48th Session
ESCAP/WMO Typhoon Committee
22- 25 February 2016
Honolulu Hawaii, USA

Performance of tropical cyclone forecast in western North Pacific in 2015

Xiaotu LEI, Guomin CHEN, Xiping ZHANG, Peiyan CHEN, Hui YU

(*Shanghai Typhoon Institute / China Meteorological Administration, Shanghai, China*)

CONTENTS

1. Introduction

An important key to making better predictions of tropical cyclone (TC) is having an understanding of the forecast errors in current predictions. Subjective and objective verification of 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 report is primarily about *the performance of typhoon forecast over western North Pacific in 2015*. 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 subjective methods, global models and regional models. And then, we will evaluate the cyclone track, genesis, intensity forecast. In last part, the track forecast performance of seven ensemble prediction system will be evaluated.

2. Best track datasets

Currently, four agencies provide their own TC best track analyses for the WNP region: 1) Shanghai Typhoon Institute of China Meteorological Administration (STI/CMA, dataset can be found at http://tcdata.typhoon.gov.cn/en/index.html), 2) the Japan Meteorological Agency (JMA) Regional Specialized Meteorological Center (RSMC) in Tokyo (RSMC-Tokyo, dataset can be found at http://www.jma.go.jp/jma/jma-eng/jma-center/rsmc-hp-pub-eg/besttrack.html), 3)Joint Typhoon Warning Center (JTWC, dataset can be found at http://www.usno.navy.mil/NOOC/nmfc-ph/RSS/jtwc/best_tracks/), 4)Hong Kong Observatory (HKO, dataset can be found at http://www.weather.gov.hk/publica/pubtc.htm). Table.1 provides 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. As the typhoon center in RSMC-Tokyo is the regional center that carries out specialized activities in analysis and forecasting of WNP TCs within the framework of the World Weather Watch (WWW) Program of WMO, in this verification report, we'll use RSMC-Tokyo best track-dataset as the reference.

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, the TC position and intensity forecast results from five subjective methods, six global models and three regional models are evaluated. These totally 14 methods are deterministic forecast guidance, detail explanations including their abbreviations, short description and source agencies are listed in Table.2.

Table.2. Details of deterministic forecast guidance

Category	Abbreviation	Full name or short description	Source agency
Subjective method (5)	**CMA-sub**	*China Meteorological Administration*	CMA
	JMA-sub	*Japan Meteorological Agency*	JMA
	JTWC-sub	*Joint Typhoon Warning Center*	JTWC
	KMA-sub	*Korea Meteorological Administration*	KMA
	HKO-sub	*Hong Kong Observatory*	HKO
Global NWP model (6)	**CMA-T639**	*Global spectral model of CMA at a resolution of T639L60*	CMA
	ECMWF-IFS	*Integrated Forecasting System of ECMWF*	ECMWF
	KMA-GDAPS	*Global Data Assimilation and Prediction System of KMA*	KMA
	JMA-GSM	*Global Spectral Model of JMA*	JMA
	NCEP-GFS	*Global Forecast System of NCEP*	NCEP
	UKMO-MetUM	*Unified Model system of UKMO*	UKMO
Regional NWP model (3)	**BOM-ACCESS**	*Tropical cyclone model in the Australian Community Climate and Earth-System Simulator Numerical Weather Prediction systems*	BOM
	STI-GRAPES	*Regional TC-forecasting model based on the Global/Regional Assimilation and PrEdiction System (GRAPES)*	STI/CMA
	CMA-TRAMS	*Tropical Regional Atmosphere Model for the South China Sea based on GRAPES*	ITMM/CMA

4. TC track forecast verification

4.1 Deterministic forecast

TC track error is defined as the great-circle difference between a TC's forecast center position and the best track position at the verification time. TC Track errors typically are presented as mean errors for a large sample of TCs, as in Fig.1, which shows mean track errors for each subjective methods, global models and regional models at the lead time levels of 24, 48, 72, 96 and 120h. The detail numerical values of track error which related to Fig. 2 are list in Table.3. It shows that the track errors at the lead time of 24h are generally less than 100km for most methods and some global models' mean track error at 120h are less than 300km, such as ECMWF-IFS and UKMO-MetUM.

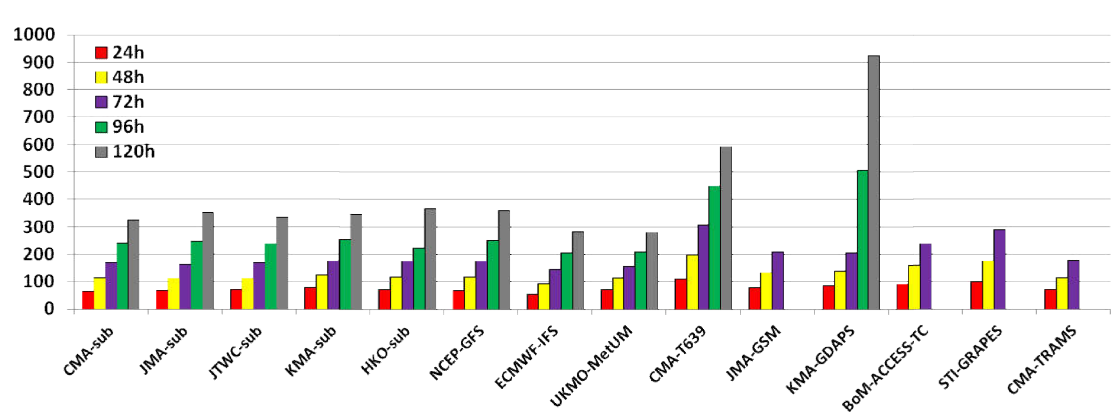
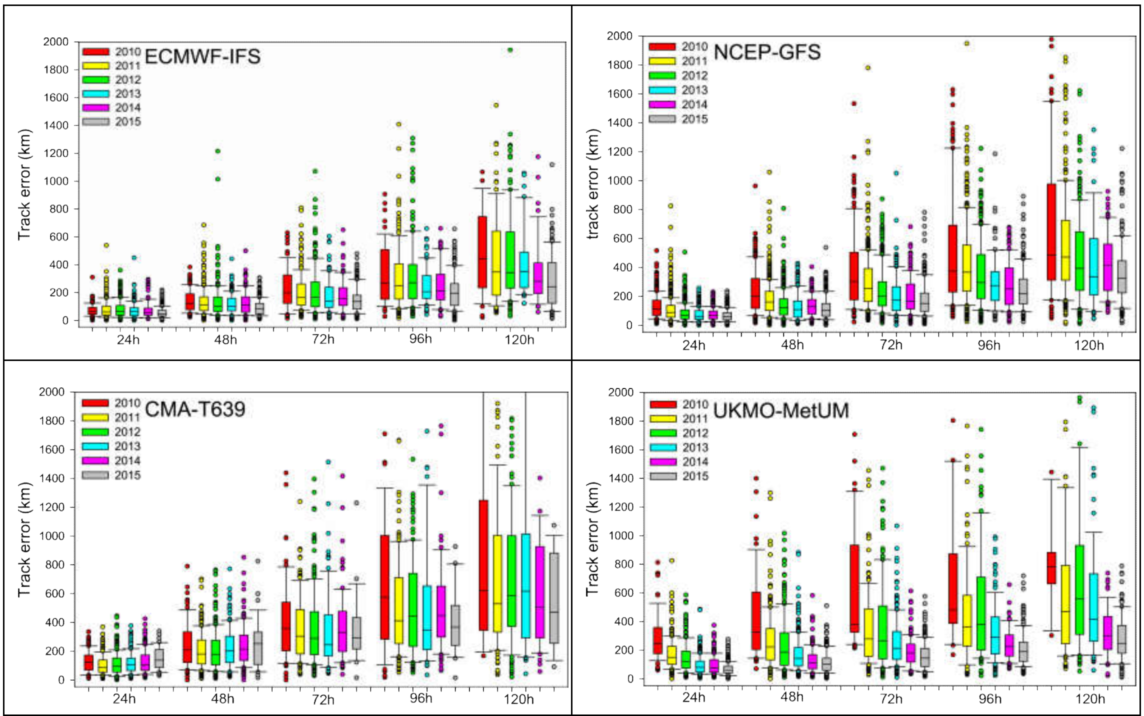


