



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

석사학위논문

수치모델의 태풍 강도 예측성능
평가 및 컨센서스 기법 연구

제주대학교 대학원

해양기상학 협동과정

오 유 정

2015년 8월

A Study on Performance Evaluation of Typhoon
Intensity Prediction Models and their Consensus
Technique

You Jeong Oh

(Supervised by professor Il-Ju Moon)

A thesis submitted in partial fulfillment of the requirement for the
degree of Master of Science

2015. 06.

This thesis has been examined and approved.

Department of Marine Meteorology
GRADUATE SCHOOL
JEJU NATIONAL UNIVERSITY

목 차

LIST OF FIGURES	ii
LIST OF TABLES	iv
ABSTRACT	vi
I. 서론	1
II. 자료 및 분석방법	4
1. 태풍 수치모델 특징	4
2. 태풍 최적경로 자료	8
3. 연구방법	11
III. 수치모델의 태풍 강도 예측성능	15
IV. 태풍 강도 예측 컨센서스 기법	31
1. 훈련단계	31
2. 예측단계	36
3. 예측성능 검증	40
V. 요약 및 결론	47
참고문헌	52

LIST OF FIGURES

Fig. 1. Model intensity forecast for Typhoon Bolaven (2012). The black line indicates RSMC-Tokyo best-track. Other lines indicate each model output. (a) GFS, (b) JGSM, (c) HWRF, and (d) KWRF.	9
Fig. 2. Maximum wind speed (m/s) comparison between the RSMC-Tokyo and the JTWC from 2008 to 2013.	10
Fig. 3. Scatter plot of maximum wind speed (0~72h) between all models and RSMC-Tokyo best-track from 2013 to 2014. The numbers in the bracket indicate ranking of models associated with verification method.	17
Fig. 4. MAE of 48-h forecast result for all models from 2008 to 2014. (a) Maximum wind speed (m/s), (b) Minimum Sea Level Pressure (hPa), (c) Track (km).	20
Fig. 5. Six Fuzzy clusters of 15377 TC track case (0~120h) from 1982 to 2014 and all the tracks before using FCM. The thick track are the cluster mean tracks.	28
Fig. 6. Geographical distribution of average maximum wind speed (m/s) relative to RSMC-Tokyo best-track within 10 Lat./Long. Boxes.	30
Fig. 7. Errors with forecast time and total errors (12~72h) for best individual model and CON1, CON2 from 2012 to 2014.	37
Fig. 8. Scatter plot of maximum wind speed (12~72h) between all models and RSMC-Tokyo best-track from 2013 to 2014. (a)-(l) all individual models, (k) CON1, (l) CON2. The number in the bracket indicate ranking of models associated with verification method.	41
Fig. 9. 48h forecast MAE (m/s) of 4 consensus schemes associated with using number of models from 2013 to 2014. The number in figure	

indicates mean error of 4 consensus schemes. 45

Fig. 10. Same as in Fig. 9 except for 72h forecast MAE (m/s). 45

LIST OF TABLES

Table 1. List of the numerical model used for this study.	6
Table 2. MAE of maximum wind speed and minimum sea level pressure with lead time for all models from 2013 to 2014. The number in bracket is ranking of errors. Shaded boxes represent from 1 st model to 3 rd model.	18
Table 3. MAE of maximum wind speed (0~72h) associated with month for all models from 2013 to 2014. The number in bracket is ranking of errors. Shaded boxes represent from 1 st model to 3 rd model.	21
Table 4. Same as in Table 3 except for category.	22
Table 5. Same as in Table 3 but except for rapidly intensity change case.	24
Table 6. Same as in Table 3 but except for approaching korea peninsula case.	25
Table 7. Correlation coeff., MAE, and Bias of maximum wind speed (6~72h) from 2013 to 2014. The number in bracket is ranking of errors. Shaded boxes represent from 1 st model to 3 rd model.	27
Table 8. Same as in Table 3 except for cluster group.	29
Table 9. Errors of each consensus scheme with forecast time in training period (2011). Shaded boxes indicates selecting model.	34
Table 10. Ranking of individual models with forecast time for training period (2011). Based on (a) MAE, (b) Corr., Shaded area shows the cases that CON1 scheme were used. The models inside thick line box were applied by the CON2 scheme.	35
Table 11. Ranking of individual models with forecast time for forecast period (2012). Based on (a) MAE, (b) Corr.,	38

Table 12. Compare with errors associated with training period for 2013 forecast.	42
Table 13. Same as in Table 12 except for 2014.	42
Table 14. Compare with errors associated with fixed training period for 2013 forecast.	44
Table 15. Same as in Table 14 except for 2014.	44
Table 16. 48h and 72h forecast errors of 4 consensus scheme (4) from 2013 to 2014.	46

Abstract

For improvement of TC (Tropical Cyclone) intensity predictability, we have investigated the performance of sixteen TC intensity guidance models in the western North Pacific during 2008–2014 and tested various consensus techniques on the basis of the results. From the evaluation using total errors up to 72 hours, we found that HWRF (Hurricane Weather Research and Forecasting model) outperforms all others overall although GFS (Global Forecast System), KWRF (Korea Meteorological Administration WRF) and JGSM (Japan meteorological agency Global Spectral Model) also shows a good performance in some forecast times. HWRF, in particular, shows the highest performance in predicting the intensity of strong TCs above Category 3, which may be attributed to its highest spatial resolution (~ 3 km). GFS was the most improved model from 2008 to 2014. For initial intensity error, two Japanese models, JGSM and TEPS (Typhoon Ensemble Prediction System), had the smallest error. In track forecast, the ECMWF (European Centre for Medium-Range Weather Forecasts), including recent GFS model, outperformed others.

On the basis of evaluation of 16 numerical models, we have tested two consensus techniques according to the number of model used, training period, rank criteria (correlation or absolute error), and averaging methods (equally or weighted mean). Prediction tests for 2012–2014 reveals that (i) the CON2 (using one consensus method showing the smallest error up to 72 hours during training period) generally outperformed CON1 (using the best performance consensus methods at individual forecast time during training period), (ii) 1-year training period showed lower errors than 2-year, (iii) the consensus using 4 or 5 models showed the best performance, (iv) weighted averaging matched correlation criteria while equally averaging matched absolute error criteria in terms of absolute error.

Keyword: tropical cyclone, intensity prediction, consensus, guidance models.

I. 서론

태풍은 강풍, 집중호우, 해일 등을 동반하며 많은 인명 및 재산 피해를 주는 자연재해 중 하나이다. 우리나라는 자연재해로 발생한 피해들 중 50% 이상이 태풍의 직·간접적 영향으로 발생한다(NEMA, 2011). 최근 기후변화와 더불어 한반도에 상륙하는 태풍의 강도는 지난 55년간 꾸준히 증가하였고, 특히 2000년에 들어 그 증가 경향이 더욱 뚜렷해졌다(Choi and Moon, 2008). 태풍에 의한 피해를 저감하기 위해서는 정확한 태풍 예측이 필수적이다. 최근 수치모델의 발달은 태풍의 진로 및 강도 예측기술의 정확도 향상에 많은 기여를 하였다(Knaff et al., 2003; DeMaria et al., 2005; Roger et al., 2006). 특히, 여러 수치모델의 다양한 결과를 조합하는 컨센서스(Consensus) 기법(Cangialosi and Franklin, 2013)은 개별적인 태풍 수치모델들의 무작위적인 오차들을 상쇄시켜 태풍의 진로예측 향상에 큰 기여를 한 것으로 알려져 현재 국내·외 태풍예보기관에서 널리 사용되고 있다(Elsberry, 2014; WMO, 2007).

수치모델들을 이용한 컨센서스 방법에는 단순평균방법과 가중치평균방법으로 나눌 수 있다. 단순평균방법은 여러 수치모델들의 결과를 같은 비중으로 평균하는 간단한 방법지만 개별모델의 예측성능보다 평균적으로 뛰어난 것으로 알려져 있다(Goerss, 2000). 실제로 Goerss (2000)는 3개의 전구 모델로 만든 진로 컨센서스와 2개의 지역모델로 만든 진로 컨센서스를 이용하여 개별모델 보다 태풍 진로 예측의 정확도를 향상시킬 수 있음을 보였다. 가중치평균방법은 컨센서스에 포함되는 개별 수치모델들의 과거성능을 반영하여 각 수치모델들에 가중치를 주어 평균하는 방법이다. Kumar et al., (2003)는 1998~2000년 태평양에서 발생한 태풍의 진로 및 강도에 대하여 모델 별로 가중치를 고려하는 슈퍼앙상블 방법이 개별 모델 성능들 보다 우수함을 보인바 있다. Kotal et al. (2011)는 북인도양에서 2008년부터 2009년까지 발생한 태풍에 대해 5개의 수치모델을 고정하고 예측 시간 별 가중치를 달리하는 컨센서스 방법을 사용하여 태풍 진로 향상을 이루었다. 우리나라 국가태풍센터에서도 다중모델을 단순 평균하는 방법(CONB1)과 가

중치를 주어 평균하는 방법(CONB2)을 병행하여 사용하고 있다(Jun et al., 2015).

국외 태풍예보 기관에 대해 살펴보면, 다수의 수치모델들을 보유하고 있는 미국 NHC(National Hurricane Center)에서는 단순평균방법, 보정하는 방법, 슈퍼앙상블(Florida State University Super Ensemble, FSSE)방법 등이 예보의 첫 단계로 사용되고 있다. 북서태평양 도쿄 지역특별기상센터(Regional Specialized Meteorological Centers-Tokyo, RSMC-Tokyo)에서는 단순평균 컨센서스를 사용하고 있다(JMA, 2012). 미국합동 태풍경보센터(Joint Typhoon Warning Center-Honolulu, JTWC)에서는 보유하고 있는 수치모델 중 이용 가능한 모델이 2개 이상일 때, 이 수치모델들을 단순평균하는 방법(CONW)를 사용하고 있다.

지금까지 태풍 진로에 대한 컨센서스 기법에 대해서는 많은 연구가 있었지만 태풍 강도에 대한 연구는 거의 없었다. 본 연구에서는 국가태풍센터에서 사용 중인 수치예보 모델을 이용하여 2011년부터 2014년까지 북서태평양에서 발생한 태풍 강도 예측을 위한 최적의 컨센서스 기법을 찾고자 한다. 이를 위해해서는 먼저 컨센서스 기법 개발에 사용될 수치모델(즉 가이드스 모델)들의 태풍강도 예측 성능을 자세히 평가할 필요가 있다.

최근 태풍 역학모델은 해상도 및 물리과정 개선 등을 통해 그 성능이 해마다 바뀌고 있으며 사용가능한 가이드스 모델의 수도 점점 많아지고 있다. 따라서 매년 가이드스 모델들의 성능을 평가하고 그 경향성을 분석할 필요가 있다. 실제로 미국 합동태풍경고센터와 도쿄 지역특별기상센터와 같은 태풍 예보기관에서도 사용하는 가이드스 모델에 대한 성능 평가를 매년 실시하여 모델들의 특징을 파악하고 컨센서스 기법에 근거자료로 활용하고 있다.

본 연구에서는 2008년부터 2014년까지 최근 7년간 국가태풍센터에서 사용하고 있는 16개의 태풍 강도 가이드스 수치모델들(Table. 1)에 대한 강도 예측성능을 평가하고, 컨센서스 기법 개발에 대한 기초 근거 자료로 활용하였다. 수치모델의 태풍강도 예측성능평가는 기본적인 통계오차분석 외에도 강도별 예측정확도, 급격한 강도변화 예측능력, 한반도 영향태풍의 예측능력, 강도변화 예측능력, 이동경로별 강도예측 능력 등과 같은 기존 연구에 다루지 않은 다양한 방법으로 분석을 실시하고자 한다. 또한 현업 이용 시에 참고할 수 있도록 분석 항목마다 모델별 성능 순위를 매기고자 한다.

최적의 컨센서스 기법은 다음의 네 가지 방법을 사용하여 조사하고자 한다. 첫째, 선행연구들에서 모델 성능에 대한 순위를 결정하는데 있어 오차만을 기준으로 사용한 것과는 달리 본 연구에서는, 태풍 가이드선스 모델들 중 오차는 크지만 태풍 강도 변화의 경향성을 잘 모의하는 모델이 있음을 고려하여, 상관도를 기준으로 순위를 결정하는 방법을 시도하고자 한다. 둘째, 단순평균방법(AVE)과 모델들의 과거 성능을 반영한 가중치 평균방법(WEI)을 모두 적용하여 최적의 기법을 조사하고자 한다. 셋째, 예측시간별로 최적의 컨센서스 방법을 따로 사용하는 방법과 모든 예측시간 동안 가장 성능이 좋은 한 가지 방법만을 사용하는 방법에 대한 성능을 비교해 보고자 한다. 마지막으로 사용한 모델 개수를 달리하여 최적의 컨센서스 기법을 찾고자 한다.

제 2장에서는 본 연구에서 사용된 태풍 강도 가이드선스 모델들과 태풍 최적경로 자료 및 분석방법에 대해 상술한다. 제 3장에서는 다양한 방법으로 분석된 수치모델의 성능평가 결과를 기술한다. 제 4장에서는 제 3장의 평가를 바탕으로 다양한 컨센서스 기법에 대한 실험을 수행하고, 그 결과를 바탕으로 최적의 태풍강도 컨센서스 기법을 제시한다. 끝으로 제 5장에서는 본 연구의 요약 및 결론을 제시한다.

II. 자료 및 연구방법

1. 태풍 수치모델 특징

본 연구에서는 국가태풍센터에서 북서태평양 태풍 강도 가이드스 모델로 사용하고 있는 16개 모델들에 대한 예측 성능을 평가하고, 이 중 11개의 모델에 대하여 컨센서스 기법 연구에 활용하였다. 각 모델의 설명, 자료기간, 공간해상도, 자료생산 주기 등은 Table 1에 기술되어 있다.

16개의 모델 중에 우리나라에서 운용 중인 모델은 UM, UM_R, TWRF, KWRF, DLM, DBAR, KEPS이다. UM은 영국 기상청의 통합 수치모델(Unified Model, UM)을 도입하여 개발한 기상청의 전지구 자료동화·예측 시스템(Global Data Assimilation and Prediction System, GDAPS)으로 2010년부터 기상청의 현업예측 모델로 사용되고 있다(Lee et al, 2011). UM은 전구모델과 지역모델로 구성되어 있다. UM 전구모델은 4차원 변분법을 사용하며 25 km의 공간해상도와 70개의 수직 층을 가지며, 하루 4번 예측 정보를 생산한다. 자료기간은 2010년부터 2014년까지이다. UM 지역모델(UM Regional, UM_R)은 12 km의 공간해상도와 70개 수직 층을 가지며, 하루 4번의 예측 자료를 생산한다. 자료기간은 총 4년(2011~2014년)이다.

KWRF는 WRF (Weather Research and Forecasting) 모델을 기반으로 한 기상청의 지역예보시스템(RDAPS) 중의 하나이다. KWRF는 10 km의 고해상도 격자와 고급역학을 적용한 고정밀 모델로, 특히 초기 분석장 품질 향상을 위해 3차원 변분법을 이용한 자체분석을 수행할 뿐 아니라 태풍 초기장 개선을 위한 보거싱(Bogusing) 기법을 개발하여 적용하고 있다(Park et al., 2008). 하루 4번 예측자료를 생산하며, 자료기간은 총 4년(2011~2014년)이다. TWRF는 KWRF와 마찬가지로 WRF를 기반으로 하는 모델로서 태풍 전용으로 운용 중이다. 15 km의 공간해상도를 가지며 하루 2번 예측을 수행한다. 자료는 2010년부터 2014년까지 5년간의 자료를 포함하고 있지만 최대중심풍속(Maximum Wind Speed,

MWS)에 대해서는 2013년부터 자료를 제공한다.

KEPS는 2011년부터 기상청에서 운영해 오고 있는 앙상블모델로 4차원 변분법을 사용하여 전구모델로부터 초기 장을 생성하며 40 km의 공간해상도와 70개 층의 수직 층을 갖는다. 초기 섭동은 Ensemble Transform Kalman Filter (ETKF)를 사용하여 생산되고 총 24개의 앙상블 멤버로 구성되어 있다. 하루 2번 예측정보를 생산한다. DBAR는 부경대에서 개발한 이중 푸리에 급수를 이용한 전 지구 순압태풍 예보모델(Double-Fourier-Series BARotropic Typhoon Prediction Model)에 GFDL (Geophysical Fluid Dynamics Laboratory) 보거싱 기법을 접목한 태풍 예측모델이다(Lim et al., 2003). 하루 4회 예보를 하며 35 km의 해상도를 갖는다. DLM (Dynamic Linear Model)은 JMA (Japan Meteorological Agency)의 수치모델 결과를 이용한 역학-통계 모델이다(Sohn, 2005).

16개의 태풍 강도 예측 가이드스 모델 중 GFS는 미국 NCEP (National Centers for Environmental Prediction) 산하 EMC (Environmental Modeling Center)에서 개발되어 NWS (National Weather Service)에서 운영하고 있는 전 지구 예측시스템(Global Forecast System) 이다(NCEP, 2003). 수평해상도는 35 km이며, SAS (Simplified Arakawa-Schuber) 구름 모수화, 3차원 변분 GSI 자료 동화, 그리고 Vortex relocation 방식 초기화 방법을 사용한다. 하루 2회 예측자료를 생산하며, 자료기간은 2008년부터 2014년까지 7년이다. GFS_EPS는 GFS의 앙상블 모델로 23 km 수평 해상도를 가지며 21개 앙상블 멤버로 구성되어 있다.

