52st Session FOR PARTICIPANTS ONLY
ESCAP/WMO Typhoon Committee ENGLISH ONLY
25-28 February 2020
Hongkong · China

Verification of Tropical Cyclone Operational forecast in 2019

Guomin CHEN, Mengqi YANG, Xiping ZHANG, Rijin WAN
(*Shanghai Typhoon Institute/CMA*)

CONTENTS

1. Introduction

An important key to making better predictions is having an understanding of the errors in current predictions. Subjective and objective verification of tropical cyclone (TC) forecasts give evidence regarding the accuracy and performance characteristics of TC forecasts and warnings. Verification analyses diagnose and quantify the systematic and random errors so that improvements can be made to operational forecasting methodologies and to the underpinning numerical models. This process also provides users of TC forecasts with information on the reliability of the forecasts, so that they can make better decisions accordingly. Particularly, forecasters need the verification results for different numerical weather prediction (NWP) models in order to use the multiple sources of guidance in an optimal fashion.

This report is primarily about *verification of tropical cyclone operational forecast in 2019*. As the conclusion of the typhoon season, forecast results are evaluated by comparing the projected positions and intensities to the corresponding post-storm derived “best track” positions and intensities for each TC. A forecast is included in the verification only if the system is classified in the final best track as a tropical cyclone at both the forecast’s initial time and at the projection’s valid time. In this report, we start with a short discussion of best track datasets, which are the first requirement for verifying TC forecasts. The next section describes deterministic forecast methods, which will be evaluated here including official forecast guidances, global models and regional models, and ensemble prediction systems will also be depicted. Last and most important, we will evaluate the tropical cyclone track, intensity forecast, which will include deterministic and ensemble predictions.

2. Best track

With the development of modern meteorological techniques, an increasing amount of observational data became available for creating a specialized tropical cyclone database. Currently, four agencies provide their own TC best track analyses for the WNP region: 1) the Japan Meteorological Agency (JMA) Regional Specialized Meteorological Center (RSMC) in Tokyo, 2) Shanghai Typhoon Institute of China Meteorological Administration, 3) Hong Kong Observatory, 4) Joint Typhoon Warning Center. Table 1 provide the data period, characteristics and wind averaging time information of these four best track datasets. It should be noted that the TC position, intensity and structural information usually differ among those agencies due to the lack of sufficient surface observations for TCs, as well as the different techniques used to estimate the position and intensity of a TC. Thus, differences in TC forecast performance may be obtained, depending on the best-track dataset used as a reference. In this annual report, we primely used RSMC-Tokyo best track dataset as the reference. As a complement, evaluation results which refer to other best track datasets are proposed to reveal the effect on final forecast performance.

Table 1. Descriptions of western North Pacific best-track datasets.

Agency	Period	Characteristics	Wind
RSMC Tokyo	1951 to present	Includes extratropical cyclone stage, longitude, latitude, MCP and TS markers since 1951; MSW and typical severe wind radii since 1977 (without TD cases).	10 min
CMA	1949 to	Includes sub-centers, some double eyewall cases/coastal severe wind of	2 min

	present	landfalling TCs (until 2004); includes TD cases; extratropical cyclone stage; longitude, latitude, MSW and MCP since 1949.	
HKO	1961 to present	Includes TD cases; longitude, latitude, MSW and MCP since 1961 (extratropical cyclone stages are not marked).	10 min
JTWC	1945 to present	Includes TD cases; extratropical cyclone stage since 2000; longitude, latitude, and MSW since 1945; MCP and TC size parameters since 2001.	1 min

3. TC position and intensity forecast data

In this report, TC position and intensity forecast results from 5 official guidances, 6 global models and 6 regional models are evaluated. These totally 17 methods are deterministic forecast guidance, detail explanations including their abbreviations, short description and source agencies are listed in Table 2. Additional verification on position forecast of ensemble prediction system will also be show in this report. The ensemble prediction systems (EPSs) include ECMWF-EPS, NCEP-GEFS, UKMO-EPS, JMA-GEPS and MSC-CENS.

Table 2. Details of forecast guidances.

Category		**Abbreviation**	**Full name or short description**	**Source**
Deterministic	**Official**	***CMA***	*China Meteorological Administration*	CMA
		JMA	*Japan Meteorological Agency*	JMA
		JTWC	*Joint Typhoon Warning Center*	JTWC
		KMA	*Korea Meteorological Administration*	KMA
		HKO	*Hong Kong Observatory*	HKO
	Global	***ECMWF-IFS***	*Integrated Forecasting System of ECMWF*	ECMWF
		JMA-GSM	*Global Spectral Model of JMA*	JMA
		NCEP-GFS	*Global Forecast System of NCEP*	NCEP
		KMA-GDAPS	*Global Data Assimilation and Prediction System of KMA*	KMA
		UKMO-MetUM	*Unified Model system of UKMO*	UKMO
	Regional	***HWRF***	*The atmosphere-ocean coupled Hurricane Weather Research and Forecast modeling system*	NCEP/EMC
		GRAPES-TCM	*Regional TC-forecasting model based on the Global/Regional Assimilation and PrEdiction*	STI/CMA
		GRAPES-TYM	*Regional TC-forecasting model based on the Global/Regional Assimilation and PrEdiction*	CMA
		CMA-TRAMS	*Tropical Regional Atmosphere Model for the South China Sea based on GRAPES GRAPES*	ITMM/CMA
		SHANGHAI-TCM	*Regional TC forecasting model based on WRF*	STI/CMA
Ensemble		***ECMWF-EPS***	*ECMWF Ensemble Prediction System*	ECMWF
		JMA-GEFS	*JMA Global Ensemble Forecast System*	JMA
		MSC-CENS	*MSC Canada Ensemble System*	MSC
		NCEP-GEFS	*NCEP Global Ensemble Forecast System*	NCEP
		UKMO-EPS	*UKMO Ensemble Prediction System*	UKMO

4. Performance of TC track forecast

TC position error or track error is defined as the great-circle difference between a TC's forecast center position and the best track position (unless otherwise stated, the following will be used RSMC-Tokyo's best track datasets as reference) at the verification time. TC position errors typically are presented as mean errors for the samples of entire typhoon season.

4.1 Subjective deterministic forecasts

Normally, the subjective deterministic forecasts issued by official typhoon prediction agencies. In 2019, position errors from 5 official typhoon prediction agencies (JMA, CMA, JTWC, KMA and HKO) are 87.6-107.3km, 142.4-174.1km, 202.2-243.2km, 261.2-305.8km and 349.5-401.2km at the lead time level of 24, 48, 72, 96 and 120 h, respectively. The mean position errors in 2019 for the most agencies were much larger than the ones in 2018 at lead time level less than 4 days (Fig.1).

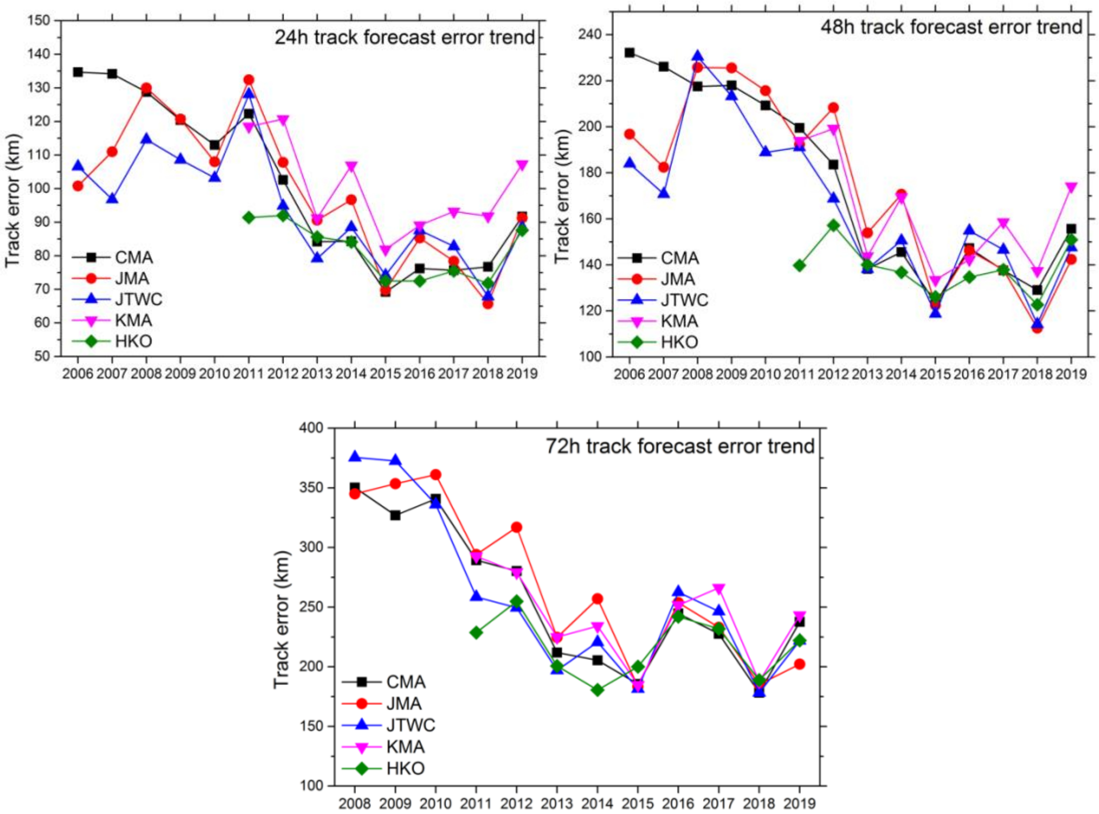


Fig.1 Track error evolutions of each official typhoon prediction agencies at the lead time of 24, 48 and 72 h.

