MAUSAM, **72**, 1 (January 2021), 87-106

551.515.2 : 551.509.314 (267)

Evolution of tropical cyclone forecasts of dynamical - Statistical Cyclone Prediction System (CPS) over the North Indian Ocean during the decade (2010-2019)

S. D. KOTAL and S. K. BHATTACHARYA

India Meteorological Department, Ministry of Earth Sciences, Lodi Road, New Delhi – 110 003, India

e mail : sdkotal.imd@gmail.com

सार – उत्तरी हिंद महासागर (NIO) में उष्णकटिबंधीय चक्रवात (TC) का पूर्वानुमान देने में पिछले एक दशक (2010-2019) के दौरान काफी सुधार हुआ है। इस सुधार का श्रेय काफी हद तक संख्यात्मक मौसम पूर्वानुमान (NWP) मॉडल और गतिकीय-सांख्यिकीय चक्रवात पूर्वानुमान प्रणाली (CPS) को जाता है। चक्रवात पूर्वानुमान प्रणाली (CPS) के पांच घटक हैं (*i*) जेनेसिस पोटेंशियल पैरामीटर (GPP) (*ii*) मार्ग पूर्वानुमान के लिए मल्टी-मॉडल एनसेंबल (MME) तकनीक (*iii*) संख्यात्मक चक्रवात तीव्रता पूर्वानुमान (SCIP) (*iv*) तेजी से तीव्रीकरण (RI) और (*v*) स्थल प्रवेश के बाद तीव्रता के पूर्वानुमान के लिए क्षय मॉडल। दशक के दौरान इस सुधार की समीक्षा करने के लिए चक्रवात पूर्वानुमान प्रणाली (CPS) और संख्यात्मक मौसम पूर्वानुमान मॉडल के प्रदर्शन का आकलन किया गया है। पूर्वानुमान विश्लेषण से पता चला है कि GPP के लिए गलत चेतावनी का अनुपात (0.13), संसूचन की उच्च संभावना (0.95) और उच्च महत्वपूर्ण सफलता सूचकांक (0.84) है। MME की औसत ट्रैक त्रुटि 24 घंटे में 67 कि.मी. से 120 घंटे में 246 कि.मी. तक हो गई है और NWP मॉडल से इसे लगभग 30% कम बताया गया है। 2010-2019 के दौरान पूर्वानुमान में एमएमई त्रुटि भी 24 से 48 घंटे के लिए 52% से 55% और 72 घंटे से 96 घंटे के लिए 41% से 24% तक कम हो गई है। औसत स्थल प्रवेश बिंदु त्रुटि 24 घंटे में 31 कि.मी. से 120 घंटे में 127 किमी के मध्य रही और स्थल प्रवेश की समय त्रुटि 24 घंटे में 2.2 घंटे की तथा 120 घंटे में 8.1 घंटे की सीमा में रही। SCIP मॉडल की औसत तीव्रता त्रुटि 24 घंटे में 8.3 kt तथा 120 घंटे में 12.6 kt तक रही। संभाव्य तेजी से तीव्रता सूचकांक (RII) ने RI पूर्वानुमान के लिए एक अच्छा बैरियर स्कोर (BS) 0.079 हासिल किया। क्षय की औसत तीव्रता में त्रुटि (स्थल प्रवेश के बाद) 6 घंटे में 6.6 kt से 24 घंटे में 3.3 kt पर आ गई। NWP मॉडलों के पूर्वानुमानों में भी सुधार हुआ है लेकिन MME ने सभी मॉडलों से बेहतर प्रदर्शन किया है। पिछले एक दशक में उत्तरी हिंद महासागर में चक्रवात के पूर्वानुमान में सुधार के परिणाम CPS की भूमिका प्रदर्शित करते हैं।

ABSTRACT. Tropical cyclone (TC) forecasting over the North Indian Ocean (NIO) has been improved significantly during past decade (2010-2019). The improvement is largely attributed to the numerical weather prediction (NWP) models and dynamical-statistical Cyclone Prediction System (CPS). The CPS has five components namely, (*i*) Genesis Potential Parameter (GPP), (*ii*) Multi-Model Ensemble (MME) technique for track prediction, (*iii*) Statistical Cyclone Intensity Prediction (SCIP), (*iv*) Rapid intensification (RI) and (*v*) Decay model for intensity forecast after landfall. Performance of CPS and NWP models is assessed to review the improvement during the decade. Forecast analysis revealed low false alarm ratio (0.13), high probability of detection (0.95) and high critical success index (0.84) for GPP. Mean track error of MME was ranged from 67 km at 24h to 246 km at 120 h and about 30% less than NWP models. The MME error has also reduced by 52% to 55% for 24 h to 48 h and 41% to 24% for 72 h to 96 h forecast during 2010-2019. Mean landfall point error was ranged from 31 km at 24 h to 127 km at 120 h and landfall time error was ranged from 2.2 h at 24 h to 8.1 h at 120 h. Mean intensity error of SCIP model was ranged from 8.3 kt at 24 h to 12.6 kt at 120 h. Probabilistic rapid intensification index (RII) achieved a good Brier score (BS) 0.079 for RI forecasting. Mean decaying intensity error (after landfall) was ranged from 6.6 kt at 6 h to 3.3 kt at 24 h. There has been improvement of forecasts for NWP models also but MME outperformed all models. Results demonstrate the role of CPS for improvement of cyclone forecast over the NIO in past decade.

Key words – Tropical cyclone, Cyclone Prediction System (CPS), Dynamical model, Forecast verification, North Indian Ocean.

1. Introduction

In recent years, human casualties due to tropical cyclones (TCs) over the densely populated coastal area surrounding the North Indian Ocean (NIO) have significantly decreased due to improved forecast accuracy. Super cyclone of Orissa (1999), the severest one during the recent time over the Bay of Bengal experienced wind speed of about 250 kmph. The massive destruction caused by the cyclone was the collapse of nearly 4 lakh houses,

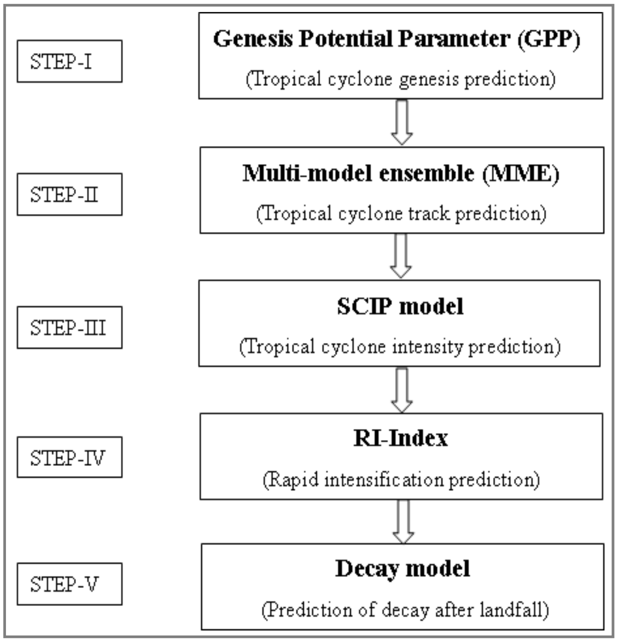


Fig. 1. Flow diagram of Cyclone Prediction System (CPS)

damage of about 19 lakh houses, took a toll of nearly 10000 human lives and affected more than 25 lakh people. This devastating cyclone illustrates the need for accurate prediction of track and intensity. Despite skilled and well experienced, the forecasters were not well equipped with sophisticated NWP system and reliable tool that presently available. It is not possible to forecast track and probable landfall point and intensity five days ahead using synoptic method without reliable outputs from NWP models.

For all operationally designated tropical cyclones in the NIO, India Meteorological Department (IMD) has the responsibility to issue official forecast of TC centre (location) and intensity (maximum 3-minute surface wind speed). Over the past decade, the performance of NWP models in forecasting TC tracks has improved significantly with sophisticated numerical techniques (new observing systems, advanced data assimilation algorithms, physical parameterizations and advancements in computing power) and by the phenomenal increase in satellite observations.

In various studies, efforts were being made (Kotal and Roy Bhowmik, 2011, 2013; Kotal *et al*., 2008, 2009; Roy Bhowmik *et al*., 2005, 2007; Roy Bhowmik, 2003) towards the development of dynamical-statistical methods to add skill to NWP models and to aid operational cyclone forecasting service over the NIO. A NWP based cyclone prediction system (CPS) (Kotal *et al*., 2014) has been introduced for operational TC forecasting. Guidance from CPS has been used in real-time forecasting. The CPS has five components, namely (*i*) genesis prediction by genesis potential parameter (GPP), (*ii*) track prediction by multi-model ensemble (MME), (*iii*) intensity prediction by SCIP model, (*iv*) probability of rapid intensification by rapid intensification index (RII) and (*v*) prediction of decaying intensity after the landfall by decay model. The flow diagram of the five-step CPS is shown in Fig. 1.

The major rationales behind forecast verification (Brier and Allen, 1951) are:

(*i*) To understand the quality of forecasts quantitably in order to assess the improvement of the forecasting system over time.

(*ii*) To understand the strengths and weaknesses of the forecast system.

(*iii*) To improve the forecasting model in its structure.

(*iv*) To understand the probable forecast errors, required to issue better guidance (Jarrell, 1978).

The purpose of this study was to carry out a systematic evaluation of real-time forecast of NWP models and CPS for TCs over the NIO during the past decade (2010-2019) to reveal the improvement of forecasts and efficacy of each component of CPS. All verifications in this paper include the depression stages and above as classified in the IMD best track data. All other stages of TCs (*e.g.*, low pressure system, well marked low pressure system) were excluded.

The paper is organized as follows: The description of data used in this study is given in Section 2. A brief history of track and intensity forecast is presented in Section 3. Section 4 briefly described the cyclone prediction system (CPS). Forecast performance of the CPS is presented in Section 5. Section 6 summarizes the main conclusions.

2. Data

Different agencies provide their own best track data for TCs over the NIO. The Joint Typhoon Warning Center (JTWC, dataset can be found at http://www.usno.navy.mil/NOOC/nmfc-ph/RSS/jtwc/best_tracks/ and the Regional Specialized Meteorological Centre (RSMC) in New Delhi dataset can be found at http://www.rsmcnewdelhi.imd.gov.in/index.php?option=com_ content &view=article&id=48&Itemid=194&lang=en. The TC position and intensity information generally differ among these agencies due to the lack of sufficient surface observations for TCs, as well as the different techniques

TABLE 1

Classification of tropical disturbances over the North Indian Ocean

S. No.	T. No.	Classification of Cyclonic Disturbance	Wind speed (Knots)	Wind criteria (Knots)	Pressure depth (Δp) hPa
1.	T1.0	Low (L)	-	<17	-
2.	T1.5	Depression (D)	25	17-27	-
3.	T2.0	Deep Depression (DD)	30	28-33	4.5
4.	T2.5	Cyclonic Storm (CS)	35	34-47	6.1
5.	T3.0	Cyclonic Storm (CS)	45	34-47	10.0
6.	T3.5	Severe Cyclonic Storm (SCS)	55	48-63	15.0
7.	T4.0	Very Severe Cyclonic Storm (VSCS)	65	64-89	20.9
8.	T4.5	Very Severe Cyclonic Storm (VSCS)	77	64-89	29.4
9.	T5.0	Extremely Severe Cyclonic Storm (ESCS)	90	90-119	40.2
10.	T5.5	Extremely Severe Cyclonic Storm (ESCS)	102	90-119	51.6
11.	T6.0	Extremely Severe Cyclonic Storm (ESCS)	115	90-119	65.6
12.	T6.5	Super Cyclonic Storm (SuCS)	127	≥120	80.0
13.	T7.0	Super Cyclonic Storm (SuCS)	140	≥120	97.2
14.	T7.5	Super Cyclonic Storm (SuCS)	155	≥120	119.1
15.	T8.0	Super Cyclonic Storm (SuCS)	170	≥120	143.3

used to estimate the position and intensity of a TC and also due to different skill of analysts (Lee *et al*., 2012; Kotal *et al*., 2018). RSMC New Delhi is a regional specialized meteorological centre for analysis and forecasting of TCs over the NIO within the framework of the World Weather Watch Programme of the World Meteorological Organization. Therefore, best-track data of RSMC-New Delhi was used for evaluation in this study.

