51st Session FOR PARTICIPANTS ONLY
ESCAP/WMO Typhoon Committee ENGLISH ONLY
26 February – 1 March 2019
Guangzhou · China

Verification of Tropical Cyclone Operational forecast in 2018

Guomin CHEN, Xiping ZHANG, Lina BAI, 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 2018*. 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 present	Includes sub-centers, some double eyewall cases/coastal severe wind of landfalling TCs (until 2004); includes TD cases; extratropical cyclone stage; longitude, latitude, MSW and MCP since 1949.	2 min
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 model	***BoM-ACCESS-TC***	*Tropical cyclone model in the Australian Community Climate and Earth-System Simulator Numerical Weather Prediction systems*	BoM
		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
		HWRF	*The atmosphere-ocean coupled Hurricane Weather Research and Forecast modeling system*	NCEP/EMC
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 2018, position errors from 5 official typhoon prediction agencies (JMA, CMA, JTWC, KMA and HKO) are 65.7-91.8km, 112.5-137.4km, 178.2-187.5km, 257.4-288.9km and 372.8-417.8km at the lead time level of 24, 48, 72, 96 and 120 h, respectively. The mean position errors for the most agencies are decreasing for the last three years as showing in the track forecast error evolutions (Fig.1).

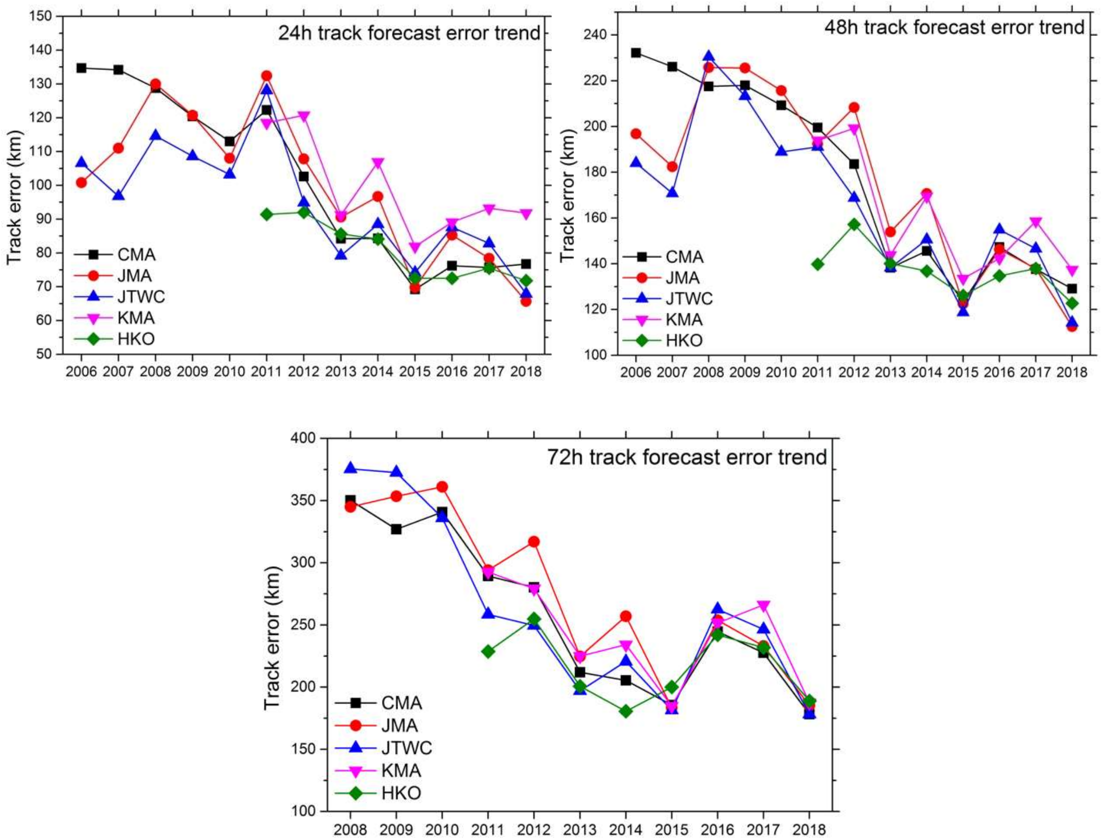


Fig.1 Track error evolutions of each official typhoon prediction agencies at the lead time of 24, 48 and 72 h.

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 2018. 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.

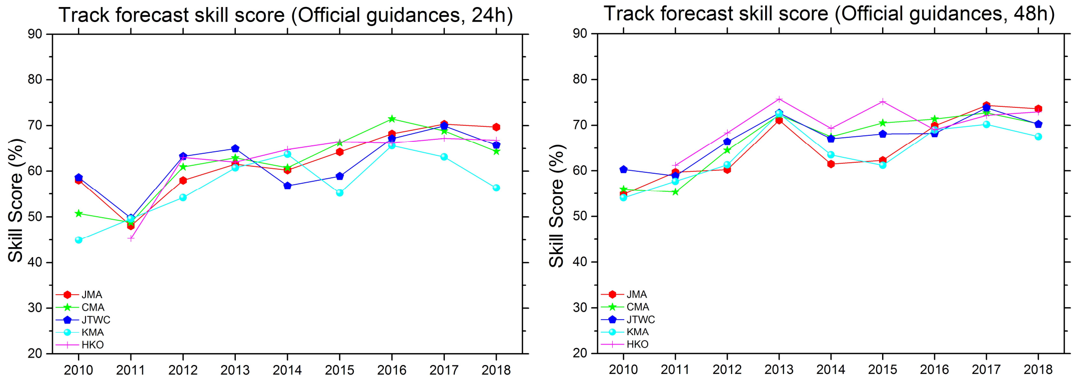


Fig.2 Track forecast skill scores evolutions of official guidances at the lead times of 24 h (left) and 48 h (right).

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 96 h in 2018.

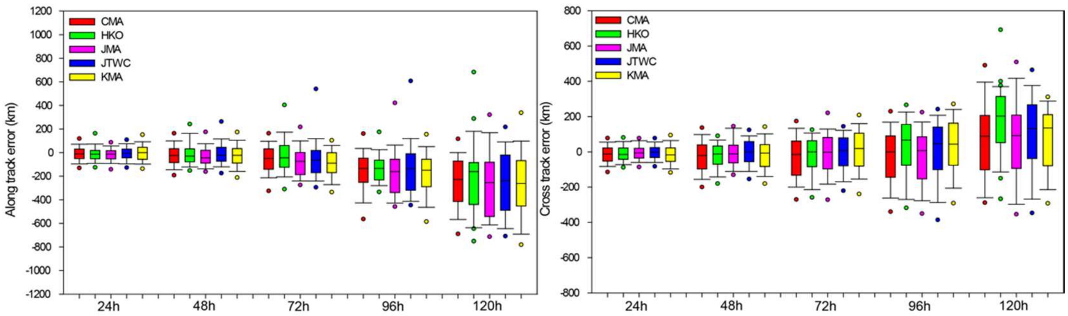


Fig.3 Along-track (left) and cross-track (right) biases for official guidances. The bar in the middle of the plot represents the median values of errors, the lower and upper ends of the boxes represent the 25th and 75th quantile values. The bars below and above the box represent the non-outlier extreme 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 2018. In the 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. The TER diagram reveals the position error distribution (both error magnitude and percentage of sample size) at 8 azimuthal directions. Take the TER diagram of JMA 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 is close to 18%, and the dominant position error range at 0-100km, 100 – 200km and 200 – 300km are about 5.5% (black), 8% (yellow) and 4% (red), respectively. However, forecast positions with large errors almost located at the north or northwest side of the OBS points.

