Five-Day Track Forecast Skills of WRF Model for the Western North Pacific Tropical Cyclones

JIHONG MOON,[a] JINYOUNG PARK,[a] DONG-HYUN CHA,[a] AND YUMIN MOON[b]

[a] *School of Urban and Environmental Engineering, Ulsan National Institute of Science and Technology, Ulsan, South Korea*

[b] *Department of Atmospheric Sciences, University of Washington, Seattle, Washington*

(Manuscript received 12 June 2020, in final form 10 May 2021)

ABSTRACT: In this study, the characteristics of simulated tropical cyclones (TCs) over the western North Pacific by a regional model (the WRF Model) are verified. We utilize 12-km horizontal grid spacing, and simulations are integrated for 5 days from model initialization. A total of 125 forecasts are divided into five clusters through the *k*-means clustering method. The TCs in the cluster 1 and 2 (group 1), which includes many TCs moving northward in the subtropical region, generally have larger track errors than for TCs in cluster 3 and 4 (group 2). The optimal steering vector is used to examine the difference in the track forecast skill between these two groups. The bias in the steering vector between the model and analysis data is found to be more substantial for group 1 TCs than group 2 TCs. The larger steering vector difference for group 1 TCs indicates that environmental fields tend to be poorly simulated in group 1 TC cases. Furthermore, the residual terms, including the storm-scale process, asymmetric convection distribution, or beta-related effect, are also larger for group 1 TCs than group 2 TCs. Therefore, it is probable that the large track forecast error for group 1 TCs is a result of unreasonable simulations of environmental wind fields and residual processes in the midlatitudes.

KEYWORDS: Hurricanes/typhoons; Numerical weather prediction/forecasting; Model errors; Regional models

1. Introduction

A tropical cyclone (TC) is a natural disaster that causes enormous property damage and casualties across a wide geographic area in a short amount of time, due to strong winds and heavy rainfall. Socioeconomic damage caused by TCs has increased rapidly over the past few decades because of an increase in population density and social development (Pielke et al. 2008; Zhang et al. 2009). Notably, the western North Pacific region experiences huge amount of damage from TCs because it has the highest frequency of TC genesis in combination with the high population densities of East Asia and Southeast Asia. In 2019, the Korean Peninsula was exposed to severe hazards from the landfalls of seven TCs, while TC Hagibis caused widespread destruction to Japan. Accurate track and intensity forecasts and appropriate preparedness and responses are necessary to minimize property damage and casualties from TCs.

The predictability of the numerical weather prediction (NWP) model plays an essential role in the decision-making process used to respond to a typhoon, since forecasters present guidance based on the predictions of the NWP model. In the past several decades, the track forecast skill of operational NWP models has been steadily improved, with a mean error at 72-h lead time of less than 200 km (Chen et al. 2019). However, NWP models still have significant errors, i.e., larger than 400 km in a 5-day forecast. With the growth in demand for accurate medium-range forecasts, the predictability of the NWP model must be improved to minimize damage.

The dynamics of mesoscale processes of a TC can be well simulated by the high-resolution model simulation that better represent the inner-core dynamics. Several studies on the proper resolution for TC forecast (Chen et al. 2007; Davis et al. 2011, 2008; Fierro et al. 2009; Gentry and Lackmann 2010; Gopalakrishnan et al. 2011; Jin et al. 2014) concluded that a grid spacing of less than 5 km is required, with 1–2 km preferred in order to provide an accurate representation of TC intensity and structure. At present, the most advanced operational global NWP models have a horizontal grid length of about 10 km, which is insufficient to resolve the small-scale processes that influence storm development (Short and Petch 2018). Given the limitations of computing resources, many TC forecasts and research studies are conducted using a regional model.

There are advantages and disadvantages of using a regional model in forecasting TC track. In comparison to global models, regional models generally have relatively larger track errors, especially for relatively long-term forecasts since their ability to resolve large-scale atmospheric waves is limited (Liu and Xie 2012; Waldron et al. 1996). On the other side, better representation of storm-scale process associated with convective heating asymmetry can improve the forecast of TC motion (Yamada et al. 2016). Moon et al. (2018) improve the track forecast of the WRF Model by using spectral nudging, which provides the large-scale features of the global model forecasts onto the regional model interior. Thus, to improve the track forecast of regional models, the characteristics of the track forecast need to be analyzed in order to understand the error sources of the model.

The TC track error in the NWP model is profoundly affected by the interaction with surrounding environmental

Corresponding author: Dong-Hyun Cha, dhcha@unist.ac.kr

DOI: 10.1175/WAF-D-20-0092.1

TABLE 1. Forecast details and initial times for the 18 TCs that occurred during 2013–14. The interval of forecast initial time is 24 h.

TC No.	TC name	Forecast initial time (interval of 24 h)	No. of cases
1307	Soulik	0000 UTC 8 Jul 2013	1
1311	Utor	0000 UTC 10 Aug 2013	1
1324	Nari	0000 UTC 10 Oct 2013	1
1326	Francisco	0000 UTC 17 Oct 2013–0000 UTC 20 Oct 2013	4
1328	Lekima	0000 UTC 21 Oct 2013	1
1331	Haiyan	0000 UTC 4 Nov 2013–0000 UTC 6 Nov 2013	3
1408	Neoguri	0000 UTC 4 Jul 2014–0000 UTC 5 Jul 2014	2
1409	Rammasun	0000 UTC 13 Jul 2014–0000 UTC 14 Jul 2014	2
1410	Matmo	0000 UTC 18 Jul 2014	1
1411	Halong	0000 UTC 29 Jul 2014–0000 UTC 5 Aug 2014	8
1418	Phanfone	0000 UTC 29 Sep 2014–0000 UTC 1 Oct 2014	3
1419	Vongfong	0000 UTC 3 Oct 2014–0000 UTC 8 Oct 2014	6
1420	Nuri	0000 UTC 1 Nov 2014	1
1422	Hagupit	0000 UTC 2 Dec 2014–0000 UTC 5 Dec 2014	4
1504	Maysak	0000 UTC 28 Mar 2015–0000 UTC 30 Mar 2015	3
1506	Noul	0000 UTC 4 May 2015–0000 UTC 6 May 2015	3
1507	Dolphin	0000 UTC 9 May 2015–0000 UTC 15 May 2015	7
1511	Nangka	0000 UTC 4 Jul 2015–0000 UTC 12 Jul 2015	9
1513	Soudelor	0000 UTC 31 Jul 2015–0000 UTC 4 Aug 2015	5
1516	Goni	0000 UTC 15 Aug 2015–0000 UTC 20 Aug 2015	6
1517	Atsani	0000 UTC 15 Aug 2015–0000 UTC 19 Aug 2015	5
1520	Krovanh	0000 UTC 16 Sep 2015	1
1521	Dujuan	0000 UTC 24 Sep 2015	1
1524	Koppu	0000 UTC 14 Oct 2015	1
1527	In-fa	0000 UTC 18 Nov 2015–0000 UTC 21 Nov 2015	4
1602	Nepartak	0000 UTC 4 Jul 2016	1
1608	Conson	0000 UTC 9 Aug 2016	1
1612	Lionrock	0000 UTC 22 Aug 2016–0000 UTC 25 Aug 2016	4
1616	Meranti	0000 UTC 10 Sep 2016	1
1618	Malakas	0000 UTC 12 Sep 2016–0000 UTC 15 Sep 2016	4
1625	Haima	0000 UTC 16 Oct 2016	1
1707	Noru	0000 UTC 22 Jul 2017–0000 UTC 2 Aug 2017	12
1720	Talim	0000 UTC 10 Sep 2017–0000 UTC 12 Sep 2017	3
1725	Lan	0000 UTC 16 Oct 2017–0000 UTC 17 Oct 2017	2
1727	Saola	0000 UTC 22 Oct 2017	1
		Total	125

fields. Carr and Elsberry (2000a,b) analyzed the forecast results of the Navy Operational Global Atmospheric Prediction System model (NOGAPS) and the Geophysical Fluid Dynamics Laboratory (GFDL) hurricane model, and suggested a conceptual model for the error mechanism in the tropical and midlatitude regions. These studies show that possible model errors can arise from errors in the location or structure of synoptic-scale features. Furthermore, Kehoe et al. (2007) showed that the errors associated with aloft TC-related anticyclone intensification could also contribute to environmental wind errors that may lead to poor TC track forecasts. In addition, Magnusson et al. (2019) reviewed studies on large error cases. In this paper, several methods including ensemble models and nudging experiment to understand the low predictability for difficult cases were introduced (Nystrom et al. 2018; Torn et al. 2018).