Forecasts were evaluated by comparing the forecast positions and intensities to the corresponding IMD best track positions and intensities for each cyclone. Satellite based "Dvorak technique" (SDT) is used to estimate the current intensity (CI) in terms of "T numbers (T. No.)" of TCs by analyzing satellite image patterns (Dvorak, 1972, 1975, 1984) and infrared cloud-top temperatures (Dvorak, 1984, 1995). The T. No. and corresponding maximum wind speed are shown in Table 1 as per convention of RSMC, New Delhi.

In addition to the five components of CPS, the TC position forecast errors from seven regional and global NWP models were also evaluated in this study. These seven regional and global models were European Center for Medium range Weather Forecasts (ECMWF), NCEP-GFS (National Center for Environmental Prediction), Japan Meteorological Agency (JMA), UK Met Office (UKMO), IMD-GFS, IMD-HWRF and IMD-WRF. Detail of these model data have already been discussed by Kotal *et al*. (2014).

Non-hydrostatic meso-scale WRF model and the Global Forecast System (GFS) are being run for operational forecasting in IMD from 2010. Therefore in this study, the forecast verification was evaluated for past one decade from 2010 to 2019. The forecast lead time of IMD-MME was extended from 72 h to 120 h in the year 2013. Therefore, the mean track forecast errors of MME from 84 h to 120 h were evaluated from 2013. Similarly, the intensity forecast errors from 84 h to 120 h were evaluated for the period 2015-2019 as intensity forecast was extended from 72 h to 120 h in 2015.

3. History of track and intensity forecasts

In early days since 1998, IMD used to operate three regional models, Limited Area Model (LAM), MM5 model and Quasi-Lagrangian Model (QLM) for short-range prediction upto 72 h. The MM5 model was run at the horizontal resolution of 45 km with 23 sigma levels in the vertical for 72 h using initial and boundary conditions from the NCEP Global Forecast System (NCEP GFS) at the resolution of 1° × 1° latitude/longitude. The LAM was run up to 72 h at the horizontal resolution of 0.75° × 0.75° latitude/longitude with 16 sigma levels in the vertical using the initial and boundary conditions provided by the

T-254 Global operational model run at NCMRWF (National Center for Medium Range Weather Forecast).

The Quasi-Lagrangian Model (QLM), a multilevel fine-mesh primitive equation model with a horizontal resolution of 40 km and 16 sigma levels in the vertical, was run for tropical cyclone track prediction up to 72 h. The integration domain size was 4440 × 4440 km^2 centered on the initial position of the cyclone. For the day-to-day weather forecasting, IMD also used of NWP products prepared by some other operational NWP Centers like, NCMRWF (National Center for Medium Range Weather Forecast), ECMWF (European centre for medium range weather forecast), NCEP GFS. The QLM model for track prediction of tropical cyclone over the NIO discarded after 2012. The average track forecast errors of QLM during the period 1998-2012 was quite high and the errors were 146 km, 233 km, 403 km at 24 h, 48 h and 72 h respectively but no specific model was used for intensity forecasting.

IMD official (OFCL) track and intensity forecast over the NIO were subjectively generated largely based on persistence up to 24 h only till 2008 and track and intensity forecast errors were 155 km and 14.9 kt respectively at 24 h. The dynamical-statistical MME technique was developed and introduced for operational track forecasting in 2009 using sophisticated NWP models as mentioned in section 2 above. The dynamical-statistical model SCIP was developed and introduced for operational intensity forecasting in 2008 using NWP model output. Both the MME and SCIP generate track and intensity forecast up to 72 h and OFCL track and intensity forecasts also extended to 72 h from 2009. Further, forecast hours of MME and SCIP were extended up to 120 h and so as the operational forecasts also. In addition to the track and intensity component, other components of CPS like genesis, rapid intensification and decay after landfall are briefly described in the following section.

4. Dynamical-Statistical Cyclone Forecast System (CPS)

A detailed description of CPS was provided by Kotal *et al*. (2014). A brief description of CPS is presented in this section and performance of all five components of CPS is presented in the following section. The quantitative performance statistics on the probable forecast errors will provide the operational forecasters better guidance for better monitoring the TCs. The basic objectives of the CPS are:

(*i*) To improve the skill of dynamical model forecasts by statistical post processing.

(*ii*) To generate a consensus single forecast among different forecasts from different NWP models.

(*iii*) To develop a collective approach for improving various components of cyclone forecasting.

The five components of CPS for the operational forecasting of TCs over the NIO are briefly described below.

4.1. *Component - I : Genesis Potential Parameter (GPP)*

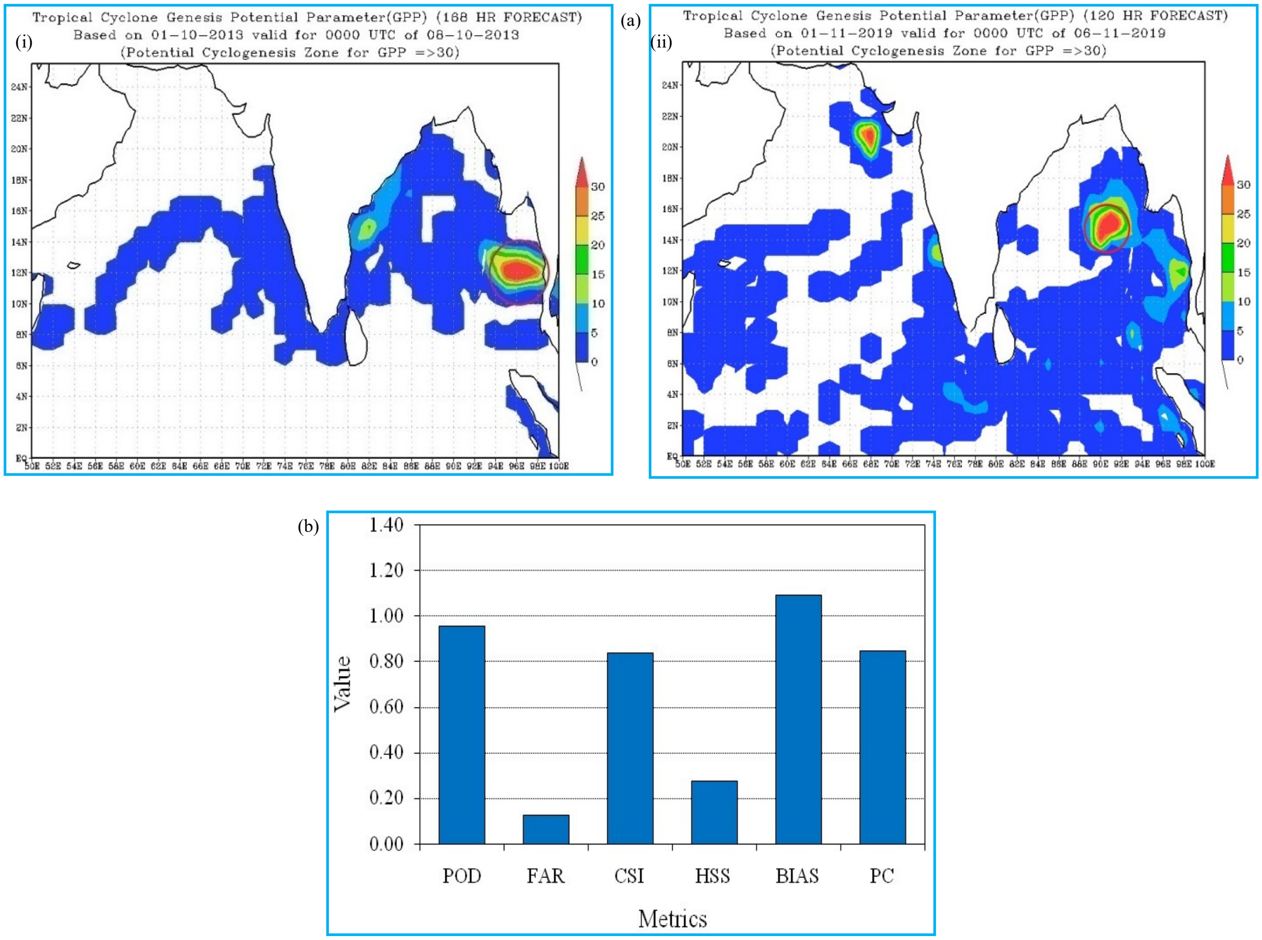
A genesis potential parameter (GPP) was developed for the NIO by Kotal *et al*. (2009). The GPP estimates the potential of a low pressure system for intensification into a tropical cyclone at the early stages of development. The parameter has been used operationally at IMD since 2008 for distinguishing non-developing and developing systems at the early stage of development. The spatial distribution of the parameter has also been used to identify the most favourable area of cyclogenesis over the Sea (Kotal and Bhattacharya, 2013).

4.2. *Component - II : Multi-model ensemble (MME) technique for track prediction*

The objective of this component was to generate a consensus track forecast, as there were variations of track forecasts among different NWP models. The MME track is generated from the model forecast positions by collective bias correction with respect to the observed position of TCs using multiple linear regression based minimization principle (Kotal and Roy Bhowmik, 2011). The MME technique has been used operationally at IMD since 2009 for real-time track prediction of TCs. The predictors selected for the ensemble technique are forecasts of latitude and longitude positions at 12 h interval up to 120 h of five global models (IMD-GFS, NCEP-GFS, ECMWF, UKMO and JMA).

4.3. *Component - III : Statistical Cyclone Intensity Prediction (SCIP) model*

The statistical cyclone intensity prediction (SCIP) model was developed by Kotal *et al*. (2008). The various parameters selected as predictors were determined from the forecast fields of NWP models for multiple regression analyses. Therefore, the SCIP model is principally a dynamical-statistical model. The dependent variable is intensity changes in knots (1 knot = 0.5144 ms^{-1}). The model estimates changes of intensity during 12 h, 24 h, 36 h, 48 h, 60 h, 72 h, 84 h, 96 h, 108 h and 120 h intervals. The SCIP model has been used operationally at IMD since 2008 for real-time forecasting of TC intensity.



Figs. 2[a(i, ii)&b]. (a) Spatial distribution of GPP for (i) Phailin & (ii) Bulbul and (b) POD, FAR, CSI, HSS, BIAS and PC of the genesis forecasts of GPP during 2010-2019

4.4. *Component - IV : Rapid Intensification (RI) index*

The rapid intensification index (RII) was formulated using threshold (index) values of eight large scale atmospheric variables for which statistically significant differences were found between the RI and non-RI cases (Kotal and Roy Bhowmik, 2013). The RI phase is defined as an increase of intensity by 30 kt (15.4 ms^{-1}) or more during 24 h. The RII technique estimates the probability of RI over the subsequent 24 h. The technique has been used at IMD since 2011 for real-time forecasting of RI.

4.5. *Component - V : Decay of intensity after the landfall*

Considering the destructive potential and impact on human activities during landfall and after landfall, Roy Bhowmik *et al*. (2005) developed a Decay model for prediction of decaying intensity of TCs after landfall at 6 h interval up to 24 h. A correction method for updating forecasts after 6 h of landfall is also applied in the decay model using the decay rate of first 6 h. Roy Bhowmik *et al*. (2005) have explained the details of correction method. The decay model has been used at IMD for real-time forecasting since 2008.