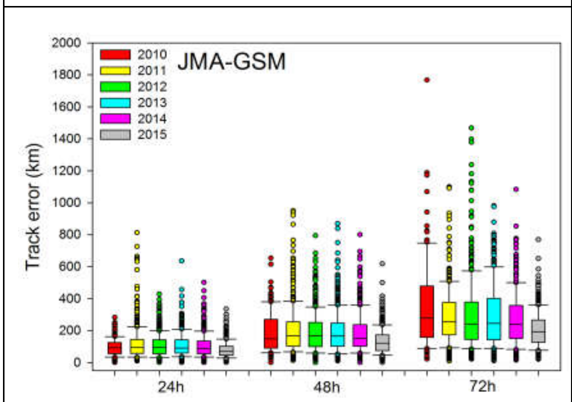
Fig.1. Mean track errors of subjective methods, global models and regional models at the lead time levels of 24h(red), 48h(yellow), 72h(violet), 96h(green) and 120h(gray) in 2015.

Table.3. Average track error for each method at 24, 48, 72, 96 and 120h lead time levels in 2015 (Unit: km)

Method	Lead times	24h	48h	72h	96h	120h
Subjective Methods	CMA-sub	65.0(640)	114.9(537)	170.3(441)	241.2(359)	326.2(287)
	JMA-sub	66.7(639)	112.6(539)	162.6(447)	248.7(361)	354.0(287)
	KMA-sub	78.3(634)	125.6(536)	175.7(441)	254.2(360)	346.9(287)
	JTWC-sub	72.1(612)	112.7(512)	169.7(426)	238.1(347)	334.1(273)
	HKO-sub	69.4(251)	120.0(196)	176.5(142)	223.3(96)	367.2(58)
Global NWP Models	ECMWF-IFS	56.3(305)	93.2(261)	146.2(212)	206.4(171)	283.0(135)
	NCEP-GFS	66.6(414)	119.1(355)	176.5(289)	251.0(234)	360.0(188)
	JMA-GSM	79.9(638)	133.2(540)	209.1(443)	/	/
	CMA-T639	109.7(46)	198.5(36)	307.1(28)	449.6(20)	593.6(15)
	KMA-GDAPS	84.3(207)	138.9(173)	206.5(142)	505.4(116)	923.6(81)
	UKMO-MetUM	69.6(315)	114.5(269)	158.1(223)	209.2(179)	281.0(143)
Regional NWP Models	BOM-ACCASS	92.3(300)	161.9(254)	239.1(208)	/	/
	CMA-TRAMS	71.0(252)	115.1(211)	178.4(169)	/	/
	STI-GRAPES	99.8(423)	175.6(211)	288.5(274)	/	/

An alternative approach to examining the average errors is to consider the distributions of errors, as in Fig. 2. In this example, box plots are used to summarize the distributions of errors in track forecasts from 2010 to 2015 for five global models. Such a track error distributional 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 progress of each global model. This methodology was developed to evaluate the uncertainty in verification measures through confidence intervals and paired statistical tests. And it can provide a consistent set of results that allowed the forecasts from the various models to be compared and fairly evaluated. In Fig.2, it clearly shows that stepped decreases in the values of each quantile were made at every lead time level from 2010 to 2015, and the forecast accuracy at 48h (72, 96 and 120h) in 2015 were almost close to or beyond the forecast accuracy at 24h (48, 72 and 96h) in 2010. Anyway, it should be noted that this is not necessarily a conclusive comparison because the storms in 2010 and 2015 were not the same.





Box plots show specific quantiles and other statistics used to represent a distribution. The bar in the middle of the plot represents the median value, 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 specific circles represent the outliers.

Fig.2. Box plots for representing the distributions of track errors for TC track forecast from 2010 to 2015.

Fig.3 shows the track forecast skill score at the lead time levels of 24 and 48h for subjective method, global and regional models from 2010 to 2015. All the forecast methods obtained positive skill indicating the forecast accuracy of subjective methods, global and regional models are better than climatic persistence method in the last 6 years.

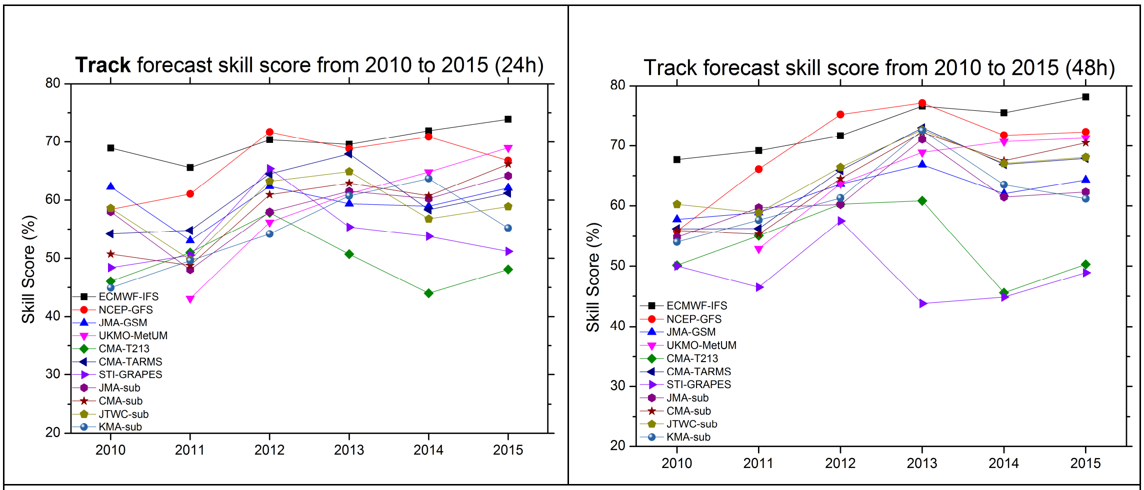


Fig.3. Track forecast skill score trend at the lead time level of 24h (left) and 48h (right) for subjective methods, global models and regional models.

A new approach called the Track Forecast Integral Deviation (TFID) integrates the track error over an entire forecast period (Yu et al., 2013). Fig.4 shows the TFID annual distributions from 2010 to 2015 for ECMWF-IFS, NCEP-GFS, CMA-T639 and UKMO-MetUM at the lead time levels of 24, 48, 72, 96 and 120h. These TFID diagrams show a clearly decrease trend for most global models, indicate that the TC forecast tracks became increasingly similar to the observation.