TC track is determined by the influence of the large-scale flow around it, which is called the "steering effect." Several studies have been conducted to predict TC motion by analyzing the surrounding wind fields. George and Gray (1976) compared the TC track and surrounding winds, and determined that 700- or 700–500-hPa flows can represent the motion of TC. Chan and Gray (1982) defined the steering vector by averaging large-scale wind at 700–500 hPa with a radius of 5°–7° of the center of a TC. There have been many studies which have sought to determine the steering vector since then, but it has considerable uncertainty regarding the steering flow depth and radius for individual TCs (Aberson 2010; Dong and Neumann 1986; George and Gray 1976; Velden and Leslie 1991). Galarneau and Davis (2013) suggested the optimal steering vector, which was defined by optimized depth and radius for each TC case, and quantitively analyzed the error of TC motion. Understanding the steering vector can be useful for error analysis because it can represent the characteristics of the surrounding wind field.

TC track forecast error also varies depending on the location or basin of the TC. Notably, the track error for the western North Pacific TCs tends to be larger than that of other basins (Hodges and Emerton 2015). Since the western North Pacific has a large area for TC activity and therefore TCs tend to

sustain for longer, various types of TCs coexist. Several studies have classified TCs using various methods in order to understand the characteristics of TCs (Camargo et al. 2008, 2007a,b; Elsner 2003; Elsner and Liu 2003; Hall and Jewson 2007; Harr and Elsberry 1991, 1995a,b; Hodanish and Gray 1993; Kim et al. 2011; Lander 1996; Nakamura et al. 2009). Cluster analysis is one of the classification methods, and is relatively objective. Studies using cluster analysis are typically conducted using the *k*-means clustering method (Camargo et al. 2007a,b; Elsner 2003; Elsner and Liu 2003) and the fuzzy clustering method (Kim et al. 2011).

In this study, we analyzed TC simulation errors using the Weather Research and Forecasting (WRF) Model, which is a suitable regional model for simulating TCs (Fierro et al. 2009; Raju et al. 2011; Skamarock et al. 2008). To figure out the error sources, optimal steering vector analysis was conducted. Regional differences in forecast skill were also studied through *k*-means clustering analysis. Model configuration and analysis methods are described in section 2, and model results and errors are analyzed in section 3. Finally, a summary and conclusions are presented in section 4.

2. Data and methods

a. Model configuration and data

We used the WRF Model version 4.0.0 to analyze its track forecast skill for the western North Pacific TCs. The horizontal grid spacing of the forecast domain was 12 km, and the number of horizontal grids was 421 and 371 for the west–east and south–north direction, respectively. The domain contained 35 vertical levels from the surface to the top of the atmosphere at 50 hPa, with the sigma coordinate. All simulations were 120 h forecasts from model initiation. The center of the domain varied with the center of the observed TC and was designated as 10° north and west of the observed TC center when the latitude of the TC center is below 20°N, and 10° north when the TC center is north of 20°N. Due to the change in the domain center for each forecast, the domain was able to cover the entire track of the simulated TCs for five days. Real-time global analysis data and forecast data from the Global Forecast System (GFS) of the National Centers for Environmental Prediction (NCEP) were used to provide the initial and boundary conditions for the WRF Model with a grid spacing of 0.5°. The GFS real-time forecast is available every 6 h. This 6-hourly forecast data are used as the initial and boundary forcing data through the WRF preprocessing system (WPS), and the model sea surface temperature is updated with the GFS.

TC vortex initialization methods were not applied. The model utilized the Yonsei University planetary boundary layer scheme (Hong et al. 2006; Noh et al. 2003), WRF single-moment 6-class microphysics scheme (Hong and Lim 2006), Kain–Fritsch cumulus parameterization scheme (Kain 2004), Dudhia shortwave radiation scheme (Dudhia 1989), and Rapid Radiative Transfer Model longwave radiation scheme (Mlawer et al. 1997). We conducted 125 forecasts of 37 TCs formed on the western North Pacific in 2013–17, as listed in Table 1. The cases which maintain the intensity stronger than tropical storm for five days are chosen for test cases.

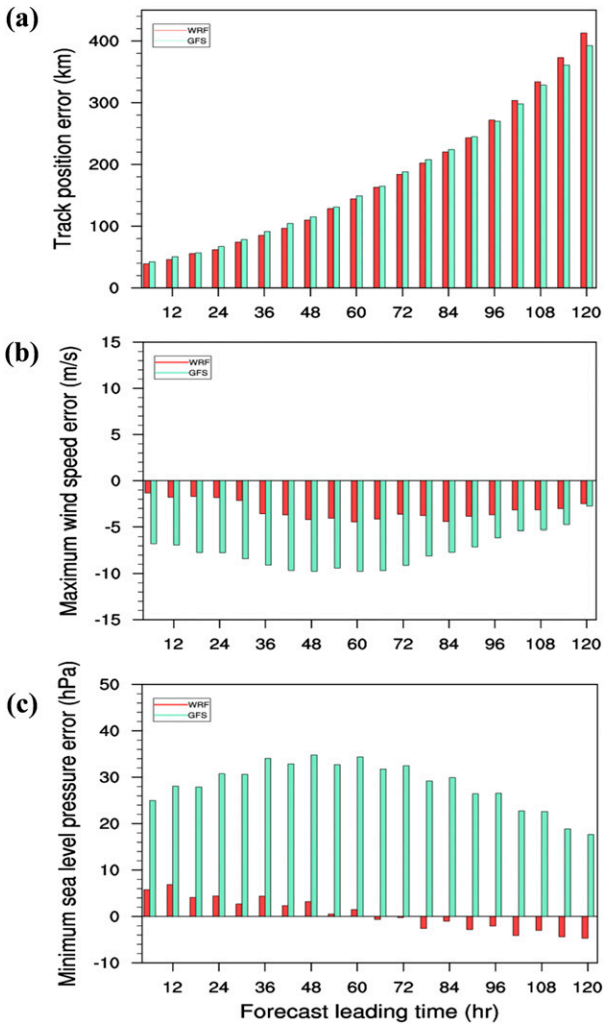

FIG. 1. Mean (a) track error, (b) maximum wind speed error, and (c) minimum sea level pressure error of the WRF Model and GFS for 125 forecasts. The sample size is the same at each forecast time.

Cluster analysis and verification are conducted using Joint Typhoon Warning Center (JTWC) best track data. NCEP GFS analysis data are used for the observation data for optimal steering vector analysis. The GFS analysis data are available every 6 h with a horizontal grid spacing of 0.5°. The simulated track and intensity of each experiment are compared with the JTWC best track data. The location of the TC center is defined by the simple method of identifying the minimum pressure in the sea level pressure field (Feser and von Storch 2008). The maximum wind speed is defined as the highest wind speed value within a radius of 500 km from the TC center. Track error is defined as the great circle distance between the TC center of the best track data and the simulated center (Neumann and Pelissier 1981; Powell and Aberson 2001), and it can be calculated as follows:

$$\text{track error (km)} = 111.11 \cos^{-1}[\sin\varphi_0 \sin\varphi_s + \cos\varphi_0 \cos\varphi_s \cos(\lambda_0 - \lambda_s)],$$