5. Evaluation of performance of CPS

5.1. *Forecast skills of cyclogenesis*

All low pressure systems over the NIO do not intensify into TCs. Therefore, it is imperative to estimate the potential of a low pressure system for intensification into a TC at its early stages. The GPP has been used for distinguishing developing and non-developing systems at their early stages (*T* - number 1.0, 1.5, 2.0) of development. The spatial distribution of GPP for two typical cyclone cases of Phailin [Fig. 2a(i)] and Bulbul [Fig. 2a(ii)] show that the parameter was able to predict the most favourable zone of cyclogenesis for cyclone Phailin over the east Andaman Sea seven days ahead and for cyclone Bulbul

TABLE 2

Contingency table for categorical variables

Forecast ↓	Observation→	
	Cyclogenesis	No Cyclogenesis
Cyclogenesis	a (YY)	b (YN)
No Cyclogenesis	c (NY)	d (NN)

a = No. of Hits (predicted and observed)
b = No. of False Alarms (predicted but not observed)
c = No. of misses (observed but not predicted)
d = No. of correct predictions of no cyclogenesis (neither predicted nor observed)

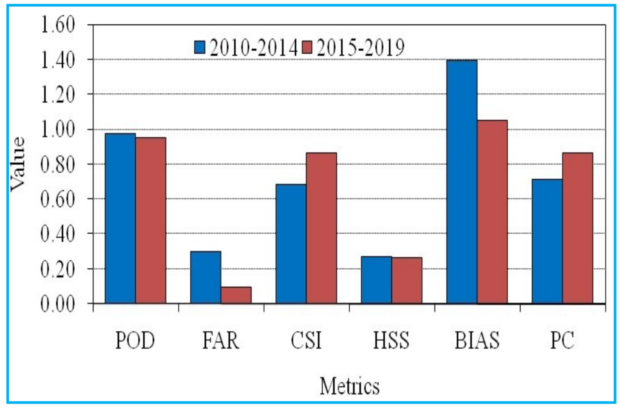


Fig. 3. POD, FAR, CSI, HSS, BIAS and PC of the genesis forecasts of GPP during 2010-2014 and 2015-2019

over the north Andaman Sea five days ahead. In general, the GPP could able to predict the cyclogenesis zone four to five days ahead (Kotal and Bhattacharya, 2013).

Six statistical metrics for cyclogenesis, such as the probability of detection (POD), false alarm ratio (FAR), critical success index (CSI), Heidke Skill Score (HSS), frequency bias (BIAS) and proportion correct (PC) have been computed to evaluate the skill (Brown and Brandes, 1997) of the GPP for genesis forecasts issued during the period 2010-2019.

The statistical metrics are computed using 2 × 2 contingency table (Table 2) as follows:

Qualitative verification has been done with the help of following scores.

(*i*) *Probability of Detection (POD)* : A measure of discrimination, POD is defined as the number of hits divided by the total number of observed events.

POD = a/(a+c). Range : 0 to 1. Perfect score : 1.

(*ii*) *False Alarm Ratio (FAR)* : A measure of reliability, FAR is defined as the number of false alarms divided by the total number of forecast events.

FAR = b/(a+b). Range : 0 to 1. Perfect score : 0.

(*iii*) *Critical Success Index (CSI)* : A value of warnings combines Hit Rate and False Alarm Ratio into one score. It is calculated as follows:

CSI = a/(a+b+c). Range : 0 to 1. Perfect score : 1.

(*iv*) *Heidke Skill Score (HSS)* : A measure of the fractional improvement of the forecast over the standard forecast.

$$HSS = \frac{2(ad - bc)}{(a+c)(c+d)+(a+b)(b+d)}$$

Range : -α to 1. Perfect : 1

(*v*) *Percent Correct (PC)* : The percent correct is the percent of forecasts that are correct.

PC = (a+d) / (a+b+c+d)

Range : 0 to 1. Perfect score : 1.

(*vi*) The frequency bias (BIAS) gives the ratio of the forecast event frequency to the observed event frequency. The frequency bias (BIAS) is given by:

BIAS = (a+b)/(a+c)

Range : 0 to infinity. Desirable value for bias is 1.

The desirable value for bias 1 indicates an unbiased forecast, bias greater than 1 indicates the event is over forecast and bias less than 1 indicates the event is under forecast.

Fig. 2(b) depicts the statistical metrics of the GPP forecasts for cyclogenesis. The figure shows that the POD of the GPP was 0.95, the FAR was 0.13, CSI was 0.84, HSS was 0.27, BIAS was 1.09 and PC was 0.84 for 847 forecast events during 2010-2019. The much higher POD than FAR, near desirable value of BIAS and high CSI and PC indicate that the GPP was skillful for cyclogenesis prediction. Comparison of the metrics (Fig. 3) for first five years (2010-2014) and second five years (2015-2019) suggest that FAR and BIAS has reduced and CSI and PC have increased in recent five years.

TABLE 3

Number of track forecasts verified at different lead time during the period 2010-2019

Lead time →	12 h	24 h	36 h	48 h	60 h	72 h	84 h	96 h	108 h	120 h
ECMWF	339	319	284	251	204	176	122	101	75	60
NCEP-GFS	358	339	306	265	221	183	127	109	81	64
IMD-GFS	333	316	286	253	208	176	125	104	77	61
UKMO	299	289	265	233	190	165	133	106	79	61
JMA	333	317	287	252	208	174	129	-	-	-
IMD-HWRF	512	477	432	378	324	276	235	194	146	115
IMD-WRF	335	322	294	257	212	176	-	-	-	-
IMD-OFCL	690	608	522	444	367	295	195	149	105	67
IMD-MME	359	344	311	272	224	190	138	112	83	66

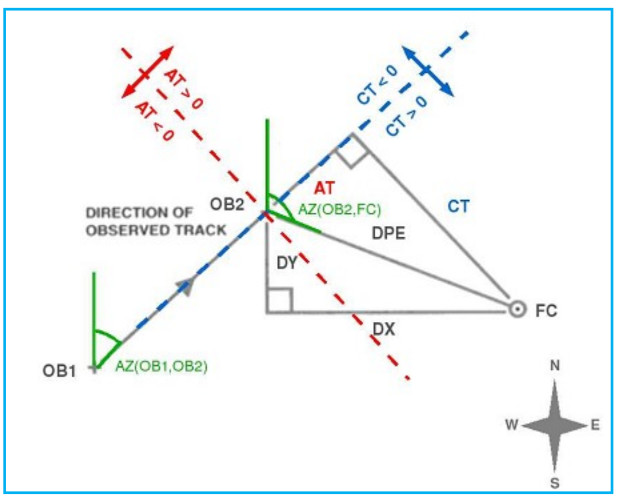


Fig. 4. Types of positional forecast errors. DPE represents the direct positional error, CTE is the cross track error and ATE the along track error [Adapted from Heming (1994)]

5.2. *Track forecast errors*

In this study, track forecast verification has been done for the tropical cyclones for their stages classified at least as Depression (winds greater than 16 kt) both at the initial time and validity time of the forecast. The direct position error (DPE) generally considered as track forecast error is defined as the shortest distance between the forecast position and best-track position of the centre of TC. The DPE gives an indication of location forecast error, but gives no information on whether the forecast errors were due to slow or fast movement of TCs. The DPE also do not give the actual deviation of forecast track from the best-track of TC. In order to review this, the along track error (ATE) and cross track error (CTE) are also evaluated. The ATE occurs due to differential speed between forecast TC and actual TC. The deviation of forecast positions from TC track is measured by the CTE. Detailed TC track forecast verification technique (Fig. 4) was demonstrated by Heming (1994). In this study, DPE, ATE, CTE of NWP models and MME forecasts are analyzed. Number of forecasts verified at different lead time is given in Table 3. The importance of sample size is well known. Precise and more accurate conclusions can be drawn with an appropriate sample size. A smaller sample may not be sufficiently powered to detect difference between the results and very large sample size also has different consequences. However in this study, average track errors are computed to assess an overview of the performance by considering all available data during 2010-2019.

5.2.1. *Direct Position Error (DPE)*

In this section, a summary of the DPE of NWP models and MME during the decade 2010-2019 for the forecast lead time at 12 h intervals up to 120 h is presented. The track forecast was extended from 72 h to 120 h from 2013. Therefore the errors from 84 h to 120 h were computed for the period 2013-2019. Fig. 5 shows the average DPE of IMD-GFS, NCEP-GFS, ECMWE, UKMO, JMA, IMD-HWRF, IMD-WRF, IMD-OFCL and IMD-MME during the period 2010-2019. The average DPE of consensus forecast (IMD-MME) during 2010-2019 were 53 km, 67 km, 84 km, 103 km, 120 km, 143 km, 160 km, 188 km, 220 km and 246 km for the forecast lead time 12 h, 24 h, 36 h, 48 h, 60 h, 72 h, 84 h, 96 h, 108 h and 120 h respectively (Fig. 6). The average DPE of IMD OFCL forecasts during 2009-2013 were 124 km,

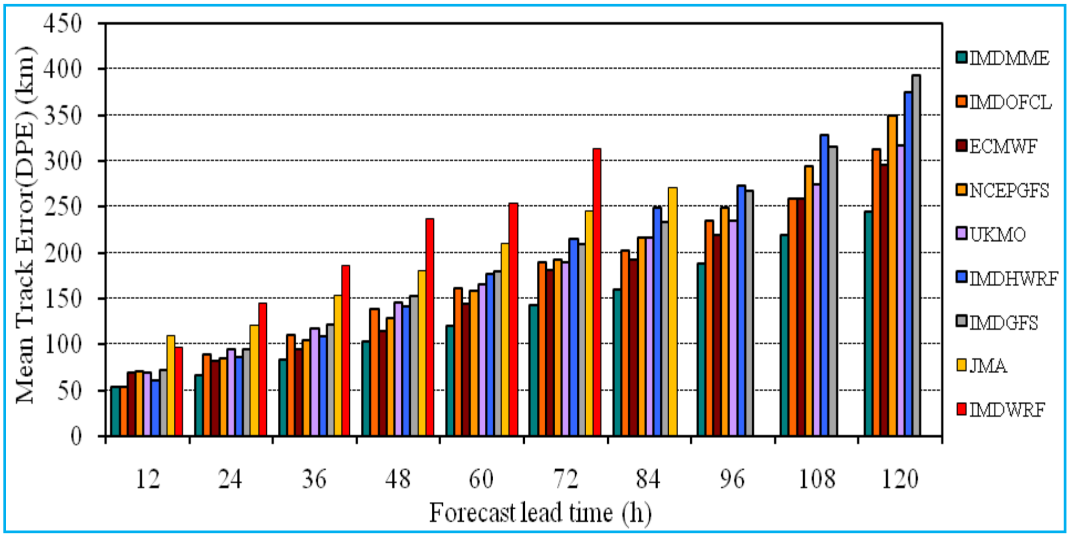


Fig. 5. Mean track forecast errors (km) (DPE) during 2010-2019 (84 h to 120 h error from 2013-2019)

TABLE 4

Relative errors of MME (the percentage difference between MME and model error of DPE) with respect to IMD-OFCL, ECMWF, NCEP-GFS, UKMO, IMD-HWRF, IMD-GFS, JMA and IMD-WRF

Lead time →	12 h	24 h	36 h	48 h	60 h	72 h	84 h	96 h	108 h	120 h
IMD-OFCL	-0.7	-25.1	-24.4	-25.3	-25.3	-24.5	-21.1	-20.3	-15.3	-21.5
ECMWF	-22.4	-18.2	-12.3	-9.9	-16.8	-21.0	-17.1	-14.7	-15.2	-17.1
NCEP-GFS	-24.7	-21.5	-20.7	-19.8	-24.4	-25.6	-26.0	-24.7	-25.3	-29.8
UKMO	-23.6	-28.7	-28.6	-28.9	-27.5	-24.3	-26.0	-20.1	-20.0	-22.7
IMD-HWRF	-12.3	-22.8	-23.6	-27.1	-31.9	-33.5	-35.9	-31.4	-33.0	-34.4
IMD-GFS	-26.1	-29.2	-31.0	-32.4	-33.3	-31.5	-31.7	-30.1	-30.2	-37.5
JMA	-51.3	-44.5	-45.3	-42.5	-42.7	-41.6	-40.9	-	-	-
IMD-WRF	-44.5	-53.4	-54.8	-56.2	-52.4	-54.2	-	-	-	-

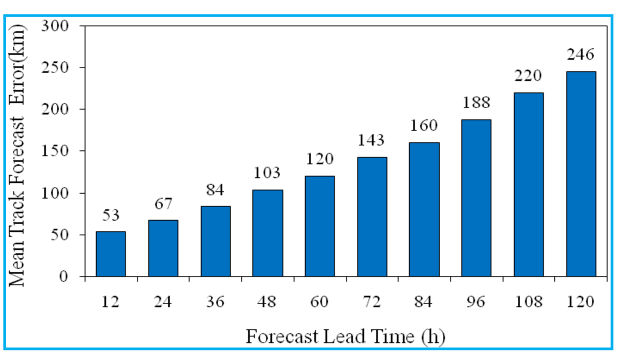


Fig. 6. Mean MME track forecast errors (km) during 2010-2019 (84 h to 120 h error from 2013-2019)

202 km, 268 km, 251 km and 296 km and these were 86 km, 132 km, 178 km, 260 km and 285 km during 2014-2018 for the forecast lead time 24 h, 48 h, 72 h, 96 h and 120 h respectively (Mohapatra and Sharma, 2019). The ECMWF, NCEP-GFS, UKMO, IMD-GFS, IMD-HWRF model and IMD-OFCL forecast errors were ranged from 69 km to 296 km, from 71 km to 350 km, 70 km to 318 km, from 72 km to 393 km, from 61 km to 374 km and from 54 km to 313 km for forecast lead time 12 h to 120 h respectively during 2010-2019. It is noted that IMD-MME outperformed all the forecasts as shown in the Fig. 5. It is also noted that among individual models, ECMWF performed better than all other models.