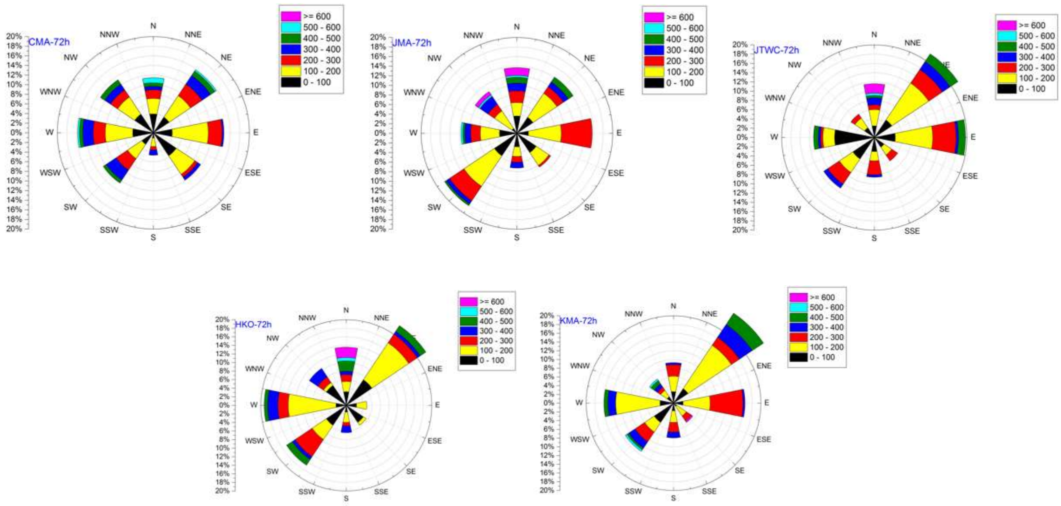


Fig.4 Track error rose (TER) diagrams for official guidances at the lead time of 72 h.

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 2018. (Calculated by difference reference data, Unit: km)

Agency	reference	24 h	48 h	72 h	96 h	120 h
	RSMC-Tokyo	65.7(435)	112.5(329)	185.3(248)	288.9(180)	417.8(129)
	CMA	67.5(481)	109.0(349)	183.8(264)	293.7(192)	422.1(140)
JMA	JTWC real-time	77.7(378)	122.8(244)	192.4(178)	277.5(127)	400.2(91)
	KMA real-time	70.3(425)	109.7(329)	181.0(247)	284.2(180)	421.5(130)
	HKO real-time	66.5(352)	108.9(229)	178.8(160)	287.3(106)	405.4(65)
	RSMC-Tokyo	76.7(479)	129.1(375)	178.2(279)	263.5(200)	381.7(136)
	CMA	74.8(544)	125.7(419)	175.4(316)	265.5(226)	393.3(150)
CMA	JTWC real-time	83.4(350)	134.8(275)	191.2(199)	275.3(140)	403.8(90)
	KMA real-time	78.9(418)	125.9(326)	184.8(245)	272.3(177)	404.8(127)
	HKO real-time	76.2(389)	130.9(276)	181.8(192)	262.5(128)	348.9(70)
	RSMC-Tokyo	67.9(342)	114.3(263)	178.3(198)	257.4(139)	372.8(99)
	CMA	74.8(382)	114.3(306)	185.4(232)	285.4(167)	397.6(114)
JTWC	JTWC real-time	89.8(392)	129.6(313)	191.7(237)	273.4(172)	398.6(117)
	KMA real-time	72.7(351)	115.4(273)	180.8(208)	275.2(150)	408.8(107)
	HKO real-time	79.4(294)	119.2(223)	195.3(155)	316.5(105)	365.1(65)
	RSMC-Tokyo	91.8(413)	137.4(317)	187.5(226)	285.3(165)	394.0(119)
	CMA	92.2(425)	136.5(331)	191.4(236)	292.1(174)	410.0(128)
KMA	JTWC real-time	96.6(318)	144.1(242)	193.9(167)	287.1(123)	417.2(88)
	KMA real-time	98.3(427)	136.4(328)	184.3(235)	293.0(176)	427.2(43)
	HKO real-time	94.9(300)	147.7(213)	181.0(139)	272.4(94)	364.3(59)
	RSMC-Tokyo	71.8(254)	122.7(181)	188.9(125)	262.2(72)	416.3(44)
	CMA	73.2(304)	121.5(231)	197.2(171)	287.0(108)	410.4(70)
HKO	JTWC real-time	81.0(211)	124.8(154)	200.2(103)	250.5(61)	383.2(34)
	KMA real-time	70.5(221)	114.1(161)	194.5(112)	284.7(69)	447.6(43)
	HKO real-time	73.8(317)	122.1(233)	183.9(164)	262.4(111)	371.3(73)

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 5% - 20% track error varieties while using different observation data as reference to evaluate official guidances' track error at different lead time levels in 2018.

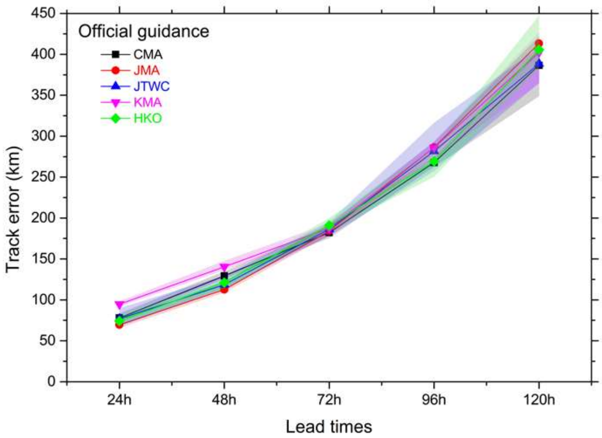


Fig.5 Variation intervals of track errors for official guidances by referring to different TC best tracks or real-time operational position.