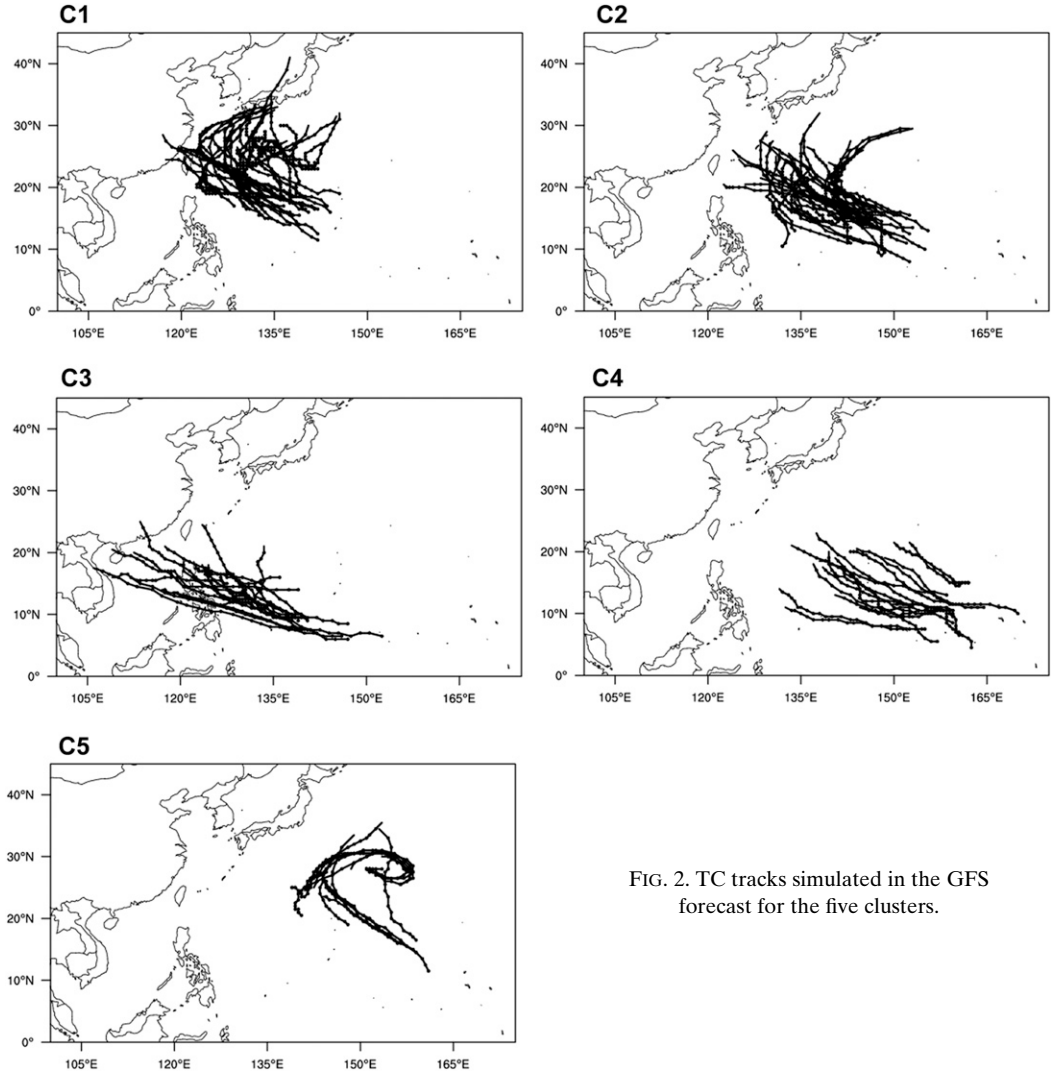

FIG. 2. TC tracks simulated in the GFS forecast for the five clusters.

where λ_0 and φ_0 are the longitude and latitude of the TC center in the best track data, respectively; and λ_s and φ_s are longitude and latitude of simulated TC center, respectively.

b. *Cluster analysis*

Cluster analysis is a data mining technique that categorizes a set of data into several groups by considering the characteristics of each data point. Members of the same group show greater similarity in characteristics to each other than to the members of other groups. There are several cluster analysis methods depending on how groups are classified. In this study, the *k*-means cluster algorithm was used to classify the simulation cases based on the active TC region (MacQueen 1967). The grouping method of *k*-means works in a way that minimizes the variance of distance differences between the members in a group. From Applied Statistics-136 (AS-136) advanced *k*-means clustering algorithm (Hartigan and Wong 1979), the sum of the squared distances in the clusters is calculated as follows:

$$S = \sum_{i=1}^{k} \sum_{x \in C_i} |x - \mu_i|^2, \tag{1}$$

where x is a data point, k is the total number of clusters, C_i is the set of points that belong to cluster i, and μ_i is the cluster center of C_i.

The *k*-means is the clustering method that aims to find the positions of μ_i that minimize the sum of the squared distance S. First, the centers of the k clusters are randomly initialized. Next, each data point is attributed to the closest cluster center, and new cluster centers are set to the mean of all points belonging to each cluster. The convergence points by several iterations of this process are the centers of each cluster, and the peripheral points are the member of each cluster. The 5-day mean latitude and longitude are used for data points in this study.

The optimum number of clusters is determined by employing the Krzanowski–Lai (KL) index (Krzanowski and Lai 1988). The KL index is computed as

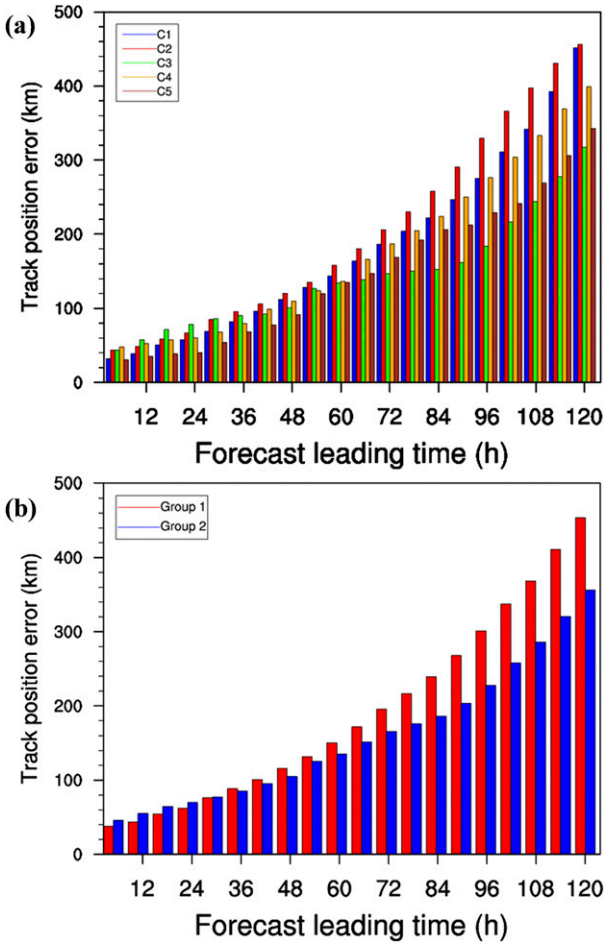

FIG. 3. Mean track error of (a) each cluster and (b) group 1 and 2 TCs.

$$\text{Krzanowski–Lai index} = \frac{|\text{diff}_i|}{|\text{diff}_{i+1}|}, \tag{2}$$

$$\text{diff}_i = (i-1)^{2/D}\text{SSW}_{i-1} - i^{2/D}\text{SSW}_i, \tag{3}$$

where i is the number of clusters, D is the number of datasets, and SSW is the within-group sum of squares for all datasets in the ith cluster. The optimum number of clusters is the i which maximizes the KL index.

c. *Optimal steering vector analysis*

TC motion is largely governed by the surrounding flow called the steering flow (Chan 2005). The steering flow is obtained by averaging environmental winds, which are the residual wind resulting from the removal of TC vortex related features (Galarneau and Davis 2013). To calculate environmental winds from total wind field, the vortex removal method from Davis et al. (2008) is used. The vortex is removed at the layer from 850 to 200 hPa with an interval of 50 hPa within a given radius using the Poisson equation for streamfunction (4) and velocity potential (5) with homogeneous boundary conditions applied at the edge of the computational domain. The boundary value problem on each pressure level can be stated as

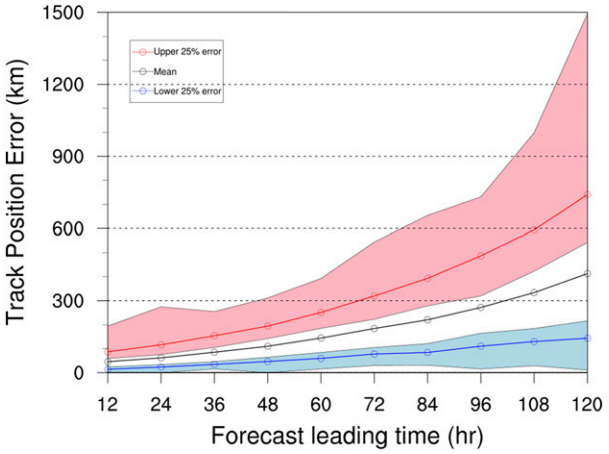

FIG. 4. Mean track error for all cases (black line), upper 25% error cases (red line), and lower 25% error cases (blue line) with the spread of upper and lower error cases (red and blue shading, respectively).

$$\nabla^2\psi = \begin{pmatrix} \zeta & \text{for} & r \le r_0 \\ 0 & \text{for} & r > r_0 \end{pmatrix}, \tag{4}$$

$$\nabla^2\chi = \begin{pmatrix} \delta & \text{for} & r \le r_0 \\ 0 & \text{for} & r > r_0 \end{pmatrix}, \tag{5}$$

where ψ is the streamfunction, χ is the velocity potential, $\psi = \chi = 0$ on the lateral boundaries of the computational domain, ζ is the relative vorticity, δ is the divergence, and r_0 is the radius of TC removal. Given the solutions for streamfunction and velocity potential from (4) and (5), we can determine the nondivergent and irrotational wind vectors from

$$\mathbf{V}_\psi(x,y,p) = \mathbf{k} \times \nabla\psi, \tag{6}$$

and

$$\mathbf{V}_\chi(x,y,p) = \nabla\chi, \tag{7}$$

where $\mathbf{k}$ is the vertical unit vector.

