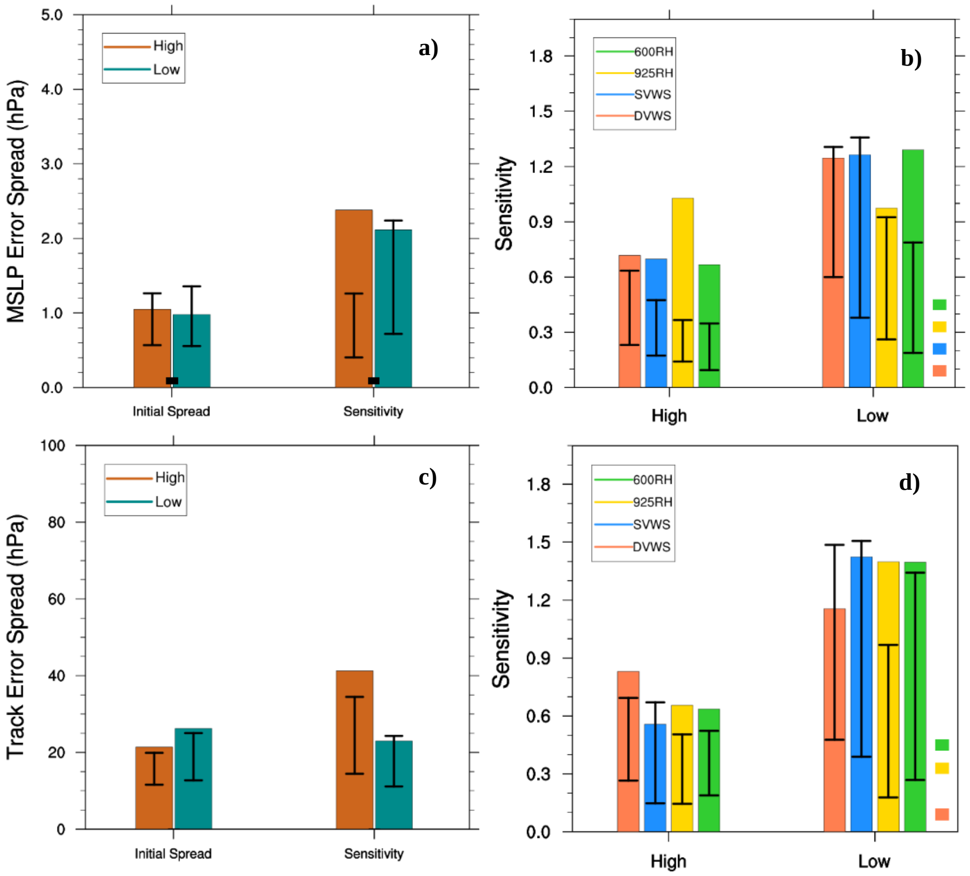

Figure 5.7. Analysis spread in SLP and maximum absolute ensemble sensitivity of MSLP error to SLP, DVWS, SVWS, 925RH and 600RH (top panel); analysis spread in track and maximum absolute ensemble sensitivity of position error to SLP, DVWS, SVWS, 925RH and 600RH (bottom panel). The bars represent the mean value corresponding to each group and the vertical lines are the ranges between the upper and lower quartiles of each group. The small black box at the bottom and the coloured box at the right end indicates significantly different pairs.

5.4.3 Perturbation experiment

In this section, perturbation initial condition experiments are performed on all cyclones to understand the error growth characteristics. For this purpose, we estimate the error growth between the unperturbed (CNTL) and perturbed (PERT) simulations. The error growth is estimated by finding the time evolution of domain-integrated difference total energy (DTE) per unit mass as in Zhang, Snyder and Rotunno (2003). The DTE is defined at any grid point and time as

$$DTE = \frac{1}{2}\sum[(\delta u)^2 + (\delta v)^2 + \kappa(\delta T)^2]$$

where δu, δv and δT are the difference in wind components and difference in temperature between the unperturbed and perturbed simulations. The constant $\kappa = C_p/T_r$, with $C_p = 1004 JK^{-1}kg^{-1}$ being the specific heat at constant pressure and

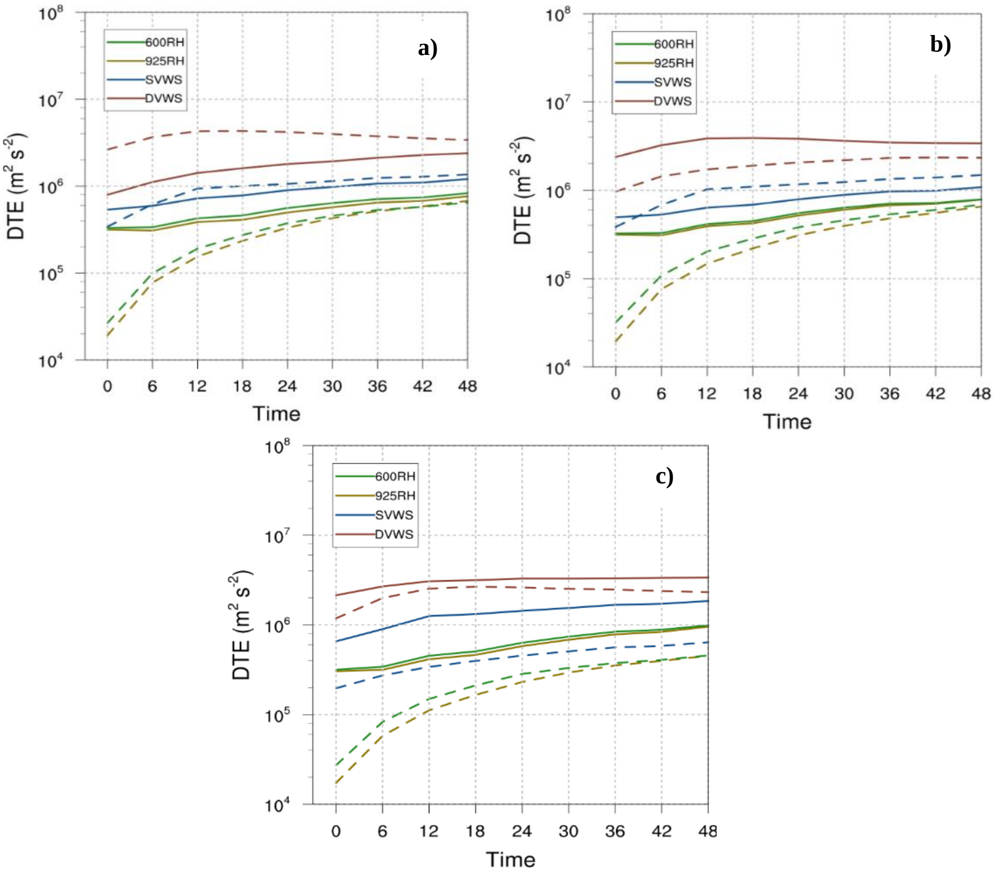

Figure 5.8. Time series of domain integrated difference total energy for the perturbation experiments. DTE time series for the various categories such as (a) weak (solid line) and intense (dashed line), (b) slow (solid line) and fast moving (dashed line), and (c) south (solid line) and north (dashed line) falling cyclones.

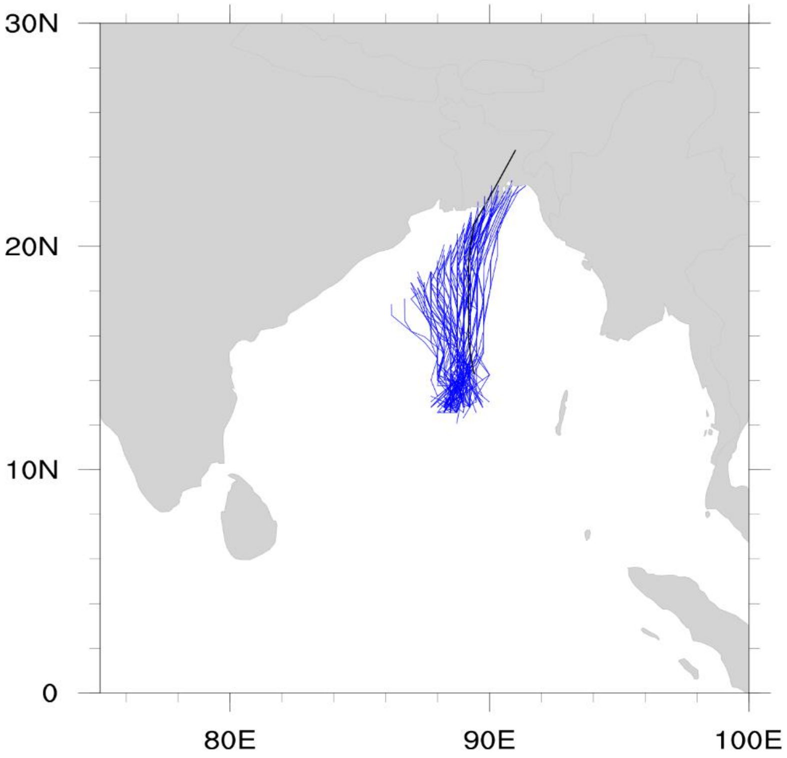

Figure 5.9. Time evolution of the track of TC Sidr. The JTWC best track is denoted in black line and the ensemble forecasts in blue lines.

$T_r = 270K$ being the reference temperature. Figure 5.8 shows the domain-integrated DTE time series of the vertical wind shear and humidity for the different categories of cyclones described in section 5.4.1. As can be seen from Fig. 5.8a the error growth is rapid up to forecast hour 12 for all the fields of the intense cyclones compared to that of the weak cyclones and slows steadily over the rest of the forecast hours. Moreover, errors are found to rapidly increase in the first 6 h for the moisture fields compared to other fields. Similarly, from Figs. 5.8b and c, atmospheric fields associated with fast and north falling cyclones exhibit comparatively rapid error growth than slow and south falling cyclones. This proves that small amplitude perturbations will introduce intrinsic error growth in the atmospheric fields particularly moisture fields at the core, that could affect the predictability of the cyclones.

5.4.4 Tropical cyclone Sidr

Tropical cyclone Sidr is one of the most devastating tropical cyclones that caused several thousand deaths and substantial damages. Sidr is found to exhibit large

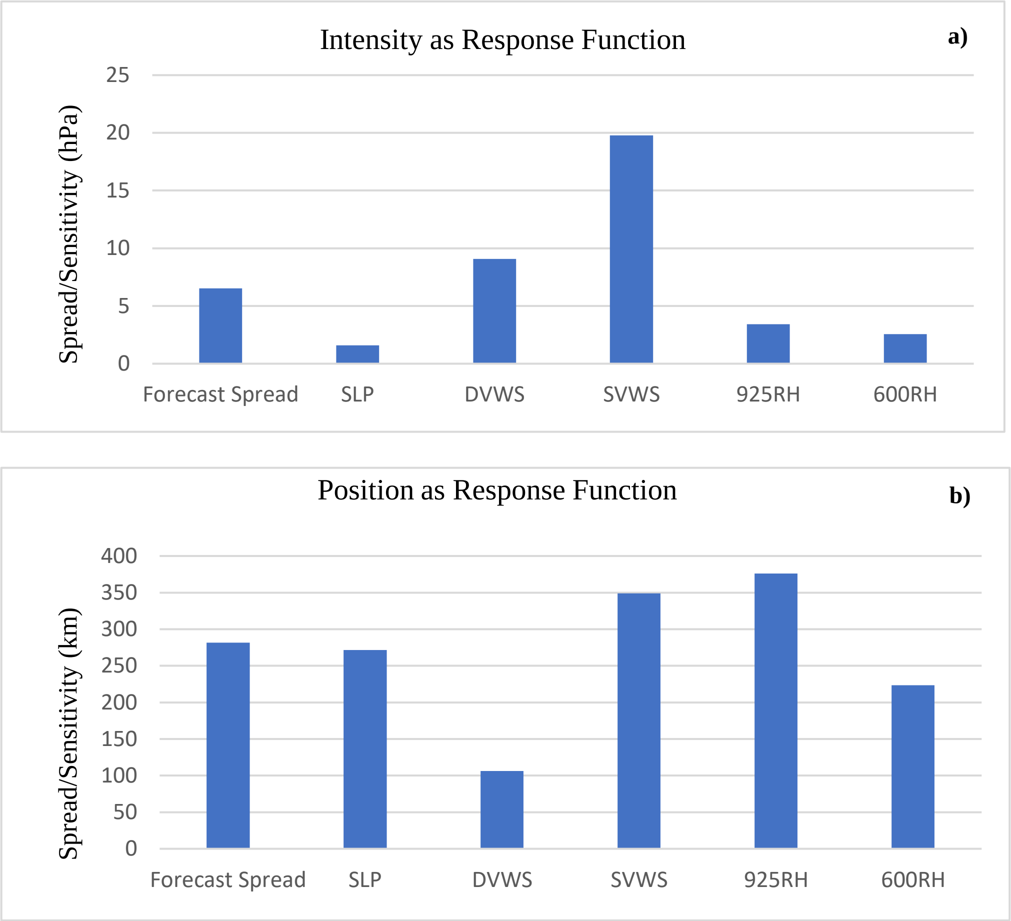

Figure 5.10. Forecast spread in central pressure and maximum absolute ensemble sensitivity of MSLP error to SLP, DVWS, SVWS, 925RH and 600RH (top panel); forecast spread in track and maximum absolute ensemble sensitivity of position error to SLP, DVWS, SVWS, 925RH and 600RH (bottom panel) for TC Sidr.