To assess subjective track forecast skills, the track forecast error is compared to the error of a persistent climatology model with no information on the state of the atmosphere during the storm. Fig.2 shows the track forecast skill scores at the lead times of 24 and 48 h for official guidances from 2010 to 2019. All the forecast methods have positive skill scores indicating that over the past eight years, these forecast accuracies are better than the climatic persistence method. However, track forecast skill scores for official guidances were decreasing since 2017.

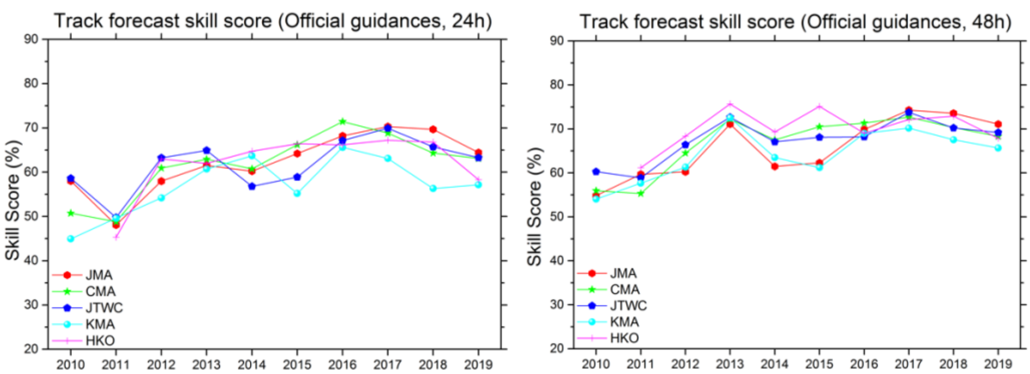


Fig.2 Track forecast skill scores evolutions of official guidances at the lead times of 24 h (left) and 48 h (right).

The along-track and cross-track bias of official guidances from 24 to 120 h are showed in Fig.3. The figures show that with increased forecast lead times, the forecasted TCs propagated, on average, too slow for most official guidances. There are not obvious leftward or rightward biases for official guidances at the lead time levels less than 48 h in 2019.

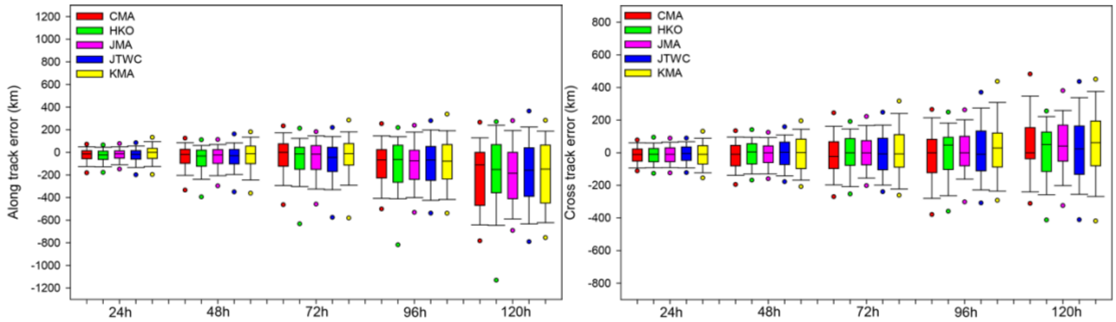


Fig.3 Along-track (left) and cross-track (right) biases for official guidances. The bar in the middle of the plot represents the median values of errors, the lower and upper ends of the boxes represent the 25th and 75th quantile values. The bars below and above the box represent the non-outlier extreme values, and the upper and lower circles represent the 95% and 5% quantile values.

Fig.4 are track error rose (TER) diagrams which is a useful tool to evaluate the systematic bias of a track forecast method. TER uses the same concept as a "wind rose" diagram. Fig.4 shows TER representations of the distributions of direction and magnitude position-errors for five official guides at 72 h lead time in 2019. In TER diagrams, each color bar represents a different magnitude of position error, and the length of each alignment of the color bars represents the proportion for each azimuthal angle. TER diagram reveals the position error distribution (both error magnitude and percentage of sample size) at 8 azimuthal directions. Take the TER diagram of CMA at lead time of 72 h as an example, their forecasted TC positions most concentrate on southwest at 72 h, with the percentage of sample size at southwest direction are close to 25%, and the dominant position error range at 100 – 200 km, 300 – 400 km and 200 – 300 km are about 8% (yellow), 5% (blue) and 4% (red), respectively. As we can see that, for the five official guidances, most forecast TC position were concentrated at southwest side compare to OBS position, and largest track errors were also appeared at southwest directions.

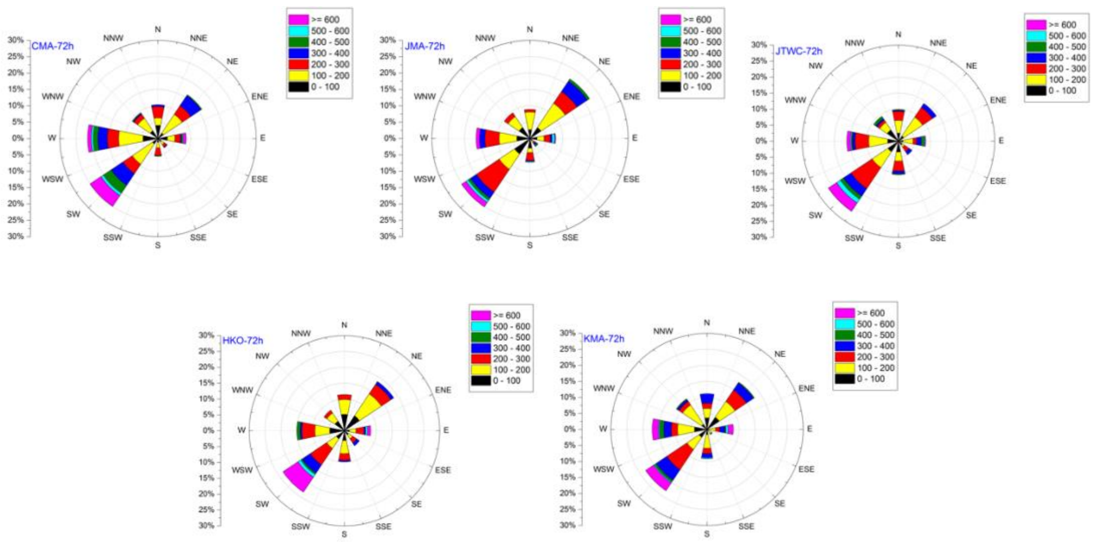


Fig.4 Track error rose (TER) diagrams for official guidances at the lead time of 72 h.

Generally, due to the limitations of different technologies, there exist large variations in TC center position estimates from different operational agencies. In order to demonstrate how different reference TC data, which called TC “true” position may influence the final verification results, this report also provide the track errors of all above 5 official guidances, which recalculate by referring to difference TC best tracks (RSMC-TOKYO and CMA) or real-time operational TC position (JTWC, HKO and KMA) at lead time levels of 24, 48, 72, 96 and 120 h (table 3).

Table 3. Mean position error of official agencies in 2019. (Calculated by difference reference data, Unit: km)

Agency	Reference	24 h	48 h	72 h	96 h	120 h
JMA	RSMC-Tokyo best	91.3(720)	142.4(486)	202.2(370)	261.2(269)	365.2(192)
	CMA best	73.6(781)	133.6(549)	193.6(396)	279.3(258)	374.3(164)
	JTWC real-time	84.5(431)	141.9(314)	211.3(235)	263.3(161)	374.8(110)
	KMA real-time	75.3(432)	122.6(325)	180.8(237)	247.2(165)	346.6(113)
	HKO real-time	81.0(371)	134.5(242)	195.7(165)	232.1(103)	311.3(65)
CMA	RSMC-Tokyo best	91.7(608)	155.7(467)	237.7(348)	305.8(266)	401.2(184)
	CMA best	77.0(791)	137.1(595)	217.1(412)	287.9(276)	373.7(176)
	JTWC real-time	87.2(401)	148.7(313)	236.3(233)	277.4(161)	351.8(107)
	KMA real-time	83.4(424)	137.1(321)	208.9(233)	250.5(161)	342.1(111)
	HKO real-time	85.1(346)	155.5(252)	237.9(168)	277.6(105)	340.2(65)
JTWC	RSMC-Tokyo best	88.5(527)	147.6(424)	222(353)	293.7(272)	374.1(196)
	CMA best	72.5(404)	124.4(317)	189.5(238)	260.7(167)	363.1(105)
	JTWC real-time	86.6(459)	144.5(361)	221.2(280)	290.5(203)	343.7(142)
	KMA real-time	76.9(401)	125.6(314)	185.4(233)	248.6(163)	352.9(113)
	HKO real-time	77.8(334)	135.1(243)	196.4(168)	241.9(105)	313.2(68)
KMA	RSMC-Tokyo best	107.3(476)	174.1(376)	243.2(304)	299.2(208)	390.7(138)
	CMA best	91.1(431)	154.7(319)	219.2(235)	281.3(151)	362.9(97)
	JTWC real-time	103.9(399)	173.5(314)	239.9(235)	272.7(150)	357.6(103)
	KMA real-time	91.7(434)	153.0(327)	203.8(238)	253.4(155)	350.7(103)
	HKO real-time	102.0(327)	175.9(238)	239.5(162)	260.9(92)	320.7(59)
HKO	RSMC-Tokyo best	87.6(352)	150.9(279)	222.1(193)	284.3(128)	349.5(72)
	CMA best	77.5(291)	135.3(205)	192.7(136)	265.8(87)	373.0(51)
	JTWC real-time	88.0(275)	153.7(194)	237.6(127)	301.9(79)	316.4(47)
	KMA real-time	81.2(271)	138.2(189)	190.7(125)	207.8(77)	281.8(48)
	HKO real-time	85.5(320)	145.2(230)	211.4(147)	257.7(91)	315.9(50)

Fig.5 shows the variation intervals of official agencies’ track errors by referring to different TC best track or real time operational position. The solid lines with different colors and symbols in Fig.5 indicate the mean track errors by referring to different best track or operational TC position. The upper and lower shaded areas indicate the maximum and minimum track errors by referring to different best track or operational TC position. Fig.5 shows that there may exist 3% - 12% track error varieties while using different observation data as reference to evaluate official guidances’ track error at different lead time levels in 2019.