By subtracting (6) and (7) from the total wind as follows:

$$\mathbf{V}_{\text{env}}(x,y,p) = \mathbf{V}(x,y,p) - \mathbf{V}_\psi(x,y,p) - \mathbf{V}_\chi(x,y,p), \tag{8}$$

the environmental wind $\mathbf{V}_{\text{env}}$ is obtained for each pressure level within a given radius.

TABLE 2. Number and ratio of upper and lower 25% error cases for group 1 and 2 TCs.

	Group 1 (75)	Group 2 (36)
Upper 25% error cases	22 (29.3%)	7 (19.4%)
Lower 25% error cases	14 (18.7%)	13 (36.1%)

The steering vector for observation ($\mathbf{V}_O$) and model ($\mathbf{V}_M$) is computed by averaging the environmental wind within a removal radius as follows:

$$\mathbf{V}_O = \frac{1}{\pi r_o^2}\int_0^{2\pi}\int_0^{r_o}\mathbf{V}_{\text{obs}}\, r\, dr\, d\Theta, \tag{9}$$

$$\mathbf{V}_M = \frac{1}{\pi r_o^2}\int_0^{2\pi}\int_0^{r_m}\mathbf{V}_{\text{wrf}}\, r\, dr\, d\Theta, \tag{10}$$

where $\mathbf{V}_O$ and $\mathbf{V}_M$ are the area-average environmental wind and r_o and r_m are the TC removal radii for the observation and model TC, respectively. The area-averaged environmental winds are computed for eight different radii ranging from 1° to 8° from the TC center. Then, the area-average environmental winds of each radius are vertically averaged for layers of increasing depth ranging from the shallowest layer of 850–800 hPa to the deepest layer of 850–200 hPa with an interval of 50 hPa. A total of 104 steering vectors are computed for eight averaging radius and thirteen averaging depth. The optimal steering vector is defined as the best matched to the actual TC motion among the 104 steering vectors for each averaging radius and depth. Actual TC motion is calculated from the position difference between 12 h before and after the analysis time to reduce the short-term variation. The example of a specific calculation process is revealed in chapter 3b. We applied this approach to analyze the track errors in the model.

3. Results

a. Cluster analysis

Figure 1 shows the mean track error, maximum wind speed bias, and minimum sea level pressure bias of all experiments against the JTWC best track data. The error of GFS is also analyzed for comparison. Overall, WRF and GFS have similar mean track errors until the middle of the forecast, but the error in WRF tends to be slightly larger after 96 h forecast time (Fig. 1a). In terms of intensity forecast, the maximum wind speed and minimum sea level pressure are more realistically predicted by the WRF (Figs. 1b,c). The WRF Model simulates the mean maximum wind speed and the minimum sea level pressure with biases not exceeding 5 m s^{-1} and 7 hPa, respectively, across the entire forecast time. On the other hand, the GFS tends to significantly underestimate TC intensity compared to the WRF Model. As shown in previous studies (Cha et al. 2011; Moon et al. 2018), intensities of TC are better simulated by the model since the structure of TCs is more realistically represented with the parameterization of the WRF Model. In this study, we focus on the track forecast skill and analyze the track forecast of the WRF Model in detail through the clustering analysis.

The optimized number of clusters is determined by calculating the KL index. This index has the highest value when the number of clusters is five. At this point, TC cases are classified into five clusters using the *k*-means clustering algorithm. Figure 2 shows the tracks of GFS forecast for each cluster. Each cluster is classified according to the 5-day mean latitude and longitude of each case. The 5-day mean locations of cluster 1 (C1) to cluster 5 (C5) are 23.3°N, 131.4°E; 18.1°N, 139.0°E; 13.4°N, 127.9°E; 13.2°N, 150.2°E; and 25.4°N, 148.9°E, respectively. In C1, most TC forecasts are initialized at the subtropical regions, and the TCs move northwestward toward the midlatitude continents. C2 TCs are initialized at the lower latitudes and move northwestward. Both C3 and C4 TCs are initialized at the low latitude and move nearly westward, but C3 TCs are located to the west of the C4 TCs. The C5 TCs are

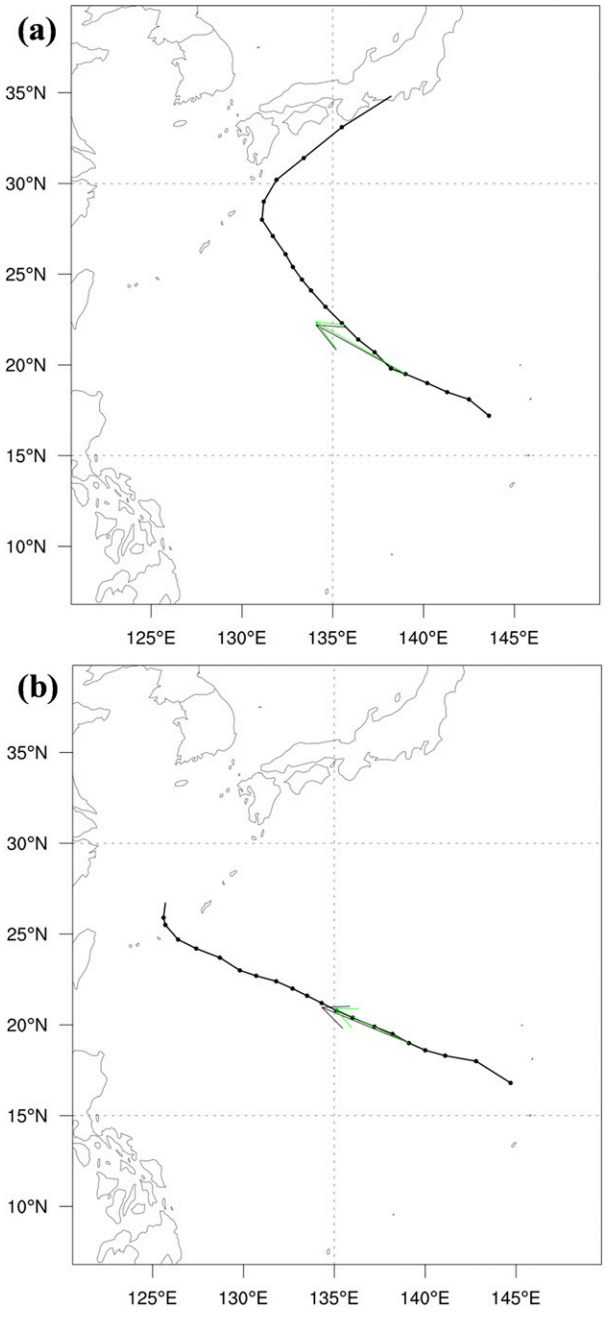

FIG. 5. TC track (dotted line), actual motion vector (black arrow), and optimal steering vector (green arrow) of (a) analysis data and (b) model simulation at 24 h forecast initialized at 1 Oct 2014.

TABLE 3. Magnitude of residuals for TC Phanfone initialized at 0000 UTC 1 Oct 2014. The residuals of the model and analysis data are calculated from the difference between actual motion and optimal steering vector according to the removal radius and averaging depth. The minimum value is bolded.