HWRP (Hurricane WRF)는 미국 NCEP에서 운용 중인 WRF 기반의 지역 허리케인 예측모델이다. GFS 전지구 예측장을 초기 및 경계 값으로 사용하며, 향산 된 보거싱 방법과 물리과정 모수화 방법을 사용한다. 태풍 중심을 따라 이동하는 등지 격자 시스템을 사용하여 태풍 중심 부근에서 3 km까지 고해상도로 예측이 가능하다(Vijay et al., 2014). NOGAPS (Navy Operational Global Atmospheric Prediction System)는 미 해군에서 개발·운영 중인 55 km 수평 해상도와 30개의 수직 층으로 구성된 전지구 모델이다. 3차원 자료동화 방법을 사용하며, 자료기간은 2008년부터 2014년까지 이다.

Table 1. List of the numerical model used for this study.

No	Model Name	Description	Data Period	Spatial Resolution	Fcst interval	Fcst # per day	Fcst data
1	*UM	KMA Operational Global Model (Unified Model)	2010-2014	25km	6h	4	x,y,w,p
2	*UM_R	KMA Operational Regional Model (Unified Model)	2011-2014	12km	3h	4	x,y,w,p
3	*TWRF	WRF-based Typhoon Model	2010-2014	15km	6h	2	x,y,w,p
4	*KWRF	WRF-based KMA Operational Regional Model	2011-2014	10km	6h	4	x,y,w,p
5	DLM	BATS model data based Dynamic Linear Model	2008-2014	-	6h	4	x,y
6	DBAR	Double-Fourier-Series BARotropic	2008-2014	35km	3h	4	x,y
7	*JGSM	Typhoon-Prediction Model JMA Operational Global Spectral Model	2010X-2014	20km	6h	4	x,y,w,p
8	*TEPS	JMA Operational Typhoon Ensemble Prediction System	2009-2014	60km	6h	2	x,y,w,p
9	*NOGAPS (NAVGEM)	Navy Operational Global Prediction System	2008-2014	55km	6h	2	x,y,w,p
10	*GFS	NCEP Global Forecast System	2008-2014	1°	6h	2	x,y,w,p
11	*ECMWF	European Center for Medium-range Weather Forecasting	2008-2014	0.5°	24h	2	x,y,w,p
12	*KEPS	KMA Ensemble Prediction System (24 members)	2012-2014	40km	6h	2	x,y,w,p
13	ECMWF _TIGG	ECMWF Global Model received from TIGGE	2011-2014	16km	6h	2	x,y,w,p
14	ECMWF _EPS	ECMWF Global Ensemble Model (51 members)	2011-2014	32km	12h	2	x,y,w,p
15	GFS_EPS	NCEP Global Ensemble Model (21 members)	2012-2014	23km	6h	4	x,y,w,p
16	*HWRF	NCEP EMC WRF-based Regional Model	2012-2014	27/9/3km	6h	4	x,y,w,p

ECMWF (European Centre for Medium-Range Weather Forecast)는 유럽 28 개국을 회원국으로 구성되어 있는 독립적인 정부간(intergovernmental) 조직에서 운용중인 전지구 예측모델이다. ECMWF는 최신의 역학코어(Dynamic core), 자료동화(4차원 변분), 물리과정 모수화를 사용하며, 가장 예측 성능이 우수한 전지구 모델 중의 하나로 알려져 있다. 25 km 수평격자와 91개의 수직 층을 가지며 하루 2번 예측정보를 생산한다. 자료기간은 2008년부터 2014년까지 이다. ECMWF_EPS는 이 모델의 저해상도 버전을 사용하여 61개의 앙상블 멤버를 생산하는 앙상블 예측시스템이다. 수평해상도는 32 km이고 2011년부터 자료가 존재한다.

JGSM (Japan Global Spectral Model)은 일본 기상청의 전지구 예측모델로 하루 4번 예측정보를 생산한다. 수평해상도는 18 km이며, 4차원 변분 자료동화 방법을 사용한다(Nakagawa, 2009). TEPS (Typhoon Ensemble Prediction System)는 일본 기상청의 전지구 앙상블 모델로 수평해상도는 56 km이며, 11개의 앙상블 멤버로 구성되어 있다(Yamaguchi and Komori, 2009). TIGGE (The Observing System Research and Predictability Experiment Interactive Grand Global Ensemble)는 세계기상기구(WMO) 산하 세계 날씨 연구 프로그램(World Weather Research Programme, WWRP)에서 인간 생활에 큰 영향(high-impact)을 미치는 날씨예보의 1일에서 15일 예측정확도를 높이기 위하여 운용 중인 앙상블 모델이다. 본 연구에서는 16 km 해상도의 ECMWF_TIGG를 사용한다.

사용된 수치 모델들은 8개의 전구모델, 4개의 앙상블(Ensemble) 모델, 4개의 지역모델로 구성되어 있다. 이 중 컨센서스 기법 연구 활용에는 11개의 모델(UM, UM_R, TWRF, KWRF, JGSM, TEPS, NOGAPS, GFS, ECMWF, KEPS, HWRF)이 이용되었다.

2. 태풍 최적경로 자료

연구지역에 해당하는 북서태평양에서 발생하는 태풍은 도쿄 지역특별기상센터와 미국 합동경보센터와 같은 태풍 예보기관에서 매년 태풍들의 최적경로 자료를 제공하고 있다. 이 중 도쿄 지역특별기상센터는 WMO의 세계 날씨 감시(World Weather Watch, WWW) 프로그램의 일환으로 북서태평양 태풍의 대한 분석과 예보를 하고 있으며, 태풍의 최적경로 또한 제공하고 있다. 이 자료는 연도별 북서태평양 태풍의 태풍 번호와 6시간 간격의 최소중심기압(Minimum Sea Level Pressure, MSLP), 위도, 경도, 그리고 열대저기압의 유형이 정리 되어 있으며 최대풍속 자료는 1977년부터 제공하고 있다.

도쿄 지역특별기상센터에서 제공하는 최대풍속, 중심기압, 그리고 태풍중심 위치자료를 16개의 태풍 수치모델들의 예측성능 평가와 컨센서스 기법에 활용하기 위한 기준 값으로 사용되었다(Fig. 1). 도쿄 지역특별기상센터 최적경로 자료는 최대풍속의 단위를 knot로 제공하고 있지만 대부분의 수치모델들의 단위가 m/s 인 것을 고려하여 단위를 변환하여 수치모델과 비교하였다.

최대풍속의 정의에 있어 도쿄 지역특별기상센터에서는 WMO권고 기준인 10분 평균 최대풍속을 사용하지만 미국 합동태풍경고센터는 1분 평균 풍속을 사용한다. Simiu and Scalon (1978)에 따르면 10분 평균 풍속은 1분 평균 풍속의 약 88%에 해당되며 풍속이 증가할수록 그 차이는 더 커지는 것으로 보고하고 있다. 따라서 미국 합동태풍경고센터 기준을 사용하여 초기장을 생성하고 모델 편차조정을 하는 모델들(GFS, HWRF 등) 중에서 도쿄 지역특별기상센터 자료로 성능 평가를 할 경우 상대적으로 태풍강도를 관측보다 강하게 모의하는 경향이 발견된다. 이와 같은 경향을 보이는 모델들의 최대풍속을 보정하기 위해, 2008년부터 2013년의 도쿄 지역특별기상센터 최대풍속과 미국 합동태풍경고센터 최대풍속과의 관계를 알아보았다. Fig. 2는 이와 같은 관계를 보여주고 있고, 식 (1)과 같이 나타낼 수 있다. 이 결과에 따라 앞서 언급한 경향을 보이는 모델에 대해서는 도쿄 지역특별기상센터와 미국 합동태풍경고센터의 관계식을 이용하여 최대풍속을 보정하여 분석에 이용하였다.

$$MWS_{RSMC-Tokyo} = 1.43 \times MWS_{JTWC} - 9.25 \quad (MWS_{JTWC} > 21.65 \text{ m/s}) \quad (1)$$

식 (1)은 2008년에서 2013년까지 도쿄 지역특별기상센터와 미국 합동태풍경고 센터의 최대풍속 자료를 이용하여 구해졌으며, 21.65 m/s 이하의 적용할 경우 오히려 도쿄 지역특별기상센터 풍속보다 더 낮게 모의하므로 풍속이 21.65 m/s 이상인 경우에만 적용된다.

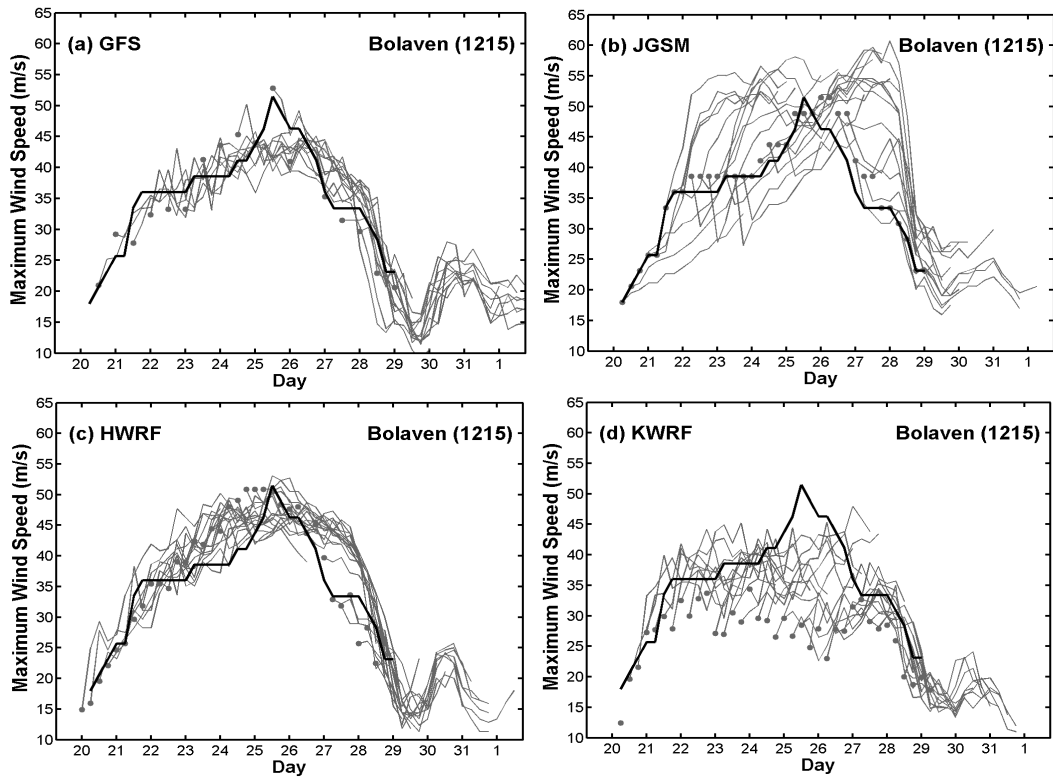


Fig. 1. Model intensity forecast for Typhoon Bolaven (2012). The black line indicates RSMC-Tokyo best-track. Other lines indicate each model output. (a) GFS, (b) JGSM, (c) HWRF, and (d) KWRF.

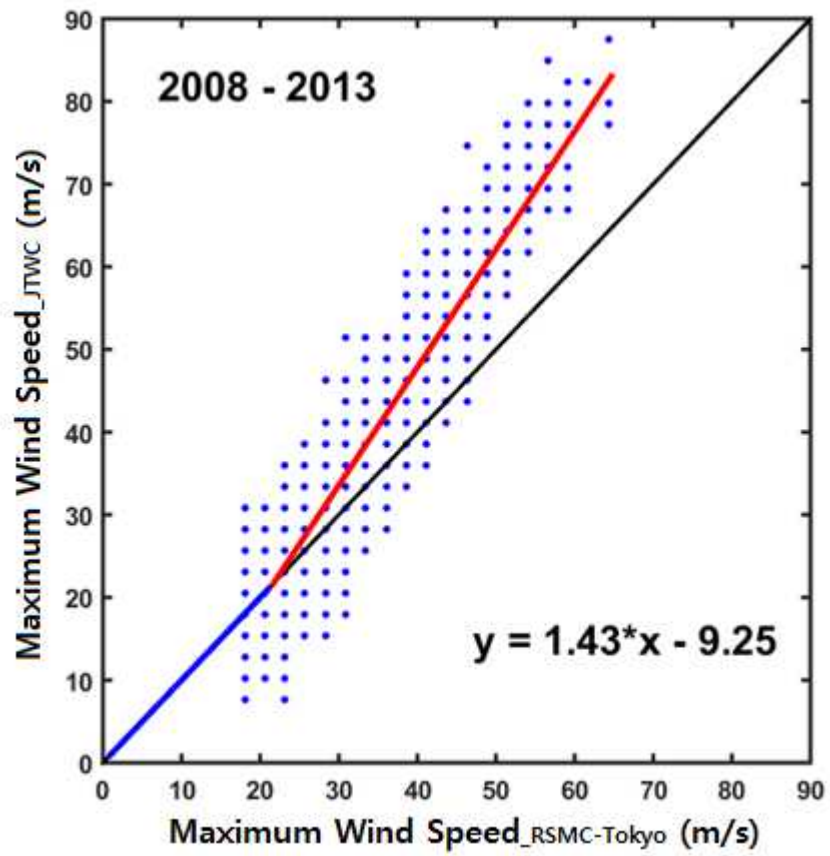


Fig. 2. Maximum wind speed (m/s) comparison between the RSMC-Tokyo and the JTWC from 2008 to 2013.

3. 연구방법

연구에 이용되는 수치모델들 중 대부분의 모델들은 하루 2회 예측자료를 생산하지만, 일부모델들의 경우는 하루 4회 자료를 생성한다. 동일한 조건으로 통계 분석을 하기 위하여 하루 2회 00 UTC와 12 UTC 예측 결과만을 이용하였다. 분석은 2008년부터 2014년까지 기간에 대해 실시되었지만 모든 모델의 자료가 존재하는 2013년부터 2014년에 대해 대부분의 통계분석이 이루어졌고 나머지 시기에는 연도별 오차변화 경향만이 조사되었다.

태풍강도 가이드스 모델들의 예측정확도 분석은 주로 최대풍속(m/s) 자료를 이용하여 이루어졌다. 태풍 강도의 또 다른 지표인 태풍 중심기압(hPa) 그리고 태풍 진로예측 결과는 최대풍속 자료와 비교·분석을 위해서만 제한적으로 사용되었다. 통계분석은 모델과 최적경로자료 사이의 평균절대오차(MAE), 편차(Bias), 그리고 상관계수(Correlation Coefficient)를 이용하여 수행되었다(식 (2)와 (3) 참고).

$$MAE = \frac{1}{n} \sum_{i=1}^n (|F_i - O_i|) \quad (2)$$

$$Bias = \frac{1}{n} \sum_{i=1}^n (F_i - O_i) \quad (3)$$

여기서 F_i 와 O_i 는 모델 값과 관측 값(최적경로 자료)를 각각 의미하며, n 는 자료의 총 개수를 의미한다.

태풍 수치모델들의 강도 예측성능평가에서는 먼저 최근 2년간(2013~2014년) 상기 평가항목을 예보시간별, 연도별, 월별, 태풍 강도별, 급격한 강도 변화 시, 한반도 영향태풍의 경우, 태풍 강도변화(Δ MWS) 별, 이동경로별, 그리고 해역별로 나누 평가하였다. 예보시간별 평가는 모든 모델들이 공통적으로 자료가 존재하는 72시간 (6시간 간격)까지만 이루어졌다. 사용 된 모델은 Table 1에 별표(*)로 표시되어 있는 11개의 모델이다. 모델 성능의 연변화 경향은 모델별로 2008년

부터 2014년까지 자료가 존재하는 기간에 대해 48h 예측 값을 이용하여 실시하였다. 이 분석에서는 강도뿐만 아니라 진로에 대한 성능 평가를 함께 실시하였다. 월별로는 약 70% 이상의 태풍 예측자료들이 8~10월에 속해있기 때문에 상대적으로 표본수가 적었던 4~7월(AMJJ)을 묶고, 표본수가 많았던 8월(Aug), 9월(Sep), 그리고 10~11월(ON)으로 나누어 성능평가를 실시하였다. 태풍 강도등급별 평가는 국가태풍센터에서 사용하고 있는 태풍 강도기준에 따라 약($17 \text{ m/s} \leq \text{MWS} < 25 \text{ m/s}$, TS급), 중($25 \text{ m/s} \leq \text{MWS} < 33 \text{ m/s}$, STS급), 강($33 \text{ m/s} \leq \text{MWS} < 44 \text{ m/s}$, TY급), 매우 강($44 \text{ m/s} \leq \text{MWS}$, STY급)으로 나누어 분석하였다.

Elsberry et al., (2014)는 급격하게 강도 변화를 할 경우를 24시간 동안 30 knot 증가하였을 때로 정의하여 사용하였다. 이에 따라 강도변화 크기에 따른 분석에서는 24시간 동안 강도의 증가 또는 약화가 약 15.43 m/s (30 knot) 이상인 급격한 강화(Rapid Intensification, RI)와 급격한 약화(Rapid Weakening, RW)로 분류하여 분석하였다.

한반도 영향태풍에 대한 분석은 국가태풍센터에서 발간한 2013년과 2014년 태풍분석보고서에 기술된 한반도 영향 태풍 6개(1304, 1315, 1324, 1408, 1411, 1419)에 대해 실시하였다. 태풍 진로별 강도오차 분석은 Kim et al., (2011)에서 제안하였던 Fuzzy c-Means Clustering Method (FCM)을 이용하여 태풍진로를 6개의 그룹으로 분류하였고, 각 그룹별로 평가항목을 분석하였다. FCM 방법은 태풍의 진로와 같이 군집 경계가 애매한 자료에 적합하다고 알려져 있고, 군집 간 태풍의 멤버 사이의 상관계수가 높은 경우로 분류된다. 해석별 오차 분포를 확인하기 위해 북서태평양을 위·경도 10도 간격으로 나누어 평균오차를 확인하였다. 이 방법에 따른 연구 결과는 제 3장 수치모델의 태풍강도 예측성능평가에 서술하였다.