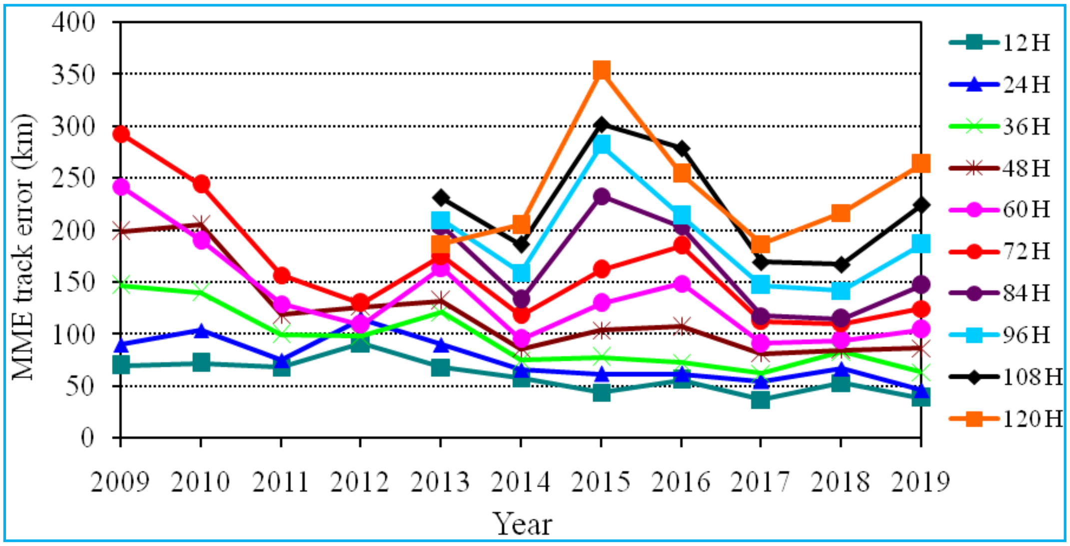


Fig. 7. Yearly mean track forecast errors (DPE) of MME during 2009-2019

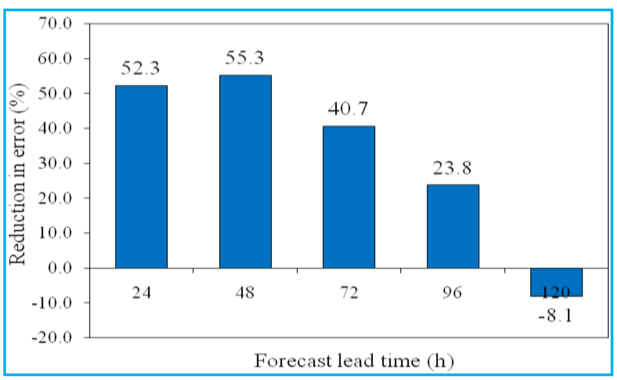


Fig. 8. MME track forecast error (DPE) reduction during 2010-2019

The relative error (RE) of MME compared to all other NWP models is shown in Table 4. The RE of MME with respect to NWP models (the percentage difference between MME and model track forecast error) is defined as:

$$RE(\%) = \frac{ME_{MME} - ME_{MODEL}}{ME_{MODEL}} \times 100$$

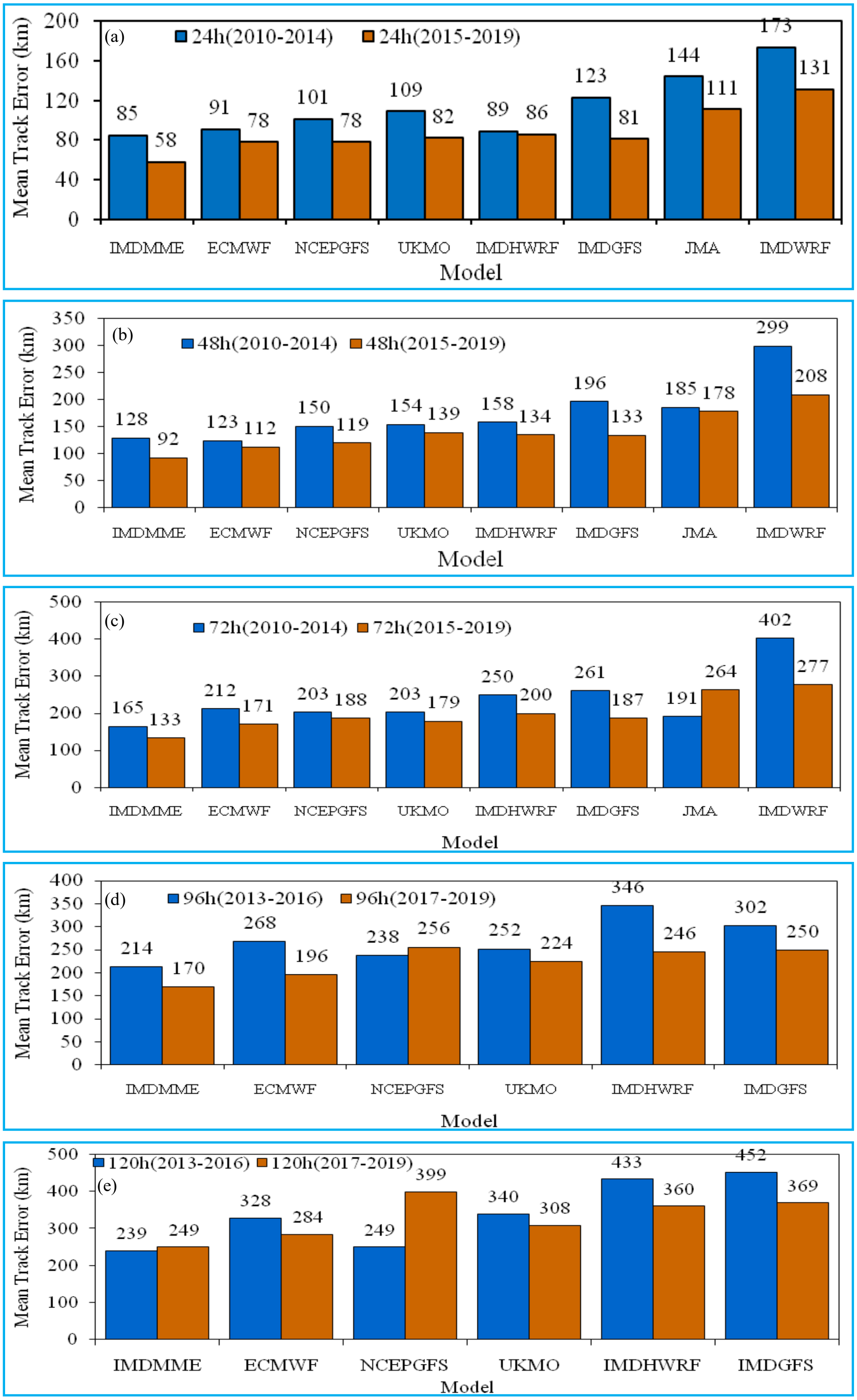
where, ME is the mean track error (DPE).

It is found from the Table 4 that under the scenario of multiple forecasts provided by the different models, the MME forecast errors were lowest compared to all the member models at all forecast hours. The MME forecasts were less by about 18-53%, 10-56%, 21-54%, 15-31% and 17-37% than all individual models and IMD-OFCL for forecast lead time 24 h, 48 h, 72 h, 96 h and 120 h respectively. MME forecasts were also less than IMD-OFCL forecasts by about 16%, 31%, 30% at 24 h, 48 h, 72 h respectively during 2009-2013 and less by about 28%, 31%, 27% at 24 h, 48 h, 72 h respectively during 2014-2018 (Heming *et al.*, 2019). Year wise mean track forecast errors of MME during 2009-2019 is shown in Fig. 7. The MME track forecast errors show increase in the errors in 2015. It is to be mentioned that the increase in the errors in 2015 is due to large errors of member models of few systems over the Arabian Sea. Similarly, for the year 2014 also, therefore increasing trend from 2013-2015. Further trend analysis of track error shows that during the period 2010-2019, DPE of MME has reduced significantly by about 41% to 55% for forecast lead time 24 h to 72 h (Fig. 8).

TABLE 5

Improvement (%) of DPE during (2015-2019) compared to (2010-2014)

Lead time →	24 h	48 h	72 h	96 h	120 h
IMD-MME	-31.8	-28.4	-19.1	-20.5	4.1
ECMWF	-14.3	-9.0	-19.5	-26.9	-13.5
NCEP-GFS	-22.8	-20.4	-7.5	7.6	60.3
UKMO	-24.8	-9.7	-11.8	-11.2	-9.3
IMD-HWRF	-3.4	-14.8	-19.9	-29.0	-16.9
IMD-GFS	-34.1	-32.1	-28.3	-17.3	-18.3
JMA	-22.9	-4.0	38.2	-	-
IMD-WRF	-24.3	-30.4	-31.1	-	-



Figs. 9(a-e). Comparative mean track forecast errors (km) of MME during 2010-2014 and 2015-2019 for (a) 24 h, (b) 48 h, (c) 72 h, (d) 96 h and (e) 120 h forecast

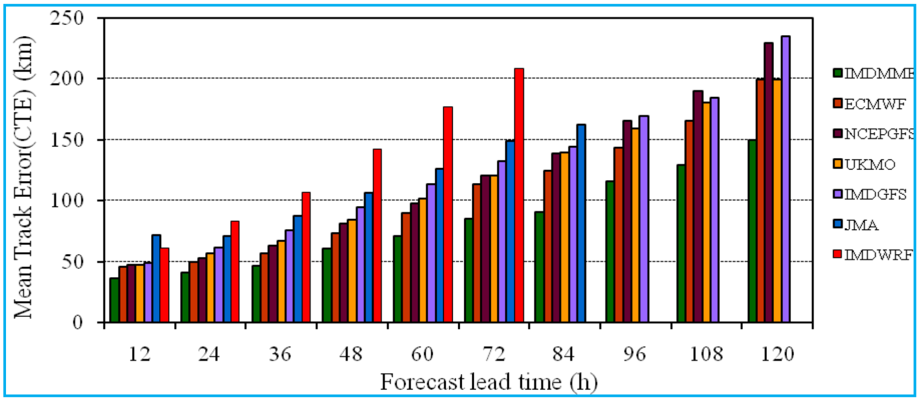


Fig. 10. Mean track forecast errors (km) (CTE) during 2010-2019 (84 h to 120 h error from 2013-2019)