	Depth (hPa)												
Radius (°)	50	100	150	200	250	300	350	400	450	500	550	600	650
							$\lvert\mathbf{V}^{O}_{\text{act}} - \mathbf{V}_{O}\rvert$						
1	1.80	1.08	1.04	1.44	1.84	2.14	2.29	2.22	1.99	1.62	1.19	0.89	0.95
2	1.70	1.40	1.44	1.51	1.61	1.76	1.89	1.92	1.89	1.80	1.65	1.51	1.56
3	2.54	2.36	2.24	1.98	1.69	1.51	1.36	1.24	1.15	1.11	1.03	0.93	0.91
4	2.68	2.44	2.26	1.98	1.68	1.45	1.24	1.02	0.76	0.51	0.34	0.30	0.22
5	3.10	2.86	2.66	2.37	2.03	1.73	1.48	1.21	0.89	0.56	0.31	0.22	**0.17**
6	3.02	2.88	2.69	2.39	2.03	1.71	1.47	1.23	0.97	0.70	0.52	0.40	0.26
7	2.73	2.66	2.51	2.26	1.98	1.76	1.61	1.43	1.21	0.96	0.75	0.59	0.36
8	2.25	2.27	2.20	2.08	1.96	1.90	1.85	1.77	1.59	1.37	1.13	0.95	0.67
							$\lvert\mathbf{V}^{M}_{\text{act}} - \mathbf{V}_{M}\rvert$						
1	8.99	8.23	7.52	7.04	6.75	6.41	5.97	5.48	4.97	4.49	4.10	3.83	3.55
2	6.55	6.28	5.92	5.53	5.15	4.80	4.46	4.10	3.61	2.90	2.13	1.60	1.48
3	3.53	3.69	3.95	4.14	4.27	4.33	4.33	4.26	4.08	3.66	2.96	2.08	1.27
4	2.69	2.83	3.08	3.31	3.49	3.63	3.71	3.74	3.66	3.43	2.94	2.20	1.39
5	2.16	2.36	2.70	3.00	3.20	3.30	3.32	3.28	3.20	3.04	2.67	2.07	1.44
6	2.06	2.41	2.75	2.98	3.07	3.04	2.97	2.85	2.72	2.57	2.28	1.79	1.27
7	1.89	2.35	2.64	2.77	2.75	2.66	2.56	2.46	2.37	2.29	2.12	1.69	**1.16**
8	1.72	2.13	2.36	2.43	2.39	2.30	2.24	2.23	2.23	2.26	2.19	1.84	1.29

located on the eastern boundary of the active TC region of the western North Pacific. The number of members in C1 to C5 is 39, 36, 19, 17, and 14, respectively.

Figure 3a shows the mean track error at different forecast hours for each cluster against the JTWC best track data. There are significant differences in track predictability among the clusters. C1 and C2 that includes many TCs move northward have the highest track error at 120 h while C3, in which TCs are active on the southwest region of the western North Pacific basin, has the smallest track error. The track error of C1 TCs is approximately 200 km at 72 h and approximately 450 km at 120 h, while that of C3 is approximately 140 km at 72 h and about 320 km at 120 h. To understand these regional characteristic differences in track error, we organized two groups according to the 5-day track error: group 1 (C1 and C2) and group 2 TCs (C3 and C4). Figure 3b shows the difference in the track forecast error between the two groups is evident. We excluded the C5 from groups since TC cases in C5 move on abnormal course and different characteristics from other clusters. The difference between two groups is not significant before 48 h, but it increases rapidly as model integration progresses. The track error of group 1 TCs at 120 h is 453 km, which is approximately 28% greater than that of group 2 TCs (355 km).

Figure 4 shows the spread of track forecast errors for all WRF simulations. It shows that a large number of excessive error cases exist in the upper 25% error cases of the WRF forecasts. Those cases predominantly increase the mean track error value. Notably, the highest error cases have a track error of approximately 1500 km, which is more than three times larger than the mean track error at 120 h. Improving the simulations that have these extremely high track errors is an important task to reduce the mean track error. Table 2 shows the number and ratio of large and small error cases according to the direction of transition. For group 1 TCs, more TCs are included in the large error cases than the small error cases

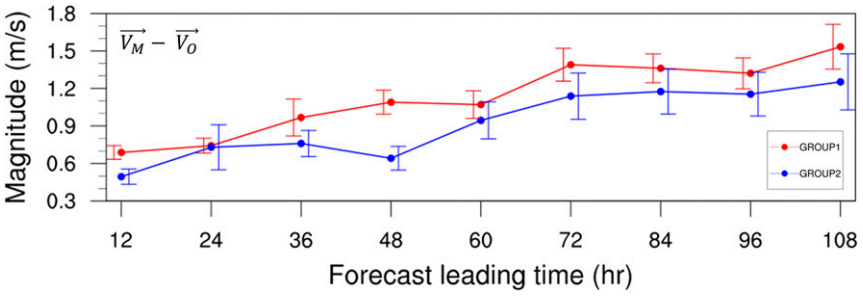

FIG. 6. Time series of the difference between the optimal steering vector of the model and analysis data for group 1 TCs (red) and group 2 TCs (blue). The bars mean the 90% confidence interval.

TABLE 4. Histogram of percentages of upper-level boundary for model and analysis data steering vector calculation for each group. The sample size is 657 and 279 for group 1 and 2, respectively.

	Upper-level boundary (hPa)												
	200	250	300	350	400	450	500	550	600	650	700	750	800
WRF_G1	37.7	11.7	9.7	8.8	7.2	5.0	3.3	2.6	2.6	2.4	1.5	2.0	5.3
WRF_G2	33.7	14.3	10.4	8.6	5.0	5.7	4.7	5.7	2.5	1.8	1.4	1.8	4.3
ANL_G1	41.4	12.8	11.0	5.3	7.8	5.6	3.8	3.2	0.9	1.2	1.7	1.8	3.5
ANL_G2	28.7	5.4	13.3	16.8	9.7	6.5	2.5	3.6	1.1	1.8	1.8	4.3	4.7

(27% and 19.4%, respectively). Conversely, for group 2 TCs, there is a large percentage of small error cases (36.1%) and a smaller percentage of large error cases (19.1%). Moon et al. (2018) also showed that the WRF Model has lower track predictability compared to the GFS for TCs located in the northeast region of the western North Pacific TC basin. Larger error cases in group 1 TCs indicate that the WRF Model has a weakness in forecasting the tracks of TCs moving to the midlatitudes.

b. *Optimal steering vector analysis*

To investigate regional differences in track forecast skill, we calculate the optimal steering vectors both for the model and for analysis data. For example, the calculation of the optimal steering vector for TC Phanfone, initialized at 0000 UTC 1 October 2014, is shown in Fig. 5. The set of steering vectors of the model and analysis data are compared with the actual motion vectors and residuals, and the magnitudes of difference between the steering vector and the actual motion are shown in Table 3. For the analysis data (model), the residual is at the minimum with a radius of 5° (7°) and the upper averaging boundary at 200 (200) hPa. Therefore, the optimal steering vector of the model and analysis data can identify the environmental winds that are best matched to the actual TC motion (Galarneau and Davis 2013).

Figure 6 shows the mean magnitudes of difference in the optimal steering vector between analysis data and the model for group 1 and 2 TCs. The difference in the steering vector between model and analysis data is related to the error of the environmental winds in the model simulation. The steering vector difference is similar between the two groups at 24 h. The steering vector difference for group 1 TCs increases to 1.1 m s^{-1} at 48 h while that for group 2 TCs slightly decreases. Although the steering vector difference for group 2 TCs also increases after 60 h, group 1 TCs still have a larger steering vector difference compared to group 2 TCs. It is notable that the steering vector difference is always great for group 1 TCs.

Since TC motion is highly influenced by the surrounding environmental winds, the larger difference in the steering wind could have induced the larger track error. The track error has correlation with the steering wind difference (not shown). This result corresponds that the mean track error difference between two groups is increasing after 48 h (Fig. 3b). Although the difference for group 2 TCs also increases after 72 h, the track error of group 1 TCs is large for longer forecast time. Thus, it seems that the wind error is continuously induced from unrealistic simulation of midlatitude pressure system, resulting in a large track error in the 5-day forecast.

Group 1 TCs have a higher upper boundary than group 2 TCs in calculating optimal steering vector for analysis data (Table 4). It is because group 1 TCs tend to be related to the midlatitude trough as well as the western North Pacific subtropical high (WNPSH). The percentage of the 200- and 250-hPa upper boundary for group 1 is 54% while that for group 2 is 34%. For the model, the percentage of the 200- and 250-hPa upper boundary is 49% and 48% for group 1 and 2, respectively. There is an inverse correlation between 200-hPa geopotential height forecast skill and TC track error (not shown) indicating that the TC track forecast error can be increased when model is not simulating the upper-level atmosphere correctly.