제 4장에서는 제 3장의 결과들을 바탕으로 태풍 강도 예측 컨센서스 기법에 대한 결과를 보여준다. 태풍 예측하는 수치모델은 예측 당시 환경에 따라 태풍 예측 정확성의 변동성이 크고 그만큼 불확실성을 가지고 있다. 이런 무작위적인 편차를 줄이고자 하는 방법이 앙상블 평균 방법이다. 앙상블 평균 방법에는 하나의 수치모델에 다양한 초기조건을 주어 얻은 결과를 평균하는 단일 앙상블 평균

방법 또한 있지만(Lee et al., 2011), 본 연구에서는 다중 앙상블 평균방법에 대해서만 초점을 두고 분석하였고, 다중 앙상블평균 방법인 단순평균방법(AVE)과 가중치 평균방법(WEI)을 모두 이용하여 최적의 예측 값을 얻고자 한다.

Kotal et al., (2011)은 다중회귀방법(Multiple Linear regression)을 이용한 컨센서스 기법을 2008년, 2009년 북인도양에서 발생하는 태풍의 진로 예측에 적용한 바 있다. 이 연구에서는 5개의 역학모델[ECMWF, NCEP, MM5, QLM (the Quasi-Lagrangial model), JMA 모델]을 고정하고 예측시간 별 위, 경도의 가중치를 달리 주어 진로 컨센서스 모델을 개발하였다. 태풍 강도 예측 컨센서스 기법을 개발하기 위하여 가중치 평균방법(WEI)을 이용한 컨센서스 방법으로는 위 연구와 같은 다중선형회귀모델을 사용하였다. 다중 선형회귀 방법은 독립변수의 수가 2개 이상인 경우의 선형회귀모형을 말한다.

$$y = a_0 + \sum_{i=1}^k a_i x_i \quad (4)$$

다중 선형회귀모델은 식 (4)와 같이 표현하고, 태풍 수치모델의 예측 값을 독립변수로서 넣어 최종 예측 값(종속변수)를 얻을 수 있다. 여기서 y 는 예측하고자 하는 값, a_0 는 회귀상수를 의미하며, k 는 사용한 태풍수치모델의 수(단, $2 \leq k$), x 는 태풍 수치모델의 예측 값, a 는 모델들의 가중치를 의미한다. 가중치(weight)와 회귀상수를 산출하는 방법은 최소 절대편차 회귀법(Least Absolute Deviation regression, LAD)을 사용하였다. LAD방법은 전체 기간 동안 절대오차의 합이 최소화가 되게 하는 가중치와 회귀상수를 계산하는 방법으로 Kim et al., (2011)에서 여름철 태풍 계절예측에 사용되었던 방법이다. Kotal et al., (2011)에 연구에 따라 예측시간 별 가중치를 다르게 주는 방법을 사용하였다. 그러나 다른 방법으로 사용하는 모델을 고정하지 않고 예측시간 별로 수치모델들의 순위를 정하여 상위 k 개의 모델을 사용하는 방법을 이용하였다. 순위를 결정하는 방법은 평균절대오차(MAE) 뿐만 아니라, 오차가 크지만 태풍 강도 변화의 경향성을 잘 모의하는 모델들을 고려하여 상관도(R)를 기준으로 하는 2가지 방법을 사용하였다.

최적의 태풍 강도 컨센서스 기법을 개발하기 위하여 연구기간을 훈련기간과 예측기간, 즉 2단계로 나누어 분석하였다. 훈련기간은 2011년부터 2014년까지 각 1년으로 하고 그 이듬해(1년)에 대해서 예측기간으로 두고 분석 하였다. 훈련기간에는 앞서 분석 된 제 3장의 수치모델들의 예측성능 평가를 바탕으로 예측시간 별(12, 24, 36, 48, 60, 72, 84, 96, 108, 120h, 총 10개)로 순위와 가중치를 산출하였으며, 이를 통해 산출 된 순위와 가중치는 단순평균방법(AVE), 가중치 평균 방법(WEI)을 이용한 컨센서스 모델을 만들었다. 훈련기간은 컨센서스 방법을 검증하기 위해 교차타당성 검사(Cross-validation)을 실시하였다. Williford et al. (2003)은 교차타당성 검사를 이용하여 1998년의 TS 예측을 검증하였다. 이는 예측 대상이 되는 자료를 제외한 나머지를 가지고 훈련단계를 거치는 과정이다.

예측시간 별 가장 좋았던 컨센서스 모델을 사용한 컨센서스 기법과 모든 예측시간에 대해 한 가지 컨센서스 모델로 고정하였을 때의 예측성능을 비교하기 위하여, 교차타당성 검사를 통해 예측시간 별로 가장 낮은 오차를 보였던 컨센서스 방법과 사용 모델 개수를 선택하여 새로운 컨센서스 방법인 CON1과 12~72h의 누적 오차(total errors)가 가장 낮았던 컨센서스 방법인 CON2를 제시하여 예측기간에 발생한 태풍에 대해 예측을 실시하고 개별모델과 강도예측성능을 비교하였다. 즉, 최적의 태풍 강도 예측을 위한 컨센서스 기법으로 ‘예측시간 별 가장 낮은 오차를 보였던 컨센서스 모델들의 조합’을 CON1와, ‘72h 누적 오차(total errors)가 가장 낮았던 컨센서스 모델’을 CON2로 제시하고 연구를 진행하였다.

Ⅲ. 수치모델의 태풍 강도 예측성능

최근 2년간(2013~2014년) 예측자료가 존재하는 11개 가이던스 모델의 태풍 강도 예측 성능을 평가하기 위하여 산포도, 상관계수(R), MAE, 그리고 Bias를 구하였다(Fig. 3). 그림에 있는 괄호 안의 값은 각각의 평가요소에 대한 순위를 나타낸다. 각 가이던스 모델별로 성능을 비교해보면, HWRF, GFS, TEPS가 모든 통계 분석에서 3위 안에 드는 성능을 보였으며 Bias에서도 1.37m/s로 상위 2번째를 차지하였다. GFS는 상관도(0.87)와 MAE(4.34 m/s)에서 상위 2번째 그리고 Bias (-1.58 m/s)로 상위 3번째를 차지하여 HWRF 다음으로 좋은 성능을 보였다. TEPS는 Bias 평가에서 0.09 m/s로 가장 Bias가 적었고 상관도와 MAE는 상위 3번째를 차지하여 좋은 예측성능을 보인 모델로 평가된다. 반면, KEPS와 UM은 상관도(0.07)가 매우 낮을 뿐 아니라 MAE (17 m/s 이상)도 높아 두 평가에서 모두 최하위 권에 머물렀다. Bias를 보면 HWRF, JGSM, TEPS, ECMWF, TWRF를 제외한 모델들이 음의 값을 보여 대부분의 모델이 최대풍속을 과소 모의함을 알 수 있었다. 특히, UM과 KEPS 모델의 경우 최대풍속 예측 값이 대부분 30 m/s 이하로 강한 태풍을 전혀 모의하지 못하였다.

예보시간별로 모델의 오차를 분석해보면 대부분의 모델이 예보시간이 길어질수록 최대풍속의 오차가 증가하는 경향이 나타났다(Table 2). 특히, 일본 현업 모델인 TEPS와 JGSM 초기시간에 최대풍속의 MAE가 0.92 m/s로 매우 낮았으나 예보시간이 길어질수록 오차가 크게 증가하여 24시간 후부터는 다른 모델에 비해 오히려 오차가 더 커지는 경향이 나타났다. UM과 KEPS의 경우 초기시간부터 MAE가 약 16 m/s 이상의 큰 오차를 가졌을 뿐 아니라 예보시간이 길어지면서 오차가 더 증가하였다. 반면 NOGAPS와 GFS, HWRF는 전체 예보시간에서 9 m/s 이하의 낮은 MAE를 보여 평균적으로 최대중심풍속을 잘 모의하였다. TEPS와 JGSM은 중심기압에서도 초기시간에 1.52 hPa로 가장 낮은 오차를 보였다. 그러나 48h 이후부터 급격하게 오차가 커지며 상위권에 들지 못 하였다. 특히 72h에는 약 18 hPa 이상의 오차를 보이며 최 하위권에 머물렀다. HWRF는 초기시간(0h)에서 최대중심풍속 오차가 2.68 m/s로 3위를 하였지만 초기시간을

제외한 나머지 예측시간에서는 약 5 m/s 이하의 MAE를 보이며 11개의 모델 중 가장 MWS를 잘 모의하였다. 최소중심기압에서는 예보시간이 길어질수록 오차가 증가하였지만 11개의 모델 중 상위권(1~5위 내)을 유지하였다.

HWRF와 함께 NCEP에서 현업으로 사용하고 있는 GFS 또한 초기시간에는 최대중심풍속 오차가 3.89 m/s의 MAE를 보이며 11개의 모델 중 4번째로 낮은 오차를 보였고, 24시간 이후부터는 HWRF에 뒤를 이어 2번째로 MWS를 잘 모의하였다. 그러나 최소중심기압은 큰 오차를 보이며 하위권에 머물렀다. 특히 초기시간에는 20.37 hPa의 오차를 보이며 11개의 모델 중 10위를 하며 최 하위권에 머물며 중심기압을 잘 모의하지 못 하는 것을 확인할 수 있었다. 기상청 지역 모델인 KWRF는 MWS 평가에서 초기시간에 MAE (8.50 m/s)는 비교적 컸지만 48h 이후에는 3위를 하였다. 최소중심기압 또한 초기시간에는 16.87 hPa의 큰 오차를 보이며 11개의 모델 중 8위를 하였지만, 24h 이후에는 상위권(1~3위 내)을 유지하였다.

ECMWF는 MWS평가에서 모든 예보시간에서 9.91 m/s 이상의 오차를 보이며 하위권(8~11위 내)에 머물렀지만, 중심기압에서는 6위안에 드는 성능을 보여주었고 특히 48h, 72h 에서는 각각 11.43, 8.72 hPa의 MAE로 11개의 모델 중 3번째로 중심기압을 잘 모의하였다. 이와 유사하게 UM_R은 최대중심풍속 평가에서는 전체 예보시간에 5~7위를 하였지만, 중심기압에서는 5위안에 드는 예측성능을 보여주었다. 특히 48h, 72h 에서는 1~2위를 하며 중심기압을 가장 잘 모의하였다.

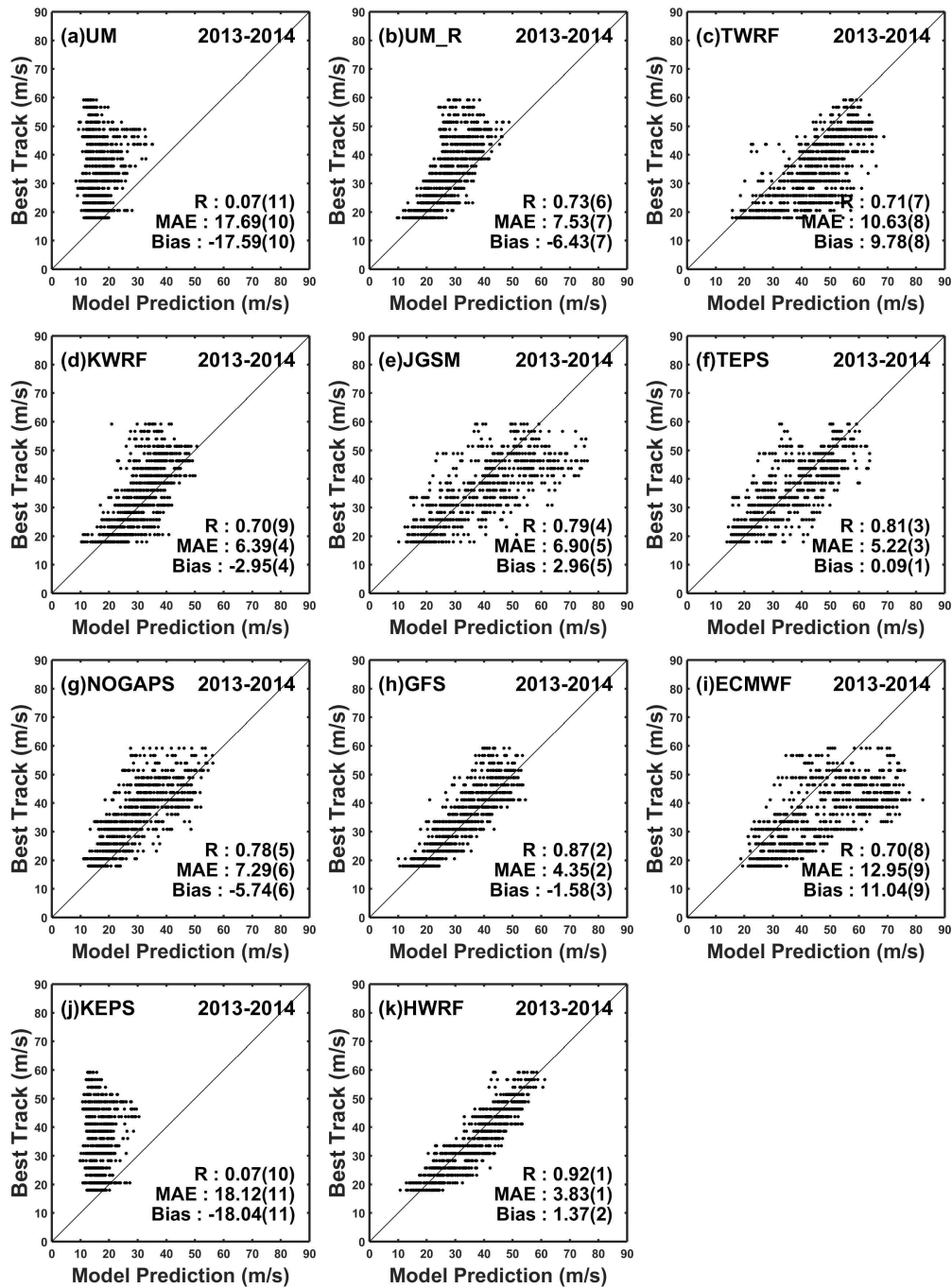


Fig. 3. Scatter plot of maximum wind speed (0~72h) between all models and RSMC-Tokyo best-track from 2013 to 2014. The numbers in the bracket indicate ranking of models associated with verification method.

Table 2. MAE of maximum wind speed and minimum sea level pressure with lead time for all models from 2013 to 2014. The number in bracket is ranking of errors. Shaded boxes represent from 1st model to 3rd model.

Forecast time Model	MWS				MSLP			
	0h	24h	48h	72h	0h	24h	48h	72h
UM	16.46 (10)	17.99 (10)	18.64 (11)	17.51 (9)	15.08 (7)	17.95 (7)	18.05 (8)	13.50 (6)
UM_R	7.68 (5)	8.15 (7)	7.00 (5)	5.32 (5)	12.57 (4)	13.06 (5)	10.75 (2)	7.43 (1)
TWRF	11.62 (9)	11.97 (8)	8.39 (7)	7.19 (6)	20.36 (9)	18.79 (9)	16.75 (7)	14.08 (7)
KWRF	8.50 (6)	7.03 (5)	5.07 (3)	4.92 (3)	16.87 (8)	12.29 (3)	10.20 (1)	8.60 (2)
JGSM	0.92 (1)	6.38 (4)	10.79 (8)	12.28 (8)	1.52 (1)	12.69 (4)	20.71 (10)	23.88 (11)
TEPS	0.92 (1)	5.01 (3)	7.37 (6)	9.32 (7)	1.52 (1)	10.23 (1)	13.95 (5)	18.47 (10)
NOGAPS	8.78 (7)	7.45 (6)	6.21 (4)	4.93 (4)	26.66 (11)	25.07 (11)	20.91 (11)	14.94 (8)
GFS	3.89 (4)	4.51 (2)	4.45 (2)	4.88 (2)	20.37 (10)	18.18 (8)	15.56 (6)	10.78 (4)
ECMWF	9.91 (8)	12.68 (9)	15.39 (9)	17.81 (11)	13.33 (5)	13.18 (6)	11.43 (3)	8.72 (3)
KEPS	-	18.44 (11)	18.60 (10)	17.56 (10)	14.90 (6)	20.11 (10)	20.06 (9)	15.36 (9)
HWRP	2.68 (3)	4.15 (1)	4.28 (1)	4.83 (1)	5.51 (3)	10.27 (2)	11.94 (4)	12.03 (5)

i) 연도 별 예측성능

연도별 가이드스 모델의 성능 변화는 2008년부터 2014년까지 48h 최대풍속, 중심기압 그리고 태풍중심위치에 대한 MAE를 이용하여 조사하였다(Fig. 4). 최대중심풍속은 NOGAPS가 2008년에 13.12 m/s에서 2014년에 3.94 m/s로 MAE가 크게 줄었다(Fig. 4a). GFS 또한 2008년에 12.55 m/s에서 2014년에 3.70 m/s로 MAE가 낮아지며 지속적인 향상이 이루어지고 있는 것을 확인할 수 있었다. 이 두 모델은 중심기압 예측에서도 가장 큰 성능 향상을 보였다(Fig. 4b). KWRF는 2011년부터 2013년까지 가장 낮은 MAE를 보이며 최대풍속은 가장 잘 모의하였

지만, 2014년에는 6.15 m/s의 MAE를 보이며 14개의 모델들 중 7위를 차지하였다. HWRF는 연도별 성능 변화는 크지 않지만 전체 해에서 MAE가 약 5 m/s 이하로 평균적으로 가장 좋은 성능을 보였다. JGSM과 ECMWF는 최근 들어 최대풍속의 오차가 오히려 증가하는 경향을 보였다. 특히, JGSM은 MWS와 MSLP 평가에서 모두 2008년에는 1~2위를 하며 상위권을 차지 하였지만 해가 거듭할수록 오차가 증가하여 2014년에는 14개의 모델 중 하위권(11~14위 내)에 머물렀다. TEPS의 경우 JGSM과 마찬가지로 MWS, MSLP 평가에서 모두 해가 거듭할수록 오차가 증가하는 경향을 보였다.