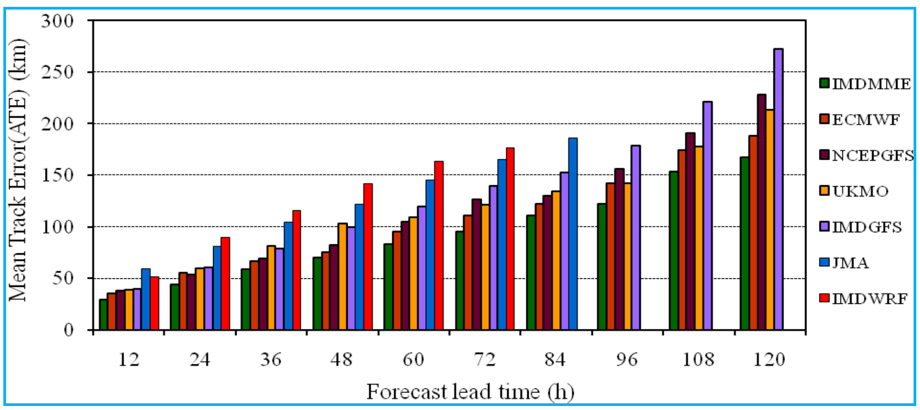
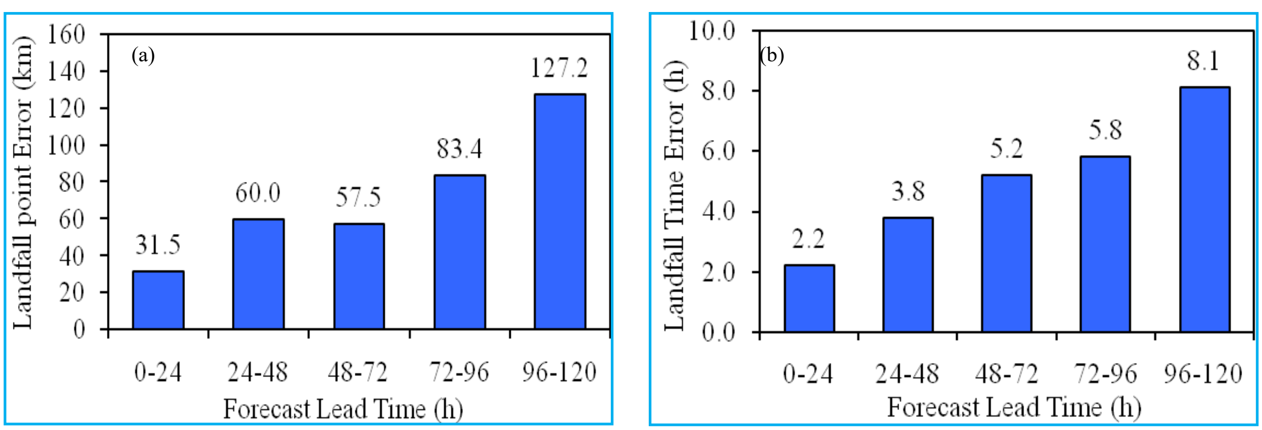
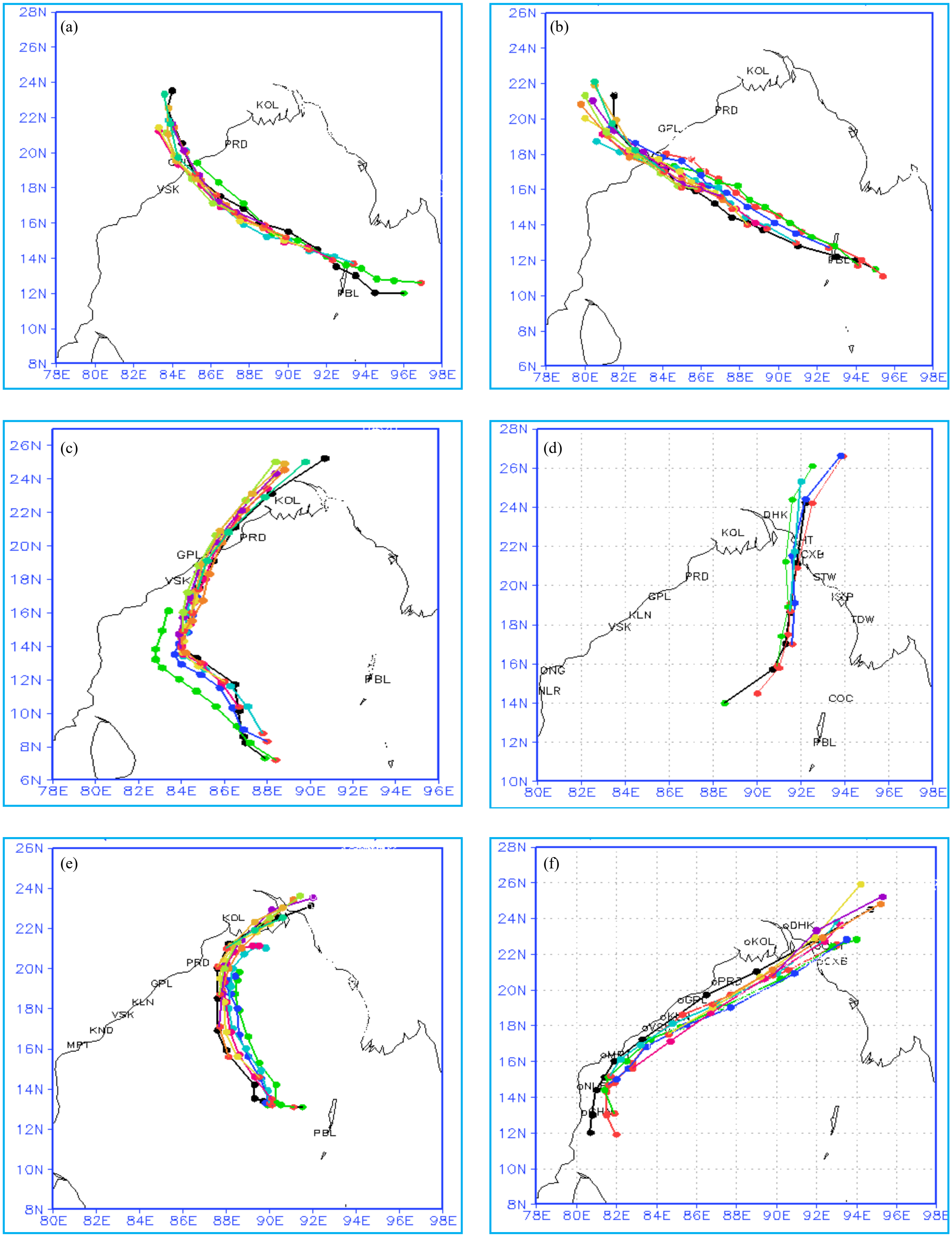


Fig. 11. Mean track forecast errors (km) (ATE) during 2010-2019 (84 h to 120 h error from 2013-2019)



Figs. 12(a&b). (a) MME landfall point error (km) and (b) MME landfall time error (hr) during 2010-2019

The average track forecast errors during the recent past five years (2015-2019) compared to previous five years (2010-2014) for forecast hours 24 h, 48 h and 72 h are shown in Figs. 9(a-c). The average track forecast error of MME during 2010-2014 were 85 km, 128 km, 165 km and these were 58 km, 92 km, 133 km during 2015-2019 for the forecast lead time 24 h, 48 h and 72 h respectively. As track forecast was extended from 72 h to 120 h, the average track forecast errors during the recent past three years (2017-2019) were compared to previous four years (2013-2016) for forecast hours 96 h and 120 h [as shown in Figs. 9(d&e)]. The average track forecast error of MME during 2013-2016 were 214 km, 239 km and these were 170 km, 249 km during 2015-2019 for the forecast lead time 96 h and 120 h respectively. The improvement during these two periods for all the forecasts is shown in Table 5.



Figs. 13(a-f). Plots of all MME forecast tracks (colour tracks) along with observed track (black colour) in each panel for TCs (a) Phailin, (b) Hudhud, (c) Fani, (d) Mora, (e) Bulbul and (f) Roanu

TABLE 6

Observed maximum sustained wind strengths and coast of landfall of cyclonic storms Phailin, Hudhud, Roanu, Mora, Fani and Bulbul

S. No.	Name (Period)	Year	Max. Wind Speed (kt)	Coast of Landfall
1.	PHAILIN (8-14 October)	2013	115	Odisha
2.	HUDHUD (7-14 October)	2014	100	Andhra Pradesh
3.	ROANU (17-22 May)	2016	45	Bangladesh
4.	MORA (28-31 May)	2017	60	Bangladesh
5.	FANI (26 April - 4 May)	2019	115	Odisha
6.	BULBUL (5-11 November)	2019	75	West Bengal

It is found that there has been improvement of forecasts for all models at all forecast hours except IMDGFS for 96 h and 120 h. The maximum improvement of ECMWF model was ranged from about 14% at 24 h to 27% at 96 h, for NCEP GFS the ranges was 23% at 24 h to 20% at 48 h, for UKMO model the range was 25% at 24 h to 9% at 120 h, for IMDHWRF model the ranges was 3% at 24 h to 20% at 72 h, for IMDGFS the range was 34% at 24 h to 18% at 120 h and for IMDWRF model the range was 24% at 24 h to 31% at 72 h. The overall maximum improvements of all models were 34%, 32%, 28%, 29% and 18% at 24 h, 48 h, 72 h, 96 h and 120 h respectively. The improvements of MME are found to be about 32%, 28%, 19% and 20% at forecast hours 24 h, 48 h, 72 h and 96 h respectively. Although there was no improvement for MME at 120 h but errors were less than all other forecasts for both the periods (2015-2019) & (2010-2014).

5.2.2. *Along and cross track errors (ATE and CTE)*

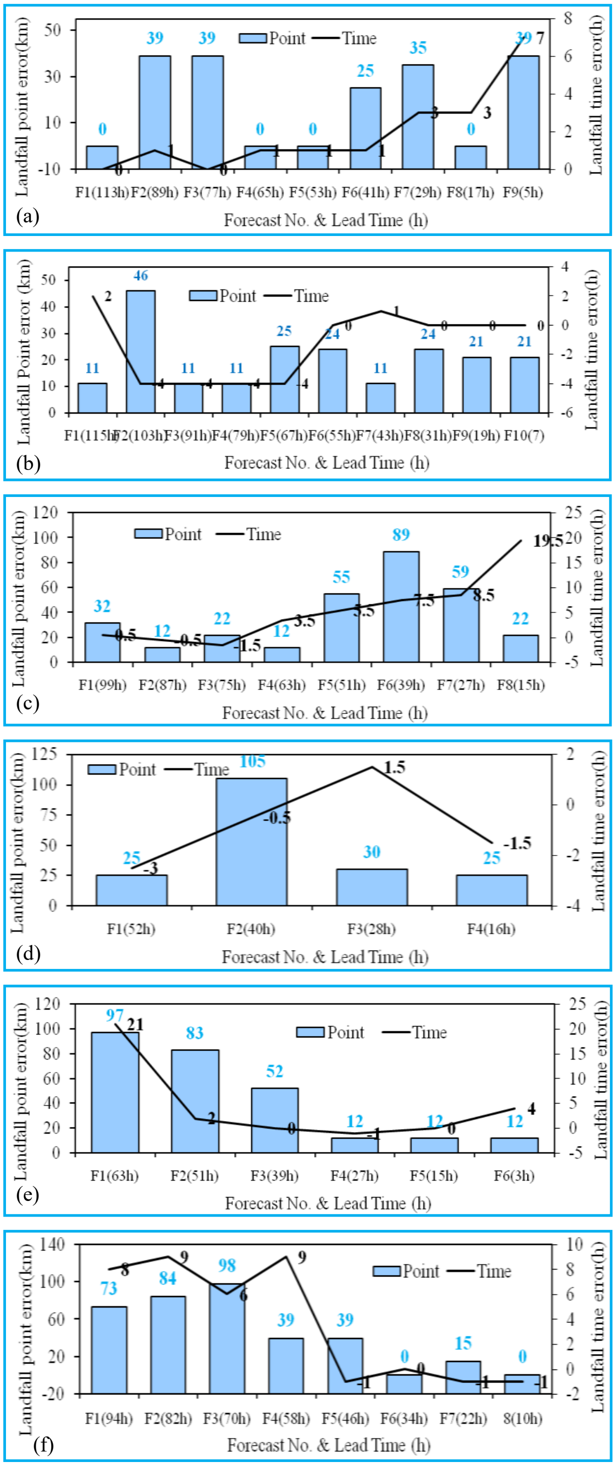
Figs. 10&11 depict the mean ATE and CTE respectively of MME and NWP models during the period 2010-2019. The mean ATE of MME was 30 km, 44 km, 59 km, 70 km, 84 km, 95 km, 111 km, 122 km, 154 km and 167 km for forecast lead time 12 h, 24 h, 36 h, 48 h, 60 h, 72 h, 84 h, 96 h, 108 h and 120 h respectively. While the mean CTE of MME was 36 km, 41 km, 47 km, 61 km, 71 km, 86 km, 91 km, 116 km, 129 km and 150 km for forecast lead time 12 h, 24 h, 36 h, 48 h, 60 h, 72 h, 84 h, 96 h, 108 h and 120 h respectively. Like DPE, the ATE and CTE of MME are found to be outperformed the all model forecasts for all forecast hours. The CTE actually measures the deviation of forecast positions from TC track. The CTE statistics of MME which shows it was less than 100 km up to 84 h and equal or less than 150 km from 96 h to 120 h and are found to be very useful considering the area of impact of TCs.