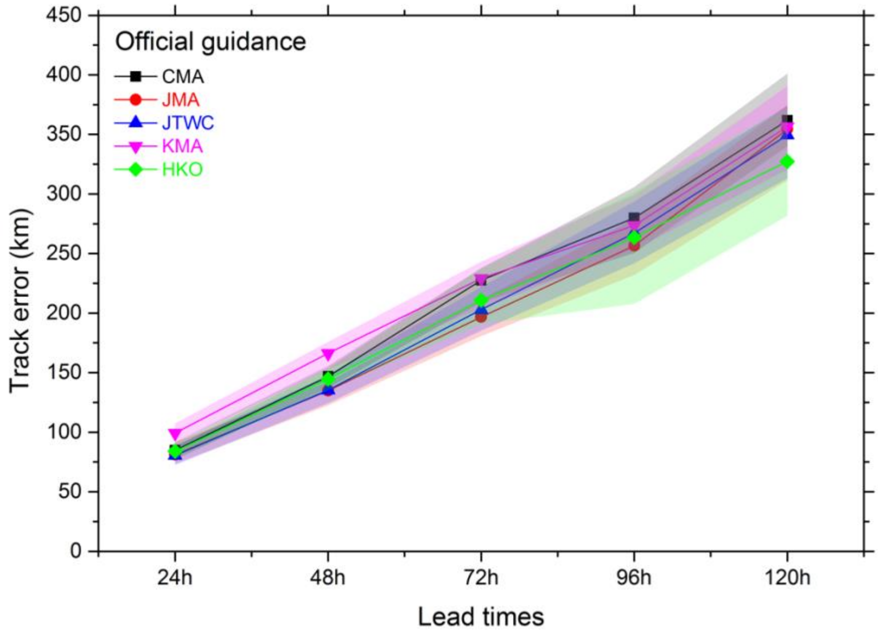


Fig.5 Variation intervals of track errors for official guidances by referring to different TC best tracks or real-time operational position.

4.2 Objective deterministic forecasts

In 2019, position errors for 5 global models are on intervals of 75.1 – 99.9 km, 128.0 – 181.3 km, 187.8 – 259.6km, 252.7 – 346.5 km and 282.1 – 465.7 km, and for 5 regional models are on intervals of 78.0 – 101.9 km, 136.4 – 180.8 km and 219.7 – 284.8 km at the lead time level of 24, 48 and 72 h, respectively. Fig.6 shows track error trends for most of the global and regional models at the lead time of 24, 48 and 72 h.

Table 4. Mean position error of objective deterministic methods in 2019. (Unit: km)

Method	Lead time	24h	48h	72h	96h	120h
Global Model	ECMWF-IFS	75.1(218)	128.0(178)	187.8(140)	252.7(104)	310.0(75)
	JMA-GSM	95.2(567)	159.7(490)	233.4(395)	318.0(300)	465.7(218)
	NCEP-GFS	79.9(365)	144.7(329)	223.0(276)	346.5(221)	432.8(161)
	KMA-GDAPS	99.9(21)	181.3(19)	259.6(15)	306.7(13)	282.1(11)
	UKMO-MetUM	86.6(288)	149(235)	217.2(184)	272.8(143)	375.2(110)
Regional Model	SHANGHAI-TCM	78.0(199)	136.4(159)	219.7(123)	/	/
	GRAPES-TCM	101.9(409)	180.8(363)	271.0(286)	/	/
	GRAPES-TYM	86.7(511)	154.3(431)	253.6(337)	412.9(241)	566.7(164)
	CMA-TRAMS	94.8(243)	151.7(204)	254.1(160)	/	/
	HWRF	96.2(294)	176.5(262)	284.8(212)	465.0(163)	634.0(115)

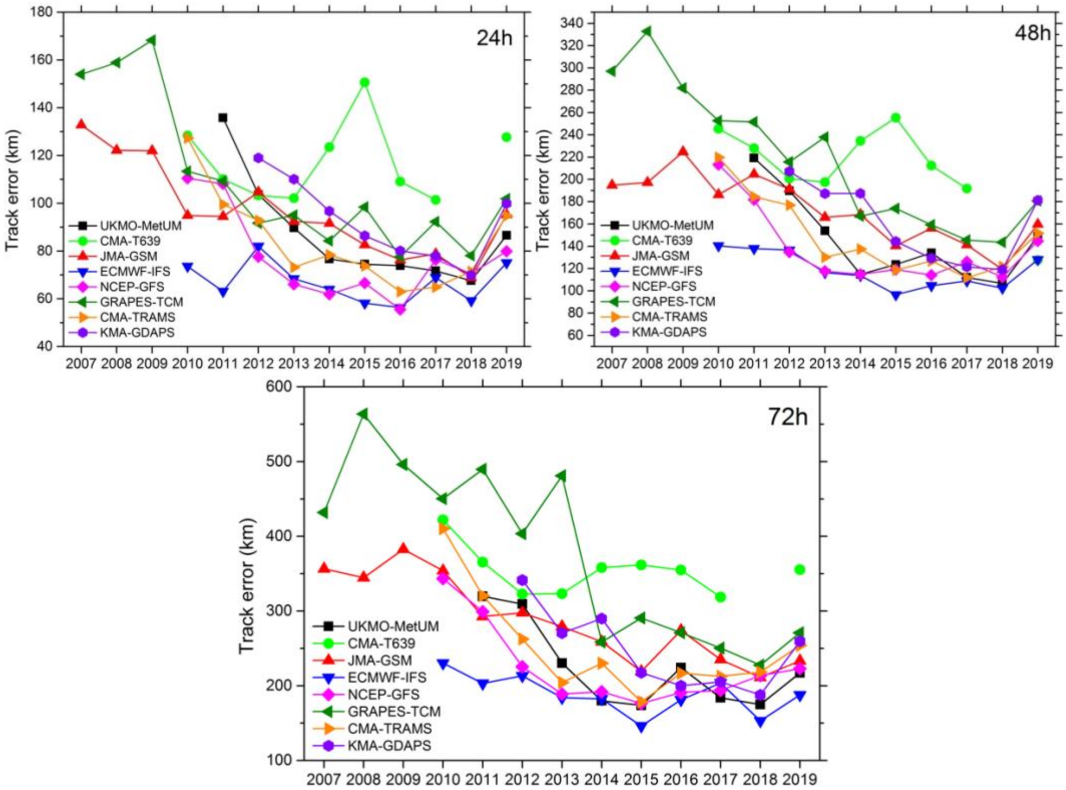


Fig.6 Same as fig.1 but for objective deterministic forecasts.

Fig.7 shows the track forecast skill scores at the lead times of 24 and 48 h for regional and global models from 2010 to 2019. Delightedly, all the models had positive skill scores.

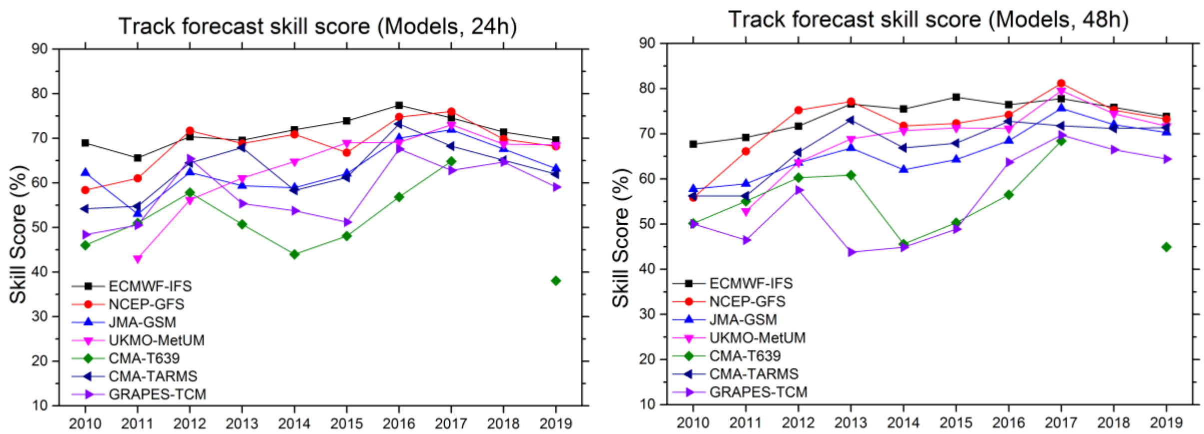


Fig.7 Track forecast skill scores evolutions of global and regional models at the lead times of 24 h (left) and 48 h (right).