연도별 가이드선스 모델의 성능 변화는 강도 뿐 아니라 진로에 대한 평가도 함께 조사되었다(Fig. 4c). 진로평가는 태풍진로 자료만 제공하는 DLM과 DBAR이 포함되었다. 분석결과, 태풍진로 예측은 TWRF, DBAR, GFS_EPS 모델을 제외한 대부분의 모델들에서 성능 향상이 이루어졌다. 특히 DLM과 GFS는 2014년에 태풍진로오차가 2008년에 비해 각각 165.1 km와 168.6 km 줄었다. GFS는 2008년에 6개의 모델 중 3위로 비교적 좋은 순위를 보였을 뿐 아니라 2014년에도 가장 낮은 평균진로오차(96.4 km)를 보였다. 특히, GFS는 16개의 모델 중 평균진로 오차가 100 km 안에 드는 유일한 모델이었다. ECMWF는 진로오차가 가장 낮은 모델로 알려져 있지만 2014년에는 GFS, ECMWF_EPS, HWRF 다음으로 좋은 성능을 보였다.

북서태평양에서는 매년 평균 25.6개의 태풍이 발생하며, 이 중 8월에 평균 5.9개로 태풍이 가장 많이 발생하고, 태풍이 주로 영향을 주는 6~10월에는 전체에 약 76%에 해당된다. 태풍모델들의 월별 예측성능을 보기 위해 연구기간에 해당하는 수치모델 예측 결과들은 총 4개의 기간으로 나누어 분석하였다(Table 3). MWS에서 전반적으로 가장 낮은 오차를 보였던 HWRF는 AMJJ, Aug. 기간에서 2위를 하였지만 약 4 m/s 이하의 오차를 보이며 비교적 좋은 예측성능을 보였고 나머지 기간 에서는 11개의 모델 중 1위를 하였다. HWRF 다음으로 좋은 평가를 받았던 GFS 또한 AMJJ, Aug. 기간에서 가장 낮은 오차를 보였고, Sep., ON 기간에서 각각 3, 2위를 하였다. 그러나 대부분의 수치모델들이 4~9월 기간보다 ON 기간에 상대적으로 오차가 더 큰 것을 확인할 수 있었다.

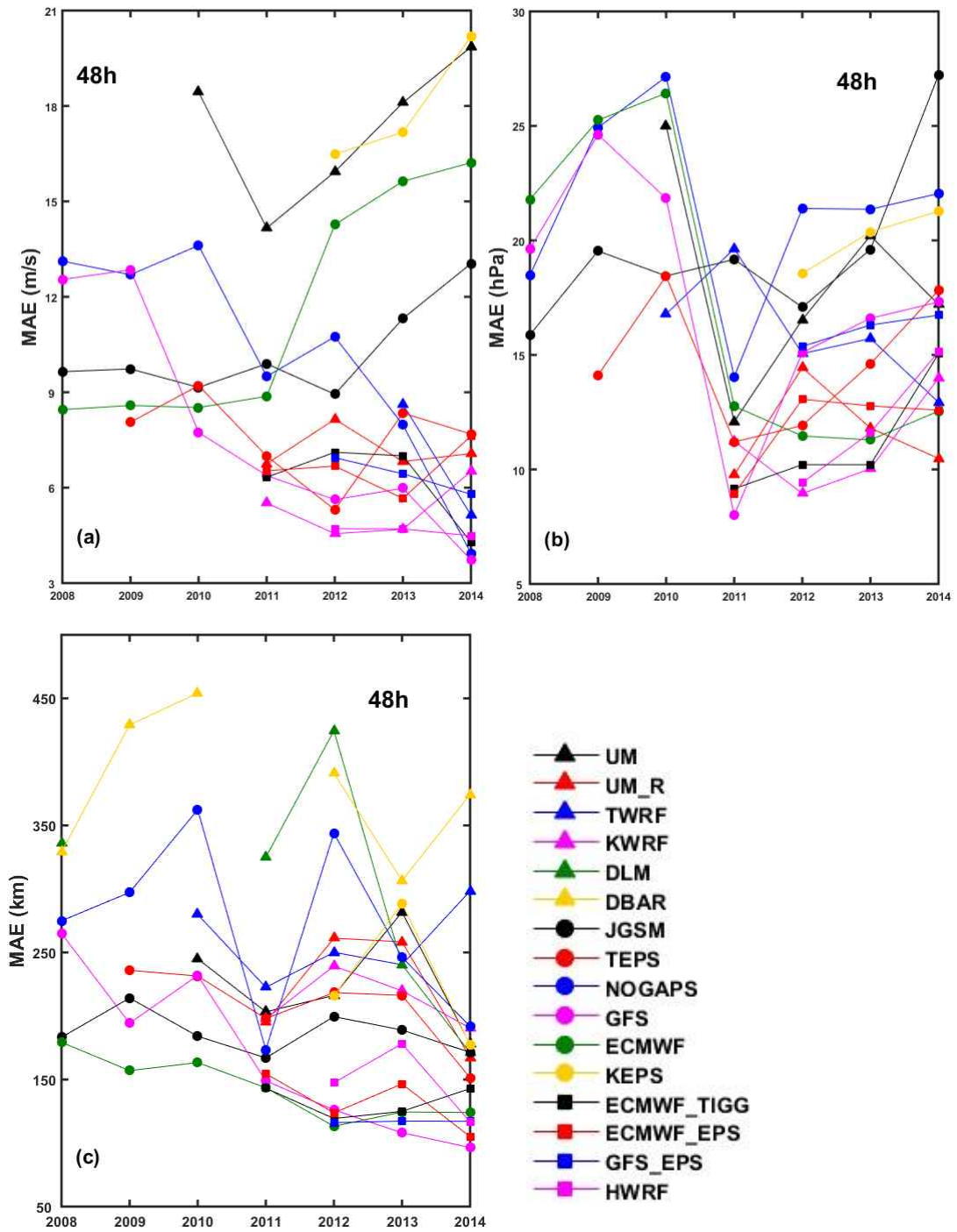


Fig. 4. MAE of 48-h forecast result for all models from 2008 to 2014. (a) Maximum wind speed (m/s), (b) Minimum Sea Level Pressure (hPa), (c) Track (km).

Table 3. MAE of maximum wind speed (0~72h) associated with month for all models from 2013 to 2014. The number in bracket is ranking of errors. Shaded boxes represent from 1st model to 3rd model.

Model	Month			
	AMJJ	Aug.	Sep.	ON
UM	16.21 (10)	15.78 (10)	12.76 (10)	22.88 (10)
UM_R	6.23 (5)	5.06 (6)	7.20 (6)	9.89 (8)
TWRF	11.64 (8)	13.48 (9)	12.51 (9)	7.21 (5)
KWRF	4.40 (3)	4.49 (5)	6.22 (4)	8.61 (7)
JGSM	6.93 (7)	4.29 (4)	7.09 (5)	6.96 (4)
TEPS	5.69 (4)	3.68 (3)	4.65 (2)	5.98 (3)
NOGAPS	6.66 (6)	6.32 (7)	8.21 (7)	7.59 (6)
GFS	3.04 (1)	2.93 (1)	4.90 (3)	5.53 (2)
ECMWF	12.60 (9)	12.42 (8)	9.70 (8)	16.38 (9)
KEPS	16.78 (11)	15.95 (11)	13.23 (11)	23.30 (11)
HWRP	3.28 (2)	3.63 (2)	4.22 (1)	3.91 (1)

ii) 태풍 강도별 예측성능

강도별로 가이드스 모델들의 2013~2014년 강도예측 성능을 MAE의 72시간 누적오차를 이용하여 비교하였다(Table 4). 태풍의 강도 구분은 국가태풍센터에서 발행한 태풍백서에 표시된 태풍 분류 방법을 따랐다. 전반적으로 태풍의 강도가 약할수록 가이드스 모델들의 오차가 작게 나타났으며 강도가 강해질수록 오차가 커졌다. 이것은 가이드스 모델들에게서 전반적으로 나타나는 과소모의 경향 때문에 강한 태풍 일수록 큰 오차가 발생하는 것으로 사료된다. 반면 TWRF

는 오히려 약한 태풍일 때 오차가 적고, 태풍 강도가 강해질수록 오차가 감소하였다. HWRF는 TS 급을 제외한 모든 강도에서 3위 안에 드는 성능을 보였다. 특히 TY, STY 급에서는 11개 모델 중 가장 낮은 오차를 보였으며 모든 카테고리에서 약 5 m/s 이하의 오차를 보였다. HWRF는 모든 모델 중에 가장 높은 공간해상도(3 km)를 가지고 있어 강한 태풍을 모의하는데 더 유리한 것으로 생각된다. HWRF 다음으로 공간해상도(10 km)가 높은 KWRF도 모든 카테고리에서 약 8 m/s 이하의 낮은 MAE를 보였다. GFS는 공간해상도가 높지 않지만(35 km) 모든 카테고리에서 약 6 m/s 이하의 낮은 오차를 보이며 3위안에 드는 성능을 보였으며, 특히 STS 급에서는 약 3.13 m/s의 MAE를 보이며 MWS를 가장 잘 모의하였다. 앞서 최대풍속을 30 m/s 이상을 모의하지 못했던 UM, KEPS는 예상한대로 강도가 강해질수록 오차는 더욱 급격하게 증가하였다.

Table 4. Same as in Table 3 except for category.

Model	Category			
	MWS			
	TS	STS	TY	STY
UM	3.78 (6)	10.01 (9)	13.32 (10)	25.06 (10)
UM_R	3.29 (4)	4.84 (6)	5.53 (5)	10.14 (8)
TWRF	13.75 (11)	15.34 (11)	11.76 (9)	7.97 (5)
KWRF	5.47 (9)	3.81 (5)	4.72 (3)	7.93 (4)
JGSM	2.24 (2)	3.51 (4)	5.90 (6)	9.42 (7)
TEPS	1.57 (1)	3.45 (2)	5.41 (4)	6.64 (3)
NOGAPS	3.49 (5)	6.40 (7)	7.17 (7)	8.57 (6)
GFS	2.73 (3)	3.13 (1)	4.44 (2)	5.09 (2)
ECMWF	10.93 (10)	7.24 (8)	10.87 (8)	15.70 (9)
KEPS	3.97 (8)	10.28 (10)	13.59 (11)	25.41 (11)
HWRF	3.88 (7)	3.47 (3)	4.22 (1)	3.77 (1)

iii) 급격한 강도 변화 시 예측성능

Knaff et al. (2005)는 태풍 강도예측 평가에서 급격하게 강도가 변하는 경우를 모델이 얼마나 잘 예측할 수 있는지를 중요한 판단 기준으로 제시하였다. 본 연구에서도 가이드스 모델들의 급격한 강도변화 예측 성능을 조사하기 위해, 2013~2014년 도쿄 지역특별기상센터 최적경로 자료를 기준으로 태풍이 RI, RW로 분류하여 24시간 강도 누적오차 분석을 실시하였다(Table 5).

KWRF, ECMWF를 제외한 대부분의 가이드스 모델들이 RI보다 RW에서, 즉 급격하게 강도가 약화 될 경우에 오차가 더 낮은 것을 확인할 수 있었다. 이 결과 또한 대부분의 수치모델들이 태풍 강도를 과소모의 하고 있는 것과 관련이 깊다. 태풍 강도가 급격하게 약화되는 경우는 미국 모델인 GFS와 HWRF가 모두 1, 2위를 차지하였다. 태풍이 급격하게 강화되는 경우는 HWRF가 가장 좋은 성능을 나타냈고, 그 뒤로 KWRF와 GFS가 각각 2, 3위의 좋은 성능을 보였다. 특히, HWRF와 GFS는 RI, RW의 모든 경우에 3위 안의 안정적인 성능을 보였다.

iv) 한반도 영향 태풍의 경우

우리나라는 평균적으로 매년 3~4개의 태풍에 직·간접적인 영향을 받는다(Lee et al., 1992). 본 연구에서는 앞서 언급한 우리나라에 영향을 주었던 6개의 태풍에 대해 도쿄 지역특별기상센터 최적경로 자료와 비교하였다. 최근 우리나라에 영향을 주었던 태풍은 1304, 1315 태풍을 제외하고 모두 STY급으로 강도가 강하여 한반도에 많은 피해를 발생시켰다. 특히 2014년에는 우리나라에 MWS가 100 knot 이상의 강한 태풍들만 상륙하였다. 1304, 1315 태풍은 발달기간 내에 최대 풍속과 최소 풍속 차이가 15 knot 이하로 강도변화가 적었던 태풍이다.

한반도에 영향을 주었던 6개의 태풍에 대하여 각 태풍 사례에 대한 모델들의 예측성능과 72시간 누적 통계오차 분석을 하였다(Table 6). 이 태풍 사례에서는 TEPS가 약 2.1 m/s 이하의 오차를 보이면서 1, 2위를 하였고 GFS와 JGSM이 그 뒤를 따랐다. 반면 TWRF는 11개의 모델 중 가장 큰 오차를 보였다. 그러나 강도 변화 및 최대 중심풍속이 가장 컸던 1419 태풍 사례에서는 4.28 m/s의 오차를 보이면서 MWS를 가장 잘 모의하였다. 종합적인 분석에서 가장 좋은 예측

성능을 보여주었던 HWRF와 GFS는 대부분의 태풍 사례에서 4위 안에 드는 순위를 보여주었다. 특히 HWRF는 최대중심풍속이 강했던 태풍사례(1324, 1408, 1411, 1419)에서 1, 2위를 하며 태풍 강도를 가장 잘 모의하였다. HWRF는 72시간 누적 통계오차 분석에서도 3.44 m/s로 가장 MWS를 잘 모의하였다. 그 다음으로 GFS와 NOGAPS가 MAE(3.99, 5.70 m/s)에서 각각 2, 3위를 하며 뒤를 따랐다.

Table 5. Same as in Table 3 but except for rapidly intensity change case.

Model	Case	MWS	
		RW	RI
UM		13.11 (10)	28.08 (10)
UM_R		5.64 (4)	15.14 (9)
TWRF		5.37 (3)	10.20 (5)
KWRF		7.90 (6)	5.81 (2)
JGSM		8.82 (7)	10.65 (6)
TEPS		10.09 (8)	13.43 (8)
NOGAPS		6.56 (5)	13.28 (7)
GFS		3.84 (1)	5.96 (3)
ECMWF		15.95 (11)	8.20 (4)
KEPS		11.83 (9)	29.48 (11)
HWRF		3.88 (2)	4.93 (1)
Average		8.45	13.20

Table 6. Same as in Table 3 but except for approaching korea peninsula case.

Model	TC		MWS					Average
	1304	1315	1324	1408	1411	1419		
UM	4.61 (7)	10.24 (10)	19.27 (10)	18.09 (9)	23.50 (10)	30.00 (10)	20.93 (10)	
UM_R	2.92 (5)	3.98 (6)	9.67 (8)	3.99 (4)	4.16 (4)	16.11 (9)	7.61 (7)	
TWRF	14.60 (11)	15.10 (11)	8.16 (6)	6.20 (6)	10.55 (8)	4.28 (1)	8.30 (8)	
KWRF	7.21 (9)	3.31 (5)	6.19 (3)	3.55 (3)	5.07 (6)	14.25 (7)	7.01 (6)	
JGSM	1.43 (2)	2.63 (3)	6.54 (4)	9.05 (8)	5.08 (7)	8.95 (5)	6.73 (5)	
TEPS	0.96 (1)	2.06 (2)	7.27 (5)	7.70 (7)	4.07 (3)	9.40 (6)	6.30 (4)	
NOGAPS	3.08 (6)	5.98 (7)	10.11 (9)	5.05 (5)	4.95 (5)	5.10 (3)	5.70 (3)	
GFS	2.12 (3)	1.77 (1)	4.81 (2)	2.66 (2)	2.10 (1)	8.04 (4)	3.99 (2)	
ECMWF	8.85 (10)	6.47 (8)	8.50 (7)	18.77 (10)	24.79 (11)	15.10 (8)	16.64 (9)	
KEPS	5.00 (8)	8.99 (9)	19.77 (11)	20.58 (11)	23.21 (9)	30.16 (11)	21.75 (11)	
HWRF	2.40 (4)	3.21 (4)	3.29 (1)	2.31 (1)	3.48 (2)	5.04 (2)	3.44 (1)	

v) 태풍 강도변화(Δ MWS) 별 예측성능

대부분의 모델들은 초기시간부터 큰 오차가 보인다. 그러나 현업예보의 측면에서 보면 초기 값은 이미 알고 있는 상태이기 때문에 초기시간 이후의 강도변화를 잘 예측하는 것이 더 중요하다. Table 7은 초기 값으로부터 태풍의 최대풍속 변화(Δ MWS)에 대한 상관도와 오차를 72시간까지 합산한 결과를 나타낸다.