The environmental wind fields are highly associated with the pressure distribution in the atmosphere. Figure 7 shows the pattern correlation of 500-hPa geopotential height between the WRF simulation and the GFS analysis. The pattern correlation of the pressure system shows a difference in aspect according to the direction of the TC movement. The pattern correlation for group 1 TCs is lower, especially at the late forecast time. Group 1 TCs are profoundly affected by the synoptic weather systems, such as the WNPSH or the midlatitude trough. It seems that lower pattern correlation for group 1 TCs arises from the unrealistic representation of environmental fields affected by the subtropical high or midlatitude trough. Since a poorer simulation of the environmental wind fields leads to a large error in track forecast, improved simulation of the synoptic fields, including the pressure distribution, is essential for producing an accurate 5-day track forecast.

Case studies for TC Lekima initialized at 0000 UTC 21 October 2013, and TC Phanfone initialized at 0000 UTC

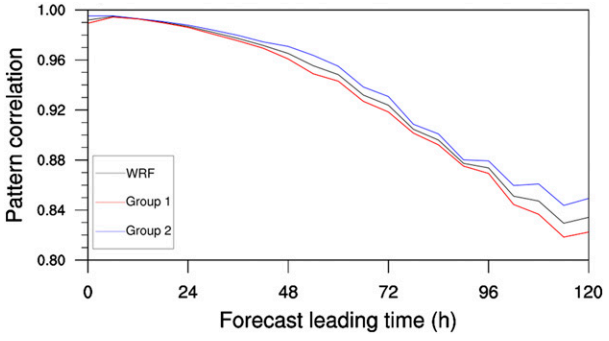

FIG. 7. Time series of pattern correlations for 500-hPa geopotential height for the WRF (black), for group 1 TCs (red) and group 2 TCs (blue).

1 October 2014 are conducted to examine the impact of the pressure system on TC track forecast. In both cases, the WNPSH is developed at the eastern boundary of the model domain, and it can induce anticyclonic flow along the 5880-gpm contour of 500-hPa geopotential height shown in Fig. 8. This flow leads the TC to track northward and curve northeastward. In the simulation of Lekima, the WNPSH is realistically simulated, and then the TC track error is small at all forecast times. In the case of Phanfone, WRF could not properly simulate the TC track moving northward south of 30°N or curving to the northeast north of that. Since the model overestimates the strength of the WNPSH, the simulated 5800-gpm line is expanded more toward the west compared to the analysis data. The stronger WNPSH does not allow the WRF-simulated TC to move toward Japan but instead pushes the TC more toward the west, as compared to the analysis data (Cha et al. 2011; Sun et al. 2017).

The deficiency in correctly simulating the large-scale pressure fields corresponds to the environmental wind errors; therefore, it can cause TC track errors. Figure 9 shows the difference in the optimal steering vector between model and analysis data for two TCs at each forecast time. For the simulation of Lekima, there seems to be little difference in the optimal steering vector between model and analysis data. The differences are less than 0.5 m s^{-1} for all forecast times, since the WNPSH and the associated wind fields are realistically simulated. For the simulation of Phanfone, the simulated WNPSH is expanded to the north of the TC initially. The expanded WNPSH inhibits the northward motion of the TC, and it drives TC to deviate to the south from the actual motion. Therefore, the optimal steering vector of WRF is directed to the south compared to that of the analysis data during the first 48 h. After 48 h, the WRF Model cannot simulate the recurving of the TC since the expanded WNPSH induces a relatively southwestward flow. Consequently, the cyclone moves southwestward compared to the actual motion, and the track error becomes large. This suggests that the difference in the optimal steering vector between the model and analysis data, or errors in the environmental winds, can be a critical cause of track error of more than 2 days for the northward moving TCs.

Figure 10 shows the mean magnitudes of the residual for both model and analysis data. The residual is the difference between the actual motion vector and the optimal steering vector, and indicate the difference between the TC translation speed and steering wind. Since the environmental wind field cannot perfectly explain the TC motion, small residual terms exist. According to Galarneau and Davis (2013), the residual contains the features not considered in the environmental wind, that is, storm-scale process, asymmetric convection distribution, or beta-related effect. A larger residual value indicates that TC track with factors other than the environmental winds. In the analysis data, the magnitude of the residual is generally larger for group 1 TCs than group 2 TCs. The residual slightly increases with corresponding forecast time but does not significantly change, with a value of approximately 0.4–0.5 m s^{-1} for two groups. The large residuals for group 1 TCs can be associated with the fact that they are driven by a more complicated process than group 2 TCs because they propagate

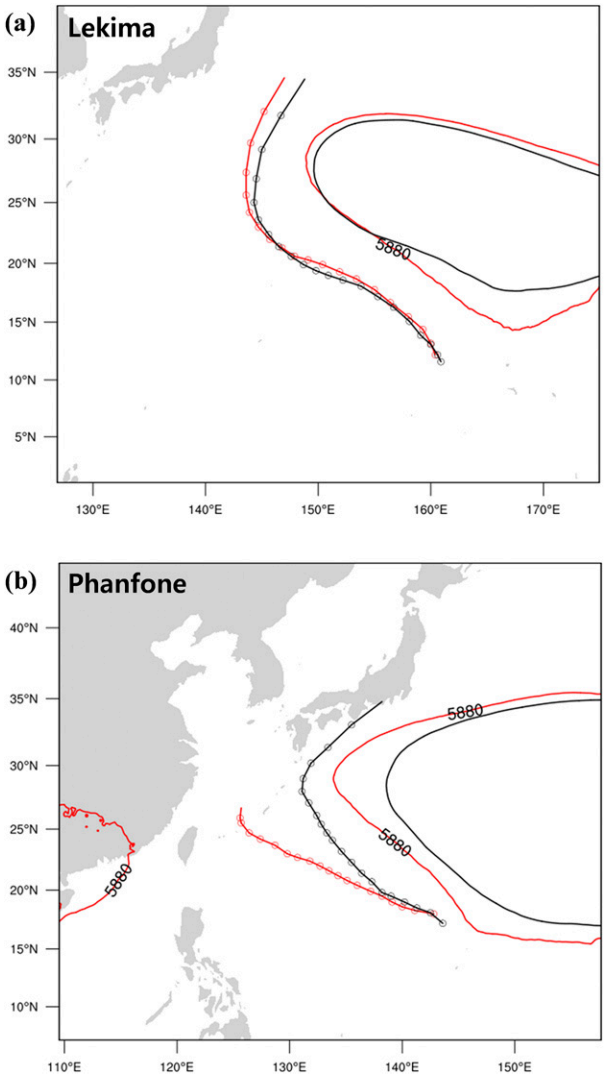

FIG. 8. The 5-day mean 5880-gpm line of 500-hPa geopotential height for the model (red contour) and analysis data (red contour), and TC tracks of the model (red dotted line) and GFS analysis (black dotted line) for (a) TC Lekima and (b) TC Phanfone initialized at 21 Oct 2013 and 1 Oct 2014, respectively.

into the baroclinic midlatitude. For the model, the mean residual value for group 2 TCs is mostly constant after 36 h (approximately 0.6 m s^{-1}) despite it being slightly larger than that of analysis data. Conversely, for group 1 TCs, the mean residual is much higher than the analysis data. It rapidly increases after 36 h and reaches 1.5 m s^{-1} at 108 h, which is the three times that of the analysis data. This means that simulated TC motion for group 1 TCs deviates more from the motion expected by the environmental wind field.