4.2 Objective deterministic forecasts

In 2018, position errors for 5 global models are on intervals of 59.2 – 70.5km, 102.4 – 119.1km, 153.1 – 213.6km, 201.3-504.2km and 272.5 – 917.9km, and for 5 regional models are on intervals of 71.4 – 100.4km, 121.8 – 170.4km and 218.2 – 247.9km 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 2018. (Unit: km)

Method	Lead time	24h	48h	72h	96h	120h
Global Model	ECMWF-IFS	59.2(204)	102.4(165)	153.1(126)	201.3(94)	272.5(65)
	JMA-GSM	68.8(428)	119.1(330)	211.3(254)	327.0(130)	485.2(97)
	NCEP-GFS	70.5(156)	112.3(124)	213.6(95)	331.4(67)	464.9(47)
	KMA-GDAPS	69.9(228)	119.0(179)	187.8(138)	278.5(102)	455.4(75)
	UKMO-MetUM	67.6(220)	106.6(169)	174.9(129)	261.3(96)	399.7(69)
Regional Model	BoM-ACCESS-TC	100.4(65)	170.4(45)	234.3(25)	/	/
	GRAPES-TCM	78.1(188)	143.6(145)	227.6(110)	/	/
	GRAPES-TYM	73.4(418)	143.2(330)	245.0(248)	378.6(183)	555.1(127)
	CMA-TRAMS	71.4(437)	121.8(343)	218.2(259)	/	/
	HWRF	79.4(295)	138.7(214)	247.9(140)	371.4(76)	550.1(36)

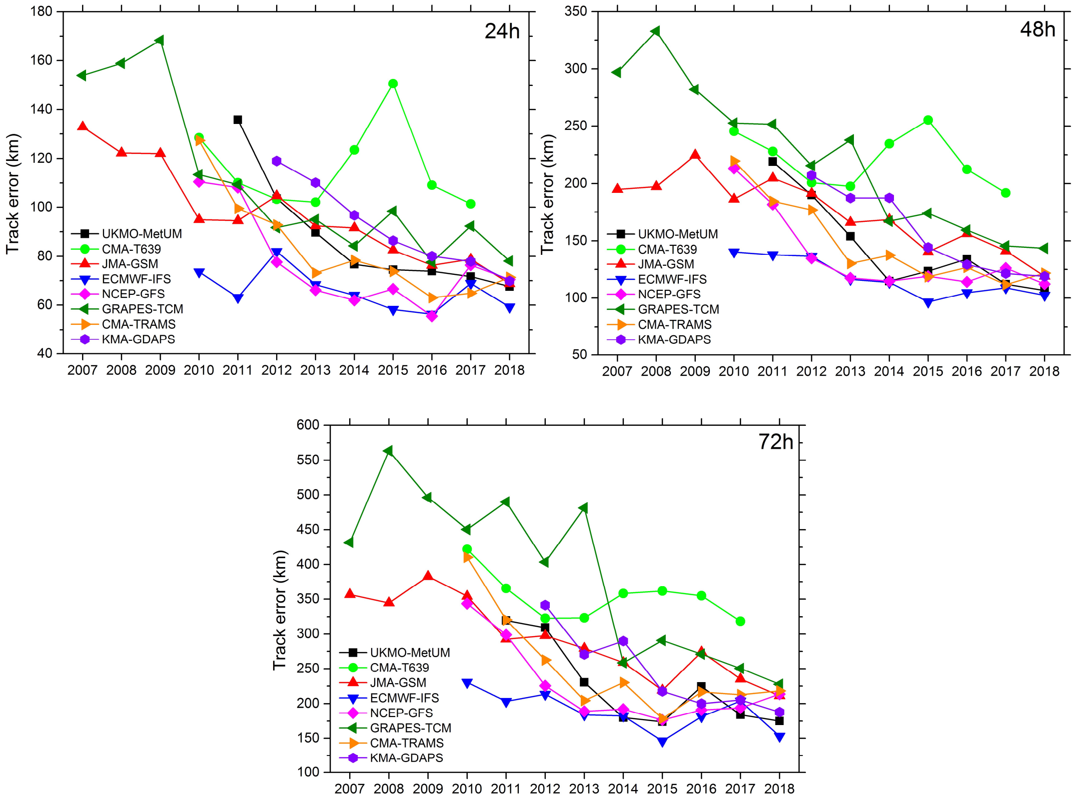


Fig.6 Same as fig.1 but for objective deterministic forecasts.

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 2018. Delightedly, all the models had positive skill scores. The overall tendency of models' skill score is generally going up during last nine years.

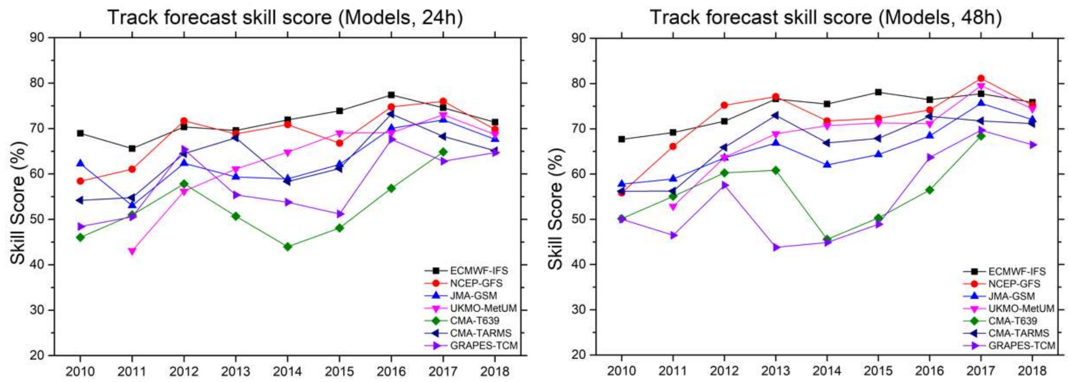


Fig.7 Track forecast skill scores evolutions of global and regional models at the lead times of 24 h (left) and 48 h (right).

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-MetUK's track forecast errors from 2010 to 2018. 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 three years.