강도변화 상관성 분석에서는 HWRF가 0.83로 1위, 그 뒤로 TWRF, GFS 순으로 나타났다. HWRF는 MAE(4.58 m/s)도 1위를 하였지만 Bias는 2.79 m/s로 비교적 적은 오차였지만 7위를 하였다. 이는 모든 모델의 강도변화 Bias가 -2.16

~6.45 m/s 오차 범위를 가지며 비교적 낮은 Bias를 보이기 때문에 사료된다. 태풍 강도의 절대값 오차 분석에서 11개 모델 중 3위 안에 드는 성능을 보였던 GFS는 태풍 강도변화에서도 3위안에 들며 태풍 강도 변화를 잘 모의하는 것을 확인할 수 있었다. JGSM의 경우 최대풍속 절대값 평가에서는 좋은 예측 성능을 보였지만, 강도변화에서는 전반적으로 좋지 않은 성능을 보였다. 이 모델의 경우 초기화 과정이 좋아 초기예측 정확도는 높았지만 이후 과정에서 태풍 강도변화를 제대로 모의하지 못하는 것으로 사료된다. TWRP는 태풍 강도 절대 값의 상관도와 오차 분석에서는 7~8위를 하였지만 강도변화 분석에서는 6위안에 포함되었다. 특히 상관도(0.76)와 Bias(-2.16 m/s)분석에서 각각 2, 3위를 하였다. KWRF는 태풍강도 절대 값 분석에서는 MAE, Bias가 4위를 하며 비교적 MWS를 잘 모의 하였지만, MWS의 변화 경향은 11개의 모델 중 10~11위로 하위권에 머물렀다.

vi) 이동경로별

태풍 수치모델들은 태풍을 예측할 때 태풍이 발생한 주변 환경과 이동경로 등에 따라 예측 정확성이 달라지기 때문에 경우에 따라 무작위적으로 오차 분포가 많아지게 된다. 본 연구에서는 1982년부터 2014년 태풍까지의 도쿄 지역특별기상센터 최적경로 자료를 이용하여 태풍의 이동경로를 6개의 군집으로 분류한 뒤 군집 별 태풍 수치모델들의 예측성능을 알아보고자 한다(Fig. 5).

분류하기에 앞서 도쿄 지역특별기상센터 최적경로 자료는 5일(~120h) 간격으로 초기시간을 분류한 후 총 15377 태풍 진로를 가지고 수행하였다. 현재 기상청에서는 동적 데이터베이스를 기반으로 하는 태풍 진로 예측 시스템인 DYTRAP (Dynamic data-base Typhoon Track Prediction)을 사용하고 있다(Lee et al., 2011). 그러나 이 방법은 주관적인 것이 포함되어 있기 때문에 결과가 주관적인 것을 포함하고 있다(Jun et al., 2015). 군집으로 나누어 평가하는 것은 이러한 주관적인 점을 보완하여 객관적인 수치모델들의 성능을 평가할 수 있다. 각 군집 별 특성을 살펴보면, Cluster1 (이하, C1)은 평균진로가 남중국해를 지나 편서풍에 영향을 받지 않고 서진하여 중국 내륙에 상륙하는 특성이 있다(Fig. 5a). C2는 우리나라 및 일본, 중국에 해당하는 동아시아에 상륙하는 태풍들로 case

number가 가장 컸다(Fig. 5b). C3은 주로 일본 동쪽 해안을 따라 태평양을 향해 북동진하는 특성을 보였다(Fig. 5c). C4는 C3과 비슷한 경향을 보이지만 C4에 속하는 태풍의 평균진로가 일본 내륙을 직접적으로 관통하는 특성을 보였다(Fig. 5d). C5는 일본에서 먼 해상에 지나가는 태풍들로 경도 140°E 이상에서 발생하여 북서진하는 특성을 가지고 있다(Fig. 5e). 분류 된 각 군집마다 연구기간의 case number를 고려하여 C5와 C6은 분석에서 제외되었다.

Table 7. Correlation coeff., MAE, and Bias of maximum wind speed (6~72h) from 2013 to 2014. The number in bracket is ranking of errors. Shaded boxes represent from 1st model to 3rd model.

Model	Δ MWS		
	R	MAE	Bias
UM	0.08 (10)	7.67 (8)	2.31 (6)
UM_R	0.47 (6)	6.41 (4)	2.25 (5)
TWRF	0.76 (2)	6.56 (6)	-2.16 (3)
KWRF	0.56 (5)	8.09 (10)	6.45 (11)
JGSM	0.37 (9)	7.68 (9)	3.35 (8)
TEPS	0.59 (4)	5.68 (3)	0.10 (1)
NOGAPS	0.46 (7)	7.43 (7)	4.61 (9)
GFS	0.72 (3)	5.06 (2)	0.25 (2)
ECMWF	0.46 (8)	8.90 (11)	5.82 (10)
KEPS	0.01 (11)	6.48 (5)	2.24 (4)
HWRP	0.83 (1)	4.58 (1)	2.79 (7)

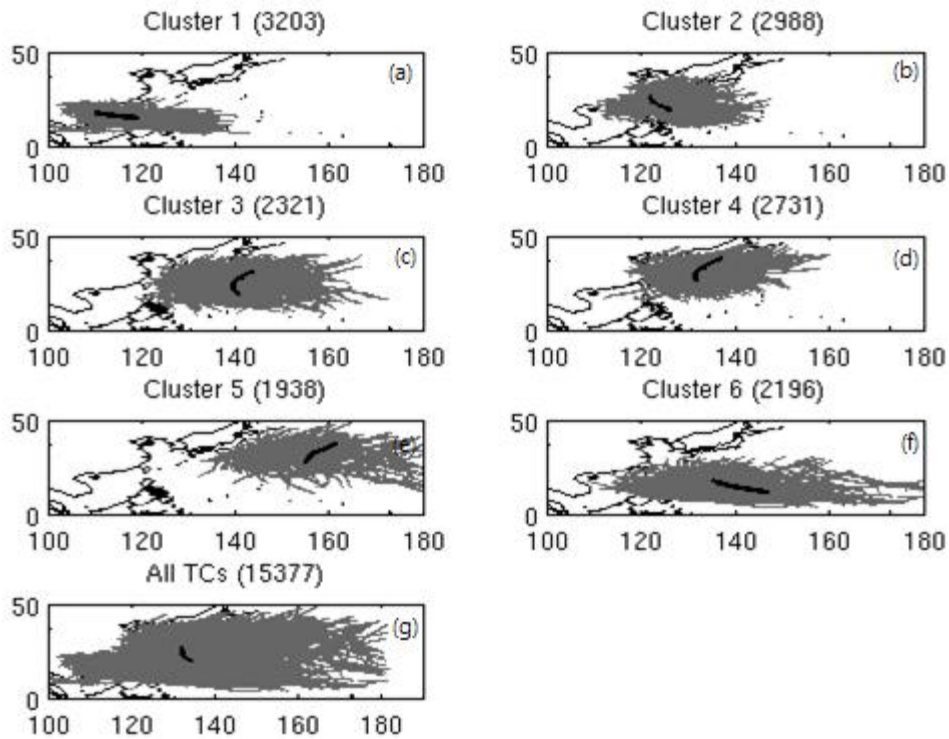


Fig. 5. Six Fuzzy clusters of 15377 TC track case (0~120h) from 1982 to 2014 and all the tracks before using FCM. The thick track are the cluster mean tracks.

대부분의 평가에서도 볼 수 있듯이 HWRF는 모든 군집에서 약 5 m/s 이하의 MAE로 가장 좋은 예측성능을 보였다(Table 8). 그 다음으로 GFS가 모든 군집에서 2위의 성능을 보였다. 일본 기상청 앙상블 모델인 TEPS는 대부분의 태풍이 중국내륙으로 상륙하는 특성을 보인 C1을 제외한 모든 군집에서 11개의 모델 중 3위를 하였다. C1에서도 5.82 m/s의 MAE를 보이며 비교적 낮은 오차를 보였다. C1에서는 HWRF, GFS 다음으로 KWRF가 4.46 m/s로 낮은 오차를 확인할 수 있었다. 태풍 전용모델인 TWRF는 C1에서는 11.75 m/s 오차를 보이면서 11개의 모델 중 가장 큰 오차를 보였다. 그러나 평균 태풍 진로가 일본 해역을 지나는 C3에서 8.37 m/s 오차로 11개의 모델 중 5위를 하며 비교적 높은 순위를

차지하였다. 반면 ECMWF는 C2~C4에서는 약 12 m/s 이상의 오차를 보이면서 8위 이상의 순위를 보였지만, C1 에서는 4위(5.56 m/s)를 하였다.

Table 8. Same as in Table 3 except for cluster group.

Model	MWS			
	C1	C2	C3	C4
UM	11.04 (9)	18.91 (11)	25.88 (11)	16.73 (10)
UM_R	6.13 (6)	8.55 (7)	10.83 (8)	5.79 (4)
TWRF	11.75 (11)	12.93 (9)	8.37 (5)	7.80 (8)
KWRF	4.46 (3)	6.66 (4)	8.88 (6)	6.04 (5)
JGSM	6.72 (7)	6.70 (5)	10.26 (7)	6.07 (6)
TEPS	5.82 (5)	5.30 (3)	6.20 (3)	4.70 (3)
NOGAPS	8.14 (8)	7.75 (6)	8.18 (4)	6.29 (7)
GFS	4.39 (2)	4.59 (2)	6.10 (2)	3.38 (2)
ECMWF	5.56 (4)	12.12 (8)	15.96 (9)	16.68 (9)
KEPS	11.48 (10)	18.86 (10)	25.69 (10)	17.24 (11)
HWRF	4.21 (1)	4.35 (1)	3.72 (1)	3.20 (1)

vii) 공간에 따른 오차분포

태풍수치모델들의 공간적 오차 분포의 경향성을 살펴보기 위해 수치모델들의 48h 예측시간에 대한 공간적 분포를 살펴보았다. 대부분의 모델들이 태풍이 많이 발생하는 위도 10~20°N, 경도 130~140°E 지역에서 MWS에 대한 오차가 컸다 (Fig. 6). 특히, 대부분의 평가에서 오차가 크게 발생하였던 UM과 KEPS는 중위

도 지역을 포함한 대부분의 해양에서 약 20 m/s 이상의 오차를 보였다. JGSM은 중국 내륙인 위도 10~20°N, 경도 100~110°E 지역에서 약 24.18 m/s의 오차로 다른 지역과 비교하여 큰 오차를 보였다. 대부분의 모델들이 저위도 지역에서 오차가 크게 발생한 반면 ECMWF는 위도 20°N 이상의 지역에서 태풍 강도를 잘 예측하지 못 하였다. HWRF와 GFS는 평균적으로 저위도 지역을 제외한 대부분의 공간에서 15 m/s 이하의 오차를 보이며 비교적 중심풍속을 잘 모의하였다.

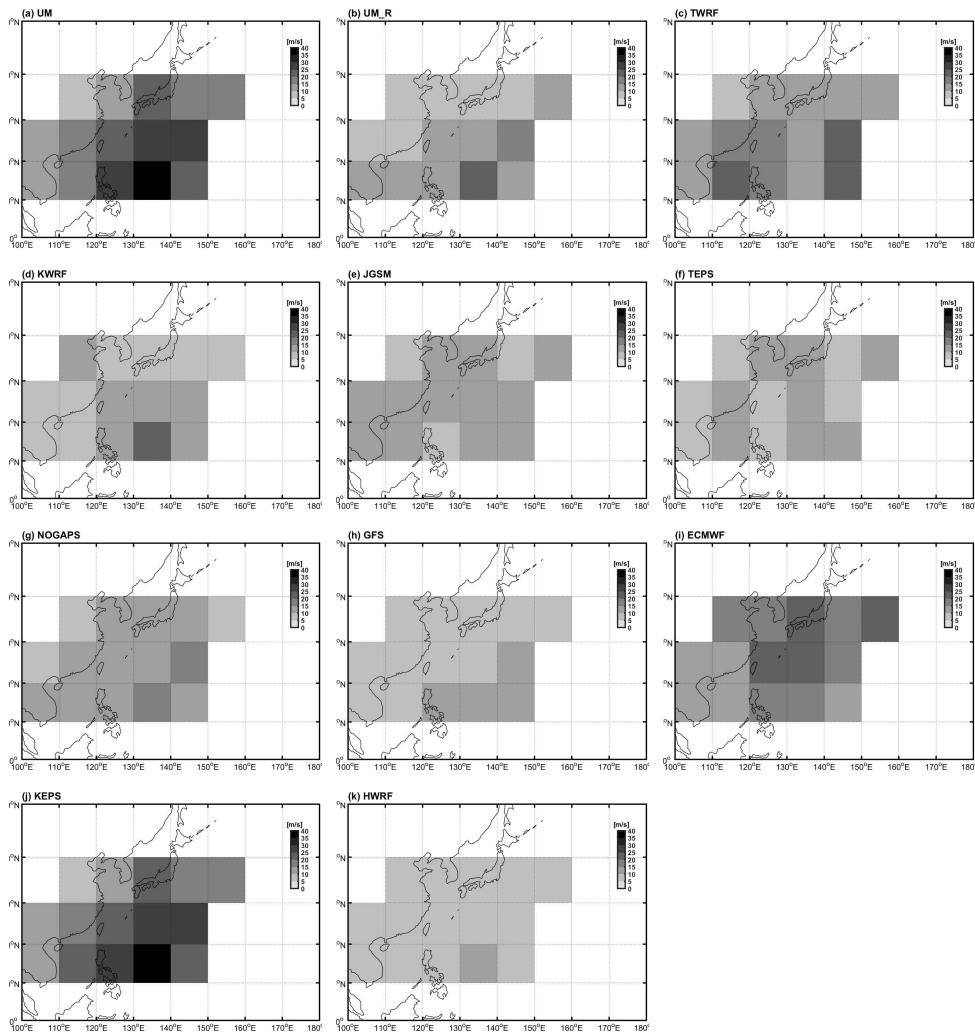


Fig. 6. Geographical distribution of average maximum wind speed (m/s) relative to RSMC-Tokyo best-track within 10 Lat./Long. Boxes.

IV. 태풍 강도 예측 컨센서스 기법

1. 훈련단계

훈련단계에는 교차타당성 검사를 통해 컨센서스 방법에 대해 검증하고 예측단계에서 이용할 컨센서스 기법(CON1, CON2)을 선정한다. 컨센서스 기법인 CON1과 CON2를 제시하기에 앞서, 다양한 컨센서스 기법을 통해 만들어진 모델들이 훈련기간 동안에 총 ‘순위 선정 방법 × 다중앙상블 방법 × 사용할 수 있는 모델 개수(n개, 상위 n위까지의 모델들)’ 만큼 만들어진다. 예를 들면 2011년에는 사용할 수 있는 모델 개수가 TWRF, KEPS, HWRF 모델을 제외 한 총 8개이다. 다중 앙상블은 최소 2개의 모델을 사용하여야 하기 때문에 2~8개(총 7번)의 수치모델을 사용할 수 있다. 그러나 일부 모델에서는 72h 까지 예측하는 경우가 있으므로 72h 이후에는 사용할 수 있는 모델 개수가 한정적이다. 72h 까지 예측하는 모델만을 고려하여 2~5개(총 4번)의 수치모델을 사용할 수 있다. 그러므로 ‘순위 선정 방법(MAE와 상관도) 2가지 × 다중 앙상블방법(AVE, WEI) 2가지 × 사용할 수 있는 모델(2~5개) 4가지’를 사용할 수 있다. 즉, 사용할 모델 개수(n개, 상위 n위까지의 모델들)가 정해진다면 ‘MAE-AVE, MAE-LAD, R-AVE, R-LAD’ 총 4개의 컨센서스 모델이 만들어지며, 사용할 모델 개수(n개)에 따라 n배 만큼 모델이 만들어진다. 이 방법은 각각 ‘순위선정방법 - 다중 앙상블 방법(사용한 모델 개수, 순위)’로 명칭 하였다. 이에 따라 2011년 훈련단계를 통해서는 총 16가지의 컨센서스 모델을 만들 수 있다(Table 9).

Table 9는 이렇게 조합할 수 있는 컨센서스 모델들의 예측시간 별 교차 타당성 검사를 한 결과를 보여준다. 모든 컨센서스 결과들은 72h 누적 상관도를 구하였을 때 0.7 이상의 높은 상관도를 보였으며 예측시간 별로 선택된 컨센서스 모델들은 모든 예보시간에서 개별 모델 중 가장 낮은 오차를 보인 모델(best model)들 보다 오차가 적은 것을 확인할 수 있었다(Table. 9). 음영부분은 각 예측시간 별 가장 낮은 오차(MAE)를 보인 컨센서스를 나타낸다. 여기서 CON1은 초기시

간(12h)부터 72h 예측시간 까지 각 예측시간 별 가장 낮은 오차를 보인 컨센서스 모델들의 또 다른 조합이다[12h : MAE-WEI(3), 24h : MAE-AVE(4), 36h : R-WEI(5), 48h : R-WEI(4), 60h : R-WEI(3), 72h : R-WEI(2)]. 6개의 time step에서 24h 예보시간을 제외하고 모든 예측시간에서 가중치평균방법(WEI)이 사용되었다. 순위결정방법에서는 예측초반에는 절대평균오차에 따른 컨센서스 모델이 선택되었고, 예측후반에서는 상관도를 기준으로 한 컨센서스 모델들이 선택되었다. CON2 역시 개별모델들 중 72h 누적 오차가 5.67 m/s로 가장 적었던 KWRF 보다 낮은 MAE를 보였다. 이렇게 교차타당성 검사를 통해 CON1, CON2를 정하고 나면, 예측을 위한 훈련기간의 수치모델 성능평가를 실시하고 가중치와 회귀상수를 구한다. 여기서 조합 된 각 컨센서스 기법들의 구성 모델들은 Table 10을 통해 확인할 수 있다.

Table 10은 2011년을 훈련기간으로 한 수치모델들의 예측시간별 순위를 나타낸다. (a), (b)는 각각 MAE, 상관도 순위를 나타낸다. 음영으로 나타낸 모델은 CON1의 예측시간 별 선택된 모델들을 나타내고, 두꺼운 선으로 된 박스 안에는 CON2에 사용 된 모델들이다. CON1 모델들을 자세히 보면 초기시간인 12, 24h 예보시간에는 앞선 3장에서의 수치모델들의 성능평가에서 볼 수 있듯이 초기시간에 예측성능이 가장 좋았던 일본 현업모델인 JGSM, TEPS가 우선적으로 선택되었다(Table 10a). JGSM과 TEPS는 초기시간에 예측성능이 다른 수치모델들과 비교하여 매우 좋았지만, 48h 이후로는 오차가 선형으로 증가하는 경향을 보였던 모델이다(Table 2). 이 분석결과에 따라 JGSM과 TEPS가 예측후반에 사용되는 횟수는 줄어들었다. 또한 연도별 변화 경향성을 봤을 때, 가장 향상을 보였던 GFS 모델이 초기시간에 JGSM과 TEPS 다음 순위로 사용되었다.