It is noted that the skill of MME was consistently higher compared to the member models. It is also found that skills of ECMWF and NCEP were higher compared to the other member models and ECMWF was slightly more skillful than NCEP in general. Among the other models, UKMO is found to be more skillful than IMD-GFS, JMA and IMDWRF.

5.2.3. *Landfall forecast errors of MME*

Average error : The accuracy of landfall point and landfall time forecasts of TCs is the most important component of TC forecast services. To assess the accuracy of landfall forecasts of MME, the forecast landfall point (at 12 h interval from 12 h to 120 h before landfall) of each case was compared with the observed landfall point from the IMD best-track data. Among all the TCs over the NIO during 2010-2019, 27 TCs made landfall surrounding the coast of the NIO. A total 162 landfall forecasts cases for different lead time are verified. Among 162 cases, 44 cases were within forecast lead time up to 24 h and 47, 36, 21 and 14 cases were within forecast lead time 24-48 h, 48-72 h, 72-96 h and 96-120 h respectively. For each case, landfall forecast error has been computed by measuring distance along the coast between predicted and observed landfall point and landfall time error has been considered as the difference between predicted and observed landfall time. Forecast lead time has been computed by the difference between forecast issue time and the landfall time of TCs from best-track data. The average landfall point errors [Fig. 12(a)] were 31.5 km for forecast lead time up to 24 h, 60 km within 24-48 h, 57.5 km within 48-72 h, 83.4 km within 72-96 h and 127.2 km for forecast lead time within 96-120 h before landfall. The average landfall time errors [Fig. 12(b)] were 2.2 h, 3.8 h, 5.2 h, 5.8 h and 8.1 h for forecast lead time within 0-24 h, 24-48 h, 48-72 h, 72-96 h and 96-120 h respectively. Landfall time analysis also shows that the MME predicted delayed landfall by 2.1 h, therefore slow movement prediction in general.

Individual cases : All forecast tracks of MME for each cyclone along with observed track for six typical individual land falling TCs (Phailin, Hudhud, Fani, Bulbul, Roanu, Mora) considering their wide diversity in nature in terms of intensity and track are shown in Fig. 13 to visualize the accuracy and consistency of MME track forecasts. In each case of the six TCs as shown in Fig. 13, colour tracks show all MME forecast tracks and black track shows observed track as per best-track data. Among six TCs, three were northwestward moving extremely severe cyclonic storms (Phailin, 2013; Hudhud, 2014; Fani, 2019), two were northeastward recurving cyclone (Bulbul, 2019; Roanu, 2016) with very severe intensity of Bulbul, and one was northward moving severe cyclonic



Figs. 14(a-f). Landfall point and time errors of MME at different forecast lead time for TCs (a) Phailin, (b) Hudhud, (c) Fani, (d) Mora, (e) Bulbul and (f) Roanu

storm (Mora, 2017). Observed maximum sustained wind strengths and coast of landfall of cyclonic storms Phailin, Hudhud, Roanu, Mora, Fani and Bulbul

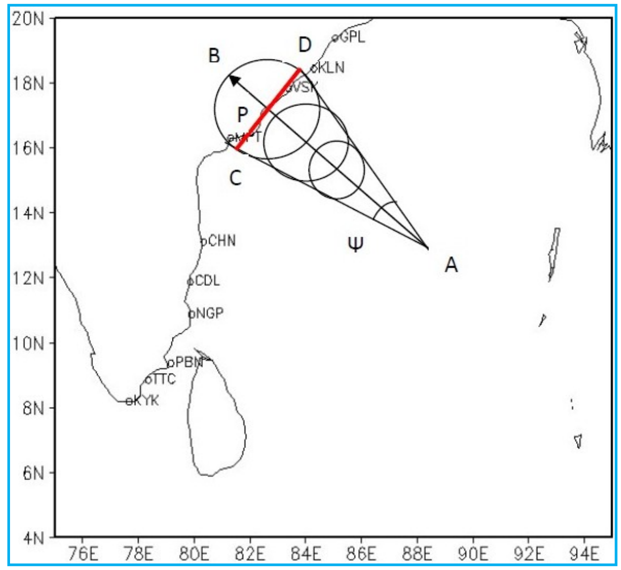


Fig. 15. Cone of uncertainty of TC Track

are shown in Table 6. Comparison of landfall point and landfall time errors at different forecast lead time for the six individual TCs is shown in Figs. 14(a-f). In the figure, *x* - axis shows forecast number (F1, F2, F3,….) along with forecast lead time in parenthesis. The result shows that the landfall point error was ranged from 11 km to 46 km with maximum landfall time error of four hours delay for forecast lead time 7 h to 115 h for Hudhud. Maximum landfall point error and landfall time error was 39 km and 7 h respectively for forecast lead time up to 113 h for Phailin and it was 99 km and 19.5 h for forecast lead time up to 99 h for TC Fani. Similarly, maximum landfall point errors were 98 km, 97 km and 105 km and landfall time errors were 9 h, 21 h and -3 h for TCs Roanu, Bulbul and Mora respectively.

The errors of six individual severe and extremely severe storms (including tracks of recurvature in nature) show that even for such cases where wide variation of intensity and track, CPS provided skillful landfall forecasts to the operational forecasters.

The results reaffirm the efficacy of the CPS for such scenario in real-time forecasting.

5.2.4. *Cone of uncertainty*

The TC track forecasts are hardly ever perfect. Predicting the future location of TCs inherently involves some level of uncertainty. In case of TC forecast map (Fig. 15), it can be seen a cone of uncertainty. In the Figure, vector AB, CD and angle Ψ typically indicates TC track, deviation of forecast track and cone of uncertainty respectively. The cone represents the probable track of

TABLE 7

Radii of circles at each forecast hour defining the cone

Forecast Hour	12 h	24 h	36 h	48 h	60 h	72 h	84 h	96 h	108 h	120 h
Radius of circle (km)	17.2	34.4	51.6	68.9	86.1	103.3	120.5	137.7	154.9	172.1

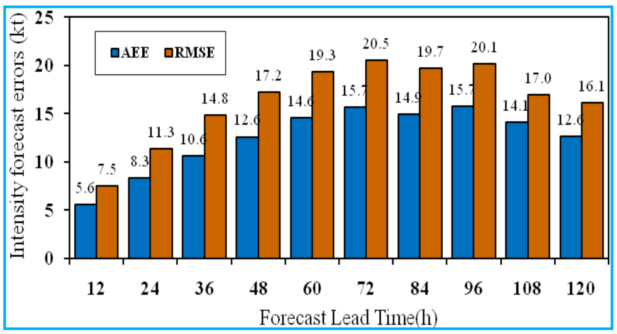


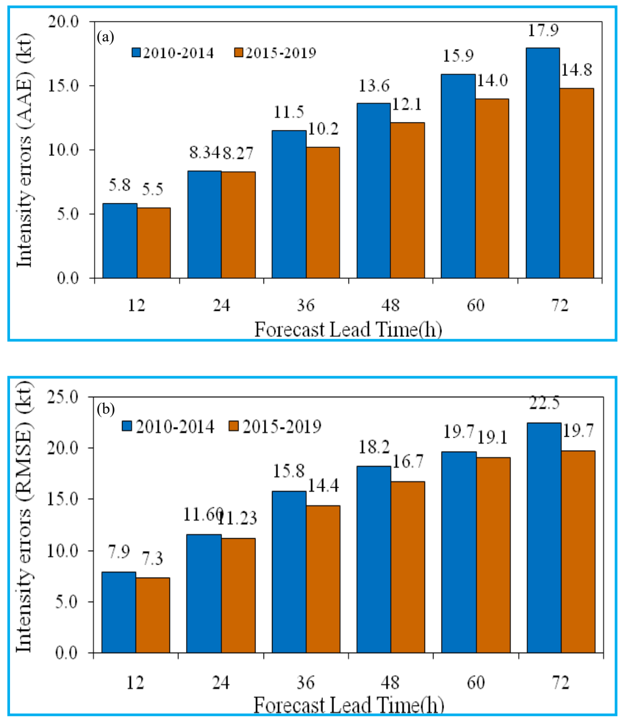
Fig. 16. Average absolute intensity forecast errors (AAE) and root mean square error (RMSE) of SCIP during 2010-2019

centre of a TC. The cone of uncertainty is prepared by plotting average CTE (as radius) of MME during the period 2010-2019 on a map using a sequence of circles along the forecast track at 12 h, 24 h, 36 h, 48 h, 60 h, 72 h, 84 h, 96 h, 108 h and 120 h and smoothing out the circle diameter (as shown in Fig. 15). The radii of circles defining the cone of MME are given in Table 7 and the cone of uncertainty (Ψ) is found to be 11.7°. The centre of TCs (2084 nos.) for all forecast hours within the cone was 86%, which means 14% cases it strayed outside of the cone. It can also be mentioned that timing error is ignored here, *i.e.*, the TCs moving slow or fast but in the expected direction still would be within the area of cone. The coastal locations within or near the cone of uncertainty four or five days out would help to keep tabs on the situation and make preparations for disaster managers.

5.3. *Intensity forecast errors of SCIP*

5.3.1. *Mean intensity forecast errors*

In this section, intensity forecast performance of the SCIP model has been analyzed for the period 2010-2019. Fig. 16 depicts the average error of intensity forecast of the SCIP model during the period. The average absolute error (AAE) was ranged from 5.6 kt at 12 h to about 12.6 kt at 120 h and root mean square error (RMSE) was ranged from 7.5 kt at 12 h to 16.1 kt at 120 h. The highest error was found to be at 72 h (AAE = 15.7; RMSE = 20.5). The sample size was 329, 307, 272, 229, 191, 157, 90, 72, 54 and 29 for forecast hours 12 h, 24 h, 36 h, 48 h, 60 h, 72 h, 84 h, 96 h, 108 h, 120 h respectively. The intensity



Figs. 17(a&b). SCIP intensity forecast error (a) AAE and (b) RMSE during 2010-2014 & 2015-2019

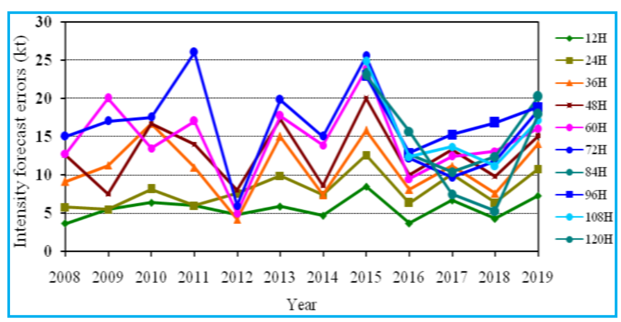


Fig. 18. Yearly mean SCIP intensity forecast error during 2008-2019 for forecast lead time 12 h, 24 h, 36 h, 48 h, 60 h, 72 h, 96 h, 108 h and 120 h (84 h to 120 h error from 2015-2019)

forecast was extended from 72 h to 120 h from 2015. Therefore the errors from 84 h to 120 h were computed for the period 2015-2019. The AAE and RMSE during the past five years 2010-2014 and previous five years 2015-2019 are shown in Fig. 17(a&b). During these periods,