uncertainty in both intensity and track, therefore the predictability of this TC is studied separately. Figure 5.9 shows the ensemble forecasts and the best track from JTWC. It should be noticed that there is large ensemble spread in the track of the TC Sidr. The magnitude of forecast spread in central pressure and track of TC Sidr is shown in Fig. 5.10 along with the magnitude of maximum absolute ensemble sensitivity. The intensity of the TC Sidr shows the largest sensitivity to shallow layer vertical wind shear at the environment and the track shows maximum absolute sensitivity to near surface humidity at the core. Fig. 5.11 shows the ensemble spread at 48 h and the ensemble sensitivity of cyclone MSLP and position to the initial sea level pressure field. The ensemble spread is the largest in the vicinity of the cyclone low with magnitude greater than 6 hPa (Fig. 5.11a). This hints that various ensemble members forecasted different cyclone positions, sea level pressure at the

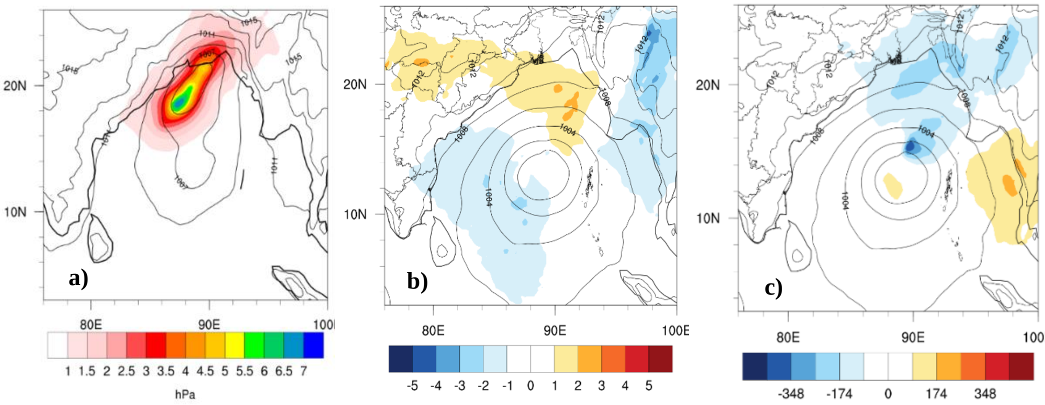

Figure 5.11. (a) Ensemble spread of sea level pressure at 48 h (shaded) and ensemble mean sea level pressure (contours, every 2 hPa) (b) Ensemble sensitivity of MSLP error at 48 h to analysis sea level pressure (shaded) and analysis sea level pressure (contours, every 2hPa).

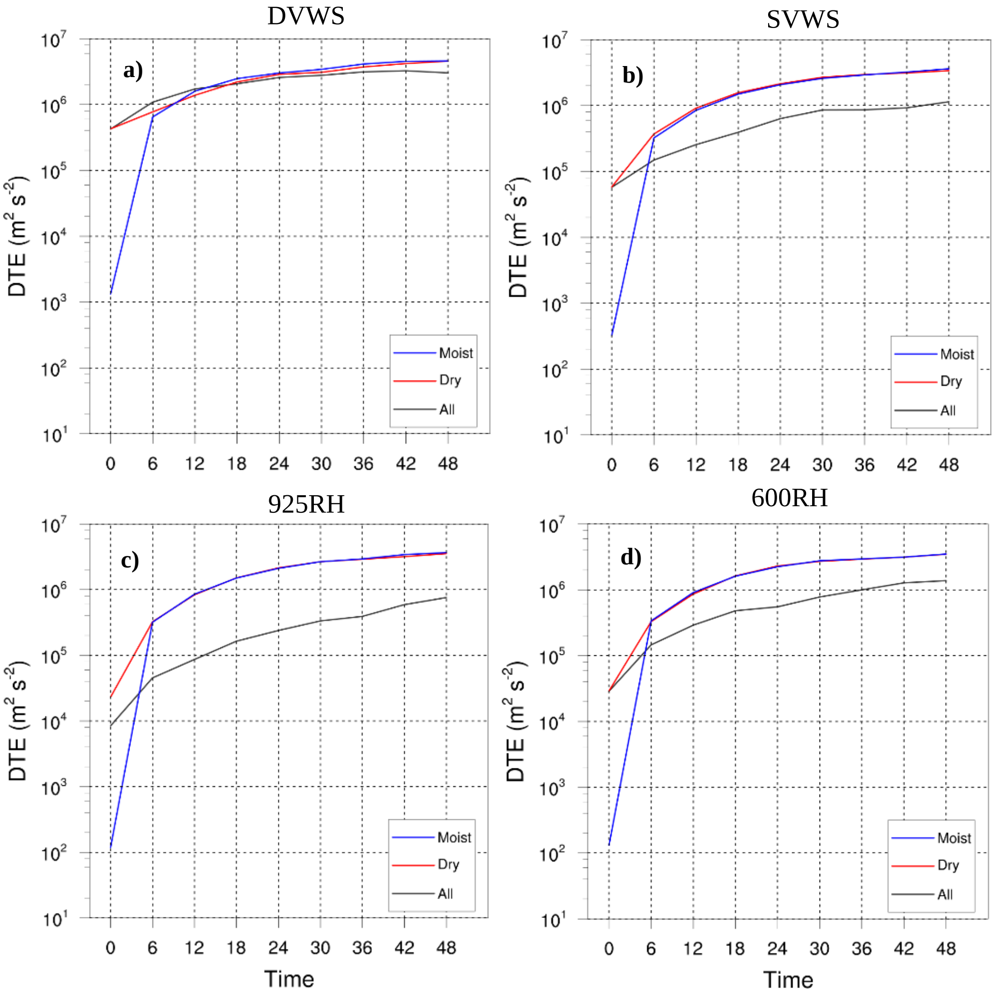

Figure 5.12. Time series of domain integrated difference total energy for the perturbation experiments. DTE time series for the All, Moist and Dry experiments based on the IC metrics (a) DVWS, (b) SVWS, (c) 925RH and (d) 600RH.

center, landfall timing or combinations of these factors. The ensemble sensitivity of central pressure to initial sea level pressure shows positive values to the north and negative values to the southwest of the initial cyclone (Fig. 5.11b). This suggests that changes to initial sea level pressure towards southwest and north of the TC Sidr are related to the intensification of the cyclone 48 h later.

Similarly, Fig. 11c shows the ensemble sensitivity of cyclone position to initial sea level pressure with negative values towards the northeast and positive values towards the east of the initial cyclone position. This means the changes in cyclone position are related to the initial sea level pressure towards northeast and east of the initial position of the cyclone.

The large ensemble spread in the intensity and track of the TC Sidr indicates that the cyclone has low predictability. The strong ensemble sensitivity suggests strong potential for perturbation growth that causes low predictability. Earlier from

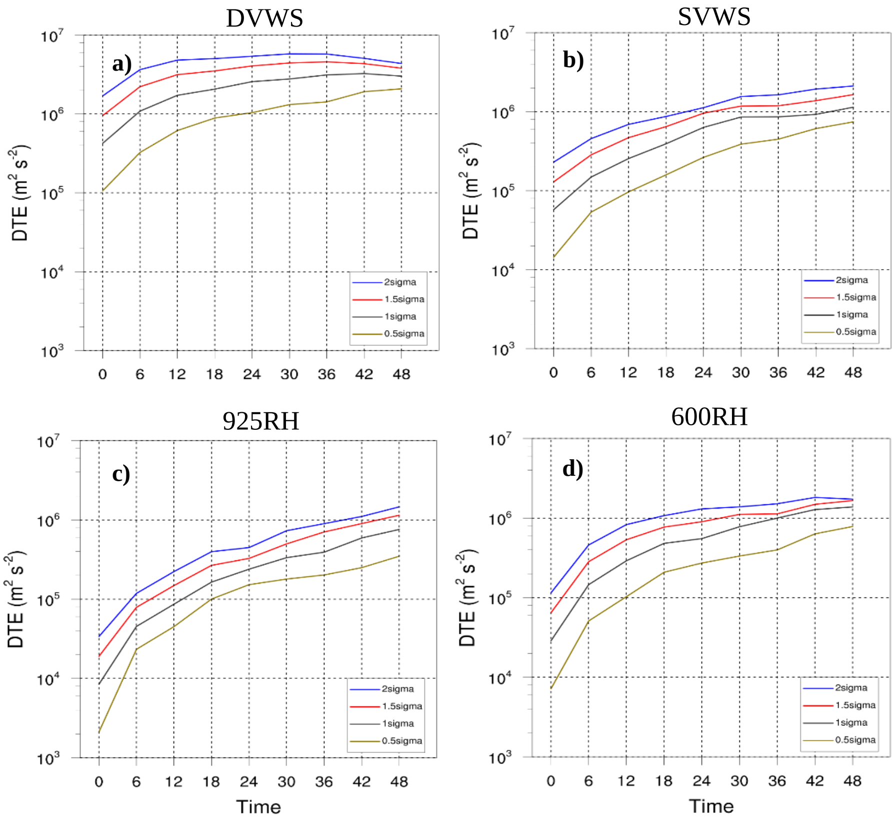

Figure 5.13. Time series of domain integrated difference total energy for the perturbation experiments. DTE time series for the All experiment with perturbation amplitude ranging from 0.5σ to 2σ based on the IC metrics (a) DVWS, (b) SVWS, (c) 925RH and (d) 600RH.

section 5.4.1 we have seen that the cyclones that exhibit low predictability in MSLP are intense, fast-moving and north-landfalling. TC Sidr is also found to be an intense, fast-moving and north-landfalling cyclone with low predictability in both intensity and track.

Perturbation experiments are performed to understand the error growth factors that affect the predictability of TC Sidr. Three different experiments are performed by considering each of the atmospheric fields as the initial condition (IC) metric (x_s) in Eq. 2.38 and 2.39 as in Rios-Berrios, Torn and Davis (2016). In the first experiment perturbations (named All) are applied to all the model state variables through Eq. 2.38, whereas in the second experiment (named Dry) the initial condition perturbations are applied to all variables except the water vapor mixing ratio. And in the third experiment (named Moist) perturbations are applied only to the water vapor mixing ratio. Figure 5.12 shows the DTE obtained for the three experiments by considering DVWS, SVWS, 925RH and 600RH as the IC metrics. From Fig. 12a, it can be seen that the total error growth is nearly the same for All and Dry experiments throughout the forecast period, however, the error growth rapidly increases till forecast hour 6 for Moist experiment and becomes steady afterwards. That is, changes in DVWS are linked to small perturbations in the water vapor mixing ratio. In addition, the DTE values in the Dry experiments are nearly identical to those of the Moist experiments after about 6-h. This suggests that the simulations of TC Sidr have finite predictability, as perturbing the initial conditions with or without the water vapor mixing ratio does not change the total error growth through 48-h lead times. It should be also noticed that the Dry and Moist experiments cause the greatest DTE growth and the All experiment causes the reduced growth for all the IC metrics except DVWS. It is also found that the error growth rate of ICs associated with moist perturbations is higher in the first 6-h. This result hints that the error growth is due to moist convection, which is also consistent with the studies by Zhang *et al.* (2007).

The perturbation experiments are repeated for different perturbation amplitude ranging from 0.5σ to 2σ with σ being the ensemble standard deviation of the respective IC metric (Fig. 13). We notice that for all the IC metrics, the smaller the magnitude of perturbations faster will be error growth. These results are similar to the findings by Lloveras, Tierney and Durran (2022). It is also observed that the error growth rate is higher for near surface humidity at the core of the TC.