예측후반에는 최대중심풍속의 변화를 잘 모의 하였던(상관도가 높은) 모델들이 컨센서스 구성모델로 사용되었다. NOGAPS는 최근 예측시간별 평가에서 절대평균오차가 초기시간에는 비교적 높았지만, 예보시간이 길어질수록 오차가 적어지는 경향을 보인 모델이었다(Table 2). 이에 따라 상관도에 따른 순위에서도 비교적 예측후반에 상위권에 머물며 컨센서스 구성모델로 사용된 것을 확인할 수 있었다. CON2는 2011년 훈련기간을 통해 생성 된 컨센서스들 중 72h 누적 오차 (total errors)가 가장 낮았던 컨센서스 모델을 말한다. 2011년에는 CON2로

R-WEI(5)가 선택되었다(Table 9). 컨센서스 모델은 R-WEI(5)모델 하나이지만, 예측시간 별로 구성모델은 다르다. 즉, 각 예측시간 별로 상관도 순위에서 상위 5개의 모델을 사용하게 된다(Table 10b). 아래의 목록은 각각 2012년, 2013년 훈련기간을 통해 선택된 컨센서스 모델들의 목록이다.

2012년 CON1 - [12h : R-WEI(3), 24h : MAE-WEI(7), 36h : R-WEI(7),
48h : MAE-WEI(7), 60h : MAE-WEI(7), 72h : R-WEI(3)]
CON2 - MAE-WEI(5)
2013년 CON1 - [12h : MAE-AVE(3), 24h : MAE-AVE(3), 36h : R-WEI(4),
48h : MAE-AVE(2), 60h : MAE-WEI(5), 72h : R-AVE(4)]
CON2 - MAE-AVE(3)

2012년에는 2011년과 유사하게 모든 예측시간에서 가중치평균방법(WEI)이 사용되었다. 다만 12h, 36h, 72h 에서는 상관도(R)를, 24h, 48h, 60h에서는 MAE를 기준으로 한 것이 다르다. 또한 사용한 모델 개수는 12h, 72h를 제외한 모든 예측시간에서 7개의 수치모델을 모두 사용하였다. 이는 2011년과 비교하여 2012년에 수치모델들이 전체적인 오차가 줄었을 뿐 아니라, 앞선 분석에서 평균 예측 성능이 좋았던 HWRF와 같은 수치모델이 포함되었기 때문으로 사료된다. 2013년에는 2011, 2012년과는 반대로 60h 제외한 모든 예측시간에서 단순평균방법(AVE)이 사용되었다. 2013년에 구성된 모델들은 대부분의 예측시간에서 제 3장에서 좋은 예측성능을 보였던 HWRF, GFS, TEPS가 선택되었다. 이는 2013년에는 여러 태풍수치모델들에게 가중치를 주어 평균하는 방법 보다는, 모델 개수는 적지만 오차가 적었던 수치모델들을 단순 평균하는 방법이 더 2013년 태풍 강도를 잘 모의한 것으로 보여진다.

위와 같은 방법으로 훈련기간을 각 1년씩 하여 2011년부터 2013년까지 총 3년에 대하여 분석을 하였다. 훈련단계에 선택된 CON1과 CON2는 다음 절에서 훈련기간의 이듬해인 2012~2014년 예측을 실시하고, 그 결과를 개별모델들과 함께 강도 예측성능 평가를 하여 새로 개발된 컨센서스 기법의 강도 예측 성능을 확인할 것이다.

Table 9. Errors of each consensus scheme with forecast time in training period (2011). Shaded boxes indicates selecting model.

Forecast Time		12h	24h	36h	48h	60h	72h	84h	96h	108h	120h	Total (~72h)	
Best individual model		3.28	4.63	4.55	5.50	6.27	6.56	6.50	6.94	6.37	5.88	5.67	
Consensus model		TEPS	TEPS	KWRF	KWRF	GFS	GFS	GFS	GFS	GFS	GFS	KWRF	
Case		116	102	63	54	49	41	37	33	28	21	425	
2	M	AVE	3.31	4.71	5.07	5.18	5.71	6.53	8.82	6.31	7.13	6.16	4.73
	A	Weight	3.41	4.40	4.95	5.86	6.77	5.70	5.14	7.86	7.12	7.14	4.80
	E	AVE	3.31	4.71	5.86	4.94	5.05	4.42	11.81	10.05	7.99	8.29	4.54
	R	Weight	3.41	4.40	4.14	4.67	4.77	4.30	5.61	4.79	5.43	5.90	4.16
3	M	AVE	3.06	4.03	4.47	6.11	5.60	5.49	5.59	5.64	5.73	4.82	4.49
	A	Weight	2.96	4.12	4.95	5.55	6.54	6.00	4.87	5.58	9.12	7.11	4.66
	E	AVE	3.96	4.40	4.20	5.40	6.07	4.82	6.81	5.33	6.20	6.59	4.64
	R	Weight	3.07	4.22	4.11	4.74	4.43	4.75	6.61	4.30	4.90	5.50	4.07
4	M	AVE	3.27	4.01	4.80	5.72	6.07	6.26	5.54	6.16	6.33	6.32	4.71
	A	Weight	3.10	4.27	4.28	6.04	6.91	6.06	4.71	4.65	8.49	6.51	4.77
	E	AVE	4.15	4.63	4.61	5.43	5.98	5.78	5.23	5.39	4.95	5.20	4.92
	R	Weight	3.43	4.58	3.80	4.66	4.56	4.66	5.78	5.94	4.46	5.32	4.18
5	M	AVE	3.79	4.28	5.30	5.99	6.63	5.56	6.90	5.08	5.02	5.02	5.00
	A	Weight	3.27	4.76	4.12	6.78	6.80	4.31	5.31	6.76	4.61	7.46	4.78
	E	AVE	4.31	4.45	4.67	5.74	5.68	6.29	4.81	5.08	5.02	5.02	4.99
	R	Weight	3.24	4.16	3.36	4.92	5.06	4.55	5.14	6.76	4.61	7.46	4.07
CON1		MAE-WEI (3)	MAE-AVE (4)	R-WEI (5)	R-WEI (4)	R-WEI (3)	R-WEI (2)						
CON2												R-WEI (5)	

Table 10. Ranking of individual models with forecast time for training period (2011). Based on (a) MAE, (b) Corr., Shaded area shows the cases that CON1 scheme were used. The models inside thick line box were applied by the CON2 scheme.

(a)

M A E	12h	24h	36h	48h	60h	72h
1	TEPS	TEPS	KWRF	KWRF	GFS	GFS
2	JGSM	JGSM	GFS	GFS	KWRF	KWRF
3	GFS	KWRF	TEPS	UM_R	UM_R	UM_R
4	KWRF	GFS	UM_R	TEPS	TEPS	TEPS
5	ECMWF	UM_R	ECMWF	ECMWF	ECMWF	NOGAPS

(b)

R	12h	24h	36h	48h	60h	72h
1	TEPS	JGSM	KWRF	NOGAPS	NOGAPS	NOGAPS
2	JGSM	TEPS	NOGAPS	KWRF	KWRF	KWRF
3	NOGAPS	NOGAPS	JGSM	JGSM	JGSM	UM_R
4	KWRF	UM_R	TEPS	UM_R	UM_R	JGSM
5	GFS	KWRF	UM_R	TEPS	GFS	ECMWF

2. 예측단계

훈련단계에서 교차타당성 검사 후 선택된 CON1과 CON2는 예측기간에 해당하는 2012~2014년까지 총 3년간의 예측을 실행한 후 도쿄 지역특별기상센터 최적 경로 자료와 비교하였다(Fig. 7). 대부분의 수치모델들의 경향성과 유사하게 CON1과 CON2 모두 예보시간이 길어질수록 오차가 증가하는 경향을 보였다. 2011년을 훈련단계로 두고 2012년을 예측하였을 때는 예측시간 별로 다른 컨센서스를 조합한 CON1이 72h 누적 오차(total errors)가 가장 낮았던 컨센서스인 CON2 보다 약 0.34 m/s 낮은 오차를 보였다. 이는 CON1이 초기시간(12~48h)에 최대풍속을 더 잘 모의하였기 때문으로 여겨진다. 컨센서스 기법의 예측 성능은 훈련단계에서의 수치모델들의 순위와 예측 단계에서의 순위가 유사할 경우 더 최대풍속을 잘 모의 할 것으로 여겨진다. 이는 예측단계에서는 훈련단계의 순위 결과를 사용하지만 실제 들어가는 자료는 예측시간의 초기 자료이기 때문에, 훈련단계에서 상위에 머물렀던 모델이 예측단계에서도 상위에 머물 경우 더 좋은 예측성능을 보여줄 수 있다. Table 10은 2011년 훈련기간에 수치모델들의 예측시간 별 순위를 볼 수 있고, Table 11은 예측기간인 2012년 수치모델들의 순위를 보여준다. 즉, Table 10과 Table 11의 순위가 유사하다면 컨센서스 기법의 결과도 효과적일 것으로 사료된다. 2011년 훈련단계에서 선택된 컨센서스 모델들을 살펴보면 CON1에 12, 24h 예측시간은 MAE에 따라 순위가 결정된 모델들이다. 이때 2011년에 모델들의 순위와 2012년에 순위의 구성은 TEPS, JGSM, GFS로 유사하였다. 반면 CON2는 상관도를 기준으로 한 R-WEI(5) 모델이었다. 상관도는 MAE에 비해 2011년과 2012년에 순위 변동이 큰 것을 볼 수 있었다(Table 10b, 11b).

NOGAPS는 2011년 훈련기간에는 상관도 순위가 상위에 위치하며 모든 예측시간에 구성 모델로서 참여하였다. 그러나 2012년 예측기간에는 4~5위에 위치하며 2012년 태풍에 대한 예측성능은 훈련기간에 비해 떨어졌다. 특히 48, 72h에는 순위 안에 포함되지도 못하였다. 2012년 예측기간에 CON1은 72h를 제외한 모든 예측시간에서 개별 모델(best model) 보다 낮은 오차를 보였다(Fig. 7a). CON2도

60, 72h를 제외한 모든 예측시간이 개별모델들 보다 오차가 적었다. 72h 누적 오차 또한 CON1, CON2는 개별 모델 중 가장 낮은 오차를 보였던 KWRF에 비해 약 1.57 m/s 낮은 오차를 보이며 개별 모델들 보다 최대중심풍속을 잘 모의함을 보였다.

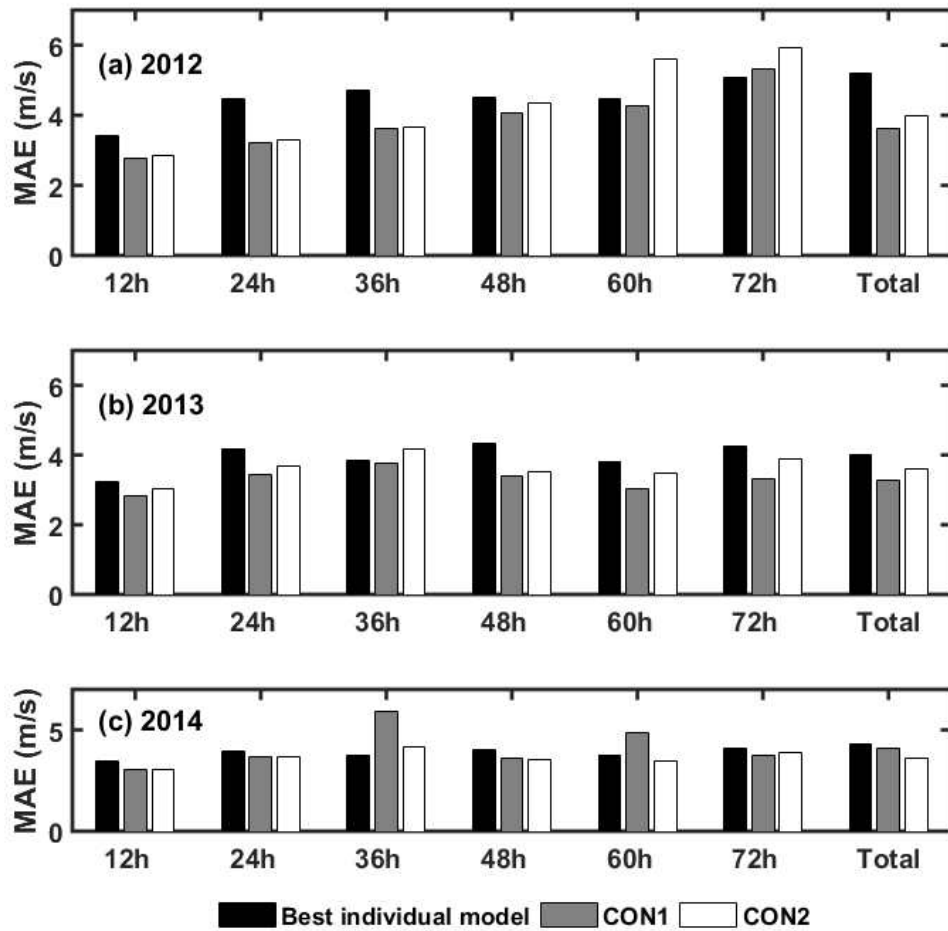


Fig. 7. Errors with forecast time and total errors (12~72h) for best individual model and CON1, CON2 from 2012 to 2014.

Table 11. Ranking of individual models with forecast time for forecast period (2012). Based on (a) MAE, (b) Corr.,

(a)

MAE	12h	24h	36h	48h	60h	72h
1	TEPS	GFS	GFS	KWRF	KWRF	KWRF
2	JGSM	TEPS	KWRF	GFS	GFS	GFS
3	GFS	JGSM	TEPS	TEPS	TEPS	UM_R
4	KWRF	KWRF	JGSM	JGSM	UM_R	TEPS
5	ECMWF	UM_R	UM_R	NOGAPS	NOGAPS	NOGAPS

(b)

R	12h	24h	36h	48h	60h	72h
1	JGSM	JGSM	JGSM	KWRF	KWRF	KWRF
2	TEPS	TEPS	GFS	JGSM	JGSM	JGSM
3	GFS	GFS	TEPS	GFS	TEPS	TEPS
4	ECMWF	NOGAPS	ECMWF	TEPS	GFS	UM_R
5	NOGAPS	ECMWF	NOGAPS	ECMWF	NOGAPS	GFS

2013년과 2014년 예측에서도 CON1과 CON2의 72h 누적 예측 오차는 2012년 예측기간과 마찬가지로 개별모델(best model) 보다 낮았다. 2012년 예측기간과 달리 2013년 예측에는 개별모델들의 성능과 순위가 2012년 훈련기간에 순위와 유사한 특징을 보였으며 오차도 2012년 예측기간에 비해 줄어든 것을 볼 수 있었다(Fig. 7b). 이는 2012년 예측 결과에서도 볼 수 있듯이 훈련기간에 컨센서스에 선정된 개별 모델들이 예측기간에도 그 성능을 유지하였느냐가 컨센서스의 성능에 영향을 주는 것으로 보여진다. 2013년에는 CON1(3.26 m/s)과 CON2(3.25 m/s)의 72h 누적 오차는 0.01 m/s 차이로 그 차이가 미미하였다. 예측시간 별

평가에서도 CON1은 모든 예측시간에서 개별모델(best model)의 MAE보다 낮았다.

2014년에는 12, 24h 예측시간에는 CON1과 CON2에 구성모델과 컨센서스 방법이 MAE-AVE(3)으로 동일하였다. 2014년 예측기간에도 2012년 예측기간과 유사하게 예측시간이 길어졌을 때 개별모델(best model)보다 오차가 더 큰 경우를 볼 수 있었다(Fig. 7c). 특히, 36h 예측시간에는 CON1, CON2 모두 개별 모델 중 가장 낮은 오차를 보인 NOGAPS보다 약 0.47~2.23 m/s 만큼 오차가 더 컸다. NOGAPS는 연변화 경향을 봤을 때 GFS와 함께 가장 향상되었던 모델이다(Fig. 4). 이때 구성 모델을 살펴보면, 2013년 훈련기간을 통해 선정된 CON1은 상관도를 기준으로 상위 4개의 모델로 구성된 가중치평균 컨센서스 이다[R-WEI(4)]. 이 컨센서스에는 순서대로 HWRF, GFS, NOGAPS, TEPS가 사용되었다. 그러나 2014년에 36h 예측시간에 상관도가 가장 높은 모델은 NOGAPS 였고, 그 뒤로 HWRF, GFS, TWRF 순이었다. TWRF는 2013년부터 최대중심풍속 자료를 제공한 모델로서 2013년 훈련단계에 처음 사용되었다. 60h 예측시간에서도 유사한 경향을 볼 수 있었는데 이는 2013년에는 하위권에 머물렀던 NOGAPS와 TWRF의 영향으로 보여진다.