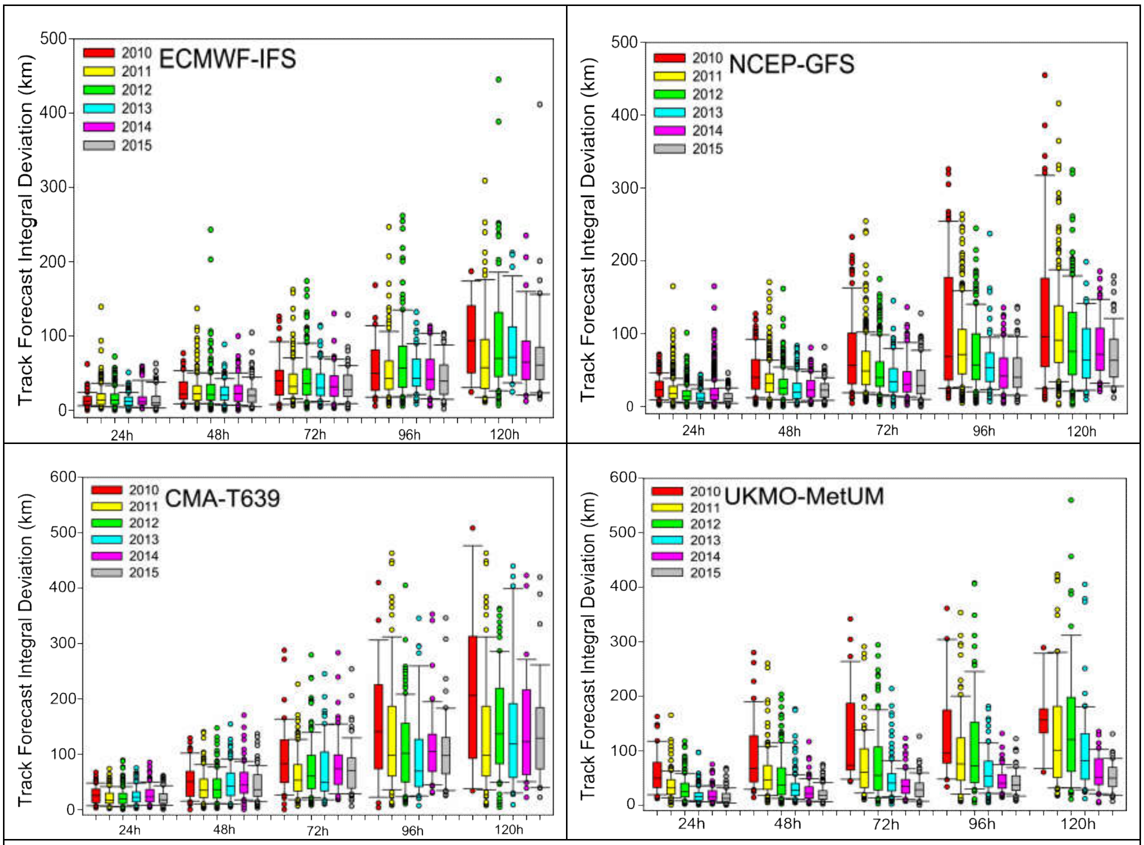


Fig.4. The same as **Fig.2**, but for track forecast integral deviation (TFID).

Fig.5 is the polar scatter plots which depicting the mean combined direction and magnitude errors around the actual storm location for global and regional models at different lead time levels in 2015. Each models' systematic biases of track forecast are showed clearly through the Fig.5. The numbers with different colors denote annual mean locations which relative to actual typhoon locations which obtain from best track dataset. Plots like those in Fig.5 provide information that is useful for pre-estimate the bias of a certain method.

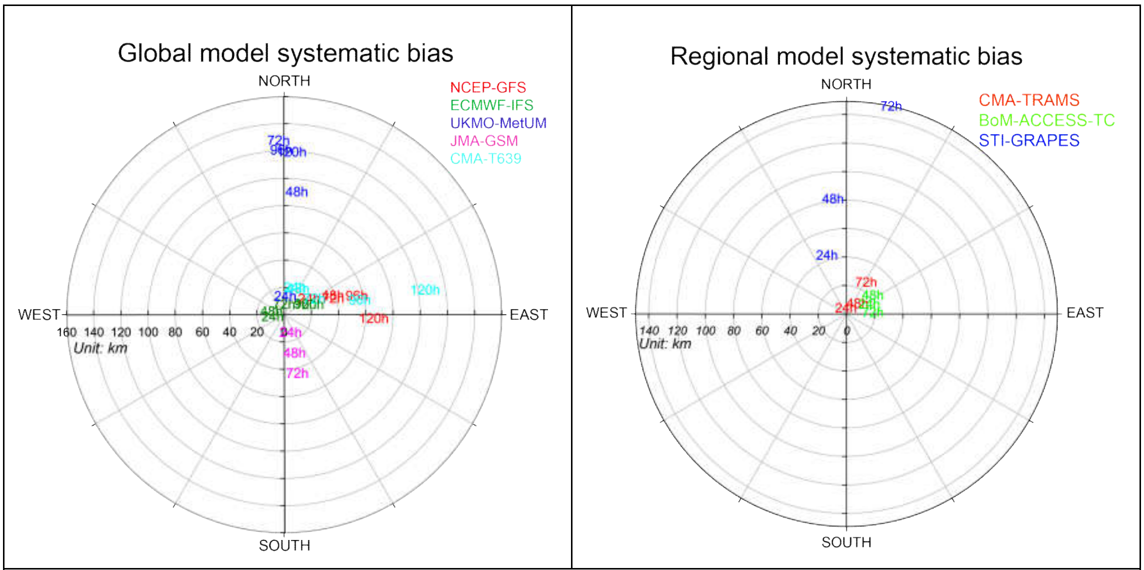
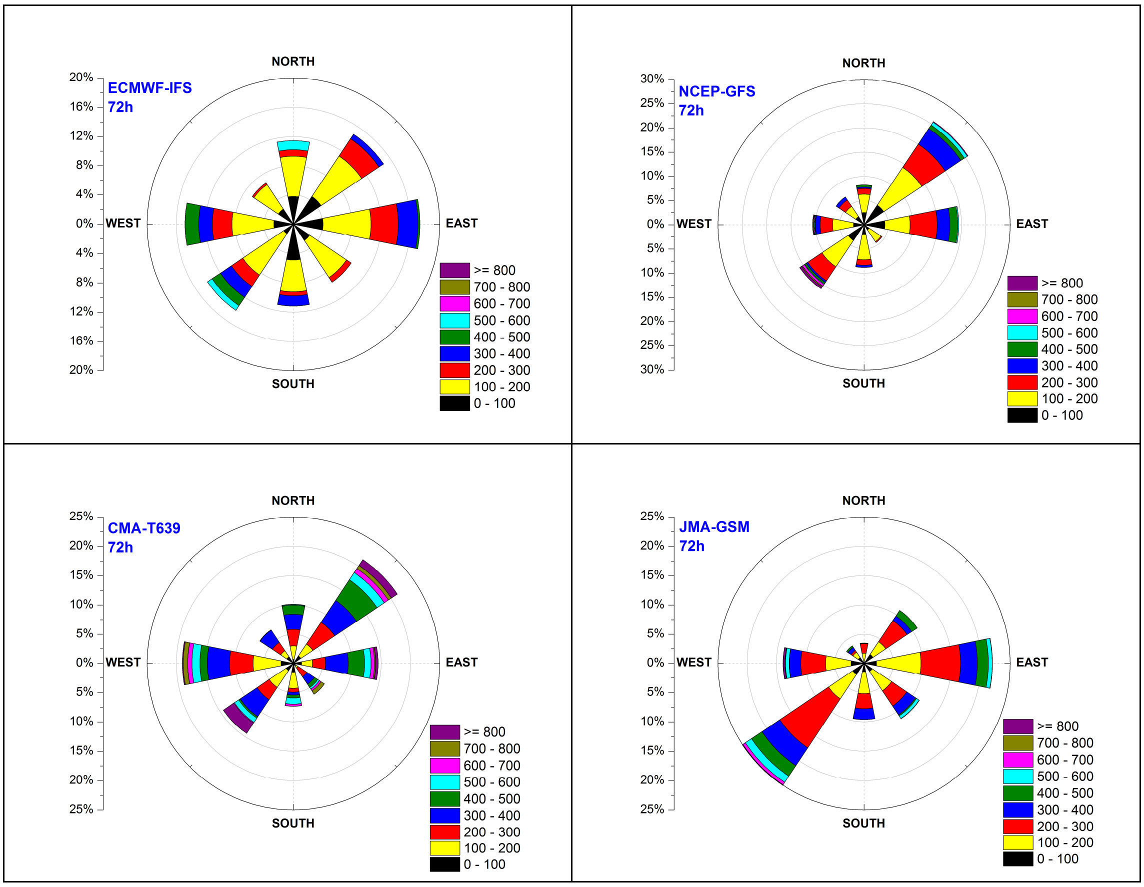