To determine the characteristic of the residual term, the deviation of the actual motion vector from the optimal steering vector for group 1 TCs are analyzed in Fig. 11. The cross and along components of the residual are calculated for the late forecast times (72, 84, 96, and 108 h) when the mean residual of

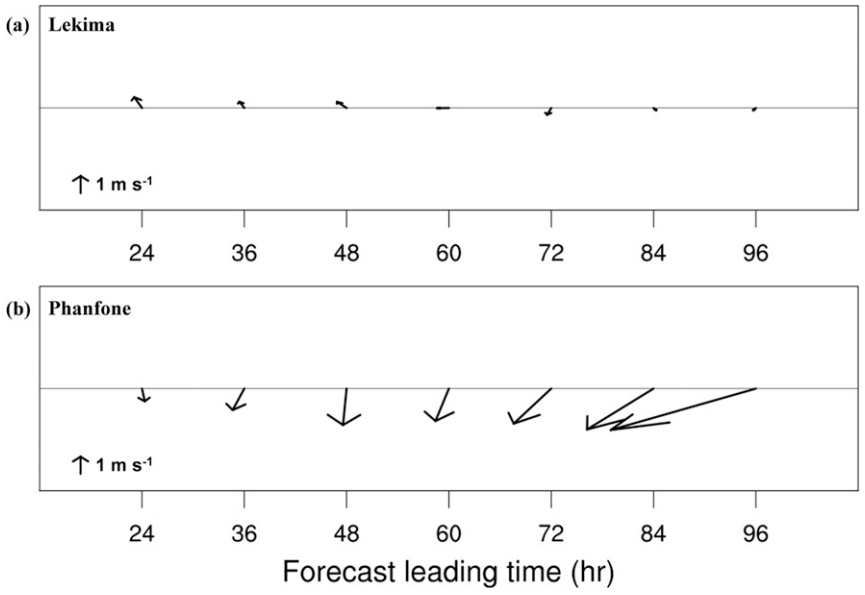

FIG. 9. Vector difference of the optimal steering vectors between the model and analysis data at each forecast time for (a) TC Lekima and (b) TC Phanfone initialized at 0000 UTC 21 Oct 2013 and 0000 UTC 1 Oct 2014, respectively.

model for group 1 TCs is significantly large. The cross (along) component is defined as the component of the residual vector perpendicular (parallel) to the transition direction of the TC. In the analysis of the cross and along component for group 1 TCs, the negative cross component cases are more dominant than the positive case. A negative value of cross component means that the actual motion is shifted to the left of the steering vector. This indicates that in a lot of cases, the TC motion deviated leftward from the steering vector in the midlatitude. It seems that the large residual value of group 1 TCs arises from the unrealistic simulation of storm-scale processes in the midlatitude baroclinic atmosphere. The considerable residual

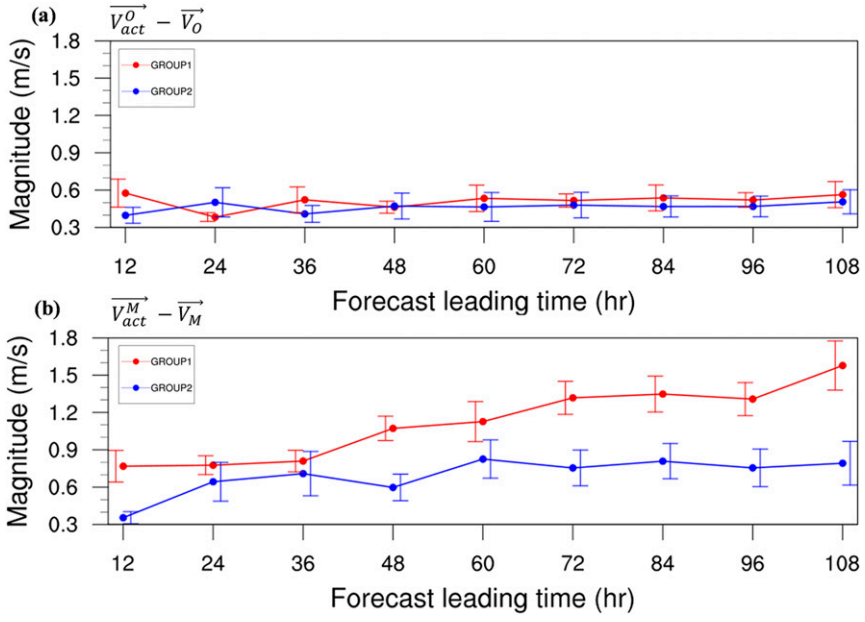

FIG. 10. Time series of the difference between the optimal steering vector and actual TC motion of (a) analysis data and (b) the model for group 1 TCs (red) and group 2 TCs (blue). The bars mean the 90% confidence interval.

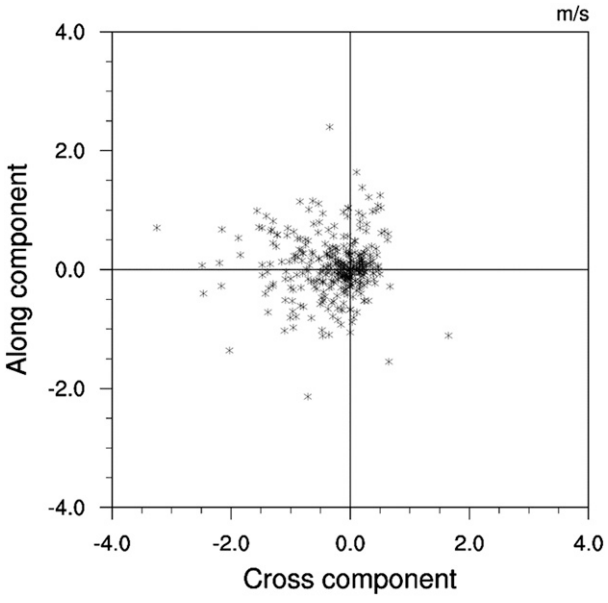

FIG. 11. Cross and along components of the vector difference of the actual TC motion vector from the optimal steering vector for group 1 TCs at 72, 84, 96, and 108 h of forecast time.

value may explain the significant track error for group 1 TCs. To improve the simulation of storm-scale processes, therefore, higher-resolution forecast or more sophisticated physical schemes could solve this issue.

4. Summary and conclusions

In this study, the track forecast error of the WRF Model for TCs over the western North Pacific was analyzed. Forecasts were conducted for the 125 TCs formed over the western North Pacific in 2013–17. The mean track error of WRF was similar or slightly larger compared to that of GFS. A cluster analysis was applied based on TC location through the *k*-means clustering algorithm, and five clusters were classified. Since there were robust differences in the track error between clusters with dissimilar transition directions, we grouped the clusters into two groups: group 1 (C1 and C2) and group 2 (C3 and C4). The two groups had apparent differences in track errors. A higher number of large error cases were included in group 1 TCs, while more small error cases were included in group 2 TCs. To understand this regional difference in track error, we analyzed the optimal steering vectors. The steering vector difference between model and analysis data, which indicates the environmental wind error, was larger for group 1 TCs, especially at the late forecast lead time. This result corresponded to the rapidly increasing track error of group 1 TCs after 48 h. The pattern correlation of 500-hPa geopotential height, which is highly related to environmental wind, was also lower for group 1 TCs. This showed that the large track error for group 1 TCs arises from the unrealistic representation of environmental fields affected by the subtropical high or midlatitude trough. The residuals, which contained the features not considered in the environmental wind, were also larger for group 1 TCs. The significant forecast error for group 1 TCs at the late forecast time can be attributed to the fact that the environmental wind error is larger for group 1 TCs. In addition, the more complicated processes of midlatitude TCs are not adequately simulated.

We utilized the WRF Model with the horizontal grid spacing of 12 km comparable to GFS forecast. This is because the WRF Model has a consistency over target years, while there is an inconsistency in the GFS forecast due to the updates of the model system. In addition, 12-km grid spacing is used to verify the compatibility of such resolution for TC forecast and figure out the limitation of it. This study suggests that higher resolution is necessary for improving storm-scale processes. For the further study, higher resolution experiments with moving nesting domain will be conducted, and the impact of increasing resolution on track and intensity forecasts will be studied. The forecast domain in this study is varies with the TC track to minimize the computational cost. The larger domain can give advantages to include more large-scale features affecting TC activities (e.g., subtropical high and midlatitude trough), but at the same time, it can lead to disadvantage such as forecast errors induced by internal factors (e.g., steep Tibetan plateau, strong air–sea interaction, and uncertain soil moisture content). This topic on optimal domain size needs further research.

Diagnosing model forecast error is an essential task in terms of identifying the problem and establishing the solution for model development. This study examined the track forecast of the WRF Model for a large number of TCs over the WNP using the optimal steering vector, and showed that different TC track patterns can lead to different sources of track error through the clustering analysis. Because the WRF Model has been widely used for TC forecasting in the East Asian countries, this study can serve as a useful reference for researchers when simulating and analyzing TCs over the WNP using the same model. However, the predictability of large-scale features in regional model is still an ongoing research issue. Previous studies have shown that improved initialization methods and the application of spectral nudging could improve the predictability of a large-scale field in a regional model. In addition, storm-scale processes are also realistically considered in the model simulation. Using smaller horizontal grid spacing and advanced physics parameterizations, as well as more realistic air–sea interaction can be helpful with this issue.