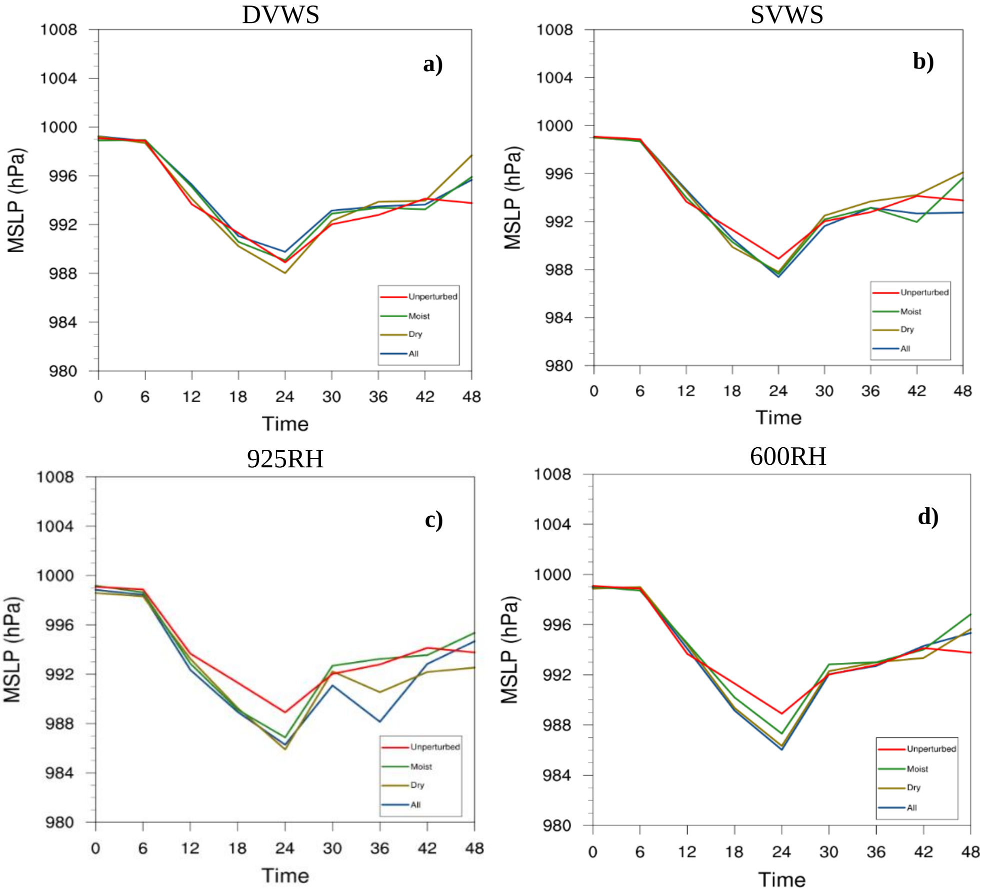

Figure 5.14. The minimum sea level pressure obtained after the All, Moist and Dry perturbation experiments for the IC metrics (a) DVWS, (b) SVWS, (c) 925RH and (d) 600RH.

The MSLP of the TC Sidr obtained after the perturbation experiments are shown in Fig. 14. Generally, there is less difference in MSLP between the All, Moist and Dry experiments. The highest change in MSLP relative to the unperturbed experiment is observed for the near-surface and midtropospheric humidity. That is, small changes to these IC metrics at the core will produce the largest change in the MSLP of TC Sidr rather than the environmental shear factor.

5.5 Summary

This study uses multivariate ensemble sensitivity analysis to understand the general predictability characteristics of landfalling TCs formed across the Bay of Bengal from 2000 to 2020. The ensemble analysis and forecasts for 46 TCs formed over the Bay of Bengal are generated using the WRF and DART-EnKF systems. In this study, the ensemble spread of cyclone central pressure and position at the 48-h forecast time is considered as a proxy for predictability. That is, TCs with large forecast spread exhibits low predictability and those with small ensemble spread shows high predictability. The large forecast spread could be either because of the large initial spread or the large potential for perturbation growth (intrinsic predictability). In this chapter, multivariate ensemble sensitivity analysis is used to obtain the intrinsic predictability. The results obtained from this study are summarized as follows.

The general predictability characteristics analysed using the ensemble sensitivity analysis show that intense, fast-moving and north-landfalling TCs exhibit large forecast spread in intensity and hence low predictability in terms of central pressure of the tropical cyclones. However, weak, fast-moving and north-landfalling TCs show large ensemble spread in track; therefore, it can be stated that TCs with these characteristics exhibit low predictability in track. Comparing analysis spread and ensemble sensitivity results show that cyclones with low predictability in intensity exhibit large potential for intrinsic error growth. Moreover, low predictability is found to be linked to the dynamical perturbation growth associated with TC core features rather than the environmental fields. In addition, it has been found that the low predictability in the track of the fast-moving and north-landfalling TCs is associated with both initial spread in sea level pressure and intrinsic error growth in TC core fields. Further analysis showed that low intensity predictability cyclones exhibit high values of DVWS and low values of SVWS in the environment of the TC. It is also found that the factors that affect the predictability of intensity and track of the TCs are the moisture fields at the core of the TCs rather than environmental shear factors. The high ensemble sensitivity of low predictability cyclones to environmental and core factors reveals the importance of dynamical perturbation growth in the low predictability of cyclones compared to high predictability cyclones. The perturbation experiments prove that small amplitude perturbations to the moisture fields at the core indeed result in low predictability of the intense, faster and north falling TCs. The large ensemble spread in the intensity and track of the TC Sidr indicates that the cyclone has the lowest predictability in the group of tropical cyclones analyzed in this study. Therefore, the predictability characteristics of TC are analysed separately in this chapter. The strong ensemble sensitivity indicates strong potential for perturbation growth that causes low predictability. The perturbation experiments suggest that the error growth rate of initial conditions associated with moist perturbations is higher. It is also found that the smaller the magnitude of initial condition perturbations, the faster will be the error growth. Nevertheless, the largest change in the central pressure of TC Sidr is observed for small changes in the near surface humidity and midtropospheric humidity at the core of the cyclone.

This study has revealed the general characteristics of TCs that exhibit large forecast spread or low predictability and the factors responsible for the low predictability. The large dynamical potential for error growth is found to be the key factor that affects the predictability of the TCs over the Bay of Bengal and such error growth cannot be solved practically. Another factor that affects the predictability is the large spread at the initial time, which can be mitigated by data assimilation methods.

CHAPTER 6

OPTIMAL LOCATIONS FOR TARGETING OBSERVATIONS IDENTIFIED USING ENSEMBLE SENSITIVITY ANALYSIS

6.1 Introduction

Providing dependable rainfall forecasts using NWP models is challenging. One factor that jeopardizes forecast accuracy is uncertainties in the initial conditions. An advanced DA system can effectively use the available observations to improve initial conditions and the subsequent forecast well in advance. However, the accuracy of forecasts is often impaired by the lack of observations. Assimilation of observations from an optimal observation network may maximize the reduction of forecast error. The flow-dependent dynamics within an NWP model play a significant role in deciding the placement of such supplement observational networks, which would produce the largest uncertainty reduction in the forecast. One approach for designing an optimal observational network is sensitivity analysis, which points to the spatial regions where the underlying error growth dynamics are significant. It is expected that additional observations on such locations reduce forecast uncertainty following data assimilation.

Several predictability studies have employed sensitivity analysis to determine how forecast features evolve within numerical models to influence chosen responses (Rabier *et al.*, 1996; Snyder, 1996; Langland, Shapiro and Gelaro, 2002). Theoretically, sensitivity analysis can be used as a tool to understand the locations of potential error growth in the initial conditions (Torn and Hakim, 2008a). Among the sensitivity studies using various methods, adjoint sensitivity and ensemble-based sensitivity analysis have been widely used for understanding error growth dynamics and observation targeting of extreme weather events (Ancell and Hakim, 2007; Torn and Hakim, 2009; Torn, 2010; Xie *et al.*, 2013; Bednarczyk and Ancell, 2015; Zhou and Cui, 2015; Limpert and Houston, 2018; Kumpf *et al.*, 2019). Ancell and Hakim (2007) indicated that the ensemble-based sensitivity analysis is useful for identifying a target region for additional observations because the analysis error statistics are included in the ensemble calculation, unlike adjoint methods. The present study addresses the question "Where should one deploy observations to curtail rapidly growing forecast errors during summer monsoon season for various target forecast domains over the Indian subcontinent?"

In recent decades, the occurrence of extreme rainfall events has increased considerably over India during the summer monsoon season (Rajeevan, Bhate and Jaswal, 2008; Roxy *et al.*, 2017). Therefore, efforts to improve short-range forecasts during the Indian summer monsoon by assimilating optimal observations have much significance in the current context.

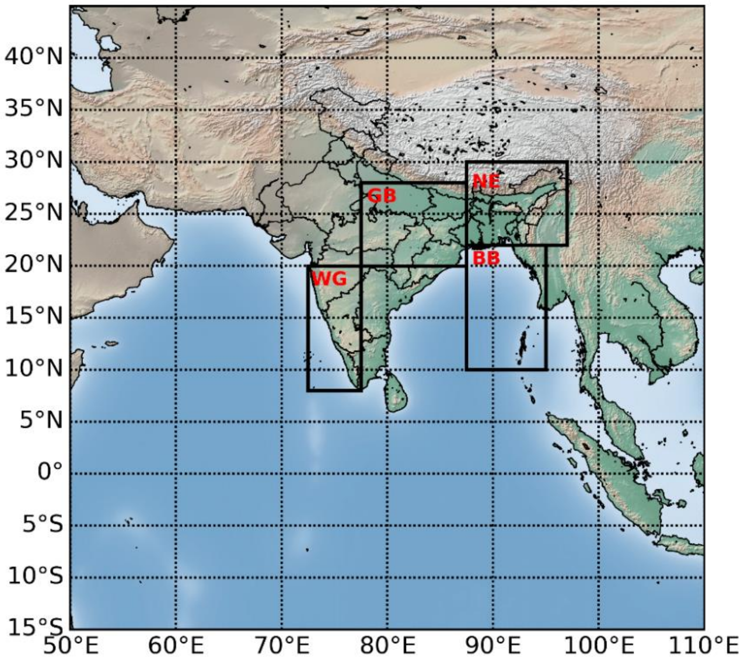

Figure 6.1. Area of study. Boxes in the figure represents the regions over which the forecast metric is area-averaged.

6.2 Data and Methods

Ensemble forecasts from The Observing System Research and Predictability Experiment (THORPEX) Interactive Grand Global Ensemble (TIGGE) are used in the computation of sensitivity analysis. For the current study, we have used the ensemble forecasts from the NCEP TIGGE data archive. This dataset consists of 21 ensemble members with an unperturbed control forecast and 20 perturbed forecasts at a horizontal grid spacing of 0.5° × 0.5°. Potential error growth regions in the initial conditions are estimated by applying ensemble sensitivity analysis using short-range ensemble forecasts initialized from 0000 UTC of 1 to 31, July. A statistically robust estimate of initial condition sensitivity is obtained by conducting the experiments for five years, starting from 2011 to 2015.

In this study, ESA uses accumulated precipitation as the forecast metric in various locations over the Indian subcontinent. These forecast metrics are averaged over four regions: the Western Ghats (WG), the Ganges basin (GB), the Bay of Bengal (BB), and the Northeast (NE) to obtain the climatological sensitivity associated with these four regions (Fig. 6.1). Climatological sensitivity is defined as the percentage of analysis cycles for which the ensemble sensitivity of the forecast metric with respect to a state variable is different from zero at a certain level of confidence. Thus, climatological sensitivity is used to determine how often a forecast metric is sensitive to changes in a state variable. Regions with a high

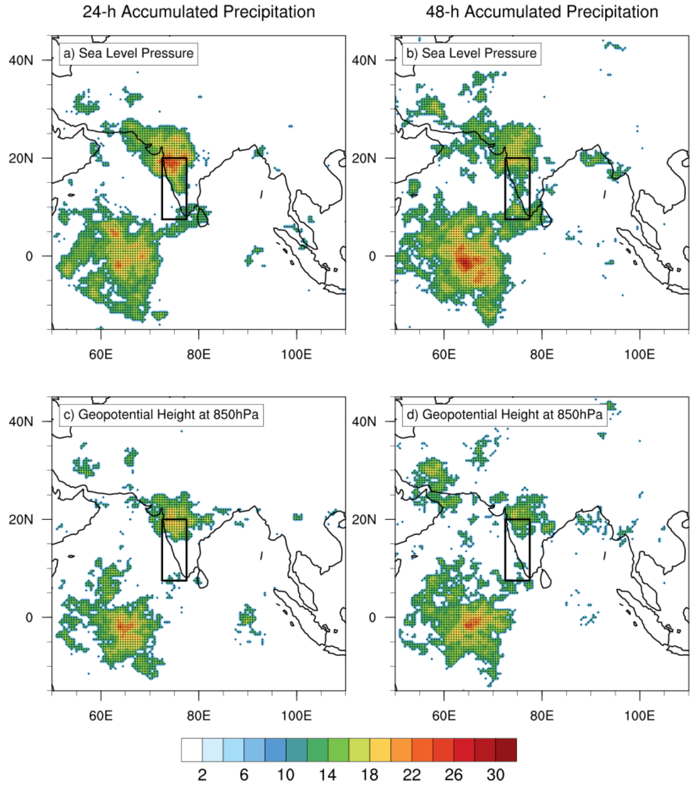

Figure 6.2. Percentage of forecast cycles with grid point sensitivity statistically significant at the 95% confidence level for the sensitivity of 24-h (left panel) and 48-h (right panel) precipitation forecasts over the Western Ghats to SLP (top panel) and 850 hPa geopotential height (bottom panel). The box indicates the region over which the forecast metric is averaged.

percentage of sensitive forecast cycles may be regarded as potential locations for new observations.