Fig.5. Polar scatter plots depicting the mean combined direction and magnitude errors around the actual storm location for each method at different lead time levels in 2015.



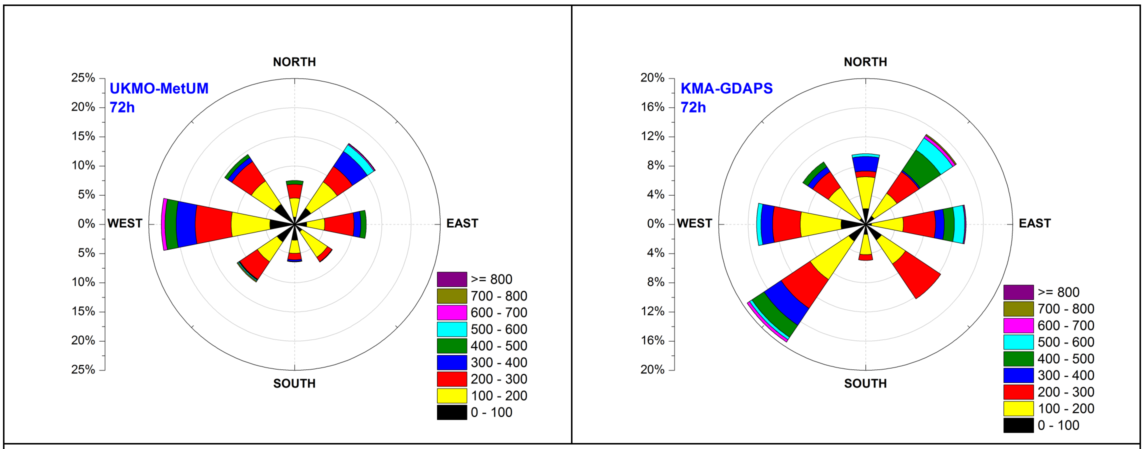


Fig.6. Examples of track error-rose diagram (TER) to represent the direction and magnitude distributions of track errors.

Another useful tool to evaluating the systematic bias of a certain objective track forecast method is name "Track Error Rose". TER uses the same conception of "wind rose" diagram as reference. Fig.6 shows the examples of track TER to represent the direction and magnitude distributions of track errors from six global models at the lead time level of 72h in 2015. In this example of TER diagram, each color bar represents different magnitude of track error, and the length of alignment of color bars represent the proportion of each azimuth angles. The TER diagram reveals the track error distribution (both the error magnitude and percentage of sample size) at each azimuth angle.

4.2 EPS forecast

Seven ensemble TC forecast systems from TIGGE are evaluated below. Table.4 listed the details description of these EPSs. To evaluate the performance of TC track forecast of each EPS, we first treat the ensemble forecasts as deterministic by summarizing the ensemble using the mean applied to the members. Fig.7 shows the ensemble mean track error for seven EPSs. Fig.7 indicates that ECMWF-EPS, UKMO-EPS and NCEP-GEFS are the top 3 EPSs in 2015. The ensemble mean track error at the lead time level of 120h for both ECMWF-EPS and UKMO-EPS are less than 300km.

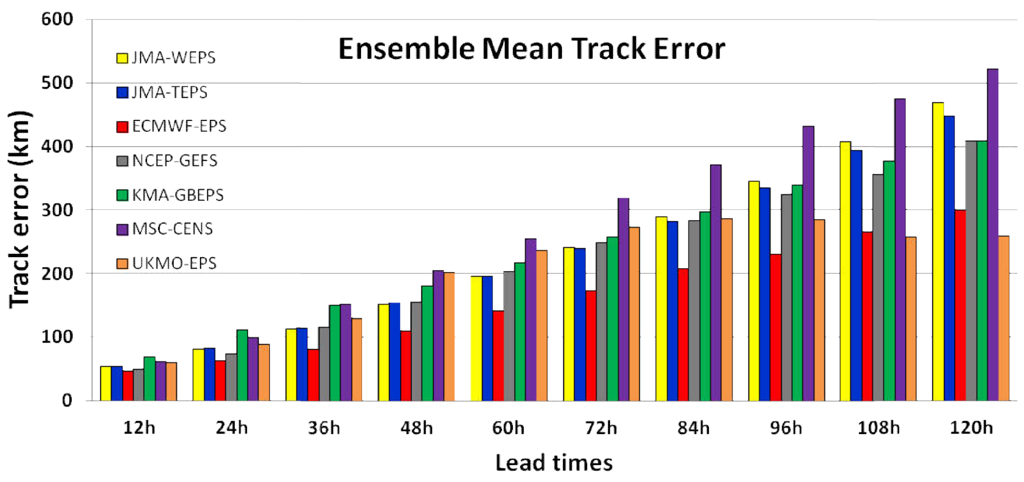


Fig.7. Ensemble mean track error for seven EPSs in 2015.

Table.4. Details of ensemble forecasts guidance

Model name	ECMWF-EPS	JMA-TEPS	JMA-WEPS	KMA-GBEPS	MSC-CENS	NCEP-GEFS	UKMO-EPS
Resolution	TL639 (0-10d) TL319 (10-15d)	TL319L60	TL319L60	T213L40	0.9°	T126L28	
Data resolution	\	0.5625°	0.5625°	0.5625°	1°	1°	
Members	51	11	51	24	21	21	24
Perturbation method	Singular Vector	SVD	SVD	Bred Vector	Ensemble Kalman	Ensemble Transform	
Forecast time	00:00 12:00	00:00 12:00	12:00	00:00 12:00	00:00 12:00	00:00 06:00 12:00 18:00	00:00 12:00
Output Interval (h)	12	6	6	6	6	6	12
Forecast hour(h)	120	132	216	120	240	240	192

The ensemble spread is an indicator of forecast uncertainties, which is not in linear relationship with mean track error. When the spread is large, the mean track error may be smaller, and vice-versa. Traditionally, researcher applied scatter plot of position error and ensemble spread to analyze the relationship between the forecast uncertainty and the error of a particular EPS. A bi-directional scatter plot is adopted here to re-analyze the traditional scatter plot. In the bi-directional scatter plot (Fig.8), the blocks in the middle of the plot represents the mean value of spread or track error. The lower (left) and upper (right) bars represent the 25th and 75th quantile values. It's found that only the spread of ECMWF-EPS are larger than track error, other 4 EPSs' spreads are smaller than track error. All the EPSs' spans of ensemble spread larger than track error, indicating that the uncertainties of ensemble spread are larger than that of track error.