An alternative approach for examining average error is to consider the distributions of errors, as shown in Fig.8. This analysis approach not only shows the entire performance of each model's track forecast at each lead time but also provides a straightforward method of understanding the annual improvements of each global model. This methodology is developed to evaluate the uncertainty in verification measures with confidence intervals and paired statistical tests, and to provide a consistent set of results to allow forecasts from the various models to be compared and fairly evaluated. In Fig.8, box plots summarize the distribution of ECMWF-IFS, NCEP-GFS and UKMO-MetUM's track forecast errors from 2010 to 2019. It clearly shows that for each lead time, decreases occur in the values of each quantile from 2010 to 2015, and the forecast accuracies at 72, 96, and 120 h in 2015 are nearly the same or better than the forecast accuracies in 2010 at 24, 48 and 72 h, respectively. However, the global models have not made significant progress in the last four years.

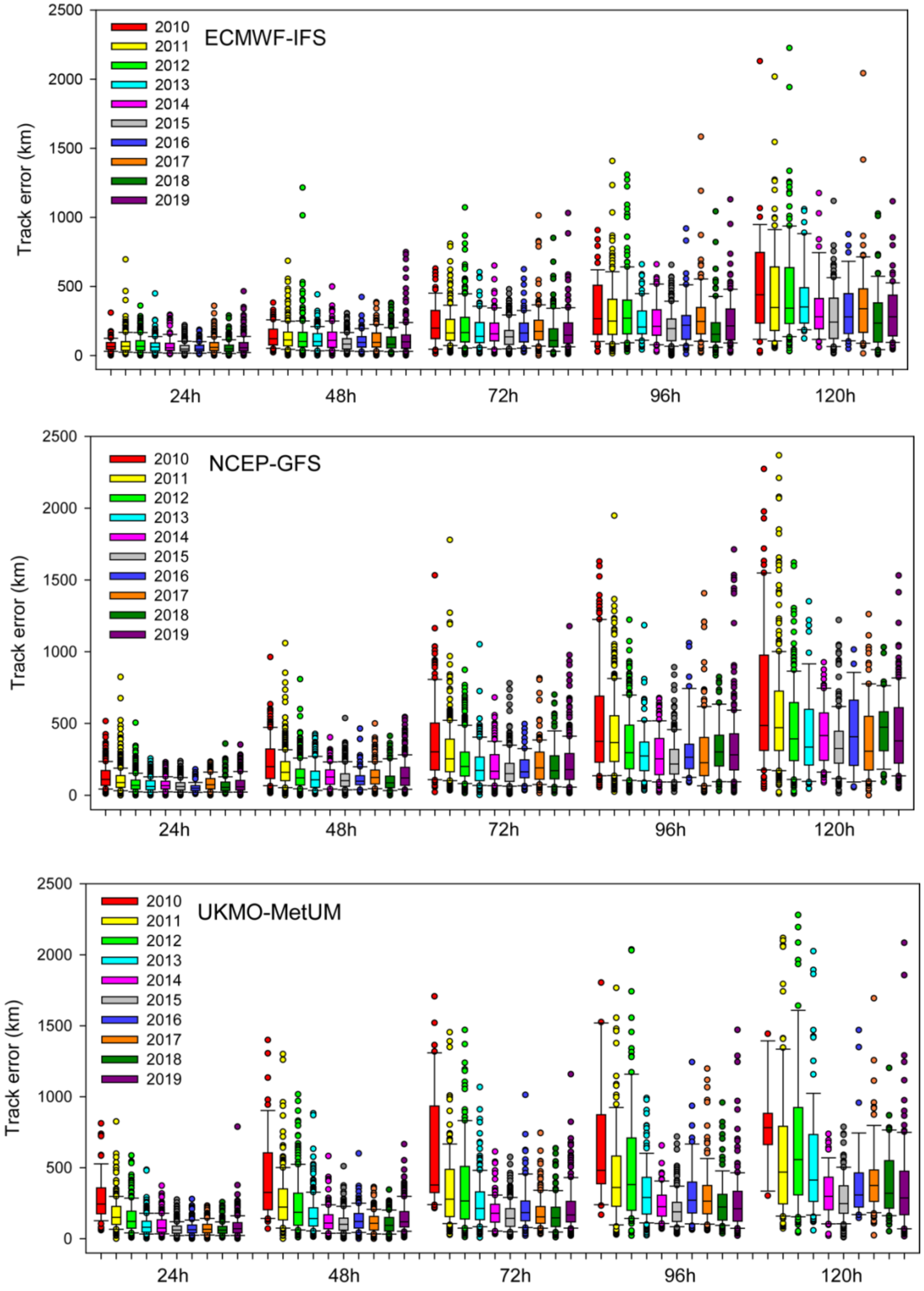


Fig.8 Box plots of position errors for ECMWF-IFS, NCEP-GFS and UKMO-MetUM in TC track forecasts from 2010 to 2019. The bar in the middle of the plot represents the median values of errors, the lower and upper ends of the boxes represent the 25th and 75th quantile values. The bars below and above the box represent the non-outlier extreme values, and the circles represent the outliers.

Fig.9 shows the along-track and cross-track biases of global models from 24 to 120 h. With the lead time increasing, the forecasted TCs from both regional and global model propagated a little slower than observations. Forecasted TCs from NCEP-GFS, KMA-GDAPS and JMA-GSM propagated, on average, distinct rightward, however, forecast TCs form ECMWF-IFS and UKMO-MetUM propagated leftward.

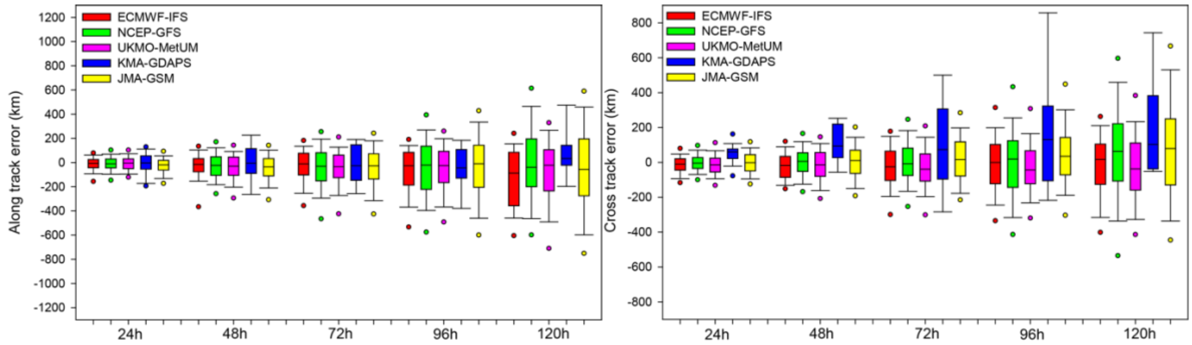


Fig.9 Along-track and cross-track biases for global models.

Fig.10 presents polar scatter plots of the mean combined direction and magnitude error relative to the actual storm locations for global and regional models at different lead times in 2019. Each models' systematic track forecast bias is clearly shown in Fig.10. The placement of lead time labels with different text colors for different models denotes the annual mean locations relative to the actual typhoon locations obtained from the best-track dataset. Fig.10 shows that the systematic bias of each global model is obviously different. With the forecast lead time increasing, JMA-GSM did not show obviously systematic bias in 2019. However, the systematic bias of ECMWF-IFS and UKMO-MetUM tended to west within the lead time 120 h. NCEP-GFS tended to west side within 96 h, but there was no bias at the lead time of 120 h. For regional models, the systematic biases of GRAPES-TCM, CMA-TRAMS and HWRF tend to west side. GRAPES-TCM tended to northeast side within 96 h, then turned to west side at lead time of 120 h. There was no obvious systematic bias for SHANGHAI-TCM in 2019. Plots like those in fig.10 provide information that is useful for the pre-estimation of the bias of a certain method.

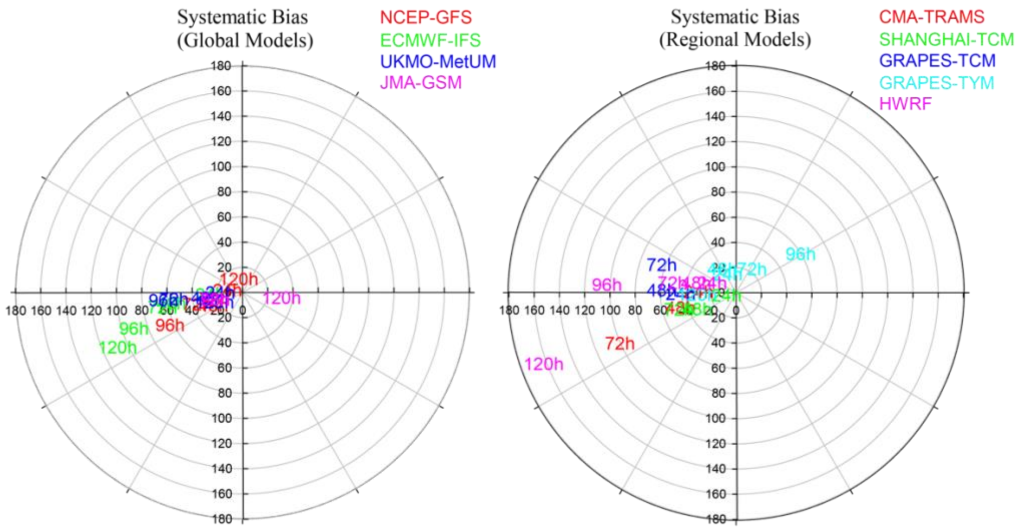


Fig.10 Polar scatter plots depicting the mean combined direction and magnitude errors relative to the actual storm location for each model at different lead times in 2019.