6.3 Results and Discussions

Figure 6.2 shows results for the sensitivity of 24-h and 48-h accumulated precipitation forecast averaged in a box to analyses of mean sea level pressure (SLP) and geopotential height at 850 hPa. It can be seen that the region with the largest percentage of forecast cycles is located far south of the forecast region in the Indian Ocean and over the northern edge of the box. Though the sensitivity region remains the same for both forecast metrics, the 48-h precipitation forecasts are more often sensitive to the region south of the Indian Ocean. (Figs. 6.2b, d). The sensitivity of 48-h precipitation forecasts to SLP and 850 hPa geopotential height analyses show similar results with the latter being less sensitive. The sensitivity of precipitation forecasts to 850 hPa wind and humidity is shown in Fig. 6.3. The rainfall over WG is frequently sensitive to 850 hPa wind to regions upstream of the response region (Figs. 6.3a, b). Similarly, the analysis of humidity depicts consistent sensitive regions to the upstream locations of the forecast box (Fig. 6.3c, d). Locations of frequent sensitivity indicate the region where additional observations are required to improve the precipitation forecast over the WG.

The precipitation forecasts in a box over the Bay of Bengal show sensitivity to SLP over regions extending to the south of the box (Fig. 6.4a). For wind at 850 hPa, the region of frequent sensitivity is seen within the box (Fig. 6.4b). The sensitivity to 850 hPa humidity shows a meridionally elongated sensitive region towards the south of the metric box.

The region of frequent sensitivity to precipitation forecasts over Gangetic basins (GB) is shown in Figs. 6.4a and 6.4d. It is noticed that for SLP the precipitation forecasts are frequently sensitive to regions within the box region and no upstream

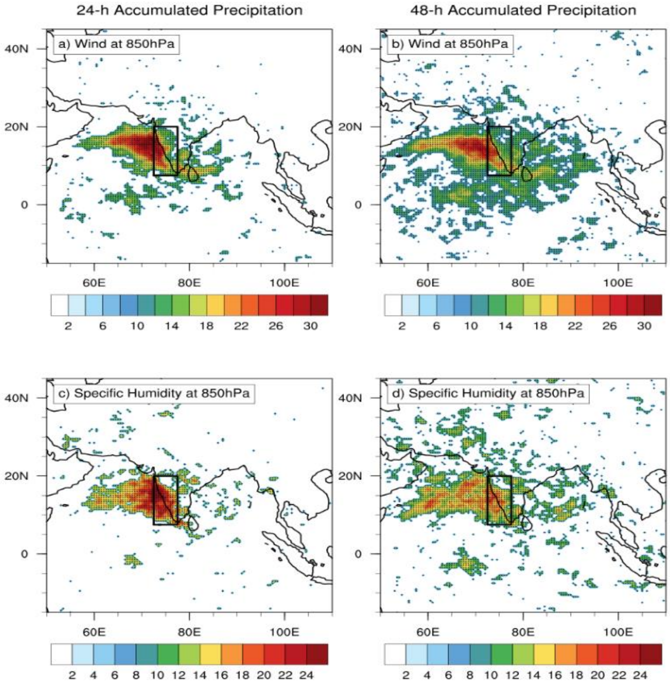

Figure 6.3. Percentage of forecast cycles with grid point sensitivity statistically significant at the 95% confidence level for the sensitivity of 24-h (left panel) and 48-h (right panel) precipitation forecasts over the Western Ghats to 850 hPa wind (top panel) and 850 hPa humidity (bottom panel). The box indicates the region over which the forecast metric is averaged.

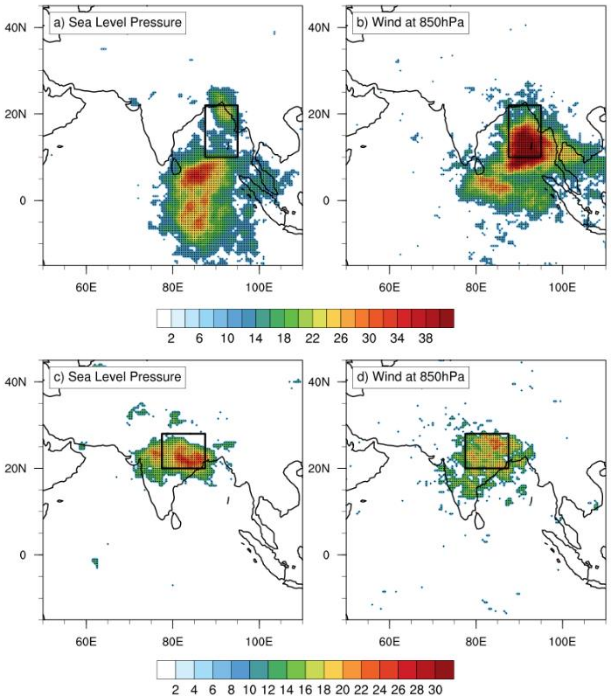

Figure 6.4. Percentage of forecast cycles with grid point sensitivity statistically significant at the 95% confidence level for the sensitivity of 24-h precipitation forecasts over the Bay of Bengal (top panel) and Gangetic basins (bottom panel) to SLP (left panel) and wind at 850 hPa (right panel). The box indicates the region over which the forecast metric is averaged.

shift in the sensitive region is found for forecasts over GB (Fig. 6.4c). A recent study by Nikumbh et al. (2021) shows that the presence of a zonally elongated monsoon trough enhances the vorticity and temperature that sets up large-scale dynamic forcing over central India for medium-range and large extreme rainfall events. The frequently sensitive region is possibly linked to the monsoon trough that extends all the way from Pakistan to head Bay. Similar to the previous result the precipitation over the GB is consistently sensitive to regions within the metric box for wind at 850 hPa, (Fig. 6.4d). Therefore, the precipitation forecasts over the GB would be benefited from additional in situ observations over central India and the zonally elongated sensitive region. Similar results are obtained for the precipitation forecasts over the northeast region of India (Fig. 6.5). A zonally elongated region over the head Bay has been identified as the frequently sensitive region to SLP and 850 hPa geopotential height (Fig. 6.5 a and b).

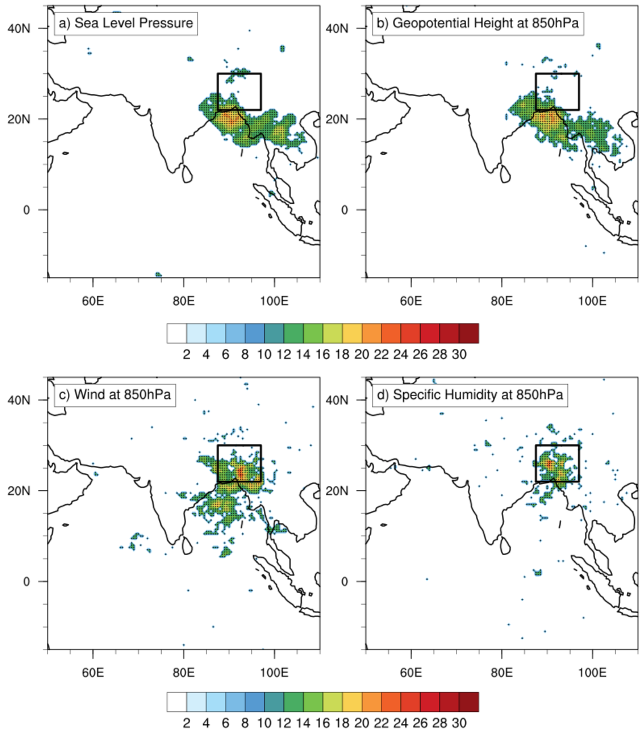

Figure 6.5. Percentage of forecast cycles with grid point sensitivity statistically significant at the 95% confidence level for the sensitivity of 24-h precipitation forecasts over Northeast to (a)SLP, (b)850 hPa geopotential height, (c) 850 hPa wind and (d) 850 hPa humidity. The box indicates the region over which the forecast metric is averaged.

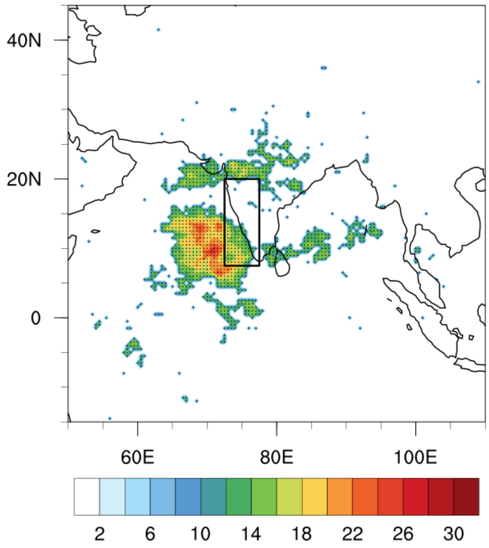

Figure 6.6. Percentage of forecast cycles with grid point sensitivity statistically significant at the 95% confidence level for the sensitivity of 24-h SLP forecasts over the Western Ghats to 850 hPa wind. The box indicates the region over which the forecast metric is averaged.

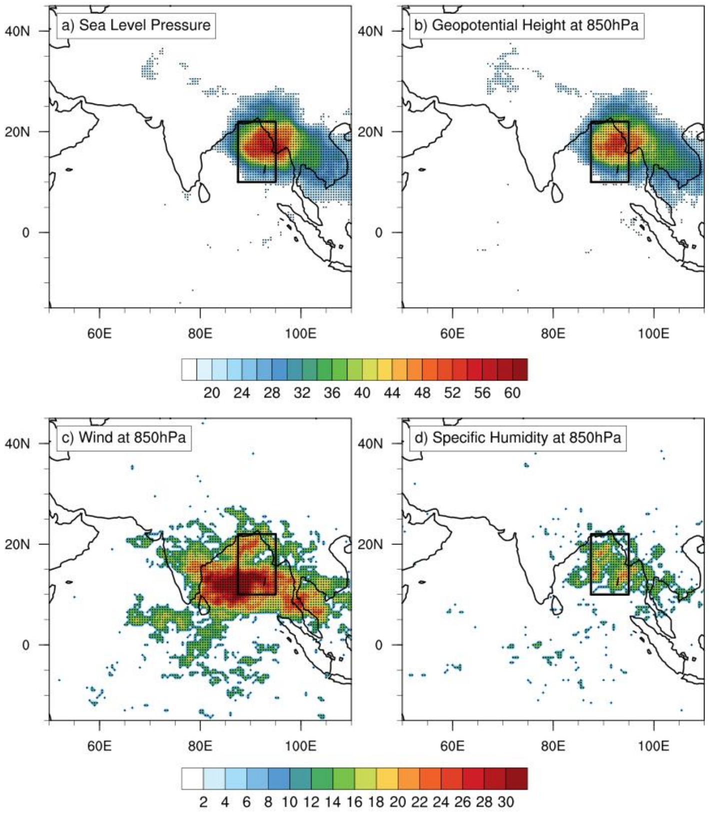

Figure 6.7. Percentage of forecast cycles with grid point sensitivity statistically significant at the 95% confidence level for the sensitivity of 24-h SLP forecasts over the Bay of Bengal to (a)SLP, (b)850 hPa geopotential height, (c) 850 hPa wind and (d) 850 hPa humidity. The box indicates the region over which the forecast metric is averaged.