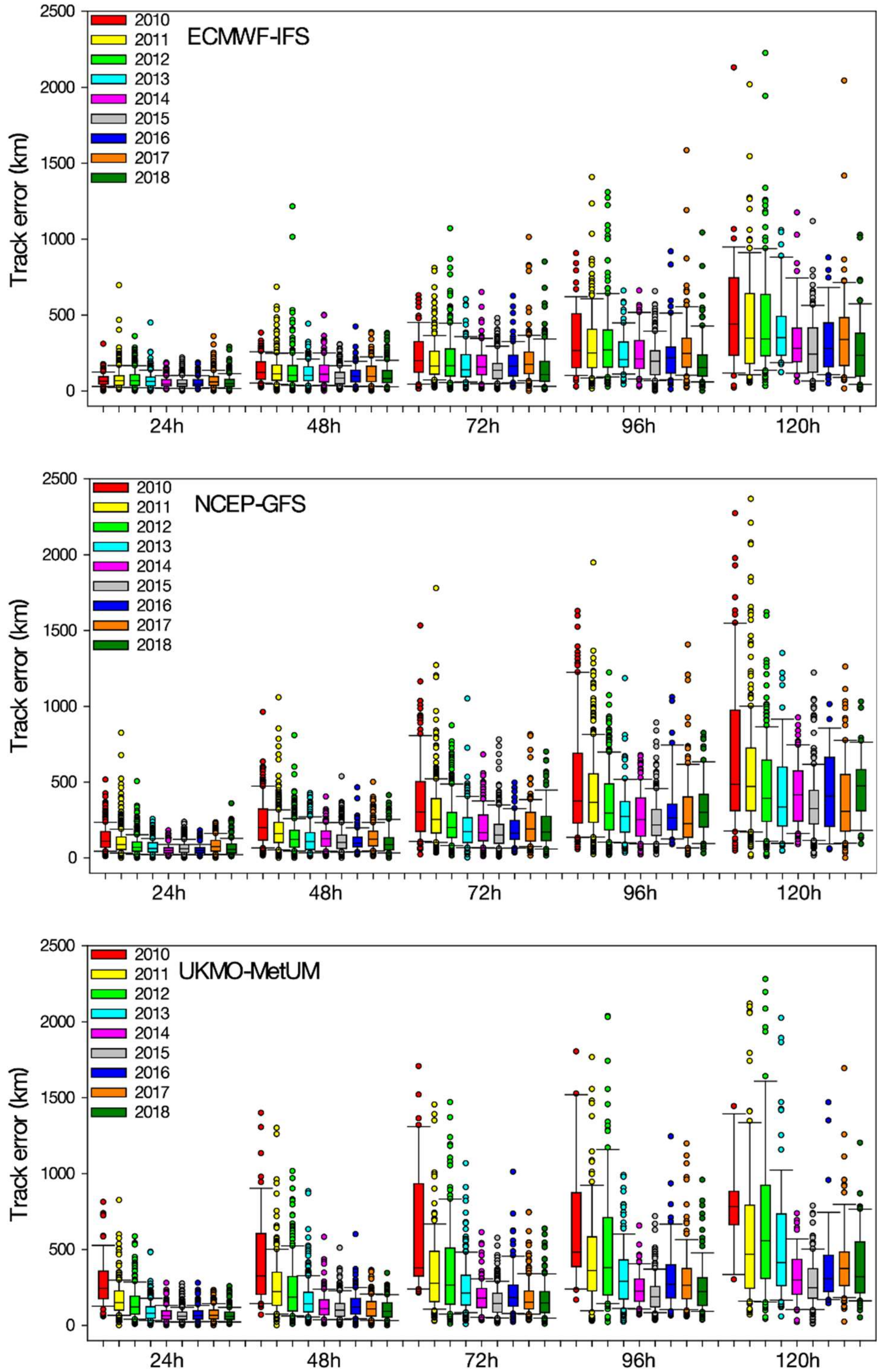


Fig.8 Box plots of position errors for ECMWF-IFS, NCEP-GFS and UKMO-MetUM in TC track forecasts from 2010 to 2018. The bar in the middle of the plot represents the median values of errors, the lower and upper ends of the boxes represent the 25th and 75th quantile values. The bars below and above the 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 and regional models from 24 to 120 h. With the lead time increasing, the forecasted TCs from both regional and global model propagated slower than observations. There are not obvious leftward or rightward biases for global models. However, forecasted TCs from HWRF and GRAPES-TYM propagated, on average, distinct leftward and rightward, respectively.

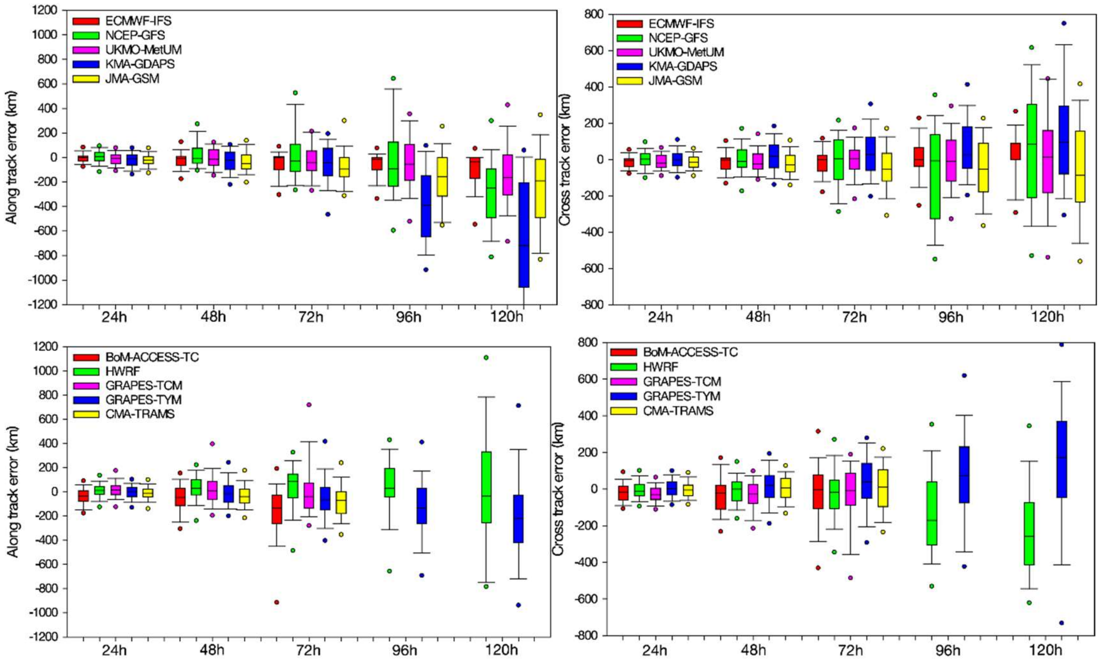


Fig.9 Along-track and cross-track biases for global and regional models.

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 2018. 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, both ECMWF-IFS and UKMO-MetUM do not show obviously systematic bias. However, the systematic bias of NCEP-GFS tends to north within the lead time 72 h, then it turns to northeast. JMA-GSM has no systematic bias within 72 h and the biases locate at southeast at the lead time of 96 h and 120 h. For regional models, the systematic biases of GRAPES-TCM and HWRF tend to north. The systematic bias of GRAPES-TYM turn towards northeast. However, CMA-TRAMS and BoM-ACCESS-TC show insignificant systematic bias. Plots like those in fig.10 provide information that is useful for the pre-estimation of the bias of a certain method.

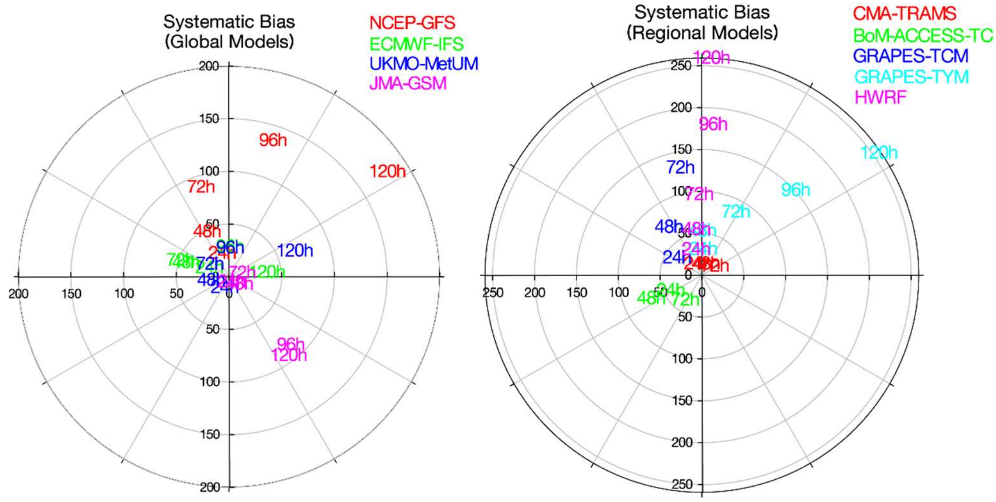


Fig.10 Polar scatter plots depicting the mean combined direction and magnitude errors relative to the actual storm location for each model at different lead times in 2018.

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.