Similar as fig.4, fig.11 show the position error rose diagrams of the distribution of direction and magnitude position errors for global and regional models at the lead time of 72 h. It is useful for model developer to further understanding the model's forecast characteristic combined using fig.10 and fig.11.

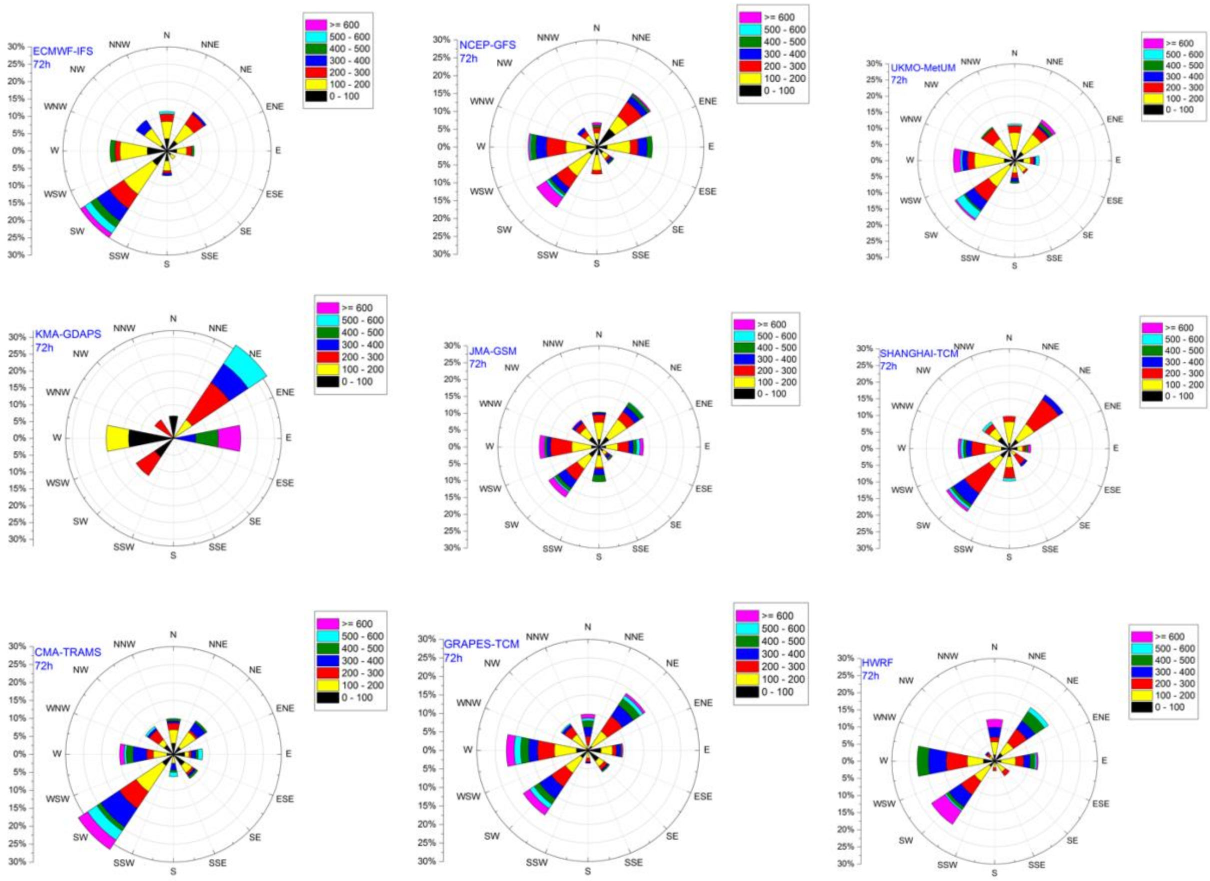


Fig.11 TER diagrams for global and regional models at the lead time of 72 h.

Fig.12 shows the variation intervals of track errors for global (left) and regional (right) models by referring to different TC best tracks or real-time operational position. It can be found from fig.12 that, by using different observation data as reference, lobal models exist 3% - 5% track error varieties with the lead time levels less than 72 h, and will increase to 5% - 8% at 120 h. The track error varieties from regional models were larger than the ones from global models.

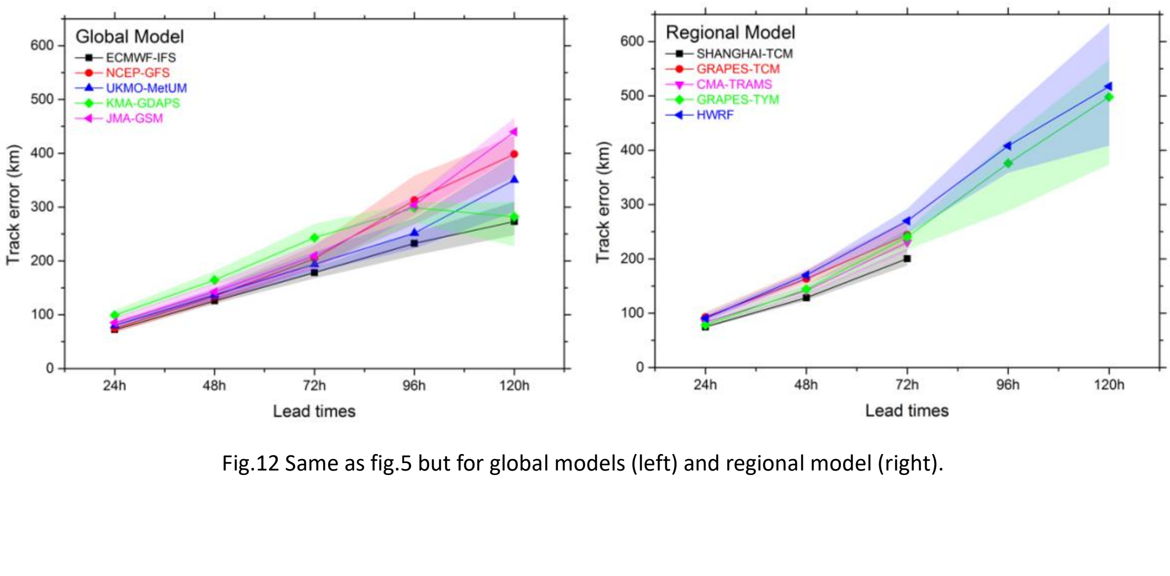


Fig.12 Same as fig.5 but for global models (left) and regional model (right).

4.3 Ensemble prediction systems

To evaluate the performance of the TC track forecasts of each EPS which mentioned in section 2, we first treated the ensemble forecasts as deterministic by summarizing the ensembles using the mean applied to the members. Fig.13 shows the ensemble mean track errors for five EPSs for both non-homogeneous and homogeneous. The detail values of each EPSs' errors at different lead time levels are listed in table 5 (non-homogeneous) and table 6 (homogeneous). Within the lead time levels less than 36 h, NCEP-GEFS got the best performance. At the lead time levels from 48 to 96 h, ECMWF-EPS was the best ensemble system. At the lead time levels beyond 96 h, UKMO-EPS outperform anyother ensemble system. The mean position error at the lead time of 120 h is for the best ensemble system approaching 400 km.

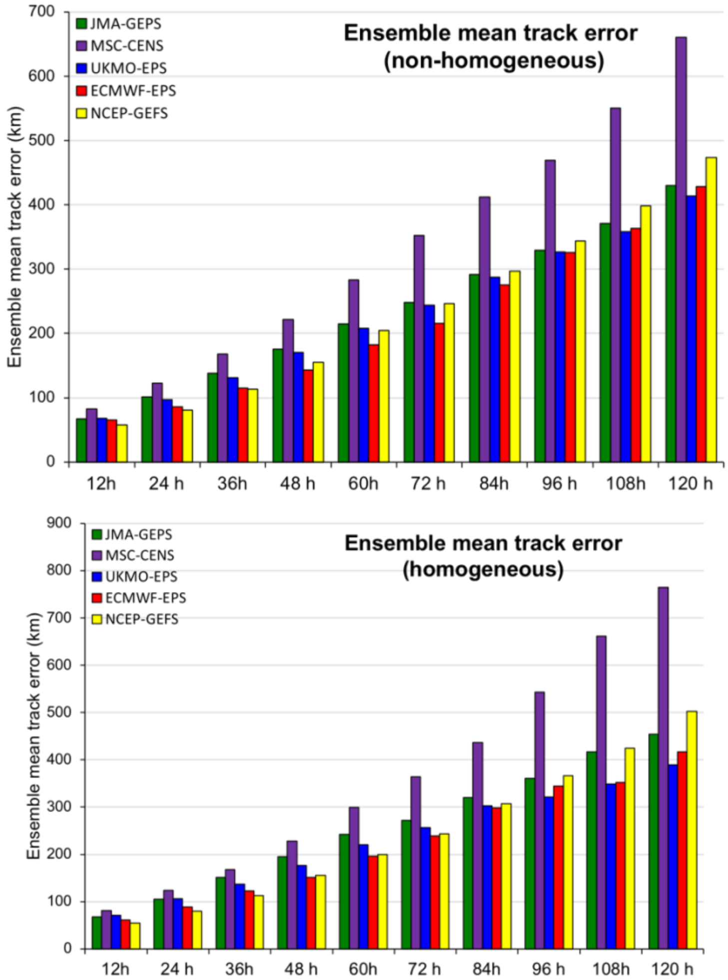


Fig.13 Ensemble mean track errors of five EPSs for both non-homogeneous (up) and homogeneous (down) comparison in 2019.