To study how the sensitivity varies with forecast metrics, the regions associated with SLP forecasts are estimated. Over the WG, the frequent sensitive region shows an upstream shift in the analysis for 850 hPa wind (Fig. 6.6). The SLP forecasts averaged over GB and NE regions show similar regions of sensitivity towards the southeast of the box. This indicates that the forecast of SLP over GB and NE may be benefitted from additional wind measurements over the Bay of Bengal. Figure 6.7a shows the sensitivity of SLP forecasts in the box over the Bay of Bengal to SLP analyses. The forecast metric is sensitive to more than 60% of the time over the region within the box. For 850 hPa wind analyses (Fig. 6.7c), the region with the largest percentage of sensitive forecast cycles is seen over the Bay of Bengal zonally extending from $80^{\circ}E$ to 100°E. Studies conducted by changing the forecast metric box do not show substantial variations in the sensitivity patterns, suggesting that the sensitive regions are independent of the dimension of the forecast metric box.

6.4 Summary

Ensemble analyses and forecasts obtained from NCEP TIGGE are used to identify locations of frequent sensitivity during the summer monsoon rainfall season over the Indian subcontinent. Locations of frequent sensitivity are determined for different forecast metrics over different domains. The 24-h and 48-h precipitation forecasts averaged over the WG are most often sensitive to far south of the target forecast region. It is also found that the assimilation of additional wind and humidity measurements to the upstream regions of the WG could improve the precipitation forecasts in this region. The region of sensitivity for the forecast domain located over the GB and NE is identified to be a zonally elongated region over central India and head Bay, respectively. Therefore, the precipitation forecasts in these regions would be benefited from the assimilation of additional observations over the monsoon trough zone. The region of consistent sensitivity for precipitation forecasts over the Bay of Bengal is found to be far south of the Bay of Bengal. According to Torn and Hakim (2008), regions of consistent sensitivity indicate assimilation of additional observations over those regions would optimally improve the forecast metric over the target domain.

Overall, this study shows that the precipitation forecasts during the Indian summer monsoon season are benefited from the assimilation of observations that are located over the upstream regions of the forecast metric box. It is to be noted that the frequently sensitive region for a target forecast region may vary depending on the flow conditions (season) as well as the NWP model used. The optimal locations for observations need to be tested by assimilating real/synthetic observations in the sensitive regions which are left for future work.

CHAPTER 7

CONCLUSIONS

Numerical Weather Prediction (NWP) is an initial-value problem, and its ability to determine the future state of the atmosphere depends on the accuracy of the initial state. The uncertainties in the initial condition and model formulations limit the predictability of weather using NWP models. Therefore, quantifying these uncertainties in the forecasts is crucial in understanding the relevant dynamics and the predictability of the weather event. One way to quantify the uncertainties in a forecast is through sensitivity analysis, which is defined as how a forecast variable responds to changes of initial conditions. Sensitivity analysis is a reliable method that is employed to quantify the uncertainties imposed by chaos in the forecast, which determines predictability of a weather event. Ensemble Sensitivity Analysis (ESA) uses sample statistics of ensemble forecasts to estimate relationships between forecast metric and initial conditions. The present thesis uses ESA to understand the dynamics and predictability of extreme weather events over the Indian subcontinent with the ensemble analyses and forecasts obtained from the Ensemble Kalman Filter (EnKF) data assimilation system along with TIGGE data from European Centre for Medium-Range Weather Forecasts (ECMWF) and National Centers for Environmental Prediction (NCEP) Ensemble Prediction System.

ESA often uses a diagonal approximation to the multivariate regression, leading it to a simple univariate regression. In Chapter 3, the forecast sensitivity of three extreme rainfall events viz. the torrential flood episode over Uttarakhand in 2013, and the catastrophic rainfall events that happened over Kerala in 2018 (KF18) and 2019 (KF19) are investigated through univariate ESA. For the heavy rainfall event over Uttarakhand in June 2013 univariate ESA indicates that the day 1 precipitation exhibits negative sensitivity to the trough over upstream regions of the storm location while on day 2, the sensitive region is found to be located over the southward intruded branch of the mid–tropospheric trough. It is found that in general, the precipitation forecasts associated with this heavy rainfall event are sensitive to the mid-tropospheric trough and moisture fields. IC perturbed experiments show that the perturbation amplitude is correlated linearly with the predicted change in precipitation, which becomes nonlinear as the forecast length increases. Univariate ESA performed on convection-permitting ensembles show that precipitation over the Uttarakhand is mostly non-convective. However, when the location of the response function box is moved northwestward of the Uttarakhand, the sensitivity patterns show signs of convection. Univariate ESA indicates that greater height fall needs to occur over the Western North Pacifc (WNP) to increase the KF18 precipitation, and any shift in the location of these features may affect the precipitation patterns over Kerala. Further analysis using ESA depicts that the existence of circulations in the WNP has played a significant role in the heavy rainfall event, KF18. Though the synoptic settings remained almost similar for KF19, the underlying dynamics that led to both events are found to be different. ESA reveals that the moisture-laden low-level flow is more substantial for KF19, which has favoured the development of deep convective clouds.

As the univariate ESA used in the previous chapter considers only the diagonal elements of the covariance matrix, it may potentially overestimate the response of a forecast metric to initial conditions due to the presence of sampling error. Therefore, in Chapter 4, the contribution of the full covariance matrix in ensemble sensitivity analysis is studied by applying a multivariate linear regression approach to an extreme rainfall event over Chennai. Both univariate and multivariate ensemble sensitivity is estimated for meso- as well as convection scale ensemble members and the effectiveness of the methods is demonstrated by validating the response against the actual response generated from the perturbed initial condition experiment. Multivariate ensemble sensitivity shows organized sensitivity patterns, while the sensitivity values are found to be broadly distributed in univariate ensemble sensitivity. The performance of univariate and multivariate methods in the convection-permitting scale is examined by using high-resolution ensemble forecasts, and it is found that the multivariate sensitivity with localization substantially improves the estimates in finer scales.

Chapter 5 investigates the general predictability characteristics of tropical cyclones (TCs) over the Bay of Bengal using the multivariate ESA. The chapter also compares the dynamical perturbation growth with the initial-time spread for the least and most predictable TCs. The forecast spread in intensity and track of tropical cyclones before landfall is considered as a proxy for predictability. A set of 90 ensemble analyses and forecasts for each of the 45 tropical cyclones over the Bay of Bengal from 2000 to 2020 is obtained using the Ensemble Kalman Filter data assimilation system. These cyclones are categorized based on the characteristics such as intensity, translational speed, and landfalling location. It is found that intense, fast-moving, and north-landfalling exhibits lower predictability than weak, slow-moving, south-landfalling tropical cyclones. It is also found that the low predictability of storms is associated with large initial condition sensitivity, which indicates a larger potential for error growth. Through the perturbation initial condition experiment, it is found that the dynamical error growth rate is higher if the perturbations are smaller. Additionally, the error growth associated with moist perturbations is higher for the less predictable tropical cyclones.

Since ESA account for uncertainties in the forecast and analysis, the approach can be used to determine where new observations should be deployed to reduce forecast errors. In Chapter 6 we have estimated the regions of potential error growth in the initial conditions and optimal observation locations are identified to curtail rapidly growing forecast errors during summer monsoon season for various target forecast domains over the Indian subcontinent. ESA is successfully applied to estimate the climatological sensitivity of short-range precipitation forecasts during the Indian summer monsoon season. Ensemble forecasts from NCEP TIGGE are used to determine a statistically robust estimate of initial condition sensitivity by conducting the experiments for five years, starting from 2011 to 2015. Results show that the 24-h and 48-h precipitation forecasts averaged over WG are most often sensitive to two regions, one to the far south over the Arabian Sea and the other to the northwest of the metric box. It is also found that additional wind and humidity measurements to the upstream regions of the Western Ghats could improve the precipitation forecasts in this region. A similar analysis is performed for a box over the Gangetic basin, the Northeast, and the Bay of Bengal. In addition, it is noticed that the magnitude of the forecast metric does not change the predictability zones, but a decrease in the percentage of cycles is noticed. To summarize, the potential regions for launching observation networks for targeted data assimilation over the Indian subcontinent have been identified using ESA. It is found that the optimal observation location may vary for each forecast domain and such regions are mostly observed upstream of the forecast domain.

Future directions

ESA is based on the underlying assumption that the relationship between the forecast and the initial state is linear. Therefore, a quantitative examination of the impact of the linearity assumption would be needed to understand the feasibility of the application of ESA for various weather events, especially in cases involving convection. The present study does not address the effect of the assimilation of real observations at sensitive locations. . A benefit of targeting based on this method is that it incorporates knowledge about the dynamical error growth as well as the analysis error, both of which are necessary for a comprehensive targeting algorithm. Hence, an extension of the present study includes targeted assimilation of observations to determine the locations and types of observations that produce the largest reduction in response function variance for various weather systems. In the present study, the predictability characteristics of only the landfalling phase of the TCs over the Bay of Bengal are studied. It would be interesting to study the predictability characteristics of the genesis and deepening phases of the tropical cyclones that develops over the Bay of Bengal. At the end of the road, ensemble sensitivity analysis is a data mining problem, and therefore, advanced machine learning algorithms could be potentially applied to solve the problems associated with the predictability of weather systems in the future.

BIBLIOGRAPHY

1. Aksoy, A., Dowell, D. C. and Snyder, C. (2010) ‘A multicase comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part II: Short-range ensemble forecasts’, *Monthly Weather Review*, 138(4). doi: 10.1175/2009MWR3086.1.

2. Ancell, B. C. and Hakim, G. J. (2007) ‘Interpreting adjoint and ensemble sensitivity toward the development of optimal observation targeting strategies’, *Meteorologische Zeitschrift*, 16(6), pp. 635–642. doi: 10.1127/0941-2948/2007/0250.

3. Ancell, B. C. and McMurdie, L. A. (2013) ‘Ensemble adaptive data assimilation techniques applied to land-falling north American cyclones’, in *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. II)*. doi: 10.1007/978-3-642-35088-7_23.

4. Ancell, B. and Hakim, G. J. (2007) ‘Comparing adjoint- and ensemble-sensitivity analysis with applications to observation targeting’, *Monthly Weather Review*, 135(12), pp. 4117–4134. doi: 10.1175/2007MWR1904.1.

5. Anderson, J. L., T. Hoar, K. Raeder, H. Liu, N. Collins, R. T. and A. A. (2009) ‘he Data Assimilation Research Testbed: A Community Facility.’, *Bulletin of the American Meteorological Society*, 90(September), pp. 1283–1296. doi: 10.1175/2009BAMS26I8.I.

6. Anderson, J. *et al.* (2009) ‘The data assimilation research testbed a community facility’, *Bulletin of the American Meteorological Society*, 90(9). doi: 10.1175/2009BAMS2618.1.

7. Anderson, J. L. (2001) ‘An ensemble adjustment Kalman filter for data assimilation’, *Monthly Weather Review*, 129(12), pp. 2884–2903. doi: 10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2.

8. Anderson, J. L. (2007) ‘An adaptive covariance inflation error correction algorithm for ensemble filters’, *Tellus, Series A: Dynamic Meteorology and Oceanography*, 59(2), pp. 210–224. doi: 10.1111/j.1600-0870.2006.00216.x.

9. Anderson, J. L. (2009) ‘Spatially and temporally varying adaptive covariance inflation for ensemble filters’, *Tellus, Series A: Dynamic Meteorology and Oceanography*, 61 A(1), pp. 72–83. doi: 10.1111/j.1600-0870.2008.00361.x.

10. Andersson, E. (1998) 'The ECMWF implementation of three-dimensional variational assimilation (3D-Var). III. Experimental results', *Quarterly Journal of the Royal Meteorological Society,* 124(550). doi: 10.1256/smsqj.55003.

11. Anthes, R. A. and Warner, T. T. (1978) 'Development of Hydrodynamic Models Suitable for Air Pollution and Other Mesometerological Studies', *Monthly Weather Review,* 106(8). doi: 10.1175/1520-0493(1978)106<1045:dohmsf>2.0.co;2.

12. Ashrit, R. *et al.* (2020) ' Prediction of the August 2018 heavy rainfall events over Kerala with high-resolution NWP models ', *Meteorological Applications,* 27(2), pp. 1–14. doi: 10.1002/met.1906.

13. Attri, S. D. and Tyagi, A. (2010) 'Climate profile of India', *Environment Meteorology, India Meteorological Department*, pp. 1–122.

14. Baisya, H. and Pattnaik, S. (2019) 'Orographic effect and multiscale interactions during an extreme rainfall event', *Environmental Research Communications,* 1(5), p. 051002. doi: 10.1088/2515-7620/ab2417.

15. Barker, D. *et al.* (2012) 'The weather research and forecasting model's community variational/ensemble data assimilation system: WRFDA', *Bulletin of the American Meteorological Society,* 93(6), pp. 831–843. doi: 10.1175/BAMS-D-11-00167.1.