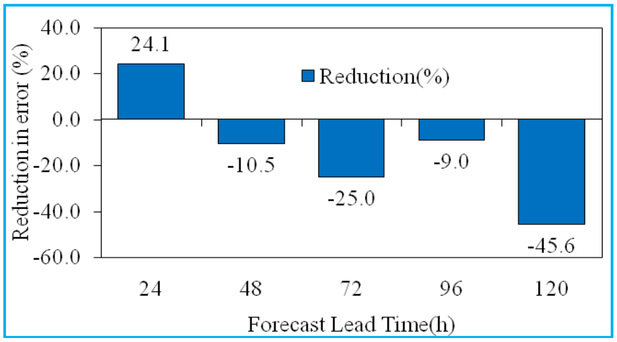


Fig. 19. Reduction of SCIP intensity forecast error during 2010-2019

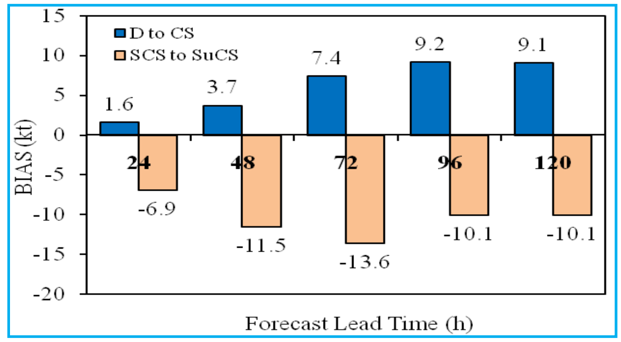


Fig. 20. Bias (kt) of intensity forecast by SCIP model for Depression (D) to Cyclonic Storm (CS) stage and Severe Cyclonic Storm (SCS) to Super Cyclonic Storm (SuCS) stage during 2010-2019

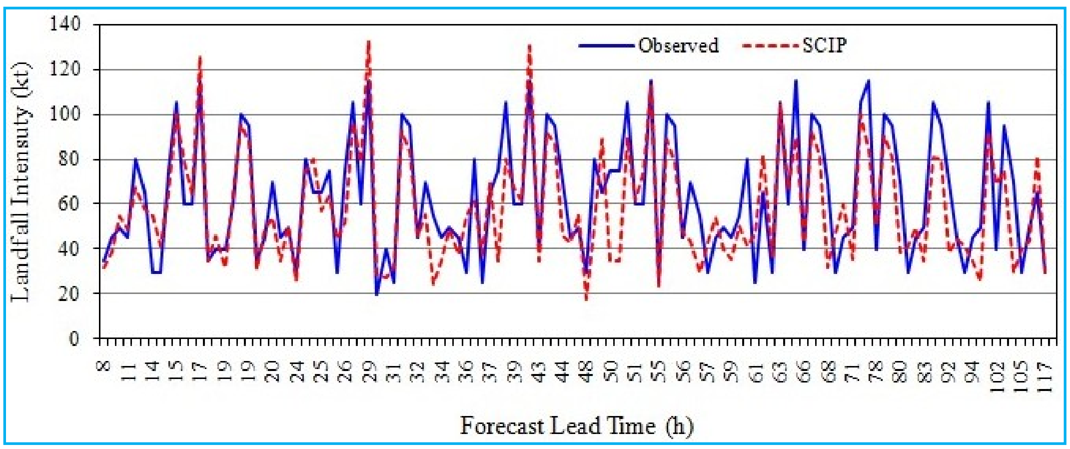


Fig. 21. Landfall intensity forecast by SCIP versus Observed intensity during 2010-2019

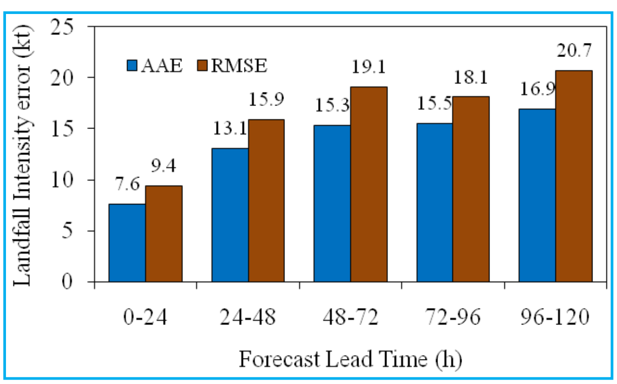
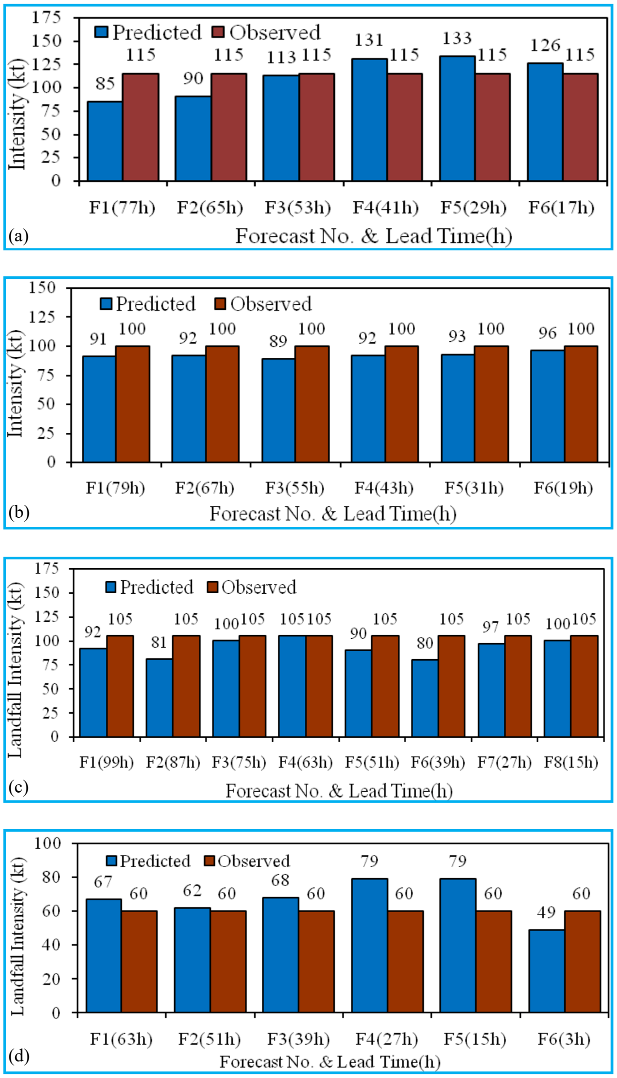


Fig. 22. Mean landfall intensity error (h) of SCIP model during 2010-2019

there was an improvement of 11-17% for 36 h to 72 h but no significant change is found at 24 h. Year to year intensity forecast error of SCIP model is shown in Fig. 18. The trend analysis during the period 2010-2019 also suggest that there was no change at 24 h but errors reduced by about 9-46% for forecast lead time 48 h to 120 h (Fig. 19). In general, the forecast bias was positive for weaker storms (depression (D) to cyclonic storm (CS) and bias was negative for stronger storms [severe cyclonic storm (SCS) to super cyclonic storm (SuCS)]. The positive bias was ranged from 1.6 kt at 24 h to 9.1 kt at 120 h and the negative bias was ranged from -6.9 kt at 24 h to -10.1 kt at 120 h with highest bias -13.6 at 72 h (Fig. 20).

5.3.2. *Landfall intensity forecast errors*

Like landfall point forecasts, landfall intensity forecasts also very important, since warning and preventive measures are also largely depends on this forecast along with landfall point forecasts. To assess the accuracy of landfall intensity forecasts of SCIP model, all landfall intensity forecasts cases (118 nos.) have been verified (shown in Fig. 21) for all tropical cyclones that struck the coasts surrounding NIO during 2010-2019. The AAE and RMSE are shown in Fig. 22. The figure shows that the AAE was ranged from 7.6 kt to 16.9 kt for



Figs. 23(a-d). Landfall intensity forecast of SCIP model versus Observed intensity at different forecast lead time for TCs (a) Phailin, (b) Hudhud, (c) Fani and (d) Bulbul

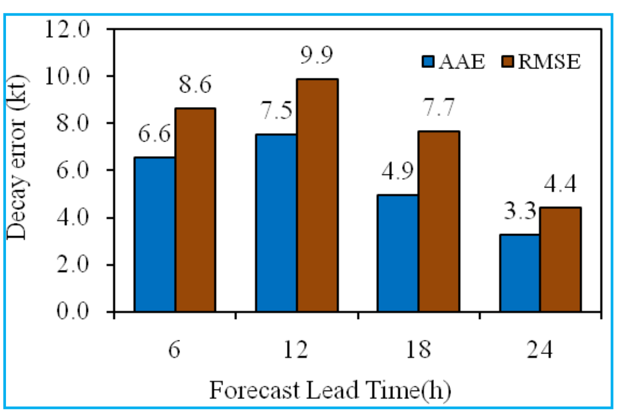
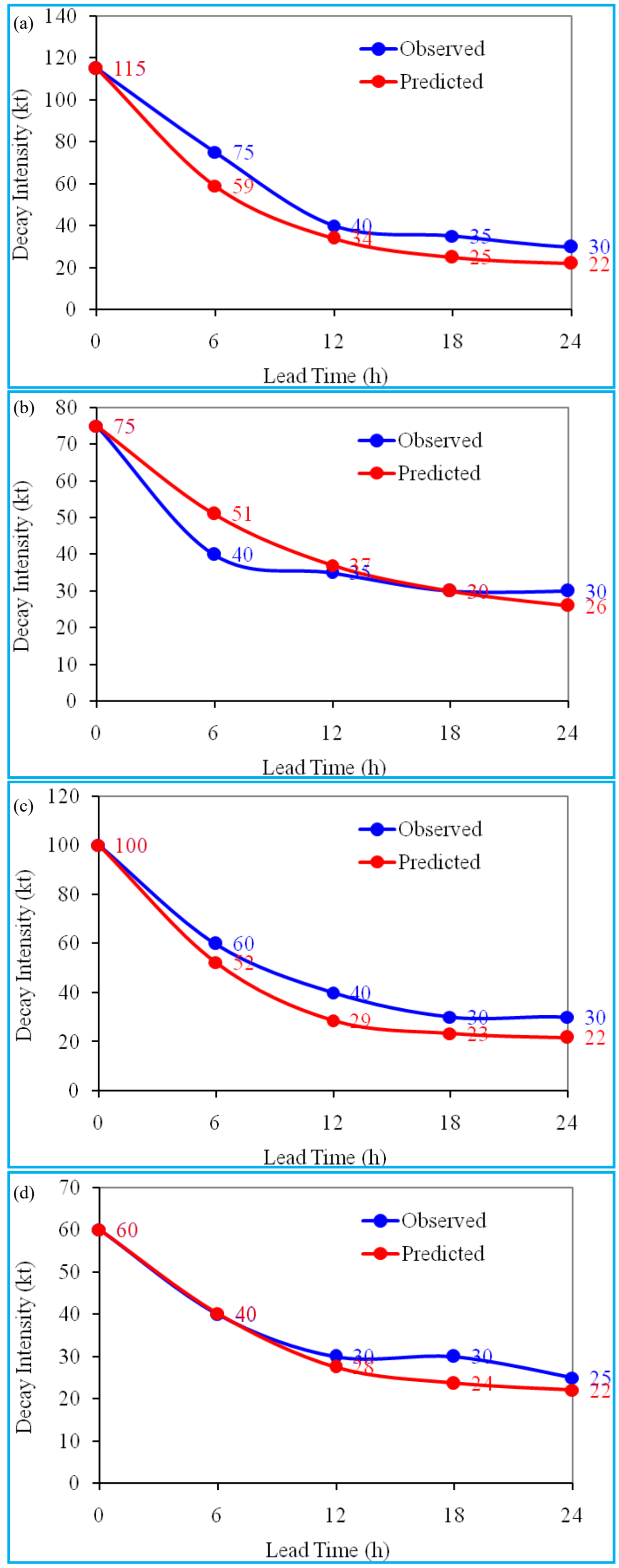


Fig. 24. Average absolute error (AAE) and root mean square error (RMSE) of decay (kt) after landfall during 2010-2019



Figs. 25(a-d). Decay forecasts by Decay model *vs* Observed decay for TCs (a) Phailin at the time of landfall, (b) Phailin updated after 6h of landfall, (c) Hudhud at the time of landfall and (d) Hudhud updated after 6 h of landfall forecast lead time 0-24 h to 96-120 h respectively and corresponding RMSE was ranged from 9.4 kt to 20.7 kt. Comparison of landfall intensity forecasts at different forecast lead time for three individual extremely severe cyclonic storms (PHAILIN, HUDHUD, FANI) and one very severe cyclonic storm (BULBUL) is shown in Figs. 23(a-d). As in the Fig. 14, *x* - axis in the Fig. 23 also shows forecast number (F1, F2, F3,….) along with forecast lead time (given in parenthesis). The average errors and four individual severe and extremely severe storm cases show that the model forecasts provided very useful guidance in real time.