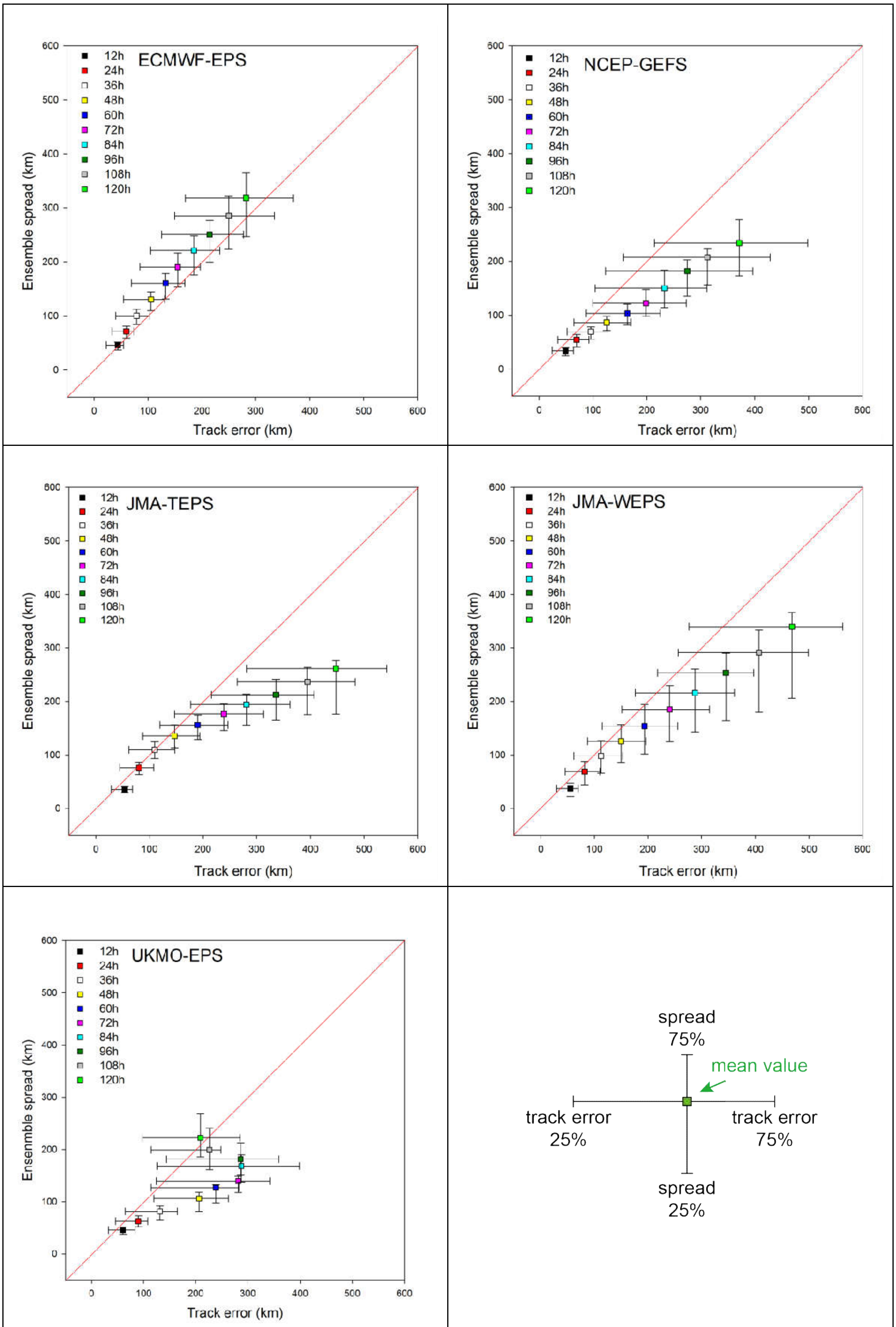


Fig.8. Bi-directional track forecast scatter plot for ECMWF-EPS, NCEP-GEFS, JMA-TEPS, JMA-WEPS and UKMO-EPS. The blocks represents the mean value of spread or track error. The lower (left) and upper (right) bars represent the 25th and 75th quantile values.

Fig.9 is schematic diagram of probability ellipse. If the observation TC location appears in the corresponding ellipse, it was taken as a hit case. The probability chosen here is 70% (Yu et al. 2012). Table.5 shows the mean hit ratio of probability ellipse for all the 7 EPSs at different prediction times in 2015. The hit ratio of ECMWF-EPS is much better than other EPSs. This may be caused by that the perturbation methods of the remaining EPSs are not as close to the realistic weather facts. The hit ratios of NCEP-GEFS and UKMO-EPS are rather poor, and it’s partly caused by the underestimate of atmospheric uncertainties.

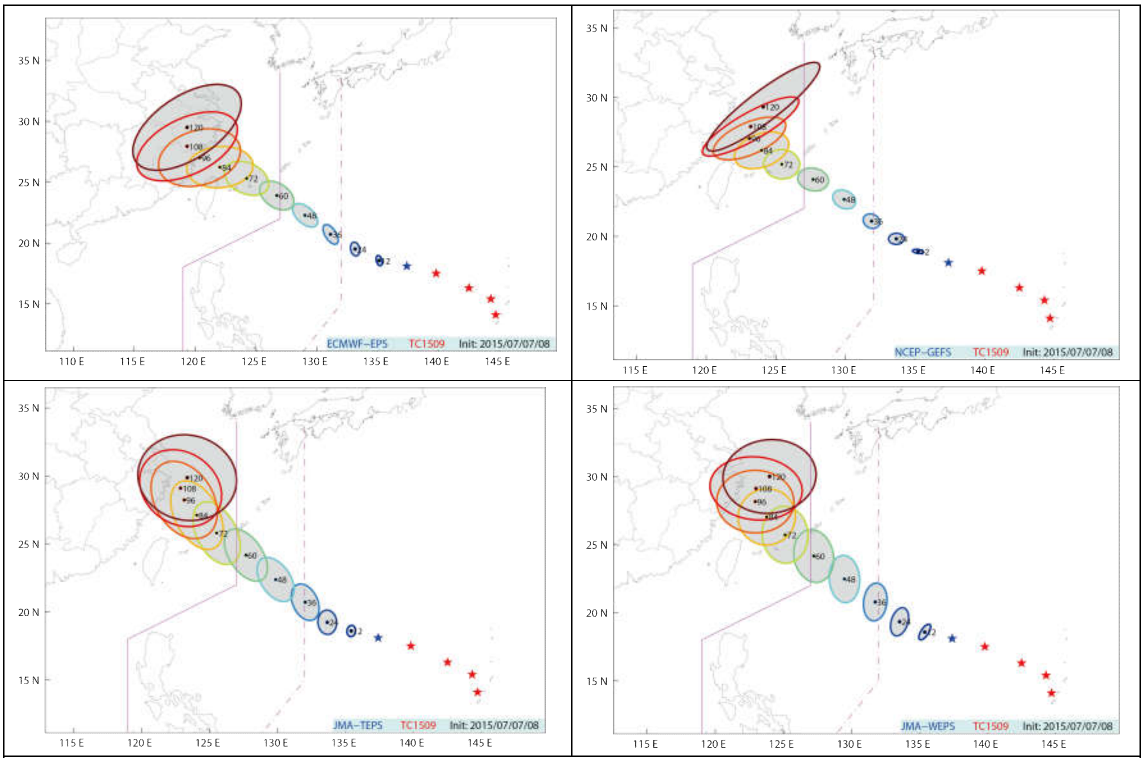


Fig.9 Schematic diagram of 70% probability ellipse for ECMWF-EPS, JMA-TEPS, NCEP-GEFS and JMA-WEPS. The stars on the figure are the best-track, and the colorful dots and ellipses are ensemble members and probability ellipses for different prediction times. On the upright corner contains the information about model name, TC number and start time, corresponding to the blue star.