16. Barker, D. M. *et al.* (2004) 'A three-dimensional variational data assimilation system for MM5: Implementation and initial results', *Monthly Weather Review,* 132(4), pp. 897–914. doi: 10.1175/1520-0493(2004)132<0897:ATVDAS>2.0.CO;2.

17. Bednarczyk, C. N. and Ancell, B. C. (2015) 'Ensemble sensitivity analysis applied to a southern plains convective event', *Monthly Weather Review,* 143(1), pp. 230–249. doi: 10.1175/MWR-D-13-00321.1.

18. Belanger, J. I. *et al.* (2012) 'Extended prediction of North Indian Ocean tropical cyclones', *Weather and Forecasting,* 27(3). doi: 10.1175/WAF-D-11-00083.1.

19. Berman, J. D. *et al.* (2017) 'Sensitivity of Northern Great Plains convection forecasts to upstream and downstream forecast errors', *Monthly Weather Review,* 145(6), pp. 2141–2163. doi: 10.1175/MWR-D-16-0353.1.

20. Berner, J. *et al.* (2009) 'A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in

the ECMWF ensemble prediction system', *Journal of the Atmospheric Sciences,* 66(3), pp. 603–626. doi: 10.1175/2008JAS2677.1.

21. Berner, J. *et al.* (2011) 'Model uncertainty in a mesoscale ensemble prediction system: Stochastic versus multiphysics representations', *Monthly Weather Review,* 139(6), pp. 1972–1995. doi: 10.1175/2010MWR3595.1.

22. Bishop, C. H., Etherton, B. J. and Majumdar, S. J. (2001) 'Adaptive sampling with the ensemble transform Kalman filter Part I: Theoretical aspects', *Monthly Weather Review,* 129(3). doi: 10.1175/1520-0493(2001)129<0420:ASWTET>2.0.CO;2.

23. Buehner, M. *et al.* (2010) 'Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part I: Description and single-observation experiments', *Monthly Weather Review,* 138(5). doi: 10.1175/2009MWR3157.1.

24. Buizza, R., Miller, M. and Palmer, T. N. (1999) 'Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System', *Quarterly Journal of the Royal Meteorological Society,* 125(560). doi: 10.1256/smsqj.56005.

25. Buizza, R. and Montani, A. (1999) 'Targeting observations using singular vectors', *Journal of the Atmospheric Sciences,* 56(17). doi: 10.1175/1520-0469(1999)056<2965:TOUSV>2.0.CO;2.

26. Carrasco, C. A., Landsea, C. W. and Lin, Y. L. (2014) 'The influence of tropical cyclone size on its intensification', *Weather and Forecasting,* 29(3). doi: 10.1175/WAF-D-13-00092.1.

27. Chakraborty, A. (2016) 'A synoptic-scale perspective of heavy rainfall over Chennai in November 2015', *Current Science,* 111(1), pp. 201–206. doi: 10.18520/cs/v111/i1/201-207.

28. Charney, J. G., FjÖrtoft, R. and Neumann, J. Von (1950) 'Numerical Integration of the Barotropic Vorticity Equation', *Tellus,* 2(4). doi: 10.3402/tellusa.v2i4.8607.

29. Chen, D. Y. C., Cheung, K. K. W. and Lee, C. S. (2011) 'Some implications of core regime wind structures in western North Pacific tropical cyclones', *Weather and Forecasting,* 26(1). doi: 10.1175/2010WAF2222420.1.

30. Chen, F. and Dudhia, J. (2001) 'Coupling and advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity', *Monthly Weather Review,* 129(4), pp. 569–585. doi: 10.1175/1520-

0493(2001)129<0569:CAALSH>2.0.CO;2.

31. Chevuturi, A. and Dimri, A. P. (2016) ‘Investigation of Uttarakhand (India) disaster-2013 using weather research and forecasting model’, *Natural Hazards*, 82(3), pp. 1703–1726. doi: 10.1007/s11069-016-2264-6.

32. Daley, R. (1991) ‘Atmospheric data analysis’, *Atmospheric data analysis*. doi: 10.4267/2042/51948.

33. DeMaria, M. *et al.* (2014) ‘Is tropical cyclone intensity guidance improving?’, *Bulletin of the American Meteorological Society*, 95(3). doi: 10.1175/BAMS-D-12-00240.1.

34. Dirren, S., Torn, R. D. and Hakim, G. J. (2007) ‘A data assimilation case study using a limited-area ensemble Kalman filter’, *Monthly Weather Review*, 135(4), pp. 1455–1473. doi: 10.1175/MWR3358.1.

35. Dowell, D. C., Wicker, L. J. and Snyder, C. (2011) ‘Ensemble kalman filter assimilation of radar observations of the 8 may 2003 oklahoma city supercell: Influences of reflectivity observations on storm-scale analyses’, *Monthly Weather Review*, 139(1). doi: 10.1175/2010MWR3438.1.

36. Dube, A. *et al.* (2014) ‘Forecasting the heavy rainfall during Himalayan flooding-June 2013’, *Weather and Climate Extremes*, 4(June 2013), pp. 22–34. doi: 10.1016/j.wace.2014.03.004.

37. Dudhia, J. (1989) ‘Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two-dimensional model’, *Journal of the Atmospheric Sciences*, pp. 3077–3107. doi: 10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2.

38. Evans, R. E. *et al.* (2000) ‘Joint medium-range ensembles from the met. office and ECMWF systems’, *Monthly Weather Review*, 128(9). doi: 10.1175/1520-0493(2000)128<3104:JMREFT>2.0.CO;2.

39. Evensen, G. (1994) ‘Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics’, *Journal of Geophysical Research*, 99(C5). doi: 10.1029/94jc00572.

40. Evensen, G. (2003) ‘The Ensemble Kalman Filter: Theoretical formulation and practical implementation’, *Ocean Dynamics*, 53(4), pp. 343–367. doi: 10.1007/s10236-003-0036-9.

41. Evensen, G. (2004) ‘Sampling strategies and square root analysis schemes for the EnKF’, *Ocean Dynamics*, 54(6). doi: 10.1007/s10236-004-0099-2.

42. Fujii, Y. *et al.* (2008) 'Application of singular vector analysis to the Kuroshio large meander', *Journal of Geophysical Research: Oceans*, 113(7). doi: 10.1029/2007JC004476.

43. Gaspari, G. and Cohn, S. E. (1999) 'Construction of correlation functions in two and three dimensions', *Quarterly Journal of the Royal Meteorological Society*, 125(554), pp. 723–757. doi: 10.1256/smsqj.55416.

44. Gelaro, R. *et al.* (1998) 'Sensitivity analysis of forecast errors and the construction of optimal perturbations using singular vectors', *Journal of the Atmospheric Sciences*, 55(6). doi: 10.1175/1520-0469(1998)055<1012:SAOFEA>2.0.CO;2.

45. Goswami, B. N. *et al.* (2006) 'Increasing trend of extreme rain events over India in a warming environment', *Science*, 314(5804). doi: 10.1126/science.1132027.

46. Guo, X. and Tan, Z. M. (2017) 'Tropical cyclone fullness: A new concept for interpreting storm intensity', *Geophysical Research Letters*, 44(9). doi: 10.1002/2017GL073680.

47. Hacker, J. P. and Lei, L. (2015) 'Multivariate ensemble sensitivity with localization', *Monthly Weather Review*, 143(6), pp. 2013–2027. doi: 10.1175/MWR-D-14-00309.1.

48. Hamill, T. M. (2006) *Ensemble-based atmospheric data assimilation, Predictability of Weather and Climate*. doi: 10.1017/CBO9780511617652.007.

49. Hamill, T. M. and Snyder, C. (2000) 'A hybrid ensemble Kalman filter-3D variational analysis scheme', *Monthly Weather Review*, 128(8 II). doi: 10.1175/1520-0493(2000)128<2905:ahekfv>2.0.co;2.

50. Hamill, T. M., Whitaker, J. S. and Snyder, C. (2001) 'Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter', *Monthly Weather Review*, 129(11), pp. 2776–2790. doi: 10.1175/1520-0493(2001)129<2776:DDFOBE>2.0.CO;2.

51. Hanley, K. E. *et al.* (2013) 'Sensitivities of a Squall Line over Central Europe in a Convective-Scale Ensemble', *Monthly Weather Review*, 141(1), pp. 112–133. doi: 10.1175/MWR-D-12-00013.1.

52. Herring, S. C. *et al.* (2016) 'Explaining Extreme Events of 2015 from a Climate Perspective', *Bulletin of the American Meteorological Society*, 97(12), pp. S1–S145. doi: 10.1175/bams-explainingextremeevents2015.1.

53. Hill, A. J., Weiss, C. C. and Ancell, B. C. (2016) 'Ensemble sensitivity

analysis for mesoscale forecasts of dryline convection initiation', *Monthly Weather Review,* 144(11), pp. 4161–4182. doi: 10.1175/MWR-D-15-0338.1.

54. Hong, S. Y., Dudhia, J. and Chen, S. H. (2004) 'A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation', *Monthly Weather Review,* 132(1). doi: 10.1175/1520-0493(2004)132<0103:ARATIM>2.0.CO;2.

55. Hong, S. Y., Noh, Y. and Dudhia, J. (2006) 'A new vertical diffusion package with an explicit treatment of entrainment processes', *Monthly Weather Review*, 134(9), pp. 2318–2341. doi: 10.1175/MWR3199.1.

56. Hou, D., Kalnay, E. and Droegemeier, K. K. (2001) 'Objective verification of the SAMEX '98 ensemble forecasts', *Monthly Weather Review,* 129(1). doi: 10.1175/1520-0493(2001)129<0073:OVOTSE>2.0.CO;2.

57. Houtekamer, P. L. *et al.* (2005) 'Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations', *Monthly Weather Review,* 133(3), pp. 604–620. doi: 10.1175/MWR-2864.1.

58. Houtekamer, P. L. and Mitchell, H. L. (1998) 'Data assimilation using an ensemble Kalman filter technique', *Monthly Weather Review,* 126(3), pp. 796–811. doi: 10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2.

59. Houtekamer, P. L. and Zhang, F. (2016) 'Review of the ensemble Kalman filter for atmospheric data assimilation', *Monthly Weather Review*. doi: 10.1175/MWR-D-15-0440.1.

60. Houze, R. A. *et al.* (2017) 'Multiscale aspects of the storm producing the June 2013 flooding in Uttarakhand, India', *Monthly Weather Review,* 145(11), pp. 4447–4466. doi: 10.1175/MWR-D-17-0004.1.

61. Hunt, K. M. R. and Menon, A. (2020) 'The 2018 Kerala floods: a climate change perspective', *Climate Dynamics*, 54(3–4), pp. 2433–2446. doi: 10.1007/s00382-020-05123-7.

62. Isaksen, L. *et al.* (2010) 'Ensemble of data assimilations at ECMWF', *ECMWF Technical Memorandum,* 636(December), pp. 1–41.

63. Jakobsen, F. *et al.* (2006) 'Cyclone storm surge levels along the Bangladeshi coastline in 1876 and 1960-2000', *Coastal Engineering Journal,* 48(3). doi: 10.1142/S057856340600143X.

64. Jayakrishnan, K. U., Kutty, G. and George, B. (2020) 'On the Predictability and Dynamics of Tropical Cyclone: Nargis (2008)', *Journal of Geophysical Research: Atmospheres,* 125(9), pp. 1–12. doi: 10.1029/2019JD032040.

65. Joseph, S. *et al.* (2015) 'North Indian heavy rainfall event during June 2013: diagnostics and extended range prediction', *Climate Dynamics*, 44(7–8), pp. 2049–2065. doi: 10.1007/s00382-014-2291-5.

66. Kain, J. S. (2004) 'The Kain–Fritsch Convective Parameterization: An Update', *Journal of Applied Meteorology*, 43(1), pp. 170–181. Available at: https://doi.org/10.1175/1520-0450(2004)043.

67. Kalnay, E. (2002) *Atmospheric Modeling, Data Assimilation and Predictability, Atmospheric Modeling, Data Assimilation and Predictability*. doi: 10.1017/cbo9780511802270.

68. Kerr, C. A., Stensrud, D. J. and Wang, X. (2017) 'Verification of convection-allowing model ensemble analyses of near-storm environments using MPEX upsonde observations', *Monthly Weather Review*, 145(3). doi: 10.1175/MWR-D-16-0287.1.