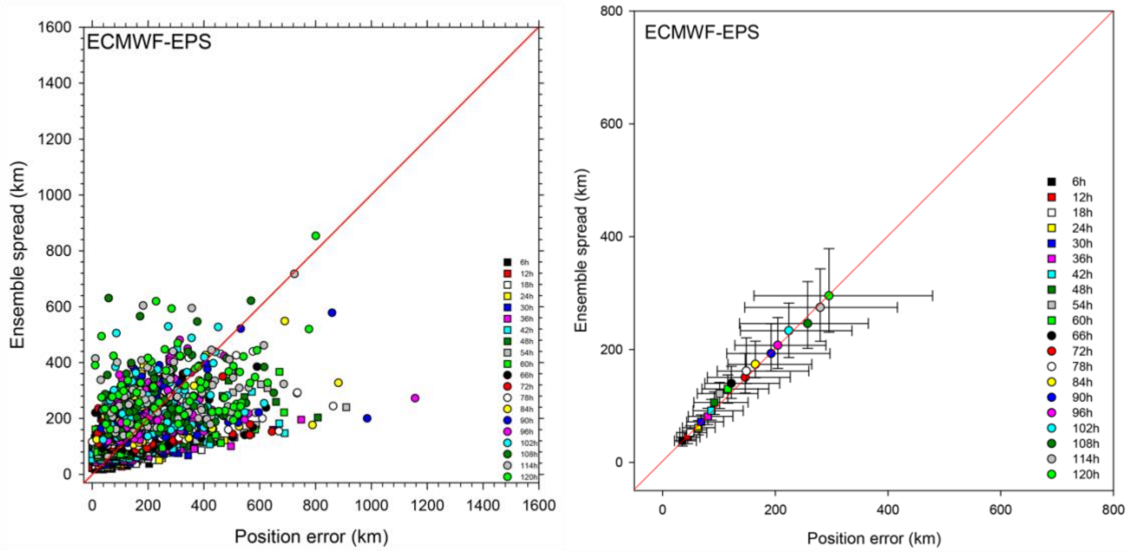
Table 5. Ensemble mean track errors in 2019 (Non-homogeneous comparison).

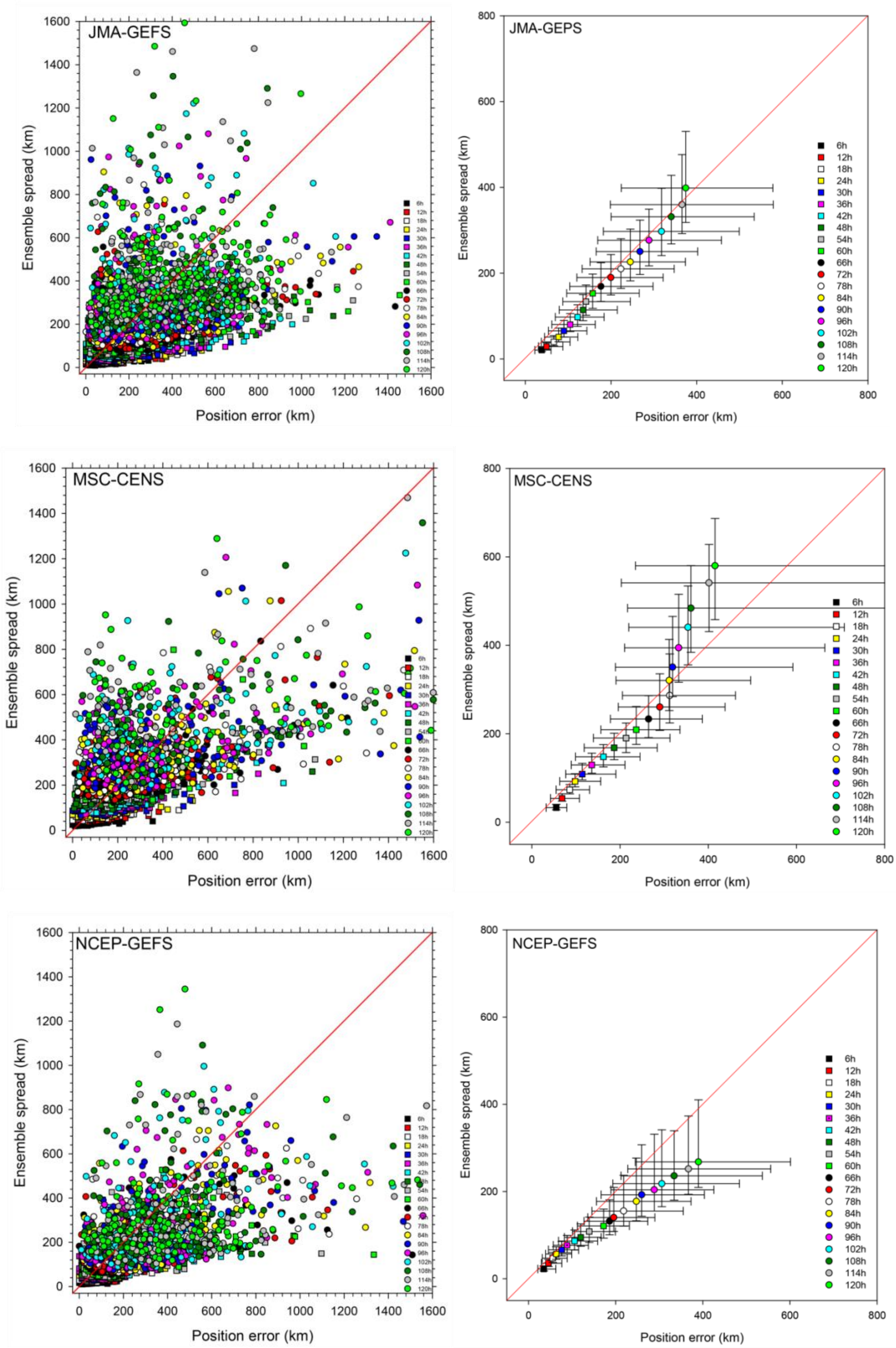
	12h	24h	36h	48h	60h	72h	84h	96h	108h	120h
JMA-GEPS	66.9(544)	101.0(523)	137.9(487)	175.6(448)	214.5(409)	247.9(368)	291.3(326)	328.9(286)	371.0(248)	430.2(212)
ECMWF-EPS	65.5(259)	85.7(243)	115.0(227)	142.8(211)	182.6(191)	215.7(171)	275.7(154)	325.7(136)	363.0(117)	428.7(101)
MSC-CENS	82.8(238)	122.8(224)	168.2(209)	222.1(192)	283.0(172)	352.4(151)	412.1(132)	469.4(113)	550.2(97)	660.5(81)
NCPE-GEFS	57.4(399)	80.8(371)	113.2(342)	155.0(309)	204.8(276)	246.7(242)	297.0(210)	343.4(179)	398.2(155)	473.2(131)
UMKO-EPS	68.0(524)	97.4(496)	131.6(461)	170.5(422)	207.7(374)	244.3(334)	287.4(295)	326.3(256)	358.6(216)	413.9(189)

Table 6. Ensemble mean track errors in 2019 (Homogeneous comparison).

	12h	24h	36h	48h	60h	72h	84h	96h	108h	120h
Sample size	(130)	(126)	(117)	(108)	(96)	(85)	(75)	(64)	(54)	(45)
JMA-GEPS	67.7	105.6	151.0	194.9	242.0	272.3	319.7	360.6	416.6	454.5
ECMWF-EPS	61.5	88.4	122.6	151.5	196.1	238.5	298.3	344.3	352.2	417.0
MSC-CENS	81.4	124.1	168.1	227.9	299.2	364.6	436.0	542.9	661.4	764.5
NCPE-GEFS	54.8	79.8	112.8	155.8	199.0	243.0	306.9	366.2	424.1	502.2
UMKO-EPS	70.7	106.0	136.4	176.8	220.4	256.3	302.6	321.1	348.6	389.0

The ensemble spread is an indicator of forecast uncertainties, which is not linearly related to mean position error. When the spread is large, the mean position error may be small and vice-versa. Traditionally, researchers apply a scatter plot of position error and ensemble spread to analyze the relationship of forecast uncertainty to the error of a particular EPS. A bidirectional scatter plot is adopted in the present report to reanalyze the traditional scatter plot. In the bidirectional scatter plot (Fig.14), the blocks in the middle of the plot represent the mean value of the spread or position error. The lower (left) and upper (right) bars represent the 25th and 75th quantile values. It is found that the median ensemble spreads and position errors are almost the same for lead times from 6 to 240 h for ECMWF-EPS and JMA-GEPS. The median ensemble position errors became larger than ensemble spreads with lead time increasing for NCEP-GEFS and UKMO-EPS. For MSC-CENS, the median ensemble position errors almost the same as ensemble spreads within 84 h, then the spreads became larger than errors from 90 h to 120 h.