3. 예측성능 검증

본 연구에서 개발 한 CON1과 CON2의 지난 2년간(2013~2014)에 예측성능을 살펴보고 개별모델과 비교하기 위하여 72시간까지 누적한 산포도, 상관계수(R), 그리고 MAE, Bias를 구하였다(Fig. 8). Fig. 8은 앞선 수치모델의 성능평가에서 Fig. 3과 다르게 12h 예보시간부터 12시간 간격으로 72h 까지 만을 누적하여 비교하였다. Fig. 8에 표시 된 (a)-(j)는 2년간 공통으로 존재하는 10개의 태풍 강도 수치모델의 산포도를 나타내며 (k), (l)은 각각 CON1과 CON2를 나타낸다. 개별 모델 중에서는 제 3장 수치모델들의 성능평가에서 볼 수 있듯이 HWRF가 상관도(0.90), MAE (4.28 m/s)가 가장 좋았고, Bias는 1.90 m/s로 세 번째로 작았다. 컨센서스 모델인 CON1과 CON2를 보면 개별 모델들 보다 상관도, MAE, Bias 오차에서 높은 순위를 차지하는 것을 확인할 수 있었다. 특히, 72h 누적 오차(total errors)가 가장 낮은 컨센서스 모델을 사용한 CON2는 상관도(0.91), MAE(3.50 m/s)로 개별모델들을 포함하여 가장 태풍 중심풍속을 잘 모의하였다. Bias에서도 -0.12로 거의 0에 가까운 Bias를 보였다. 예측시간 별로 가장 낮은 오차를 보인 컨센서스 모델들의 조합인 CON1 또한 CON2 다음으로 좋은 결과를 보여주었다.

앞 절에서 확인하였듯이 훈련기간에 이용된 개별모델들의 성능에 따라 예측단계에서 예측결과가 달라질 수 있음을 확인하였다. 이를 더 자세히 살펴보기 위해 본 연구에서는 예측기간을 2013년으로 하고, 훈련기간을 2년(2011~2012년) 동안 하여 훈련기간이 1년(2012년)인 컨센서스와 비교를 하였다. 또한 2013년뿐만 아니라 2014년에 대한 예측결과도 비교하였다. 컨센서스 결과 비교는 앞 절에 결과에 따라 2013~2014년 절대 평균 오차(MAE)가 낮고, 도쿄 지역특별기상센터 최적경로와의 상관도가 가장 높았던 CON2에 대해서 하였다.

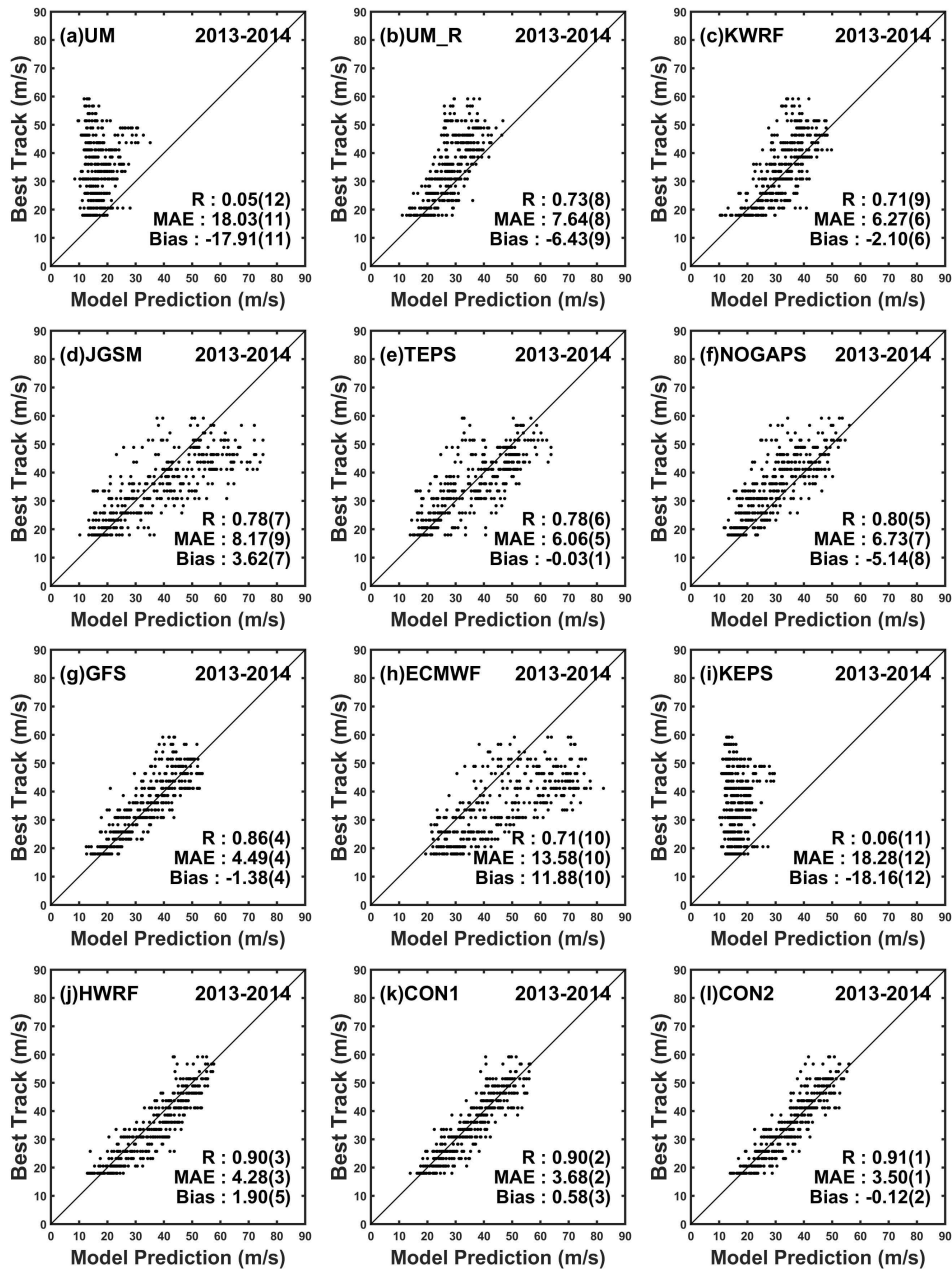


Fig. 8. Scatter plot of maximum wind speed (12–72h) between all models and RSMC–Tokyo best-track from 2013 to 2014. (a)–(l) all individual models, (k) CON1, (l) CON2. The number in the bracket indicate ranking of models associated with verification method.

2013년에 대한 예측에서는 훈련기간과 상관없이 개별모델(best model) 보다 모든 예측시간에서 오차가 적었다(Table 12). 그러나 훈련기간에 따른 CON2를 비교하였을 때는 모든 예측시간에서 가장 최근 1년(2012년)을 훈련기간으로 사용한 모델이 훈련기간을 2년 동안 실행한 모델보다 적게는 0.04 m/s, 크게는 2.07 m/s 만큼 오차가 적었다. 2014년에 대한 컨센서스 결과를 비교해보면, 2013년과 마찬가지로 72h 누적 오차는 최근 1년을 훈련기간으로 사용한 모델이 훈련기간을 2년으로 한 모델 보다 낮은 오차를 보였다(Table 13).

Table 12. Compare with errors associated with training period for 2013 forecast.

CON2	12h	24h	36h	48h	60h	72h	Total
model	3.23	4.19	3.83	4.33	3.82	4.24	4.01
Training period							
2012	2.73	3.59	3.46	3.28	3.15	3.44	3.25
2011~2012	3.06	3.63	4.36	4.18	5.22	4.56	3.96

Table 13. Same as in Table 12 except for 2014.

CON2	12h	24h	36h	48h	60h	72h	Total
model	3.49	3.94	3.71	4.04	3.77	4.11	4.30
Training period							
2013	3.02	3.69	4.18	3.52	3.49	3.88	3.60
2012~2013	3.32	3.48	3.61	3.77	3.83	4.26	3.64

이는 2011년에는 이용 가능한 모델 개수가 8개로 한정되어 있고, 2012년에 들어서면서 KEPS와 HWRF 모델의 예측 값이 추가되었으며 2013년에는 TWRF의 예측 결과도 포함할 수 있었다. 훈련기간을 2년으로 늘리면서 분석에 사용할 수 있는 모델이 한정되고, 최근 경향(2013~2014년)에 상관도와 MAE에서 가장 좋은 예측성능을 보였던 HWRF(Fig. 3)와 같은 모델들이 포함되지 않아 오차가 더

커진 것으로 보여진다. 또한 2011년에도 예측자료를 제공하던 모델들 중 NOGAPS의 경우 최근들어 최대중심풍속의 오차가 많이 향상되어 상위권을 차지한 모델이다(Fig. 4). 이는 훈련기간을 길게 하였을 때 이용 가능한 모델 개수와 개별모델들의 예측 성능 향상 등의 반영이 미흡하기 때문에 오히려 오차가 증가할 수 있음을 시사하고 있다. 그러나 이는 연구기간을 4년으로 하였기 때문에 훈련기간에 의한 컨센서스 기법의 성능을 더 자세히 파악하기 위해서는 향후 연구기간을 더 두고 살펴봐야 할 것으로 사료된다.

훈련기간에 따른 컨센서스 기법의 예측 성능을 더 자세히 분석하기 위하여 훈련기간을 2011년으로 고정하여 2014년에 대한 예측을 실시하였다. 또한 2011년뿐만 아니라 2012, 2013년을 각각 훈련하여 2014년 태풍을 예측하였다. 이는 컨센서스를 매해 변경해야 하는 필요성 여부와 개별모델들의 예측 성능 향상에 따른 결과를 확인할 수 있다. Table 14는 훈련기간을 각각 2011, 2012년으로 하고, 2013년에 대한 예측을 실시한 결과이다. 예측시간별, 72h 누적 오차 모두 2011년을 훈련기간으로 사용한 모델이 최근(2012년) 자료를 훈련기간으로 이용한 모델보다 오차가 더 크게 나타나는 것을 확인할 수 있었다. 특히, 60h 예측시간에서는 약 8.06 m/s 이상의 큰 오차를 보였다. 이는 앞 절에서도 확인할 수 있었던 2012년부터 잠재적 구성모델로서 포함 된 HWRF(최근 개별모델들 중 최대풍속 오차가 가장 적었던 모델)의 영향으로 사료된다. 2014년에 대한 예측 결과를 보면, 2013년과 마찬가지로 최근(2013년) 기간과 멀어질수록 오차가 커지는 경향을 확인할 수 있었다(Table 15). 이는 최근 개별모델의 예측성능을 많이 반영한다면 더 좋은 예측 결과를 도출해낼 수 있음을 시사한다.

Jeffries and Fukada (2002)는 진로 컨센서스에 대해 4개 또는 5개 모델을 사용한 컨센서스가 단 2개 또는 3개의 모델을 사용한 컨센서스 보다 크게 오차가 감소함을 보인 적이 있다. 본 연구에서 또한 최적의 모델 개수를 찾기 위하여 2013~2014년 48h, 72h 예보시간에 4가지 컨센서스(MAE-AVE, MAE-WEI, R-AVE, R-WEI)에 모델 개수에 따른 평균오차를 비교해보았다(Fig. 9, 10).

48h 예측오차를 보면 선행연구와 마찬가지로 모델 개수가 2개(가장 적음), 7개(가장 많음)를 사용할 때의 변동성(spread)이 컸고, 모델 개수를 3~6개를 사용하였을 때 안정적인(stable) 예측 결과를 확인할 수 있었다(Fig. 9). 비교에 사용한

4가지 컨센서스의 평균값은 모델 개수를 상위 4개를 사용하였을 때 4.10 m/s 오차로 가장 낮았다.

72h 예측시간에서도 모델 개수를 4~5개를 사용하였을 때 제일 안정적인 결과를 볼 수 있었다(Fig. 10). 그러나 그 보다 더 적거나, 많은 모델 개수를 사용하였을 때는 48h 예측시간에서 보다 MAE의 변동성(spread)이 커졌다. 48h 결과와 마찬가지로 72h 에서도 4가지 컨센서스의 평균값은 모델 개수를 상위 4개의 모델을 사용하였을 때 5.03 m/s로 가장 낮았으며 48h 예측오차 보다 변동성이 매우 조밀하였다.

Table 14. Compare with errors associated with fixed training period for 2013 forecast.

CON2	12h	24h	36h	48h	60h	72h	Total
model	3.23	4.19	3.83	4.33	3.82	4.24	4.01
Training period							
2011	3.26	4.27	5.75	5.39	8.06	6.49	5.00
2012	2.73	3.59	3.46	3.28	3.15	3.44	3.25

Table 15. Same as in Table 14 except for 2014.

CON2	12h	24h	36h	48h	60h	72h	Total
model	3.49	3.94	3.71	4.04	3.77	4.11	4.30
Training period							
2011	5.52	5.32	6.98	11.83	13.12	9.37	7.53
2012	3.67	3.55	3.58	3.78	4.07	4.08	3.75
2013	3.02	3.69	4.18	3.52	3.49	3.88	3.60

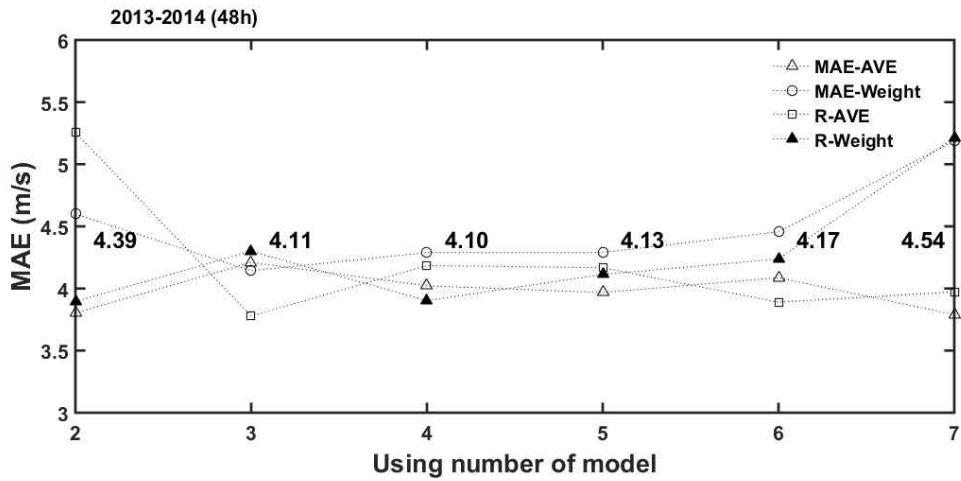


Fig. 9. 48h forecast MAE (m/s) of 4 consensus schemes associated with using number of models from 2013 to 2014. The number in figure indicates mean error of 4 consensus schemes.

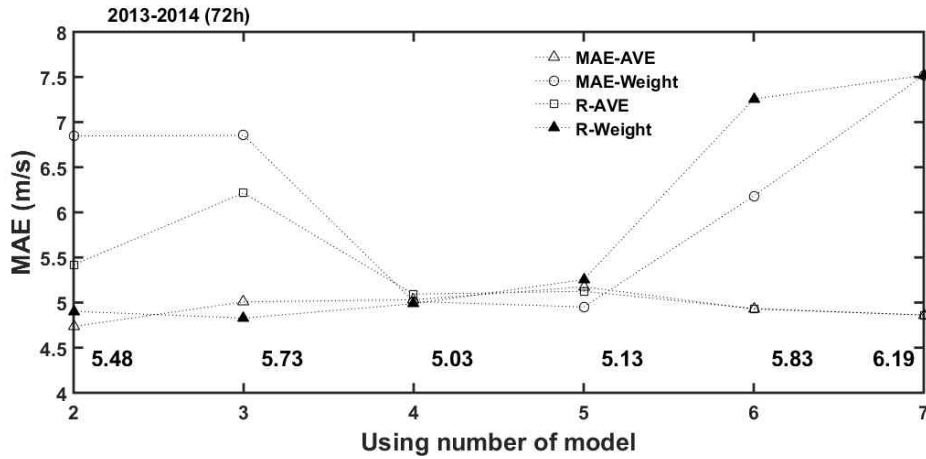


Fig. 10. Same as in Fig. 9 except for 72h forecast MAE (m/s).

앞서 훈련단계에서 CON1과 CON2를 제시할 때 사용한 컨센서스 기법으로 예측시간별로 순위를 달리하는 방법을 사용하였다. 순위를 결정하는 방법으로 절대평균오차(MAE)와 개별모델들 중 절대평균오차는 크지만 태풍 강도의 변화를

잘 모의하는 모델들을 고려하여 상관도(R), 즉 2가지 방법을 사용하였다.

순위 결정방법이 컨센서스 결과의 미치는 영향을 확인하기 위하여 앞 절에서 가장 안정적인 예측결과를 보여주었던 상위 4개 모델을 사용한 4가지 컨센서스의 2013~2014년 48h, 72h 평균 예측오차를 비교하였다(Table 16). 48h 예측오차에서는 가중치평균방법(WEI)을 사용한 모델을 비교하였을 때, 상관도를 기준으로 한 가중치평균[R-WEI(4)] 방법이 3.90 m/s로 가장 낮은 오차를 보였다. 단순평균방법(AVE)을 이용한 모델을 비교하였을 때는 가중치평균방법(WEI)과 반대로 MAE를 기준으로 한 단순평균방법[MAE-AVE(4)]이 4.02 m/s로 R-AVE(4)보다 오차가 적었다. 72h도 마찬가지로 가중치평균방법(WEI)을 사용한 모델은 상관도를 기준으로 한 컨센서스 기법이 예측오차가 더 적었고, 단순평균방법(AVE)을 사용할 경우에는 MAE를 기준으로 한 컨센서스 기법이 더 낮은 오차를 보이는 경향을 보였다.

Table 16. 48h and 72h forecast errors of 4 consensus scheme (4) from 2013 to 2014.

Forecast time \ Consensus	MAE-AVE(4)	MAE-WEI(4)	R-AVE(4)	R-WEI(4)
48h	4.02	4.29	4.18	3.90
72h	5.03	5.01	5.09	4.99

V. 요약 및 결론

본 연구에서는 2008년부터 2014년까지 북서태평양에서 발생한 태풍에 대해 국가태풍센터에서 이용하고 있는 16개 태풍 수치모델들의 강도 예측성능을 평가하고, 이를 바탕으로 최적의 태풍강도 예측 컨센서스 기법을 조사하였다. 태풍 수치모델들의 강도 예측성능 평가는 예보시간별, 태풍 강도별, 급격한 강도변화, 월별, 해역별, 태풍 이동경로 군집별, 그리고 한반도 영향태풍의 경우로 나누어 실시하였다.