69. Kivalov, S. N. and Fitzjarrald, D. R. (2018) 'Quantifying and Modelling the Effect of Cloud Shadows on the Surface Irradiance at Tropical and Midlatitude Forests', *Boundary-Layer Meteorology*, 166(2), pp. 165–198. doi: 10.1007/s10546-017-0301-y.

70. Klasa, C. *et al.* (2018) 'An evaluation of the convection-permitting ensemble COSMO-E for three contrasting precipitation events in Switzerland', *Quarterly Journal of the Royal Meteorological Society*, 144(712), pp. 744–764. doi: 10.1002/qj.3245.

71. Krishnamurthy, L. *et al.* (2018) 'Causes and probability of occurrence of extreme precipitation events like Chennai 2015', *Journal of Climate*, 31(10), pp. 3831–3848. doi: 10.1175/JCLI-D-17-0302.1.

72. Krishnamurti, T. N. *et al.* (1997) 'Physical initialization and hurricane ensemble forecasts', *Weather and Forecasting*, 12(3). doi: 10.1175/1520-0434(1997)012<0503:PIAHEF>2.0.CO;2.

73. Krishnamurti, T. N. *et al.* (1999) 'Improved weather and seasonal climate forecasts from multimodel superensemble', *Science*, 285(5433). doi: 10.1126/science.285.5433.1548.

74. Krishnamurti, T. N. *et al.* (2017) 'March of buoyancy elements during extreme rainfall over India', *Climate Dynamics*, 48(5–6), pp. 1931–1951. doi: 10.1007/s00382-016-3183-7.

75. Kumar, V. *et al.* (2020) 'Interaction of a low-pressure system, an offshore trough, and mid-tropospheric dry air intrusion: The Kerala flood of august 2018', *Atmosphere*, 11(7). doi: 10.3390/atmos11070740.

76. Kumar, V. and Krishnamurti, T. N. (2016) 'Mesoscale modeling for the rapid movement of monsoonal isochrones', *Atmospheric Science*

Letters, 17(1). doi: 10.1002/asl.617.

77. Kumpf, A. *et al.* (2019) ‘Visual Analysis of the Temporal Evolution of Ensemble Forecast Sensitivities’, *IEEE Transactions on Visualization and Computer Graphics*, 25(1), pp. 98–108. doi: 10.1109/TVCG.2018.2864901.

78. Langland, R. H., Shapiro, M. A. and Gelaro, R. (2002) ‘Initial condition sensitivity and error growth in forecasts of the 25 January 2000 east coast snowstorm’, *Monthly Weather Review*, 130(4), pp. 957–974. doi: 10.1175/1520-0493(2002)130<0957:ICSAEG>2.0.CO;2.

79. Li, J. *et al.* (2014) ‘Ensemble-based analysis and sensitivity of mesoscale forecasts of a vortex over southwest China’, *Quarterly Journal of the Royal Meteorological Society*, 140(680), pp. 766–782. doi: 10.1002/qj.2200.

80. Limpert, G. L. and Houston, A. L. (2018) ‘Ensemble sensitivity analysis for targeted observations of supercell thunderstorms’, *Monthly Weather Review*, 146(6), pp. 1705–1721. doi: 10.1175/MWR-D-17-0029.1.

81. Liu, S. *et al.* (2018) ‘Dynamics and Predictability of the Rapid Intensification of Super Typhoon Usagi (2013)’, *Journal of Geophysical Research: Atmospheres*, 123(14). doi: 10.1029/2018JD028561.

82. Lloveras, D. J., Tierney, L. H. and Durran, D. R. (2022) ‘Mesoscale Predictability in Moist Midlatitude Cyclones Is Not Sensitive to the Slope of the Background Kinetic Energy Spectrum’, *Journal of the Atmospheric Sciences*, 79(1), pp. 119–139. doi: 10.1175/JAS-D-21-0147.1.

83. Lorenz, E. (1963) ‘Deterministic Nonperiodic Flow’, *Journal of the Atmospheric Sciences*, 20(5), pp. 131–140. doi: 10.1175/1520-0469(1963)0202.0.CO;2.

84. Lynch, P. (2008) ‘The ENIAC forecasts: A re-creation’, *Bulletin of the American Meteorological Society*, 89(1), pp. 45–55. doi: 10.1175/BAMS-89-1-45.

85. Mackey, B. P. and Krishnamurti, T. N. (2001) ‘Ensemble forecast of a typhoon flood event’, *Weather and Forecasting*, 16(4). doi: 10.1175/1520-0434(2001)016<0399:EFOATF>2.0.CO;2.

86. Mishra, V. *et al.* (2018) ‘The Kerala flood of 2018: combined impact of extreme rainfall and reservoir storage’, *Hydrology and Earth System Sciences Discussions*, (2017), pp. 1–13. doi: 10.5194/hess-

2018-480.

87. Mishra, V. and Shah, H. L. (2018) ‘Hydroclimatological Perspective of the Kerala Flood of 2018’, *Journal of the Geological Society of India,* 92(5), pp. 645–650. doi: 10.1007/s12594-018-1079-3.

88. Mittermaier, M. P. and Csima, G. (2017) ‘Ensemble versus deterministic performance at the kilometer scale’, *Weather and Forecasting,* 32(5), pp. 1697–1709. doi: 10.1175/WAF-D-16-0164.1.

89. Miyoshi, T., Sato, Y. and Kadowaki, T. (2010) ‘Ensemble Kalman filter and 4D-Var intercomparison with the Japanese operational global analysis and prediction system’, *Monthly Weather Review,* 138(7). doi: 10.1175/2010MWR3209.1.

90. Mlawer, E. J. *et al.* (1997) ‘Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave’, *Journal of Geophysical Research Atmospheres,* 102(14). doi: 10.1029/97jd00237.

91. Mohandas, S. *et al.* (2020) ‘NWP perspective of the extreme precipitation and flood event in Kerala (India) during August 2018’, *Dynamics of Atmospheres and Oceans,* 91(April), p. 101158. doi: 10.1016/j.dynatmoce.2020.101158.

92. Mohanty, A. (2021) ‘Preparing India for extreme climate events’, *Hindustan Times,* (December). Available at: https://www.hindustantimes.com/ht-insight/climate-change/preparing-india-for-extreme-climate-events-101625127345593.html.

93. Molteni, F. and Palmer, T. N. (1993) ‘Predictability and finite-time instability of the northern winter circulation’, *Quarterly Journal of the Royal Meteorological Society,* 119(510). doi: 10.1002/qj.49711951004.

94. Munsell, E. B. *et al.* (2017) ‘Dynamics and predictability of the intensification of hurricane edouard (2014)’, *Journal of the Atmospheric Sciences,* 74(2). doi: 10.1175/JAS-D-16-0018.1.

95. Needham, H. F., Keim, B. D. and Sathiaraj, D. (2015) ‘A review of tropical cyclone-generated storm surges: Global data sources, observations, and impacts’, *Reviews of Geophysics*. doi: 10.1002/2014RG000477.

96. Nikumbh, A. C. *et al.* (2021) ‘Multiscale Interactions between Monsoon Intraseasonal Oscillations and Low Pressure Systems That Produce Heavy Rainfall Events of Different Spatial Extents’, *Journal of Climate*, 34(23). doi: 10.1175/JCLI-D-21-0231.1.

97. Nystrom, R. G. *et al.* (2018) 'Predictability and dynamics of Hurricane Joaquin (2015) explored through convection-permitting ensemble sensitivity experiments', *Journal of the Atmospheric Sciences*, 75(2), pp. 401–424. doi: 10.1175/JAS-D-17-0137.1.

98. Padma, T. V. (2018) 'Mining and dams exacerbated devastating Kerala floods', *Nature*. doi: 10.1038/d41586-018-06145-2.

99. Pai, D. S. *et al.* (2014) 'Development of a new high spatial resolution (0 . 25 ° × 0 . 25 °) Long Period (1901-2010) daily gridded rainfall data set over India and its comparison with existing data sets over the region data sets of different spatial resolutions and time period', 1(January), pp. 1–18.

100. Palmer, T. N. *et al.* (1998) 'Singular vectors, metrics, and adaptive observations', *Journal of the Atmospheric Sciences*, 55(4). doi: 10.1175/1520-0469(1998)055<0633:SVMAAO>2.0.CO;2.

101. Palmer, T. N. (2003) 'Predictability of Weather and Climate: From Theory To Practice – From Days To Decades', 1(1), pp. 1–18. doi: 10.1142/9789812704832_0001.

102. Palmer, T. N. *et al.* (2012) 'Development of a European Multi-Model Ensemble System for Seasonal to Inter-Annual Prediction (DEMETER)', *Bulletin of the American Meteorological Society*, 85(6).

103. Palmer, T. N. and Zanna, L. (2013) 'Singular vectors, predictability and ensemble forecasting for weather and climate', *Journal of Physics A: Mathematical and Theoretical*, 46(25). doi: 10.1088/1751-8113/46/25/254018.

104. Phadtare, J. (2018) 'Role of Eastern Ghats orography and cold pool in an extreme rainfall event over Chennai on 1 December 2015', *Monthly Weather Review*, 146(4), pp. 943–965. doi: 10.1175/MWR-D-16-0473.1.

105. Priya, P. *et al.* (2017) 'Changing monsoon and midlatitude circulation interactions over the Western Himalayas and possible links to occurrences of extreme precipitation', *Climate Dynamics*, 49(7), pp. 2351–2364. doi: 10.1007/s00382-016-3458-z.

106. Rabier, F. *et al.* (1996) 'Sensitivity of forecast errors to initial conditions', *Quarterly Journal of the Royal Meteorological Society*, 122(529), pp. 121–150. doi: 10.1002/qj.49712252906.

107. Rajeevan, M., Bhate, J. and Jaswal, A. K. (2008) 'Analysis of variability and trends of extreme rainfall events over India using 104 years of gridded daily rainfall data'. doi: 10.1029/2008GL035143.

108. Ranalkar, M. R. *et al.* (2016) ‘Dynamical features of incessant heavy rainfall event of June 2013 over Uttarakhand, India’, *Natural Hazards*, 80(3), pp. 1579–1601. doi: 10.1007/s11069-015-2040-z.

109. Ray, K. *et al.* (2021) ‘An assessment of long-term changes in mortalities due to extreme weather events in India: A study of 50 years’ data, 1970–2019’, *Weather and Climate Extremes*, 32. doi: 10.1016/j.wace.2021.100315.

110. Ray, K., Arora, K. and Srivastav, A. K. (2019) ‘Weather extremes and agriculture’, in *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*. doi: 10.5194/isprs-archives-XLII-3-W6-493-2019.

111. Ren, S. *et al.* (2019) ‘Multivariate ensemble sensitivity analysis for super typhoon haiyan (2013)’, *Monthly Weather Review*, 147(9), pp. 3467–3480. doi: 10.1175/MWR-D-19-0074.1.

112. Reshmi Mohan, P. *et al.* (2018) ‘Simulation of a heavy rainfall event over Chennai in Southeast India using WRF: Sensitivity to microphysics parameterization’, *Atmospheric Research*, 210(April), pp. 83–99. doi: 10.1016/j.atmosres.2018.04.005.

113. Rios-Berrios, R., Torn, R. D. and Davis, C. A. (2016) ‘An ensemble approach to investigate tropical cyclone intensification in sheared environments. Part I: Katia (2011)’, *Journal of the Atmospheric Sciences*, 73(1), pp. 71–93. doi: 10.1175/JAS-D-15-0052.1.

114. Rogers, R. (2010) ‘Convective-scale structure and evolution during a high-resolution simulation of tropical cyclone rapid intensification’, *Journal of the Atmospheric Sciences*, 67(1). doi: 10.1175/2009JAS3122.1.

115. Roxy, M. K. *et al.* (2017) ‘A threefold rise in widespread extreme rain events over central India’, *Nature Communications*, 8(1), pp. 1–11. doi: 10.1038/s41467-017-00744-9.

116. Sahany, S., Venugopal, V. and Nanjundiah, R. S. (2010) ‘The 26 July 2005 heavy rainfall event over Mumbai: Numerical modeling aspects’, *Meteorology and Atmospheric Physics*, 109(3), pp. 115–128. doi: 10.1007/s00703-010-0099-3.

117. Van Sang, N., Smith, R. K. and Montgomery, M. T. (2008) ‘Tropical-cyclone intensification and predictability in three dimensions’, *Quarterly Journal of the Royal Meteorological Society*, 134(632). doi: 10.1002/qj.235.

118. Schumacher, R. S. *et al.* (2013) ‘Factors influencing the development and maintenance of nocturnal heavy-rain-producing convective

systems in a storm-scale ensemble', *Monthly Weather Review*, 141(8), pp. 2778–2801. doi: 10.1175/MWR-D-12-00239.1.