Table.5 The EPS forecast hit ratio of 70% probability ellipse in 2015

	12h	**24h**	**36h**	**48h**	**60h**	**72h**	**84h**	**96h**	**108h**	**120h**
JMA-WEPS	46.9%	55.7%	53.5%	53.3%	51.6%	52.2%	55.4%	53.9%	51.0%	55.7%
JMA-TEPS	57.2%	74.9%	78.5%	81.8%	64.3%	53.4%	48.1%	43.0%	38.9%	38.6%
ECMWF-EPS	82.4%	88.7%	89.6%	91.9%	93.2%	92.4%	90.9%	89.3%	91.4%	90.9%
NCEP-GEFS	56.1%	64.5%	63.4%	64.4%	56.0%	48.8%	44.0%	50.0%	48.4%	43.4%
KMA-GBEPS	91.2%	89.4%	88.1%	82.9%	81.8%	77.4%	89.7%	82.1%	84.6%	91.7%
MSC-CENS	72.6%	80.6%	81.6%	75.9%	75.9%	87.1%	80.0%	80.4%	84.1%	84.6%
UMKO-EPS	53.2%	61.9%	48.7%	40.5%	42.4%	46.2%	54.5%	55.6%	81.3%	78.6%

5. TC genesis forecast verification

TC genesis verification		Observation	
		O=1 (Yes)	O=0 (No)
Forecast	F=1 (Yes)	Hit	False Alarm
	F=0 (No)	Miss	Correct rejection

Fig.10. Binary contingency table of observed versus forecast events.

The focus of tropical cyclone genesis is on whether a disturbance will form into a full-fledged TC. Deterministic forecasts of genesis are usually verified categorically as hits, false alarms, and misses (e.g., Elsberry *et al*. 2007). Fig.10 shows the binary contingency table which counts the number of hits, misses, false alarms, and correct rejections for a set of forecasts (e.g. TC genesis forecasts). These counts, when given as percentages and computed for a number of thresholds from small to large (for the verification of TC genesis, we use the thresholds of 200, 250 and 300km, respectively), reveal quite a bit about the nature of the errors. The verification indices are Critical Success Index (CSI), Probability of Detection (POD) and False Alarm Ratio (FAR). Fig.10 shows CSI, POD and FAR of NCEP-GFS at different lead time for 200, 250 and 300km thresholds in 2015, respectively. As shown in Fig.11, CSI and POD indices decrease according to increasing of lead times, inversely while FAR index increases according to increasing of lead times for NCEP-GFS model. It is natural due to increasing of model errors according to increasing of lead times. Fig.11 demonstrated that NCEP-GFS forecast is of fairly good quality to forecast the genesis of TCs and had been making more noticeable.

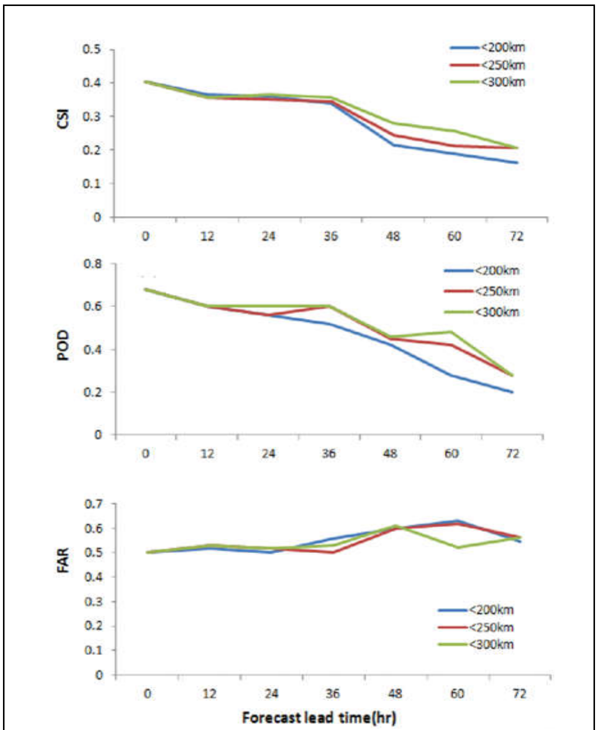


Fig.11. CSI, POD and FAR of NCEP-GFS at different lead time for 200, 250 and 300km thresholds, respectively in 2015. (Thanks to Pak Sang Il from DPRK)

6. TC intensity forecast verification

6.1 Deterministic forecast

TC intensity forecasts (i.e., maximum wind speed and minimum pressure) are typically evaluated as continuous parameters, using standard verification measures such as the Mean Absolute Error (MAE) or Mean Error (ME). MAE provides an indication of the average magnitude of the error, whereas ME measures the bias in the forecasts. Table.6 show the MAE of maximum wind speed forecast for each method at each lead time level in 2015. One thing should be remember that the wind speed of both forecast and best track were converted to 2-min average according to the WMO documentation (Harper B A. *et al*, 2010).

Table.6. Average absolute maximum wind speed error for each method at 24, 48, 72, 96 and 120h lead time levels in 2015 (Unit: m/s)

Method	Lead times	24h	48h	72h	96h	120h
Subjective Methods	CMA-sub	4.26(640)	6.36(537)	7.72(441)	8.25(359)	9.86(287)
	JMA-sub	5.08(639)	7.94(539)	9.38(447)	/	/
	KMA-sub	5.03(634)	7.49(536)	8.61(441)	9.55(360)	10.38(287)
	JTWC-sub	4.88(612)	7.17(512)	8.37(426)	8.81(347)	8.63(287)
	HKO-sub	5.04(248)	6.74(189)	7.40(135)	7.08(89)	7.73(51)
Global NWP Models	ECMWF-IFS	5.44(305)	8.42(261)	10.33(212)	11.19(171)	11.25(135)
	NCEP-GFS	6.42(414)	7.98(355)	9.60(289)	10.43(234)	10.13(188)
	JMA-GSM	6.67(638)	10.61(540)	12.60(443)	/	/
	CMA-T639	6.78(46)	9.78(36)	12.75(28)	15.40(20)	18.93(15)
	KMA-GDAPS	7.14(207)	12.07(173)	15.08(142)	17.93(116)	19.22(81)
	UKMO-MetUM	6.71(315)	10.35(269)	12.13(223)	13.65(179)	13.49(143)
Regional NWP Models	BOM-ACCASS	7.20(300)	10.20(254)	11.99(208)	/	/
	CMA-TRAMS	8.81(252)	9.76(211)	10.05(169)	/	/
	STI-GRAPES	8.05(423)	9.06(211)	10.28(274)	/	/

Some new TC intensity verification mentalities have been on trial for the last 5 year in STI, such as plots like those in Fig.12, which is called Taylor Diagram (Taylor, 2001). Taylor Diagram is introduced in the verification of TC intensity forecast to analyze the internal relationship between the standardized deviation and correlation coefficient together with center different root-mean-square. The best prediction always with highest correlation coefficient compared to “OBS”, and with standardized deviation and center different root-mean-square closed to “1”. According to Fig.12 the RMS error of both minimum surface pressure and maximum wind speed were smallest at 0h for JMA. The correlation coefficients of minimum surface pressure between observation and forecast are in the interval of 0.6 to 0.9. For the maximum wind speed forecast, the normalized standardized deviations of most global models are in the interval 0.75 to 1.25, except for KMA-GDAPS.