JMA-GEFS
Ensemble spread (km)
Position error (km)
JMA-GEPS
Ensemble spread (km)
Position error (km)
6h
12h
18h
24h
30h
36h
42h
48h
54h
60h
66h
72h
78h
84h
90h
96h
102h
108h
114h
120h
MSC-CENS
Ensemble spread (km)
Position error (km)
MSC-CENS
Ensemble spread (km)
Position error (km)
NCEP-GEFS
Ensemble spread (km)
Position error (km)
NCEP-GEFS
Ensemble spread (km)
Position error (km)

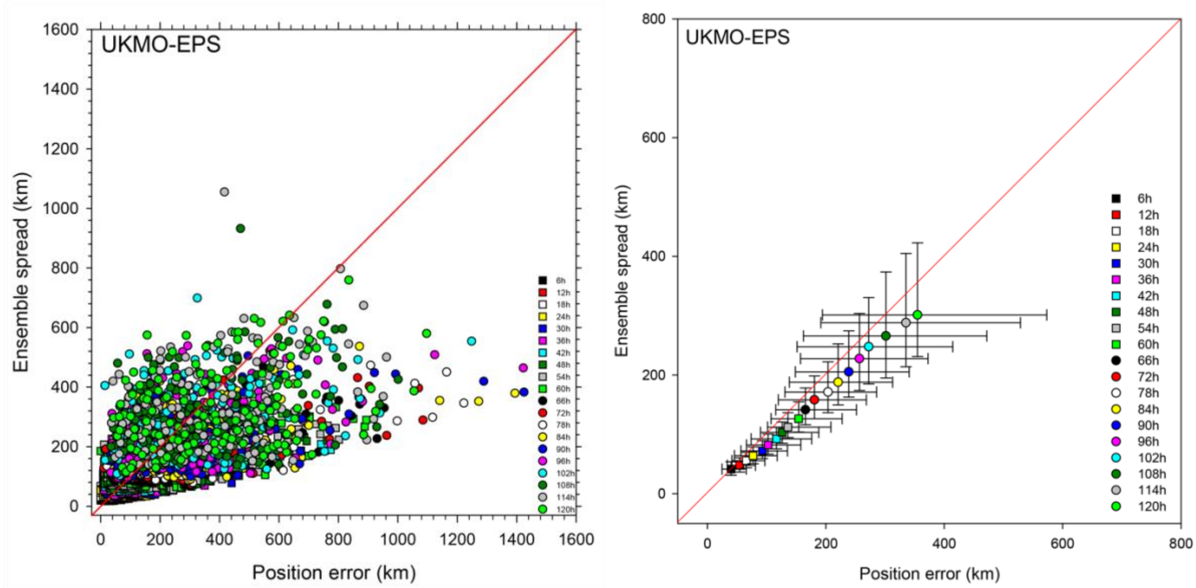


Fig.14 Scatter plot (left panels) and bi-directional track forecast scatter plot (right panels) for EPSs. The blocks in the bi-directional track forecast scatter plots represent the mean values of spread or position error, and the lower (left) and upper (right) bars represent the 25th and 75th quantile values.

5. Performance of TC intensity forecast

Forecast intensity error (i.e., maximum wind speed and minimum pressure) is defined as the mean absolute error or mean relative error of the difference between the forecast and best-track intensity for the time of forecast verification. MAE indicates the average magnitude of the error, whereas MRE measures the bias in the forecast. Table 7 shows the MAE of the maximum wind speed forecast for each method at each lead time in 2019. The wind speeds of all forecast methods are converted to 10-min averages according to the WMO documentation[1].

Table 7. Mean absolute error of maximum wind speed for each forecast method at the lead times of 24, 48, 72, 96 and 120 h in 2019. The numbers in brackets are sample sizes. (Unit: m/s)

Method	Lead time	24h	48h	72h	96h	120h
Global Model	ECMWF-IFS	7.3(218)	9.1(178)	9.8(140)	10.3(104)	11.9(75)
	JMA-GSM	7.6(567)	11.5(490)	14.0(395)	14.5(300)	14.8(218)
	NCEP-GFS	7.2(365)	9.1(329)	10.3(276)	12.1(221)	12.5(161)
	KMA-GDAPS	5.9(21)	8.5(19)	9.6(15)	11.2(13)	12.0(11)
	UKMO-MetUM	10.5(288)	11.9(235)	13.7(184)	16.0(143)	17.0(110)
Regional Model	SHANGHAI-TCM	7.1(199)	8.2(159)	9.7(123)	/	/
	GRAPES-TCM	6.9(408)	9.3(362)	10.8(285)	/	/
	GRAPES-TYM	7.9(511)	9.4(431)	11.0(337)	12.9(241)	14.8(164)
	CMA-TRAMS	7.9(241)	10.6(199)	12.1(157)	/	/
	HWRF	8.8(294)	10.7(262)	12.3(212)	13.3(163)	13.0(115)

1 Guidelines for converting between various wind averaging periods in tropical cyclone conditions. World Meteorological Organization, TCP Sub-Project Report, WMO/TD-No.1555.

5.1 Subjective forecasts

Same as table 3, table 8 show the mean intensity error of official agencies by referring to difference "OBS" data in 2019 and Fig.15 shows the variation intervals of intensity errors for official guidances by referring to different TC best tracks or real-time operational intensity. It can be found from fig.15 that, the verification results may exist a 15% -25% difference by using different observation data as reference. This indicates compare to TC position, the differences of observed TC intensity among the best tracks or real-time operational records even larger.

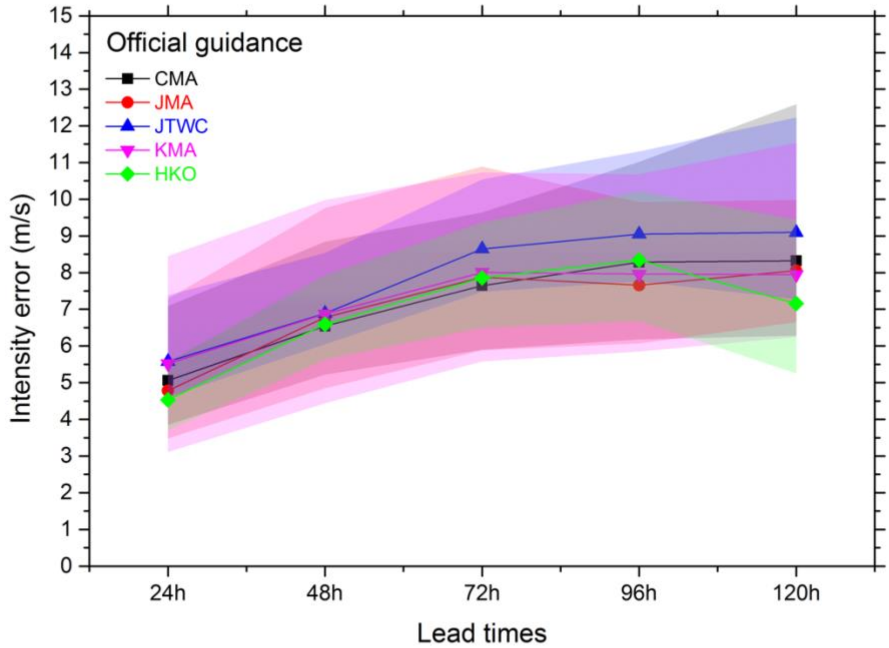


Fig.15 Variation intervals of intensity error for official guidances by referring to different TC best tracks or real-time operational intensity.

Table 8. Mean intensity error of official agencies in 2019. (Calculated by difference reference data, Unit: m/s)

Agency	Reference	24 h	48 h	72 h	96 h	120 h
JMA	RSMC-Tokyo	4.9(662)	6.9(429)	8.3(315)	8.0(203)	9.9(138)
	CMA real-time	4.1(771)	6.3(541)	7.3(387)	7.1(233)	6.6(141)
	JTWC real-time	7.2(414)	9.7(299)	10.8(223)	9.9(135)	9.0(87)
	KMA real-time	3.4(425)	4.8(321)	5.8(230)	6.0(137)	7.0(92)
	HKO real-time	4.0(362)	6.0(237)	6.9(158)	7.0(103)	7.4(61)
CMA	RSMC-Tokyo	6.1(608)	7.9(467)	9.6(348)	11.0(266)	12.5(184)
	CMA real-time	3.8(791)	5.4(595)	6.7(412)	7.4(276)	6.5(176)
	JTWC real-time	7.0(401)	8.8(313)	9.5(233)	9.5(161)	9.0(107)
	KMA real-time	4.2(424)	5.2(321)	5.9(233)	6.1(161)	6.2(111)
	HKO real-time	3.9(346)	5.3(252)	6.4(168)	7.2(105)	7.1(65)
JTWC	RSMC-Tokyo	7.3(527)	8.5(424)	10.5(353)	11.3(272)	12.2(196)
	CMA real-time	4.8(404)	6.0(317)	7.7(238)	8.4(167)	7.9(105)
	JTWC real-time	4.6(459)	7.3(361)	9.7(280)	9.8(203)	9.1(142)
	KMA real-time	6.1(401)	6.3(314)	7.4(233)	7.9(163)	8.8(113)
	HKO real-time	4.9(334)	6.1(243)	7.7(168)	7.7(105)	7.2(68)
KMA	RSMC-Tokyo	6.2(476)	7.7(376)	10.0(304)	10.6(208)	11.5(138)
	CMA real-time	5.0(431)	6.6(319)	7.5(235)	7.0(151)	6.2(97)

	JTWC real-time	8.4(399)	9.9(314)	10.7(235)	10.2(150)	8.7(103)
	KMA real-time	3.1(434)	4.4(327)	5.5(238)	5.8(155)	6.5(103)
	HKO real-time	4.7(327)	5.5(238)	6.1(162)	6.0(92)	6.6(59)
HKO	RSMC-Tokyo	5.5(352)	7.9(279)	9.2(193)	9.5(128)	9.4(72)
	CMA real-time	3.9(291)	5.6(205)	7.2(136)	7.3(87)	6.2(51)
	JTWC real-time	5.5(275)	7.9(194)	9.3(127)	10.2(79)	8.0(47)
	KMA real-time	3.9(271)	5.6(189)	6.5(125)	6.6(77)	5.2(48)
	HKO real-time	3.7(320)	5.8(230)	6.9(147)	7.9(91)	6.8(50)

In 2019, the intensity forecast skill scores of official guidances are all positive at the lead time levels of 24 h and 48 h.