119. Shutts, B. G. (2005) 'A kinetic energy backscatter algorithm for use in ensemble prediction systems', pp. 3079–3102. doi: 10.1256/qj.04.106.

120. Singh, K. S. *et al.* (2018) 'Prediction of heavy rainfall over Chennai Metropolitan City, Tamil Nadu, India: Impact of microphysical parameterization schemes', *Atmospheric Research*, 202, pp. 219–234. doi: 10.1016/j.atmosres.2017.11.028.

121. Singh, O. P. (2007) 'Long-term trends in the frequency of severe cyclones of Bay of Bengal : Observations and simulations', *MAUSAM*, 58(1), pp. 59–66. doi: 10.54302/mausam.v58i1.1131.

122. Singh, T., Mitta, R. and Upadhyaya, H. C. (2015) 'Ensemble adjustment kalman filter data assimilation for a global atmospheric model', in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. doi: 10.1007/978-3-319-25138-7_26.

123. Skamarock WC, et al. (2008) 'A description of the advanced research WRF version 3, NCAR Tech. Note, NCAR/TN-468+STR', *Natl. Cent. for Atmos. Res. Boulder, Colorado,* (June).

124. Snyder, C. (1996) 'Summary of an Informal Workshop on Adaptive Observations and FASTEX', *Bulletin of the American Meteorological Society*, 77(5), pp. 953–961. Available at: http://www.jstor.org/stable/26233173.

125. Srinivas, C. V. *et al.* (2018) 'Simulation of an extreme heavy rainfall event over Chennai, India using WRF: Sensitivity to grid resolution and boundary layer physics', *Atmospheric Research*, 210(April), pp. 66–82. doi: 10.1016/j.atmosres.2018.04.014.

126. Stensrud, D. J. *et al.* (1999) 'Using ensembles for short-range forecasting', *Monthly Weather Review*, 127(4). doi: 10.1175/1520-0493(1999)127<0433:UEFSRF>2.0.CO;2.

127. Sudheer, K. P. *et al.* (2019) 'Role of dams on the floods of August 2018 in Periyar River Basin, Kerala', *Current Science*, 116(5), pp. 780–794. doi: 10.18520/cs/v116/i5/780-794.

128. Szunyogh, I. *et al.* (2008) 'A local ensemble transform Kalman filter data assimilation system for the NCEP global model', *Tellus, Series A: Dynamic Meteorology and Oceanography*, 60 A(1). doi: 10.1111/j.1600-0870.2007.00274.x.

129. Tao, D. and Zhang, F. (2014) 'Effect of environmental shear, sea-

surface temperature, andambient moisture on the formation and predictability oftropical cyclones: An ensemble-mean perspective', *Journal of Advances in Modeling Earth Systems*, 6, pp. 384–404. doi: 10.1002/2014MS000314.Received.

130. Thompson, G., Rasmussen, R. M. and Manning, K. (2004) 'Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis', *Monthly Weather Review*, 132(2). doi: 10.1175/1520-0493(2004)132<0519:EFOWPU>2.0.CO;2.

131. Tong, M. and Xue, M. (2008) 'Simultaneous estimation of microphysical parameters and atmospheric state with simulated radar data and ensemble square root Kalman filter. Part I: Sensitivity analysis and parameter identifiability', *Monthly Weather Review*, 136(5). doi: 10.1175/2007MWR2070.1.

132. Torn, R. D. (2010) 'Ensemble-based sensitivity analysis applied to African easterly waves', *Weather and Forecasting*, 25(1), pp. 61–78. doi: 10.1175/2009WAF2222255.1.

133. Torn, R. D. (2016) 'Evaluation of atmosphere and ocean initial condition uncertainty and stochastic exchange coefficients on ensemble tropical cyclone intensity forecasts', *Monthly Weather Review*, 144(9), pp. 3487–3506. doi: 10.1175/MWR-D-16-0108.1.

134. Torn, R. D. and Hakim, G. J. (2008a) 'Ensemble-based sensitivity analysis', *Monthly Weather Review*, 136(2), pp. 663–677. doi: 10.1175/2007MWR2132.1.

135. Torn, R. D. and Hakim, G. J. (2008b) 'Performance characteristics of a pseudo-operational ensemble Kalman filter', *Monthly Weather Review*, 136(10). doi: 10.1175/2008MWR2443.1.

136. Torn, R. D. and Hakim, G. J. (2009) 'Initial condition sensitivity of Western Pacific extratropical transitions determined using ensemble-based sensitivity analysis', *Monthly Weather Review*, 137(10), pp. 3388–3406. doi: 10.1175/2009MWR2879.1.

137. Toth, Z. and Kalnay, E. (1993) 'Ensemble Forecasting at NMC: The Generation of Perturbations', *Bulletin of the American Meteorological Society*, 74(12). doi: 10.1175/1520-0477(1993)074<2317:efantg>2.0.co;2.

138. Toth, Z. and Kalnay, E. (1997) 'Ensemble forecasting at NCEP and the breeding method', *Monthly Weather Review*, 125(12). doi: 10.1175/1520-0493(1997)125<3297:EFANAT>2.0.CO;2.

139. Tracton, M. S. and Kalnay, E. (1993) 'Operational ensemble

prediction at the National Meteorological Center: practical aspects', *Weather & Forecasting*, 8(3). doi: 10.1175/1520-0434(1993)008<0379:OEPATN>2.0.CO;2.

140. Vellore, R. K. *et al.* (2016) 'Monsoon-extratropical circulation interactions in Himalayan extreme rainfall', *Climate Dynamics*, 46(11–12), pp. 3517–3546. doi: 10.1007/s00382-015-2784-x.

141. Vijaykumar, P. *et al.* (2021) 'Kerala floods in consecutive years - Its association with mesoscale cloudburst and structural changes in monsoon clouds over the west coast of India', *Weather and Climate Extremes*, 33, p. 100339. doi: 10.1016/j.wace.2021.100339.

142. Viswanadhapalli, Y. *et al.* (2019) 'A diagnostic study of extreme precipitation over Kerala during August 2018', *Atmospheric Science Letters*, 20(12), pp. 1–10. doi: 10.1002/asl.941.

143. Whitaker, J. S. *et al.* (2004) 'Reanalysis without radiosondes using ensemble data assimilation', *Monthly Weather Review*, 132(5). doi: 10.1175/1520-0493(2004)132<1190:RWRUED>2.0.CO;2.

144. Whitaker, J. S. *et al.* (2008) 'Ensemble data assimilation with the NCEP global forecast system', *Monthly Weather Review*, 136(2). doi: 10.1175/2007MWR2018.1.

145. Whitaker, J. S. and Hamill, T. M. (2002) 'Ensemble data assimilation without perturbed observations', *Monthly Weather Review*, 130(7). doi: 10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2.

146. Wilks, D. S. (2007) *Statistical methods in the atmospheric sciences, second edition, Meteorological Applications*.

147. Wobus, R. L. and Kalnay, E. (1995) 'Three Years of Operational Prediction of Forecast Skill at NMC', *Monthly Weather Review*, 123(7). doi: 10.1175/1520-0493(1995)123<2132:tyoopo>2.0.co;2.

148. Xie, B. *et al.* (2013) 'Observing strategy and observation targeting for tropical cyclones using ensemble-based sensitivity analysis and data assimilation', *Monthly Weather Review*, 141(5), pp. 1437–1453. doi: 10.1175/MWR-D-12-00188.1.

149. Xu, J. and Wang, Y. (2010) 'Sensitivity of the simulated tropical cyclone inner-core size to the initial vortex size', *Monthly Weather Review*, 138(11). doi: 10.1175/2010MWR3335.1.

150. Yu, H. and Meng, Z. (2016) 'Key synoptic-scale features influencing the high-impact heavy rainfall in Beijing, China, on 21 July 2012', *Tellus, Series A: Dynamic Meteorology and Oceanography*, 68(1). doi: 10.3402/tellusa.v68.31045.

151. Zack, J. *et al.* (2010) 'Application of ensemble sensitivity analysis to observation targeting for short-term wind speed forecasting in the Tehachapi region winter season', p. 57.

152. Zhang, F. *et al.* (2007) 'Mesoscale predictability of moist baroclinic waves: Convection-permitting experiments and multistage error growth dynamics', *Journal of the Atmospheric Sciences*, 64(10), pp. 3579–3594. doi: 10.1175/JAS4028.1.

153. Zhang, F. and Sippel, J. A. (2009) 'Effects of moist convection on hurricane predictability', *Journal of the Atmospheric Sciences*, 66(7), pp. 1944–1961. doi: 10.1175/2009JAS2824.1.

154. Zhang, F., Snyder, C. and Rotunno, R. (2003) 'Effects of moist convection on mesoscale predictability', *Journal of the Atmospheric Sciences*, 60(9), pp. 1173–1185. doi: 10.1175/1520-0469(2003)060<1173:EOMCOM>2.0.CO;2.

155. Zhang, F., Snyder, C. and Sun, J. (2004) 'Impacts of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman filter', *Monthly Weather Review*, 132(5). doi: 10.1175/1520-0493(2004)132<1238:IOIEAO>2.0.CO;2.

156. Zhang, F. and Tao, D. (2013) 'Effects of vertical wind shear on the predictability of tropical cyclones', *Journal of the Atmospheric Sciences*, 70(3), pp. 975–983. doi: 10.1175/JAS-D-12-0133.1.

157. Zhang, Z. and Krishnamurti, T. N. (1999) 'A perturbation method for hurricane ensemble predictions', *Monthly Weather Review*, 127(4). doi: 10.1175/1520-0493(1999)127<0447:APMFHE>2.0.CO;2.

158. Zhou, F. and Cui, X. (2015) 'The adjoint sensitivity of heavy rainfall to initial conditions in debris flow areas in China', *Atmospheric Science Letters*, 16(4), pp. 485–491. doi: 10.1002/asl.586.

LIST OF PUBLICATIONS

This thesis is based on the following publications:

1. George, B., Kutty, G. (2022). “Multivariate ensemble sensitivity analysis applied for an extreme rainfall over Indian subcontinent.” **Atmospheric Research** 277. 1-14.
2. George, B., Kutty, G. (2022). “Sensitivity analysis applied to two extreme rainfall events over Kerala using TIGGE ensembles.” **Meteorology and Atmospheric Physics** 134, 1–14.
3. George, B., Kutty, G. (2021). “Ensemble sensitivity analysis of an extreme rainfall event over the Himalayas in June 2013.” **Dynamics of Atmosphere and Ocean**. 93.
4. George, B., Kutty, G. “General predictability characteristics of landfalling tropical cyclones over the Bay of Bengal using multivariate ensemble sensitivity analysis.” **(under preparation)**
5. George, B., Kutty, G. “Applying ensemble sensitivity analysis for identifying the optimal locations for observations in targeted data assimilation to reduce the forecast errors during Indian Summer Monsoon.” **(under preparation)**

Spin-off publications:

1. K. U. Jayakrishnan, G. Kutty, George B. (2020). *“On the Predictability and Dynamics of Tropical Cyclone: Nargis (2008).”* **Journal of Geophysical Research: Atmospheres**. https://doi.org/10.1029/2019JD032040.

Conferences:

1. George, B., Kutty, G. “Multivariate Ensemble Sensitivity Analysis for an Extreme Weather Event Over Indian Subcontinent” paper presented at **EMS Annual Meeting**, Bonn, Germany, 4 – 9 September 2022.

2. George, B., Kutty, G. “Understanding the Dynamics of a Heavy Rainfall Event using Multivariate Ensemble Sensitivity Analysis” paper presented at **EGU General Assembly 2021**, 19 - 30 April 2021.
3. George, B., Kutty, G. “Sensitivity Analysis of Two Extreme Rainfall Events Over Southwestern Coast of India using Ensemble Statistics” paper presented at **TROPMET - 2020**, Virtual Symposium on Weather and Climate Services over Regions over Mountainous Regions, North Eastern Space Applications Centre (NESAC), Shillong, India, 14 - 17 December 2020.
4. George, B., Kutty, G. “An Ensemble-based Sensitivity Analysis of Two Extreme Rainfall Events over Kerala” paper presented at **International conference on Ensemble Methods in Modeling and Data Assimilation (EMMDA – 2020)**, National Center for Medium Range Weather Forecasting under the Ministry of Earth Sciences , Noida, India, 24 – 26 February 2020
5. George, B., Kutty, G. *“Application of Ensemble Sensitivity Analysis to a Heavy Rainfall Event over Himalayan Foothills”* paper presented at **TROPMET - 2018**, National Symposium on “Understanding Weather and Climate Variability: Research for Society”, Banaras Hindu University, Varanasi, India, 24 – 27 October 2018