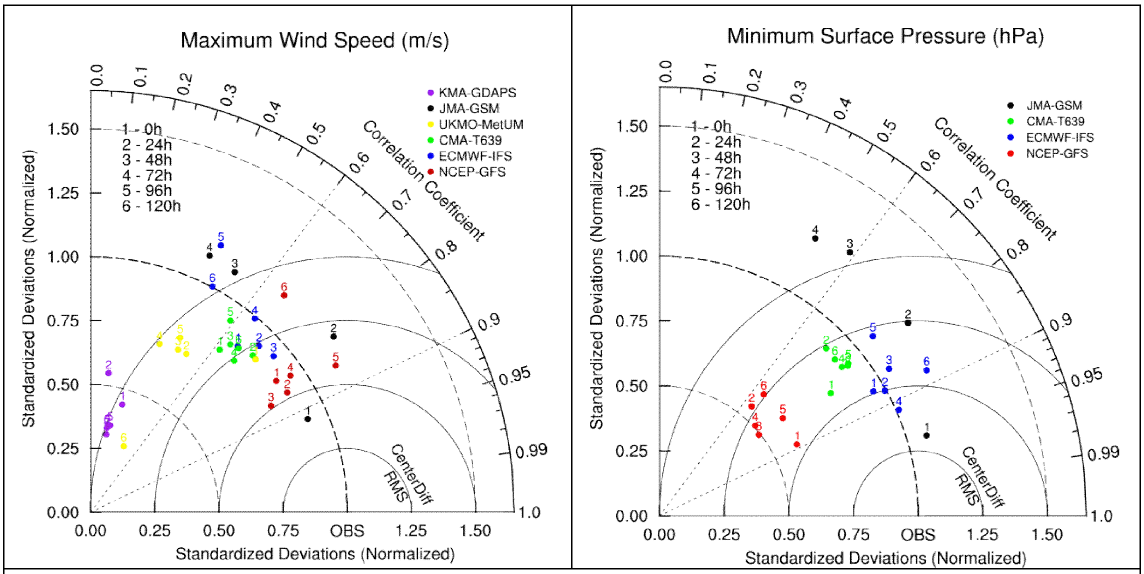


Fig.12. Taylor-diagram to evaluate global models' TC intensity forecast. Left: for maximum wind speed, right: for minimum surface pressure.

Fig.13 shows the intensity forecast skill score at the lead time levels of 24 and 48h for subjective method, global and regional models from 2010 to 2015. All the subjective methods obtained positive intensity forecast skill scores for the last 6 years, however, the skill scores were much less than track forecast skill scores. More depressing is that intensity forecast skill scores for both global and regional models were almost negative.

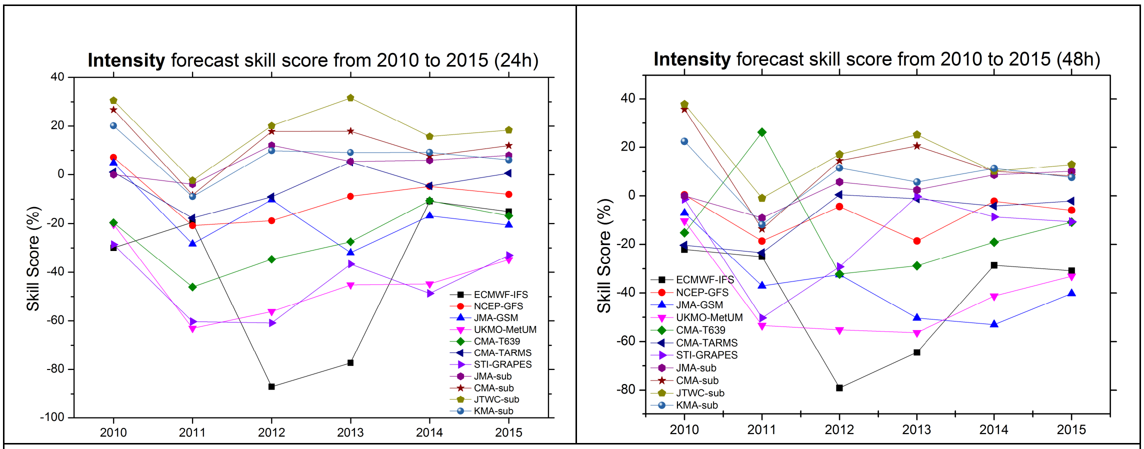


Fig.13. The same as **Fig.3** but for intensity forecast skill score.

6.2 EPS forecast

The ensemble forecasts of TC intensity from the TIGGE ensemble prediction systems as listed in Table.4 have been evaluated using history ranking analyses, Brier Score (BS), Ranked Probability Score (RPS) and forecast bias estimating since 2013 at STI. Fig.14 shows Intensity forecast bias estimating for EPS. The evaluation results show that most EPSs were under estimated the TC initial intensity, except for ECMWF-EPS.