5.4. *Probability forecast of Rapid Intensification (RI)*

The rapid intensification index (RII) is being used at IMD since 2011 for real-time forecasting of RI. In this section, the performance of RII for probability forecast of RI has been analyzed for the period 2011-2019. The skill of the RII has been assessed by computing the Brier score (BS) (Wilks, 2006). The Brier score is defined as:

$$\mathrm{BS} = \frac{1}{N}\sum (F-O)^2 \tag{3}$$

where, F is the probability of forecast, O the actual event observed (O = 0 if it did not happen and 1 if it happened) and N is the total number of forecasts. For example, when RI event is observed, a forecast probability of 75% would yield a BS = 0.0625 [*i.e.*, $(0.75\text{-}1.0)^2$]. The BS = 0 considered as the best score and 1 the worst score. For 284 forecast events during the period 2011-2019, the BS is found to be 0.079. For 24 RI cases and 260 Non-RI cases the BS were 0.494 and 0.041 respectively. The overall score (0.079) for all 284 RI and non-RI cases show RII achieved a good score for RI forecasting during the period.

5.5. *Decay after landfall*

The decaying intensity of TCs after landfall at 6 h interval up to 24 h is predicted by decay model (Roy Bhowmik *et al.*, 2005). The performance of the decay model has been analyzed for the period 2010-2019. The average forecast errors (kt) at 6 h interval valid up to 24 h during the period 2010-2019 are shown in Fig. 24. The AAE was ranged from 6.6 kt at 6 h to 3.3 kt at 24 h and corresponding RMSE was ranged from 8.6 kt to 4.4 kt with highest errors (AAE = 7.5 kt; RMSE = 9.9 kt) at 12 h. The sample size was 24, 22, 16 and 12 for 6 h, 12 h, 18 h and 24 h respectively. The decay and updated decay after 6 h of landfall of extremely severely cyclonic storms PHAILIN and HUDHUD are shown in Fig. 25. The mean error statistics and errors of two individual severe storm cases show that the model forecasts for decaying intensity after landfall were reasonably good.

6. Summary and conclusions

In this study, the evolution of forecast of dynamical-statistical cyclone prediction system (CPS) and forecast of NWP models used for real-time forecasting of TCs over the North Indian Ocean was carried out for the past decade 2010-2019. The important findings are summarized below.

(*i*) In the first component, analysis of performance of genesis potential parameter (GPP) shows that POD (0.95) was much higher than FAR (0.13) and near desirable value for BIAS (1.09) along with high CSI (0.84) and PC (0.84), which indicates that the GPP was skillful for cyclogenesis prediction. Comparison of the metrics for first five years (2010-2014) and second five years (2015-2019) suggest that FAR and BIAS has reduced and CSI and PC have increased in recent five years.

(*ii*) The consensus MME track forecast was found to be more consistent and outperformed all individual models during 2010-2019. The mean track forecast errors (DPE) of MME during the period was 53 km, 67 km, 84 km, 103 km, 120 km, 143 km, 160 km, 188 km, 220 km and 246 km for the forecast lead time 12 h, 24 h, 36 h, 48 h, 60 h, 72 h, 84 h, 96 h, 108 h and 120 h respectively. The MME forecast errors were less by 18-53%, 10-56%, 21-54%, 15-31% and 17-37% than all individual models for forecast lead time 24 h, 48 h, 72 h, 96 h and 120 h respectively. Further trend analysis of track error shows that during the period, DPE of MME has reduced significantly by about 41% to 55% for forecast lead time 24 h to 72 h. The average track forecast errors during the recent past five years (2015-2019) compared to previous five years (2010-2014) shows there has been improvement of forecasts for all models.

The most important component of cyclone forecasting is landfall point forecast. The average landfall error of MME was 31.5 km, 60 km, 57.5 km, 83.4 km and 127.2 km for forecast hours before 24 h, 48 h, 72 h, 96 h and 120 h respectively and corresponding landfall time error was 2.2 h, 3.8 h, 5.2 h, 5.8 h and 8.1 h.

Predicting the future location of TCs inherently involves some level of uncertainty represented by a cone of uncertainty, which actually represents the probable track of centre of a TC. The cone of uncertainty is found to be 11.7° and the centre of TCs for all forecast hours within the cone was 86% and 14% cases strayed outside of the cone.

(*iii*) The average intensity forecast errors AAE (RMSE) of SCIP model during 2010-2019 was 5.6 (7.5) kt, 8.3 (11.3) kt, 10.6 (14.8) kt, 12.6 (17.2) kt, 14.6 (19.3) kt, 15.7 (20.5 kt, 14.9 (19.7) kt, 15.7 (20.1) kt, 14.1 (17.0) kt and 12.6 (16.1) kt for forecast lead time 12 h, 24 h, 36 h, 48 h, 60 h, 72 h, 84 h, 96 h, 108 h and 120 h respectively. The AAE during the period 2010-2014 and 2015-2019 shows an improvement of 11-17% for 36 h to 72 h. The landfall intensity errors show that the AAE was ranged from 7.6 kt to 16.9 kt for forecast lead time up to 120 h and corresponding RMSE was ranged from 9.4 kt to 20.7 kt.

(*iv*) The performance of RII for probability forecast of rapid intensification (RI) during the period 2011-2019 was found to be skillful. The Brier score (BS) during the period was 0.079, where the BS is 0 considered as the best score and 1 and worst score. Individually, for 24 RI cases and 260 Non-RI cases the BS were 0.494 and 0.041 respectively. The overall score for 284 RI and non-RI cases show RII achieved a good score for RI forecasting during the period.

(*v*) The average absolute error (AAE) of the decay model for 6-hourly decaying intensity (after the landfall of TCs) forecasts (up to 24 h) shows that the error AAE (RMSE) were 6.6 (8.6) kt, 7.5(9.9) kt, 4.9 (7.7) kt and 3.3 (4.4) kt at 6 h, 12 h, 18 h and 24 h respectively. The mean error statistics and also errors of two individual severe storm cases show that the model forecasts were reasonably good.

The study shows that the dynamical-statistical technique can add skill to dynamical forecast. The performance analyses of NWP models and each component of CPS have brought out the probable forecast errors. Results show that track forecasts by MME outperformed all individual models and demonstrated the efficacy of the system (CPS) for improvement of cyclone forecast over the NIO in the past decade (2010-2019).

Acknowledgements

The authors are grateful to the Director General of Meteorology, India Meteorological Department, New Delhi for providing all the facilities to carry out this research work. Authors acknowledge the use of best-track data and track forecast data of tropical cyclones from the Regional Specialized Meteorological Centre (RSMC) in New Delhi and model data of NCEP, ECMWF, UKMO and JMA in this research work. Authors are also grateful to the anonymous reviewers for their valuable comments to improve the quality of the paper.

The contents and views expressed in this research paper are the views of the authors and do not necessarily reflect the views of their organizations.

References

Brier, G. W. and Allen, R. A., 1951, "Compendium of Meteorology", *Boston, American Meteorological Society, chapter - Verification of Weather Forecasts*, 841-848.

Brown, B. G. and Brandes, E., 1997, "An intercomparison of 2D storm motion extrapolation algorithms", Preprints, 28th Confon Radar Meteorology, Austin, TX, *Amer. Meteor. Soc*., 495-496.

Dvorak, V. F., 1972, "A technique for the analysis and forecasting of tropical cyclone intensities from satellite pictures", *NOAA Tech. Memo. NESS*, **36**, p15. [Available from NOAA/NESDIS, 5200 Auth Rd., Washington, DC 20333].

Dvorak, V. F., 1975, "Tropical cyclone intensity analysis and forecasting from satellite imagery", *Mon. Wea. Rev.*, **103**, 420-430.

Dvorak, V. F., 1984, "Tropical cyclone intensity analysis using satellite data", *NOAA Tech. Rep.*, **11**, p45. [Available from NOAA/ NESDIS, 5200 Auth Rd., Washington, DC 20333].

Dvorak, V. F., 1995, "Tropical clouds and cloud systems observed in satellite imagery : Tropical cyclones", *Workbook Vol. 2. NOAA/NESDIS*: 359pp. [Available from NOAA/NESDIS, 5200 Auth Rd., Washington, DC 20333].

Heming, J., 1994, "Verification of Tropical Cyclone Forecast Tracks at the UK Met Office", *NWP Gazette,* **1**, 2-8.

Heming, J., Prates, F., Bender, M. A., Bowyer, R., Cangialosi, J., Caroff, P., Coleman, T., Doyle, J. D., Dube, A., Faure, G., Fraser, J., Howell, B. C., Igarashi, Y., Taggart Cowan, R. M., Mohapatra, M., Moskaitis, J. R., Murtha, J., Rivett, R., Sharma, M., Short, C. J., Singh, A. A., Tallapragada, V., Titley, H. A. and Xiao, Y., 1019, "Review of Recent Progress in Tropical Cyclone Track Forecasting and Expression of Uncertainties", *Tropical Cyclone Research and Review*, **8**, 4**,** 181-218.

Jarrell, J. D., 1978, "Tropical cyclone strike probability forecasting", *Naval Environmental Prediction Research Facility Contractor Report,* CR 78-01, (NTIS AD-A065 545/6GA).

Kotal, S. D. and Bhattacharya, S. K., 2013, "Tropical Cyclone Genesis Potential Parameter (GPP) and its application over the North Indian Sea", *MAUSAM*, **64**, 1, 149-170.

Kotal, S. D. and Roy Bhowmik, S. K., 2011, "A Multi-model Ensemble (MME) Technique for Cyclone Track Prediction over the North Indian Sea", *Geofizika*, **28**, 275-291.

Kotal, S. D. and Roy Bhowmik, S. K., 2013, "Large-Scale Characteristics of Rapidly Intensifying Tropical Cyclones over the Bay of Bengal and a Rapid Intensification (RI) Index". *MAUSAM*, **64**, 1, 13-24.

Kotal, S. D., Bhattacharya, S. K. and Roy Bhowmik, S. K., 2014, **"**Development of NWP based objective Cyclone Prediction System (CPS) for North Indian Ocean Tropical Cyclones - Evaluation of performance", *Tropical Cyclone Research and Review*, **3**, 3, 162-177.

Kotal, S. D., Bhattacharya, S. K. and Roy Bhowmik, S. K., 2018, "Estimation of Tropical Cyclone Intensity over the north Indian Ocean - A challenge", *Meteorological Application*, **26**, 245-252**.**

Kotal, S. D., Kundu, P. K. and Roy Bhowmik, S. K., 2009, "Analysis of cyclogenesis parameter for developing and non-developing low pressure systems over the Indian Sea", *Nat. hazards*, **50**, 389-402.

Kotal, S. D., Roy Bhowmik, S. K., Kundu, P. K. and Das, A. K., 2008, "A Statistical Cyclone Intensity Prediction (SCIP) Model for Bay of Bengal", *J. Earth Syst. Sci.*, **117**, 157-168.

Lee, T. C., Knutson, T. R., Kamahori, H. and Ying, M., 2012, "Impacts of climate change on tropical cyclones in the western North Pacific basin. Part I : Past observations", *Tropical Cyclone Research and Review*, **1**, 213-230.

Mohapatra, M. and Sharma, M., 2019, "Cyclone warning services in India during recent years : A review", *MAUSAM*, **70**, 4, 635-666.

Roy Bhowmik, S. K., Kotal, S. D. and Kalsi, S. R., 2005, "An empirical model for predicting decaying rate of tropical cyclone wind speed after landfall over Indian region", *J. Appl. Meteor.*, **44**, 179-185.

Roy Bhowmik, S. K., Kotal, S. D. and Kalsi, S. R., 2007, "Operational tropical cyclone intensity prediction - An empirical technique", *Nat. Hazards*, **41**, 3, 447-455.

Roy Bhowmik, S. K., 2003, "An evaluation of cyclone genesis parameter over the Bay of Bengal using model analysis", *MAUSAM*, **54**, 2, 351-358.

Wilks, D. S., 2006, "Statistical Methods in the Atmospheric Sciences", *2nd ed. Academic Press*, p627.