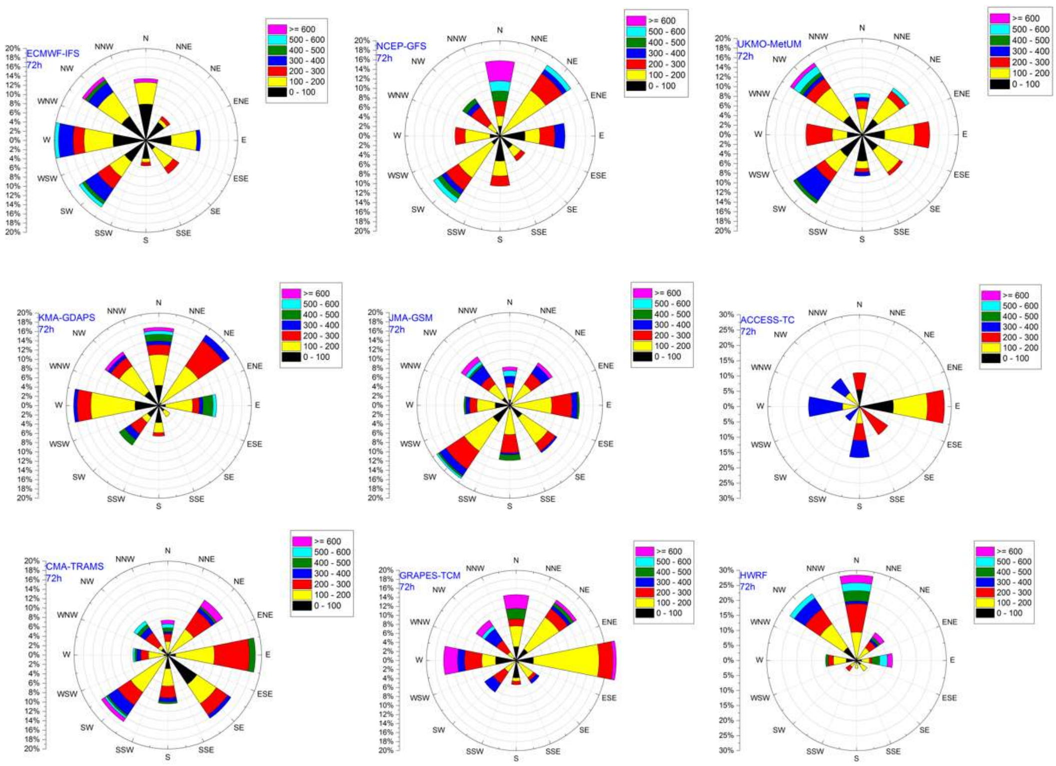


Fig.11 TER diagrams for global and regional models at the lead time of 72 h.

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, both global and regional models exist 3% - 5% track error varieties with the lead time levels less than 72 h, and will increase to 7% - 10% at 120 h.

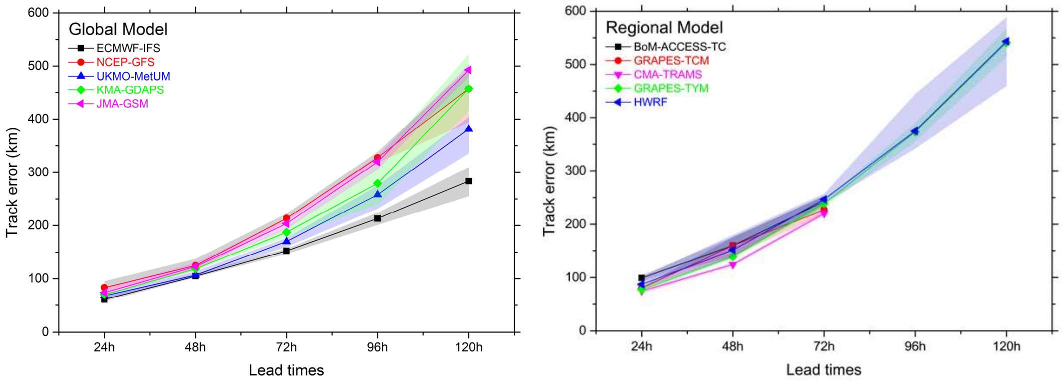


Fig.12 Same as fig.5 but for global models (left) and regional model (right).

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). ECMWF-EPS is the best EPS system, whether by non-homogeneous comparison or homogeneous comparison at any lead time levels in 2018. The ensemble mean position error of ECMWF-EPS at the lead time of 120 h is approaching 400 km.

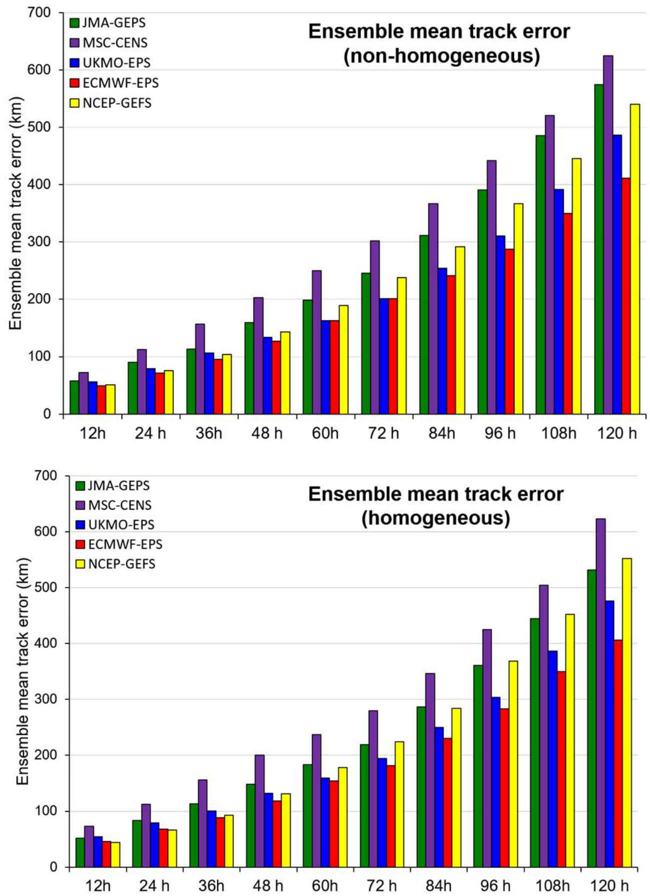


Fig.13 Ensemble mean track errors of five EPSs for both non-homogeneous (up) and homogeneous (down) comparison in 2018.

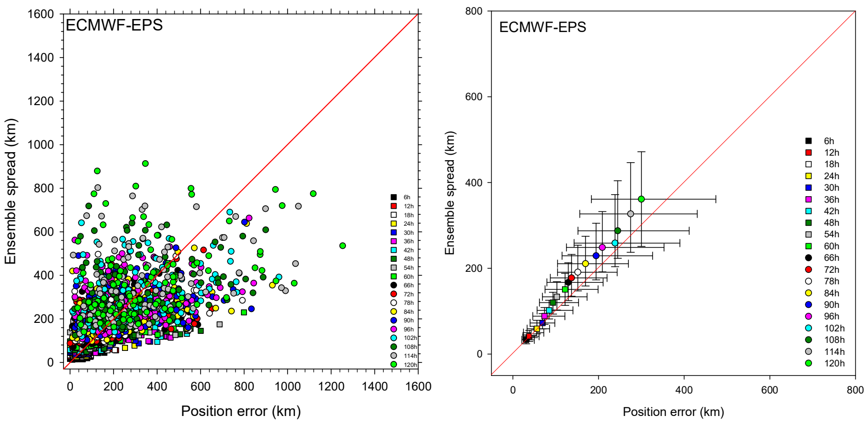
Table 5. Ensemble mean track errors in 2018 (Non-homogeneous comparison).