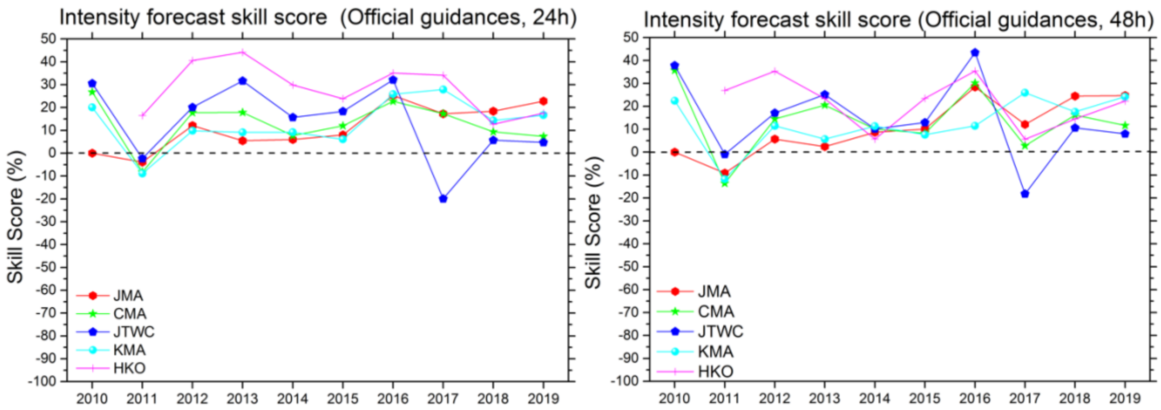


Fig.16 Intensity forecast skill scores evolutions of official guidances at the lead times of 24 h (left) and 48 h (right).

5.2 Objective forecasts

As fig.15, fig.17 shows the variation intervals of intensity errors for regional (left) and global (right) models by referring to different TC best tracks or real-time operational intensity.

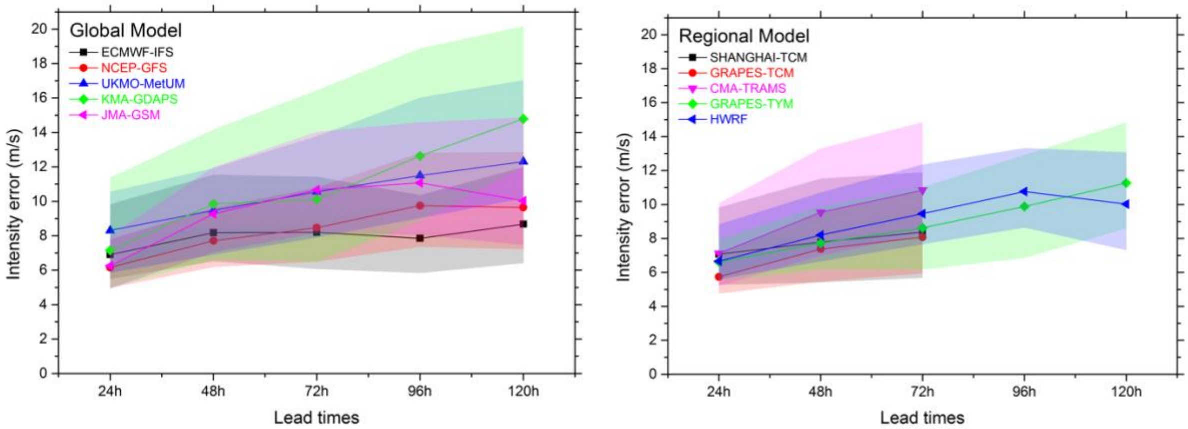


Fig.17 Same as fig.15, but for global and regional models.

Fig.18 shows the intensity forecast skill score for regional and global models at 24 h and 48 h. In 2019, NCEP-GFS and GRAPES-TCM have made positive skills at 24 h, and ECMWF-IFS, NCEP-GFS and GRAPES-TCM also have positive skill at 48 h. In general, intensity forecast skills of models are increasing year by year since 2010.

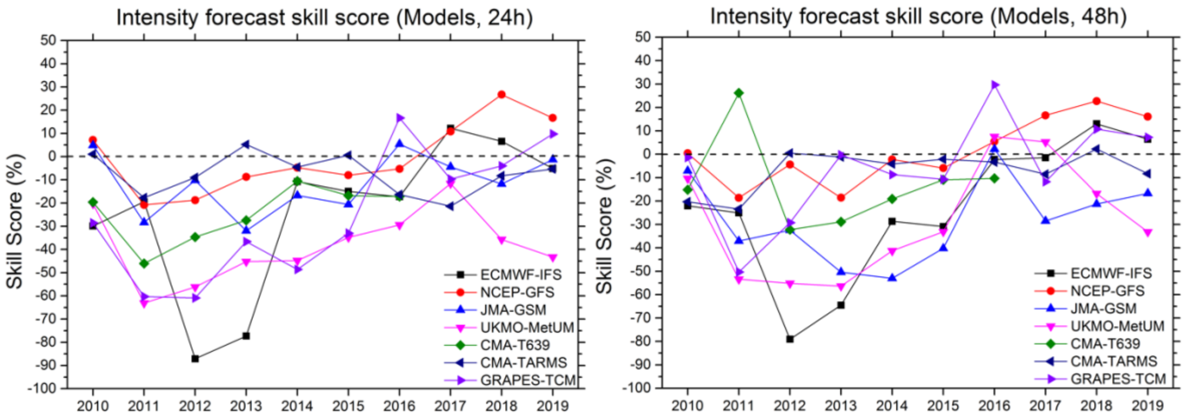


Fig.18 Same as fig.16, but for global and regional models.

6. Conclusions

Verification of TC forecasts is important for improving the NWP and official guidance that underpins the forecasts, making best use of this guidance in a forecasting context, and assisting the public, emergency managers, and other users of the TC forecasts to develop an appropriate level of confidence in the forecasts.

This report has briefly discussed the performance of typhoon forecast over western North Pacific in 2019. The verification results include TC track, and intensity for both deterministic and ensemble forecast guidance. The results show that mean position errors in 2019 for the most official agencies were much larger than the ones in 2018 at lead time level less than 4 days, and most forecast TC position were concentrated at southwest side compare to OBS position, and largest track errors were also appeared at southwest directions. Global models have not made significant progress in terms of track forecast since 2015. It is worth mentioning that there may exist large track or intensity error differences when we use different "best track" dataset as reference.

In the future, for STI, we'll not only focus on evaluation of basic TC attributes such as track, intensity and genesis, but also focus on verifying TC impact variables such as precipitation, wind and storm surge. We'll continue to develop and improve methodologies for verifying forecast aspects of TC formation, structure, evolution, and motion, particularly from high resolution and ensemble NWP which are now the foundation for most operational TC forecasts.

Acknowledgement

This verification report has been effectively supported by ESCAP/WMO Typhoon Committee, CMA, RSMC Tokyo Typhoon Center/WMO and HKO.

Contributors to this report are (in alphabetic order) lina BAI, Qing CAO, Chen CHEN, Rong GUOM Xiaotu LEI, Xiaoqin LU

Appendix: acronyms used in this report

BoM	Bureau of Meteorology (Australia)
CMA	China Meteorological Administration
MSC	Meteorological Service of Canada
ECMWF	European Centre for Medium Range Weather Forecasting
EMC	Environmental Modeling Center
EPS	Ensemble Prediction System
GEFS	Global Ensemble Forecast System
GFS	Global Forecast System
HKO	Hong Kong Observatory
ITMM	Institute of Tropical and Marine Meteorology
JMA	Japan Meteorological Agency
JTWC	Joint Typhoon Warning Center
KMA	Korea Meteorological Administration
MAE	Mean Absolute Error
ME	Mean Error
MSE	Mean Squared Error
NCEP	National Centers for Environmental Prediction
NWP	Numerical weather prediction
RMSE	Root Mean Squared Error
STI	Shanghai Typhoon Institute
TC	Tropical Cyclone
TIGGE	THORPEX Interactive Grand Global Ensemble
WMO	World Meteorological Organization