72시간까지 모든 예보시간에 대해 누적한 최대풍속의 절대오차 분석에서 가장 좋은 성능을 보인 모델은 HWRF, GFS, TEPS 순으로 나타났다. 상관도 분석에서도 이 세 모델은 같은 순서로 좋은 성능을 보였다. HWRF는 태풍전용모델, GFS는 전구모델, TEPS는 앙상블모델인 점을 감안하면, 이러한 결과는 최근 다양한 형태로 태풍강도 예측기술이 발달하고 있음을 시사한다. HWRF는 최근 고해상도, 물리과정 모수화 개선, 자료동화 향상 등으로 모델의 자체 성능이 개선되었을 뿐 아니라 HWRF 모델에 초기 및 경계 값을 제공하는 GFS 모델의 성능이 개선되어 그 성능이 더욱 향상된 것으로 사료된다. 실제로 GFS는 16개의 가이던스 모델 중에 최근 들어 가장 크게 성능이 개선된 모델 중에 하나로 평가되었다. 한편, 태풍경로 예측 성능이 뛰어난 것으로 알려진 ECMWF는 강도오차에서 11개의 비교 모델 중에 9위를 차지하였고 우리나라 전 지구모델인 UM은 10위를 차지하였다.

예보시간별로 성능을 자세히 비교해 보면, 초기시간(0h)를 제외하면 HWRF와 GFS가 최대풍속 예측에서 가장 좋은 성능을 보였다. 초기시간에서는 일본의 두 모델인 JGSM과 TEPS가 다른 모델과 큰 차이로 최대풍속과 중심기압 모두 가장 좋은 예측 성능을 보였다. 이것은 도쿄 지역특별기상센터 자료를 생산하는 일본에서 이 자료를 초기장의 자료동화에 사용한 결과로 사료된다. 앞서 72시간 누적 오차에서 TEPS가 3위 그리고 JGSM이 5위로 좋은 성능을 기록한 것도 이들 모델의 초기 오차가 낮았기 때문으로 분석된다. 실제로 TEPS와 JGSM은 예측시

간이 길어지면(48h 이상) 성능이 크게 떨어져 상위권에 들지 못하였다. 반면, KWRF는 초기시간에 정확도는 낮았지만 48h과 72h에서 최대풍속 절대오차 평가에서 3위를 기록하였다. 특히, KWRF는 초기시간만 제외하면 중심기압 예측 능력은 평균적으로 가장 뛰어난 것으로 나타났다. UM_R도 중심기압 예측능력은 상위권(1~5위 내)를 유지하였다. 한편, 최대풍속에서 좋은 예측 성능을 보인 GFS는 중심기압에서는 오차가 매우 크게 나타났다. 이러한 결과는 중심기압 예측에서는 최대풍속과는 달리 공간해상도가 비교적 높은 KWRF (10 km)와 UM_R (12 km)이 저해상도의 전구모델인 GFS (약 100 km) 보다 유리함을 나타낸다. 또한 이것은 KWRF와 UM_R가 향후 중심기압과 최대풍속 관계에 대한 모수화 방법을 개선한다면 최대풍속의 예측 능력도 개선될 여지가 있음을 시사한다.

연도별로 가이던스 모델들의 48h 최대풍속 예측성능 변화를 살펴보면, NOGAPS와 GFS는 지속적인 성능 개선이 이루어지고 있는 반면 JGSM, ECMWF, UM은 성능 향상이 거의 없거나 오히려 더 나빠지는 경향을 보인다. 이러한 성능변화는 모델별로 최근 주요 모델의 48h 최대풍속 예측능력의 차이를 유발하는 원인으로 분석된다. 이 결과는 향후 앙상블 또는 컨센서스 기법 연구에서 장기자료를 사용하는 것보다 최근 자료만을 사용하는 것이 그리고 성능개선이 큰 모델들을 더 비중 있게 고려하는 것이 예측성능을 높일 수 있음을 의미한다. 월별로 평가를 실시한 결과에서도 예측 성능이 우수한 HWRF, GFS, TEPS는 계절에 관계없이 좋은 성능을 보인 반면, 그 외의 모델은 계절에 따라 성능 차이가 뚜렷했다.

태풍 강도별로 72시간 누적 최대풍속 예측성능을 비교하면, 약한 태풍(TS 급)은 TEPS, JGSM, GFS와 같은 전구모델의 성능이 태풍전용모델이나 지역모델에 비해 우세하였지만, 강한 태풍의 경우(TY, STY 급)는 태풍전용모델인 HWRF와 고해상도 지역모델인 KWRF가 특히 좋은 성능을 보였다. 이것은 강한 태풍의 경우 공간해상도와 태풍전용 물리과정 및 역학과정 도입이 강도예측 성능을 결정하는 주요 요인으로 작용함을 나타낸다.

태풍이 급격히 강화 또는 약화되는 경우의 72시간 누적 최대풍속 오차를 비교하면 대부분 급격하게 강화되는 경우보다 약화되는 경우에 오차가 더 낮은 것을

확인할 수 있었다. 이것은 대부분의 모델들이 태풍 강도를 과소모의하고 있는 것과 관련이 깊다. 태풍 강도가 급격하게 약화되는 경우는 GFS, HWRF, TWRF 순으로 그리고 강화되는 경우는 HWRF, KWRP, GFS 순으로 가장 좋은 성능을 보였다.

한반도에 영향을 미친 태풍만 뽑아서 72시간 누적 최대풍속 예측성능을 비교하면, 앞서 전체 태풍에 대한 분석의 결과와 유사하게 HWRF와 GFS가 가장 좋은 성능을 보였다. 특히, HWRF는 강도가 약했던 1304와 1315 태풍을 제외하면 예측성능이 모두 2위 내에 분포한다. 일부 한반도 영향 태풍에 대해서는 TEPS, JGSM, KWRP, TWRF도 좋은 예측성능을 보였다.

초기 값으로부터 태풍의 최대풍속변화(Δ MWS)만을 예측한 결과(72h 누적 예측성능 평가기준)는 최대풍속 값을 예측한 결과와 대부분 큰 차이가 없었다. 그러나 JGSM의 경우 최대풍속 예측에서는 좋은 성능을 보였지만, 최대풍속변화에서는 성능이 좋지 않았다. 이것은 앞서 언급한 바와 같이 JGSM 모델의 경우 초기예측 정확도는 높았지만 이후 과정에서 태풍 강도 변화를 제대로 모의하지 못한 결과로 사료된다. TWRF는 상관도 평가에서 최대풍속의 경우 7위였지만 최대풍속변화에서는 2위를 차지하여 강도변화 경향성에 대해서는 우수한 성능을 보였다. 이 결과는 컨센서스 예측에서 모델별로 다양한 특성을 살린 기술개발이 필요함을 시사한다.

태풍 이동경로에 따라 6개 군집으로 나누어 최대풍속의 72시간 누적오차를 비교한 결과, HWRF, GFS, TEPS가 거의 모든 군집에서 1, 2, 3위를 차지하여 성능이 우수한 태풍은 군집의 특성에 의존하지 않고 좋은 성능을 보임을 알 수 있었다. 그러나 그 외 성능이 낮은 태풍들은 군집에 따라 성능 차이가 크게 나타났다. 또한 공간에 따른 오차 분포에서는 태풍 발생 빈도가 큰 지역에서 오차 분포가 크게 나타났다.

태풍 강도 예측을 위한 최적의 컨센서스 기법을 조사하였다. 다양한 방법으로 컨센서스 기법 연구에서는 두 가지 기법(CON1, CON2)이 제시되었고[(i) 예측시간 별 가장 낮은 오차를 보인 컨센서스 모델들의 조합(CON1), (ii) 72h 누적 오차가 가장 낮은 컨센서스 모델(CON2)], 총 3년간(2012~2014) 예측한 결과를 개별 수치모델들과 비교하면 대부분의 예측시간과 72h 누적오차에서 CON1과

CON2 모두 최고성능의 개별모델보다 낮은 오차를 보였다. 특히 훈련단계에 개별모델의 순위와 실제 예측단계의 개별모델들의 순위와 유사할 때 예측결과가 더 좋게 나타나며, 훈련기간에 컨센서스 모델로 선정된 개별 모델들이 예측단계에서도 그 성능을 유지하느냐가 컨센서스 성능의 영향을 주는 것을 확인하였다 (특히 CON2가 MAE와 Bias에서 모두 CON1보다도 더 낮은 오차를 보임). 이러한 결과는 태풍 수치모델들에 대해 매년 예측성능을 평가하여 컨센서스 기법에 반영해야 함을 시사한다.

훈련기간의 길이에 따른 예측성능을 조사한 결과, 최근 1년 자료를 훈련기간으로 사용하였을 때가 2년을 사용하였을 때보다 더 좋은 예측결과를 얻을 수 있었다. 이는 훈련기간을 길게 하거나 예측기간과 멀어지면 이용 가능한 모델 개수가 한정되고 수치모델들의 예측 성능 향상 등을 반영하는데 미흡하므로 오차가 증가할 수 있음을 보였다. 사용된 수치모델의 수에 따른 컨센서스 기법의 성능을 조사하기 위해 2013~2014년 기간 동안 48h과 72h 예측오차를 비교한 결과 48h과 72h 예보시간 모두 선행연구 결과와 유사하게 상위 4~5개의 모델을 사용하였을 평균적으로 가장 좋은 결과를 보였다.

상관도와 절대오차를 사용하는 두 가지 순위 결정방법에 따른 컨센서스 기법의 예측성능을 살펴보고자, 컨센서스 기법의 최적 모델개수로 사료되는 상위 4개의 모델을 이용한 컨센서스 기법들을 비교·분석하였다. 그 결과 가중치평균방법을 사용한 컨센서스 모델들은 상관도를 기준으로 하였을 때 더 적은 오차를 보였고, 반대로 단순평균방법을 사용한 컨센서스 모델들은 절대평균오차를 기준으로 한 컨센서스 기법이 더 좋은 예측성능을 보여주었다. 이것은 단순평균방법이 수치모델들의 가중치를 주는 것이 아닌 오차가 적었던 모델들의 예측 결과를 단순 평균하는 방법이기 때문에 당연히 절대평균오차를 기준으로 한 컨센서스 기법이 예측오차가 적었던 것으로 사료된다. 반대로 가중치평균방법은 개별모델들의 과거 예측성능에 따라 가중치를 계산하므로 가중치 계산에 있어 개별모델들의 오차가 포함되어 절대평균오차 보다는 최대중심풍속 변화의 경향성을 잘 모의하는 모델들이 선택되어 예측오차가 더 낮아진 것으로 사료된다.

현업에서 태풍예측은 예보관의 주관적인 판단과 태풍강도 예측모델의 성능에 대한 지식을 바탕으로 하기 때문에, 태풍 수치모델들의 평가결과는 예보자에게

는 중요한 기초자료로 활용될 수 있다. 또한 앙상블 및 컨센서스 기법 연구에서도 이러한 성능평가 결과는 컨센서스 기법 개발과 검증에 있어 필수적인 자료로 사용될 수 있다. 그러나 본 연구에서 사용된 자료는 2011~2014년까지 4년으로 한정되어 있어 향후 더 많은 경우에 대해 살펴 볼 필요가 있다. 또한 72h 누적오차 뿐 아니라 48h과 72h 예측결과를 기준으로도 컨센서스 기법을 개발한다면 예보기관에서 주로 사용하는 48h과 72h 예측오차를 줄이는 데에도 기여할 수 있을 것으로 사료된다.

참고문헌

- Bloomfield P., Steiger W.L., 1980: Least absolute deviations curve-fitting. *Journal of Scientific and Statistical Computing*, 1, 290-291, DOI:10.1131/091019.
- Cangialosi, J. P., and J. L. Franklin 2013: 2012 National Hurricane Center Forecast Verification Report. National Hurricane Center (NHC), 79 pp. [Available online at http://www.nhc.noaa.gov/verification/pdfs/Verification_2012.pdf.]
- Choi. E. S., and I. J. Moon., 2008: The Variation of Extreme Values in the Precipitation and Wind Speed during 56 Years in Korea. *Atmosphere*, 18(4), 397-416.
- DeMaria. M., M. Mainelli., L. K. Shay., J. A. Knaff., and J. Kaplan 2005: Further Improvements to the Statistical Hurricane Intensity Prediction Scheme (SHIPS). *Wea. Forecasting*, 20, 531-543.
- Elsberry, 2014: Advances in understanding and forecasting rapidly changing phenomena in tropical cyclones. *Tropical Cyclone Research and Review*, 2(1), 1-12.
- Goerss. J. S., 2000: Tropical Cyclone track forecasts using ensemble of dynamical models. *Mon. Wea. Rev.*, 128, 1187-1193.
- Jun. S-. H-, W. J. Lee., K-. R. Kang., K. Y. Byun., J. Y. Kim., and W-. T. Yun 2015: Applicability of the Superensemble to the Tropical Cyclone Track Forecasts in the Western North Pacific. *J. Atmos. Sci.*, 51(1), 39-48.
- Kim. H. -S., J., -H. Kim, C. -H. Ho, and P. -S. Chu, 2011: Pattern classification of typhoon tracks using the fuzzy c-means clustering method. *Journal of Climate*, 24(2), 488-508.
- Kim. K. S., I. J. Moon., and S. J. Kwon., 2011: Seasonal prediction of tropical

- cyclone frequency during summer in the typhoon emergency area of Korea. The Fifth Korea–Japan–China Joint Conference on Meteorology. 246pp.
- NCEP Environmental Modeling Center, 2003: The GFS Atmospheric Model. NCEP Office Note 442, Global Climate and Weather Modeling Branch, EMC, Camp Springs, Maryland.
- Knaff, J. A., M. DeMaria., C. R. Sampson., and J. M. Gross 2003: Statistical, 5-Day Tropical Cyclone Intensity Forecasts Derived from Climatology and Persistence. *Wea. Forecasting*, 18, 80–92.
- Knaff, J., C. R. Samson. and M. DeMaria, 2005: An Operational Statistical Typhoon Intensity Prediction Scheme for the Western North Pacific. *Wea. Forecasting*, 20, 688–699.
- Kotal et al., 2011: A multimodel ensemble (MME) technique for cyclone track prediction over the North Indian Sea. *GEOFIZIKE*, 28, 275–291.
- KMA NTC (Korea Meteorological Administration National Typhoon Center), 2011: Typhoon White Book.
- KMA NTC, 2013: 2013 Typhoon Analysis Report.
- KMA NTC, 2014: 2014 Typhoon Analysis Report for affecting Korean Peninsula.
- Kumar, T. S. V., T. N. Krishnamurti, M. Fiorino, and M. Nagata, 2003: Multimodel superensemble forecasting of tropical cyclones in the Pacific. *Mon. Wea. Rev.*, 131, 574–583.
- Lee. D. K., D. E. Jang., and T. K. Wee., 1992: Typhoons Approaching Korea, 1960–1989. Part I: Statistics and Synoptic Overview. *Atmosphere*, 28(2),133–147.
- Lee. J. W., S. W. Lee., S. O. Han., S. J. Lee., and D. E. Jang., 2011: The Impact of Satellite Observations on the UM-4DVar Analysis and Prediction System at KMA. *Atmosphere*, 21(1), 85–93.
- Lee, Y. J., H. J. Kwon., and D. C. Joo., 2011: Dynamic data-base Typhoon

- Track Prediction (DYTRAP). *Atmosphere*, 21(2), 209-220.
- Lim. J. H., D. K. Rha., H. B. Cheong., and H. J. Kwon., 2003: The Performance of the Double-Fourier-Series BARotrophic Typhoon-Prediction Model (DBAR). *Atmosphere*, 13(1), 428-429.
- Nakagawa, M., 2009: Outline of the High Resolution Global Model at the Japan Meteorological Agency. JMA.
- NEMA (National Emergency Management Agency), 2011: Annual Disaster Report, 986pp.
- Park, S. Y., S-. W. Joo., H. C. Shin., J. Y. Cho., 2008: Coupling of Typhoon Bogusing for KMA WRF. Proceedings of the Spring Meeting of KMS, 2008, 260-261.
- Rogers, R. and Coauthors, 2006: The intensity forecasting experiment : A NOAA mutiyear field program for improving tropical cyclone intensity forecasts. *Bull. Amer. Meteor. Soc.*, 87, 1523-1537.
- Simiu, E., and R. H. Scanlon, 1978: Wind effects on structures. Wiley-Interscience, New York, NY, 485pp.
- Sohn, 2005: Statistical Correction of Numerical Model Forecasts for Typhoon Tracks. *The Korean Communications in Statistics*, 12(2), 295-304.
- Vijay, T., and HWRF Team, 2014: Overview of the NCEP Operational HWRF Modeling System, Joint EMC/DTC HWRF Tutorial.
- Williford, C. E., T. N. Krishnamurti., R. C. Torres., S. Cocks., Z. Christidis., and T. S. Vijaya Kumar 2003: Real-Time Multimodel Superensemble Forecasts of Atlantic Tropical Systems of 1999. *Monthly Wether Review*, 131, 1878-1894.
- WMO, 2007: Sixth WMO International Workshop on Tropical Cyclone(IWTC-VI). WMO, 92pp. [Available online at http://www.aoml.noaa.gov/hrd/Landsea/WWRP2007_1_IWTC_VI.pdf.]
- Yamaguchi, M., and Komori, T., 2009: Outline of the Typhoon Ensemble Prediction System at the Japan Meteorological Agency. JMA.

Yamaguchi. M., T. Nakazawa and K. Aonashi 2012: Tropical cyclone track forecasts using JMA model with ECMWF and JMA initial conditions. The Fourth THORPEX Asian Science Workshop, Kunming, China.