	12h	24h	36h	48h	60h	72h	84h	96h	108h	120h
JMA-GEPS	57.4(591)	89.9(573)	122.9(543)	159.6(512)	198.4(471)	245.4(430)	311.6(391)	390.7(356)	485.5(319)	574.6(283)
ECMWF-EPS	49.1(319)	71.3(305)	95.7(286)	127.2(266)	162.6(243)	200.8(222)	241.7(197)	287.1(173)	349.5(152)	411.0(133)
MSC-CENS	72.7(274)	112.5(263)	156.5(249)	202.8(231)	249.5(214)	302.1(194)	367.0(172)	442.1(155)	520.3(135)	624.5(118)
NCEP-GEFS	50.7(590)	75.4(564)	103.7(535)	143.6(499)	189.6(463)	237.7(422)	292.0(376)	366.4(333)	445.7(292)	540.3(257)
UKMO-EPS	56.2(616)	79.2(589)	106.1(548)	133.8(506)	162.4(461)	201.5(419)	254.1(381)	310.3(339)	391.8(300)	486.4(263)

Table 6. Ensemble mean track errors in 2018 (Homogeneous comparison).

	12h	24h	36h	48h	60h	72h	84h	96h	108h	120h
Sample size	(235)	(228)	(211)	(196)	(176)	(160)	(143)	(126)	(109)	(93)
JMA-GEPS	51.7	83.2	113.1	148.4	183.0	219.0	286.7	361.0	444.2	531.3
ECMWF-EPS	45.5	68.3	88.4	118.5	153.9	181.7	230.5	283.0	349.9	406.5
MSC-CENS	73.5	112.5	156.1	200.3	237.4	279.4	346.6	425.1	504.4	622.5
NCEP-GEFS	44.0	66.6	92.9	131.1	178.3	224.0	283.7	368.2	452.4	552.1
UKMO-EPS	54.1	78.8	100.7	132.2	159.2	194.3	249.6	303.8	386.4	476.1

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 MSC-CENS and UKMO-EPS. However, the median ensemble spreads became larger than position errors with lead time increasing for ECMWF-EPS. JMA-GEPS and NCEP-GEFS have opposite result compare to ECMWF-EPS.



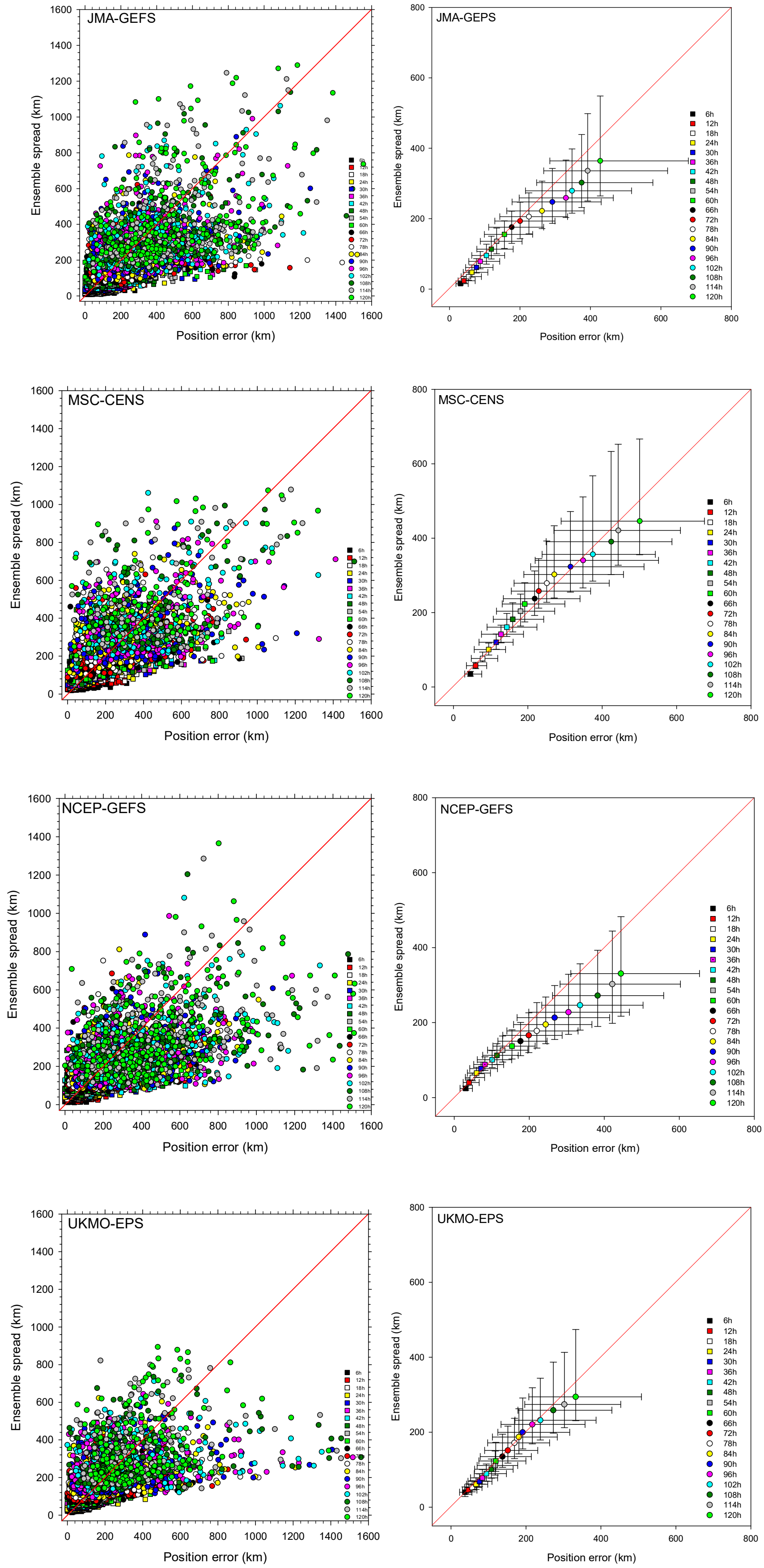


Fig.14 Scatter plot (left panels) and bi-directional track forecast scatter plot (right panels) for EPSs. The blocks in the bi-directional track forecast scatter plots represent the mean values of spread or position error, and the lower (left) and upper (right) bars represent the 25th and 75th quantile 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 2018. 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 2018. The numbers in brackets are sample sizes. (Unit: m/s)