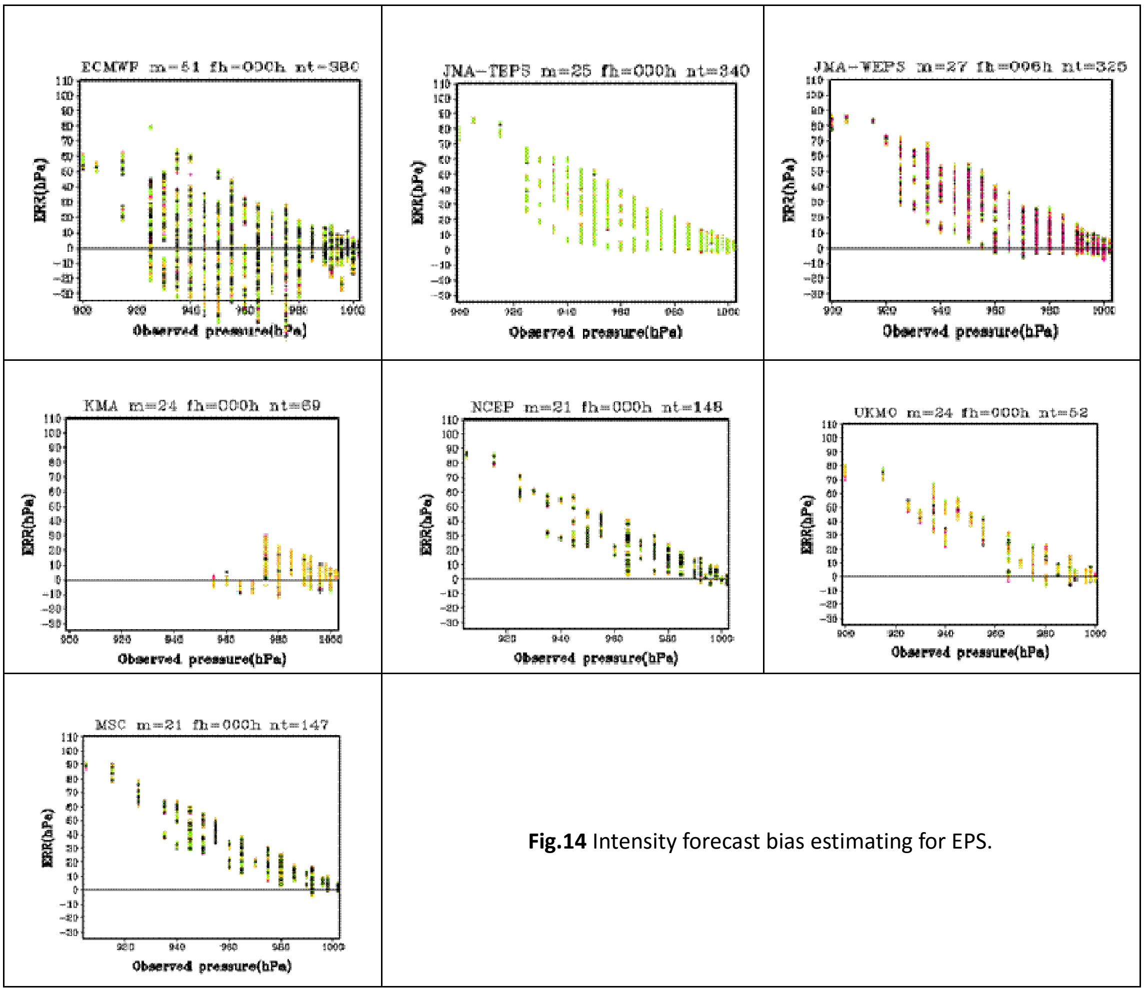


Fig.14 Intensity forecast bias estimating for EPS.

As indicated by the Brier score (Fig. 15), the ensemble system of UKMO-EPS outperforms other systems significantly at long lead times. However, BS differences in the seven EPSs have narrowed at long lead time levels. The positive contribution of initial correction degrades quickly from 6h to 30h for six of the seven systems. The effect of initial correction is in-significant or even negative for some systems after 30 h.

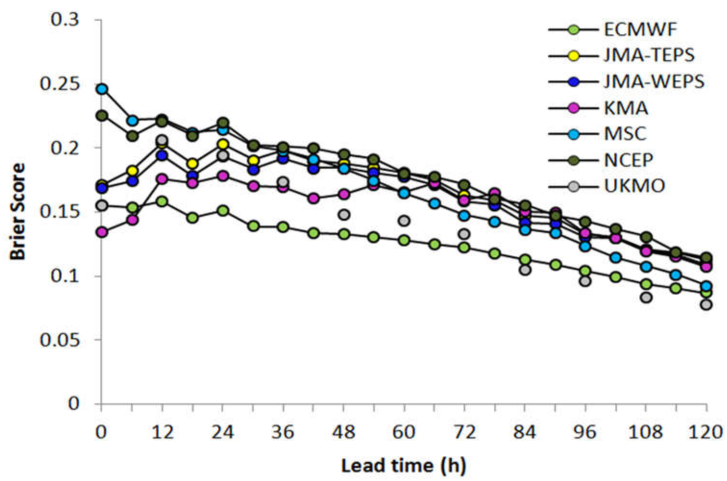


Fig.15 Brier scores for EPS intensity forecast in 2015.

7. Conclusions

Verification of TC forecasts is important for improving the NWP and subjective 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 2015. The verification results include TC genesis, track, and intensity for both deterministic and ensemble forecast guidance. The results show that stepped decreases in the values of each quantile of track errors were made at every lead time level from 2010 to 2015 for both deterministic and ensemble NWP guidance, however, intensity forecast skill for both global and regional models were almost stagnating for the last six years.

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, STI/CMA, RSMC Tokyo Typhoon Center/WMO.

Contributors to this report are (in alphabetic order) lina BAI, Barbara Brown, Qing CAO, Xiaoqin LU, Pak Sang Il, Jie TANG, Rijin WAN, Zhihua ZENG.

Appendix: acronyms used in this report

BOM	Bureau of Meteorology (Australia)
CMA	China Meteorological Administration
CMC	Canadian Meteorological Center
CSI	Critical Success Index
ECMWF	European Centre for Medium Range Weather Forecasting
EPS	Ensemble prediction system
FAR	False alarm ratio
GEFS	Global Ensemble Forecast System
GFS	Global Forecast System
JMA	Japan Meteorological Agency
JTWC	Joint Typhoon Warning Center
MAE	Mean absolute error
ME	Mean error
MSE	Mean Squared Error
NWP	Numerical weather prediction
POD	Probability of Detection
RMSE	Root Mean Squared Error
STI	Shanghai Typhoon Institute
TC	Tropical cyclone
TIGGE	THORPEX Interactive Grand Global Ensemble
WMO	World Meteorological Organization

Reference

1. Chen G. M., H. Yu, Q. Cao, Z. H. Zeng, 2013: The performance of global models in TC track forecasting over the western North Pacific from 2010 to 2012[J]. *Tropical Cyclone Res. Rev*., 2013, 2(3): 149-158.

2. Chen G. M., Hui Yu, and Qing Cao, 2015: Evaluation of operational global model on tropical cyclone forecast over the Western North Pacific in 2013. *Tropical Cyclone Res. Rev.,* (in press)
3. Elsberry, R.L., T.D.B. Lambert, and M.A. Boothe, 2007: Accuracy of Atlantic and Eastern North Pacific tropical cyclone intensity forecast guidance. Wea. Forecasting, 22, 747-762.
4. Harper B A, Kepert J D, Ginger J D. 2010, Guidelines for converting between various wind averaging periods in tropical cyclone conditions. Wold Meteorological Organization, *TCP Sub-Project Report*, WMO/TD-No.1555.

5. Lei, X.T. and H. Yu, 2014: Progress of the WMO Typhoon Landfall Forecast Demonstration Project (WMO-TLFDP). *Advances in Met. Sci. & Tech*. (in Chinese with English abstract)

6. Taylor K. E., 2001: Summarizing multiple aspects of model performance in a single diagram.*JGR*, vol 106, no. D7, 7183-7192, April 16, 2001

7. Tang, X., X.T. Lei, and H. Yu, 2012a: WMO Typhoon Landfall Forecast Demonstration Project (WMO-TLFDP) – concept and progress. *Tropical Cyclone Res. Rev.*, 1, 89-96.

8. Tang X., X.T. Lei, H. Yu, B.D. Chen, Z.Q. Chen, L.Y. Tao, 2012b: Progress report of the WMO Typhoon Landfall Forecast Demonstration Project (WMO-TLFDP). *WMO Workshop and Training Course on Operational Tropical Cyclone Forecast*, Shanghai, China.

9. Yu, H., G. Chen, and B. Brown, 2013: A new verification measure for tropical cyclone track forecasts and its experimental application. *Tropical Cyclone Res. Rev.*, 2, 185-195.

10. Yu, H., S.T. Chan, B. Brown, and Coauthors, 2012: Operational tropical cyclone forecast verification practice in the western North Pacific region. *Tropical Cyclone Res. Rev.*, 1(3), 361-372.