Method	Lead time	24h	48h	72h	96h	120h
Official guidance	CMA	4.0(544)	5.2(419)	5.3(316)	6.4(266)	6.9(150)
	JMA	3.9(467)	5.0(334)	5.7(244)	/	/
	JTWC	4.7(382)	5.4(306)	6.6(232)	8.4(167)	9.1(114)
	KMA	5.0(425)	6.6(331)	7.0(236)	7.0(174)	7.2(128)
	HKO	4.5(183)	6.7(127)	8.2(84)	10.4(53)	11.9(34)
Global Model	ECMWF-IFS	6.9(204)	7.9(165)	8.2(126)	8.4(94)	6.8(65)
	JMA-GSM	4.8(428)	6.9(330)	8.9(254)	12.8(130)	12.8(97)
	NCEP-GFS	4.8(156)	5.9(124)	5.4(95)	6.5(67)	8.9(47)
	KMA-GDAPS	7.7(228)	9.3(179)	9.4(138)	10.3(102)	11.2(75)
	UKMO-MetUM	6.9(220)	8.1(169)	8.3(129)	8.7(96)	7.9(69)
Regional Model	BoM-ACCESS-TC	6.6(65)	8.3(45)	8.9(25)	/	/
	GRAPES-TCM	5.1(188)	6.1(145)	7.2(110)	/	/
	GRAPES-TYM	5.5(418)	6.9(330)	7.9(248)	8.5(183)	8.4(127)
	CMA-TRAMS	6.1(437)	7.7(343)	7.9(259)	/	/
	HWRF	5.3(295)	5.3(214)	6.1(140)	5.7(76)	6.3(36)

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 2018 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 12% -20% 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.

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

Table 8. Mean intensity error of official agencies in 2018. (Calculated by difference reference data, Unit: m/s)

Agency	reference	24 h	48 h	72 h	96 h	120 h
JMA	RSMC-Tokyo	4.0(433)	5.2(322)	5.8(234)	/	/
	CMA	3.9(467)	5.0(334)	5.8(244)	/	/
	JTWC real-time	7.1(376)	9.7(238)	10.6(170)	/	/
	KMA real-time	3.1(416)	4.2(318)	5.0(233)	/	/
	HKO real-time	3.5(273)	4.0(134)	5.0(62)	/	/
CMA	RSMC-Tokyo	5.9(479)	7.5(375)	7.7(279)	8.4(200)	9.7(136)
	CMA	3.6(544)	5.2(419)	5.4(316)	6.5(226)	6.9(150)
	JTWC real-time	6.6(350)	8.7(275)	8.6(199)	9.3(140)	9.9(90)
	KMA real-time	5.3(416)	6.7(324)	7.2(243)	8.3(176)	8.8(126)
	HKO real-time	3.5(311)	4.6(196)	4.9(119)	6.1(65)	3.8(23)
JTWC	RSMC-Tokyo	7.1(342)	9.0(263)	11.0(198)	13.2(139)	15.9(99)
	CMA	5.0(382)	6.3(306)	7.1(232)	8.5(167)	10.7(114)
	JTWC real-time	4.9(392)	7.0(313)	8.6(237)	9.0(172)	11.9(117)
	KMA real-time	7.1(349)	8.2(271)	9.8(206)	12.0(149)	14.4(106)
	HKO real-time	3.9(182)	4.4(119)	5.6(66)	5.5(36)	3.7(15)
KMA	RSMC-Tokyo	4.4(412)	5.8(317)	6.0(226)	6.6(165)	7.9(119)
	CMA	5.0(425)	6.6(331)	7.0(236)	7.0(174)	7.2(128)
	JTWC real-time	9.1(318)	11.6(242)	11.9(167)	11.2(123)	11.6(88)
	KMA real-time	3.1(424)	4.7(326)	5.0(233)	5.3(175)	6.9(127)
	HKO real-time	4.5(186)	5.4(115)	5.5(51)	6.1(26)	6.6(10)
HKO	RSMC-Tokyo	4.5(183)	6.7(127)	8.2(84)	10.4(53)	11.9(34)
	CMA	4.0(202)	5.4(142)	5.8(96)	7.6(57)	9.4(42)
	JTWC real-time	5.5(144)	6.9(105)	6.8(67)	8.1(46)	10.8(28)
	KMA real-time	4.0(144)	5.9(103)	7.5(67)	9.1(47)	11.1(33)
	HKO real-time	2.7(247)	4.0(153)	3.6(93)	3.5(58)	6.0(26)

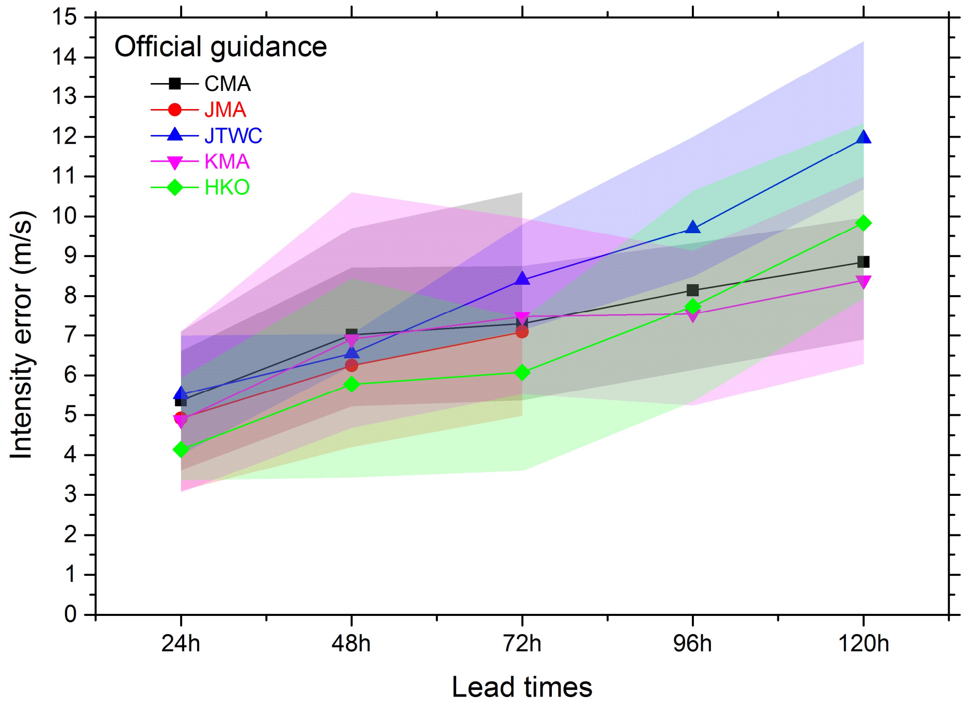


Fig.15 Variation intervals of intensity error for official guidances by referring to different TC best tracks or real-time operational intensity.

In 2018, the intensity forecast skill scores of official guidances are all positive at the lead time levels of 24 h and 48 h.

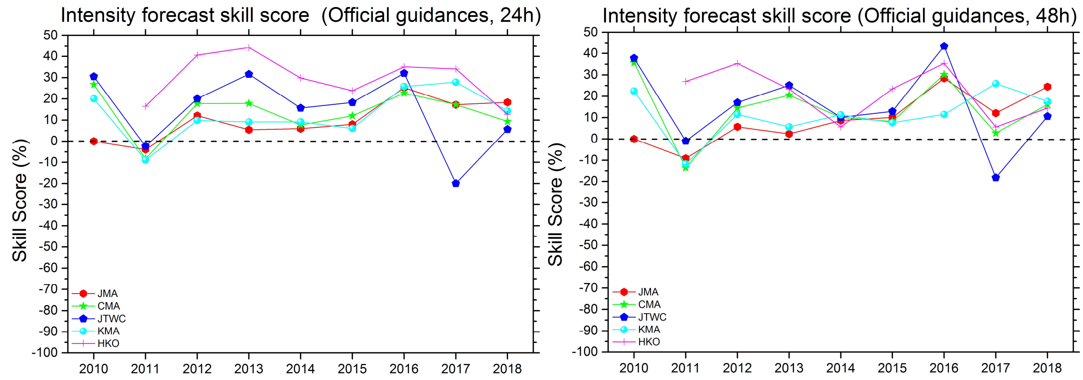


Fig.16 Intensity forecast skill scores evolutions of official guidances at the lead times of 24 h (left) and 48 h (right).

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.

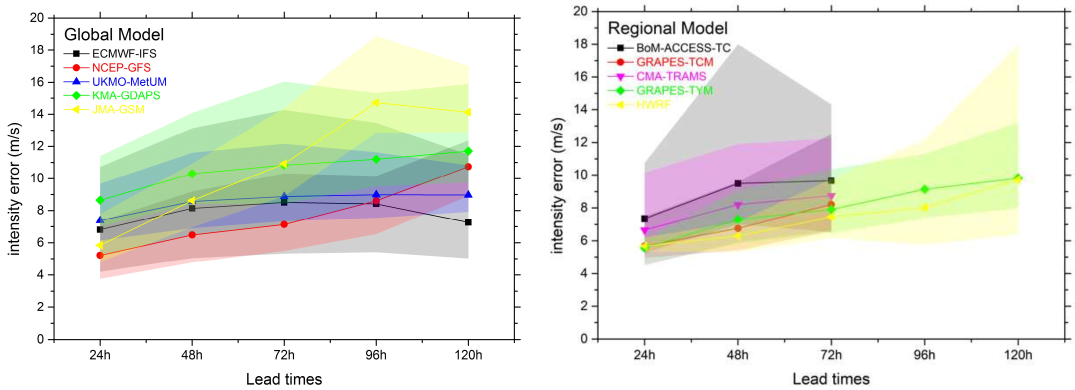


Fig.17 Same as fig.15, but for global and regional models.

Fig.18 shows the intensity forecast skill score for regional and global models at 24 h and 48 h. In 2018, ECMWF-IFS and NCEP-GFS have made positive skills at 24 h, and ECMWF-IFS, NCEP-GFS, GRAPES-TCM and CMA-TRAMS also have positive skill at 48 h. In general, intensity forecast skills of models are increasing year by year since 2010.

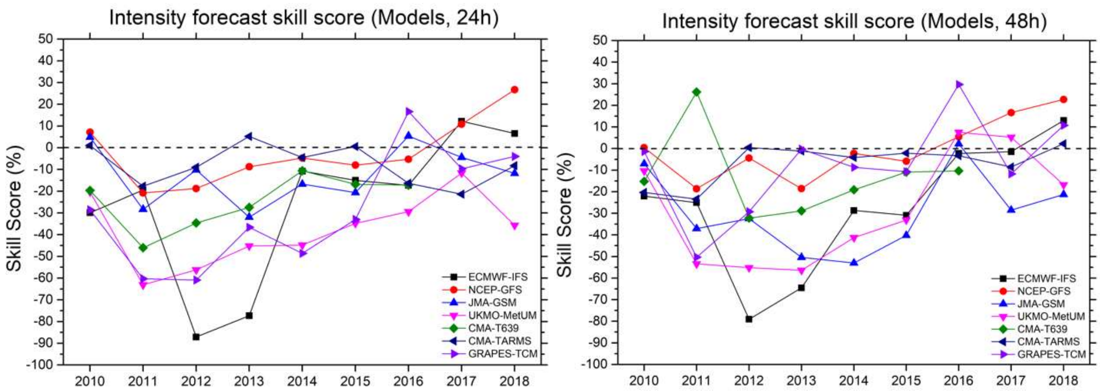


Fig.18 Same as fig.16, but for global and regional models.

Fig.19 presents two taylor diagrams to assess the performances of the maximum wind speed forecasts from both regional and global models. Taylor diagrams are used in the verification of TC intensity forecasts to analyze the internal relationships between the standardized deviation, the correlation coefficient, and the central different root-mean-square. The best prediction always has the highest correlation coefficient when with the “OBS”, and a standardized deviation and central different root-mean-square close to “1”. As showing in Fig.9, the center difference RMS errors of maximum wind speed are smallest at 0 h for JMA-GSM and GRAPES-TYM, respectively. For most global models, the correlation coefficients of the observed and forecast maximum wind speed are in the range of 0.6 to 0.9, and the standardized deviations are in the range of 0.70 to 0.95 in 2018. For the regional models, the correlation coefficients of the observed and forecast maximum wind speed are in the range of 0.65 to 0.9, and the standardized deviations are in the range of 0.75 to 0.9 in 2018. The standardized deviations of both global and regional models are less than 1.0， that indicates, in a general sense, the intensity forecasted by models are weaker than actual intensity.

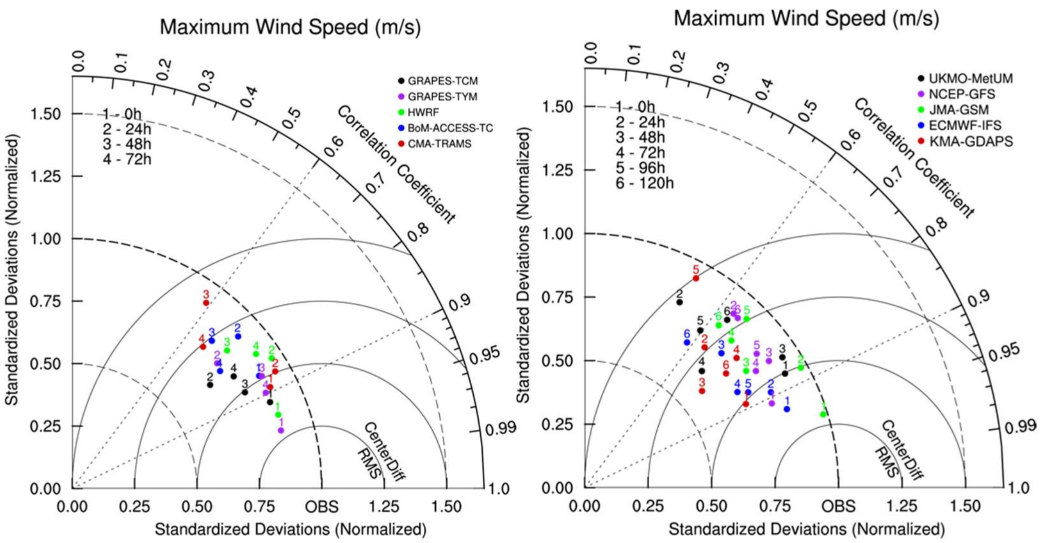


Fig.19 Taylor diagrams for the evaluation of TC maximum wind speed forecasts from models. Left: regional models, right: global models.

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