Acknowledgments. This research was supported by Next-Generation Information Computing Development Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT (NRF-2016M3C4A7952637). This work was carried out through the Numerical Weather Prediction R&D Project funded by the Republic of Korea Air force (ROKAF). YM was supported in part by the DOE Regional and Global Model Analysis (RGMA) through Grant DE-SC0016223 and NOAA Modeling, Analysis, Predictions and Projections (MAPP) through Grant NA18OAR4310276.

Data availability statement. The Global Forecast System (GFS) forecast and analysis data for this study are available

from National Oceanic and Atmospheric Administration (NOAA)'s National Operational Model Archive and Distribution System at http://nomads.ncep.noaa.gov.

REFERENCES

Aberson, S. D., 2010: 10 years of hurricane synoptic surveillance (1997–2006). *Mon. Wea. Rev.*, **138**, 1536–1549, https://doi.org/10.1175/2009MWR3090.1.

Camargo, S. J., A. W. Robertson, S. J. Gaffney, P. Smyth, and M. Ghil, 2007a: Cluster analysis of typhoon tracks. Part I: General properties. *J. Climate*, **20**, 3635–3653, https://doi.org/10.1175/JCLI4188.1.

——, ——, ——, ——, and ——, 2007b: Cluster analysis of typhoon tracks. Part II: Large-scale circulation and ENSO. *J. Climate*, **20**, 3654–3676, https://doi.org/10.1175/JCLI4203.1.

——, ——, A. G. Barnston, and M. Ghil, 2008: Clustering of eastern North Pacific tropical cyclone tracks: ENSO and MJO effects. *Geochem. Geophys. Geosyst.*, **9**, Q06V05, https://doi.org/10.1029/2007GC001861.

Carr, L. E., III, and R. L. Elsberry, 2000a: Dynamical tropical cyclone track forecast errors. Part I: Tropical region error sources. *Wea. Forecasting*, **15**, 641–661, https://doi.org/10.1175/1520-0434(2000)015<0641:DTCTFE>2.0.CO;2.

——, and ——, 2000b: Dynamical tropical cyclone track forecast errors. Part II: Midlatitude circulation influences. *Wea. Forecasting*, **15**, 662–681, https://doi.org/10.1175/1520-0434(2000)015<0662:DTCTFE>2.0.CO;2.

Cha, D. H., C. S. Jin, D. K. Lee, and Y. H. Kuo, 2011: Impact of intermittent spectral nudging on regional climate simulation using Weather Research and Forecasting model. *J. Geophys. Res.*, **116**, D10103, https://doi.org/10.1029/2010JD015069.

Chan, J. C., 2005: The physics of tropical cyclone motion. *Annu. Rev. Fluid Mech.*, **37**, 99–128, https://doi.org/10.1146/annurev.fluid.37.061903.175702.

——, and W. M. Gray, 1982: Tropical cyclone movement and surrounding flow relationships. *Mon. Wea. Rev.*, **110**, 1354–1374, https://doi.org/10.1175/1520-0493(1982)110<1354:TCMASF>2.0.CO;2.

Chen, G., X. Zhang, L. Bai, and R. Wan, 2019: Verification of tropical cyclone operational forecast in 2018. 51st Session of ESCAP/WMO Typhoon Committee, Guangzhou, China, 20 pp.

Chen, S. S., J. F. Price, W. Zhao, M. A. Donelan, and E. J. Walsh, 2007: The CBLAST-Hurricane program and the next-generation fully coupled atmosphere–wave–ocean models for hurricane research and prediction. *Bull. Amer. Meteor. Soc.*, **88**, 311–318, https://doi.org/10.1175/BAMS-88-3-311.

Davis, C., and Coauthors, 2008: Prediction of landfalling hurricanes with the Advanced Hurricane WRF Model. *Mon. Wea. Rev.*, **136**, 1990–2005, https://doi.org/10.1175/2007MWR2085.1.

——, and Coauthors, 2011: High-resolution hurricane forecasts. *Comput. Sci. Eng.*, **13**, 22–30, https://doi.org/10.1109/MCSE.2010.74.

Dong, K., and C. J. Neumann, 1986: The relationship between tropical cyclone motion and environmental geostrophic flows. *Mon. Wea. Rev.*, **114**, 115–122, https://doi.org/10.1175/1520-0493(1986)114<0115:TRBTCM>2.0.CO;2.

Dudhia, J., 1989: Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.*, **46**, 3077–3107, https://doi.org/10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2.

Elsner, J. B., 2003: Tracking hurricanes. *Bull. Amer. Meteor. Soc.*, **84**, 353–356, https://doi.org/10.1175/BAMS-84-3-353.

——, and K. Liu, 2003: Examining the ENSO-typhoon hypothesis. *Climate Res.*, **25**, 43–54, https://doi.org/10.3354/cr025043.

Feser, F., and H. von Storch, 2008: A dynamical downscaling case study for typhoons in Southeast Asia using a regional climate model. *Mon. Wea. Rev.*, **136**, 1806–1815, https://doi.org/10.1175/2007MWR2207.1.

Fierro, A. O., R. F. Rogers, F. D. Marks, and D. S. Nolan, 2009: The impact of horizontal grid spacing on the microphysical and kinematic structures of strong tropical cyclones simulated with the WRF-ARW model. *Mon. Wea. Rev.*, **137**, 3717–3743, https://doi.org/10.1175/2009MWR2946.1.

Galarneau, T. J., Jr., and C. A. Davis, 2013: Diagnosing forecast errors in tropical cyclone motion. *Mon. Wea. Rev.*, **141**, 405–430, https://doi.org/10.1175/MWR-D-12-00071.1.

Gentry, M. S., and G. M. Lackmann, 2010: Sensitivity of simulated tropical cyclone structure and intensity to horizontal resolution. *Mon. Wea. Rev.*, **138**, 688–704, https://doi.org/10.1175/2009MWR2976.1.

George, J. E., and W. M. Gray, 1976: Tropical cyclone motion and surrounding parameter relationships. *J. Appl. Meteor.*, **15**, 1252–1264, https://doi.org/10.1175/1520-0450(1976)015<1252:TCMASP>2.0.CO;2.

Gopalakrishnan, S. G., F. Marks Jr., X. Zhang, J.-W. Bao, K.-S. Yeh, and R. Atlas, 2011: The experimental HWRF system: A study on the influence of horizontal resolution on the structure and intensity changes in tropical cyclones using an idealized framework. *Mon. Wea. Rev.*, **139**, 1762–1784, https://doi.org/10.1175/2010MWR3535.1.

Hall, T. M., and S. Jewson, 2007: Statistical modelling of North Atlantic tropical cyclone tracks. *Tellus*, **59A**, 486–498, https://doi.org/10.1111/j.1600-0870.2007.00240.x.

Harr, P. A., and R. L. Elsberry, 1991: Tropical cyclone track characteristics as a function of large-scale circulation anomalies. *Mon. Wea. Rev.*, **119**, 1448–1468, https://doi.org/10.1175/1520-0493(1991)119<1448:TCTCAA>2.0.CO;2.

——, and ——, 1995a: Large-scale circulation variability over the tropical western North Pacific. Part I: Spatial patterns and tropical cyclone characteristics. *Mon. Wea. Rev.*, **123**, 1225–1246, https://doi.org/10.1175/1520-0493(1995)123<1225:LSCVOT>2.0.CO;2.

——, and ——, 1995b: Large-scale circulation variability over the tropical western North Pacific. Part II: Persistence and transition characteristics. *Mon. Wea. Rev.*, **123**, 1247–1268, https://doi.org/10.1175/1520-0493(1995)123<1247:LSCVOT>2.0.CO;2.

Hartigan, J. A., and M. A. Wong, 1979: Algorithm AS 136: A *k*-means clustering algorithm. *J. Roy. Stat. Soc.*, **28C**, 100–108, https://doi.org/10.2307/2346830.

Hodanish, S., and W. M. Gray, 1993: An observational analysis of tropical cyclone recurvature. *Mon. Wea. Rev.*, **121**, 2665–2689, https://doi.org/10.1175/1520-0493(1993)121<2665:AOAOTC>2.0.CO;2.

Hodges, K. I., and R. Emerton, 2015: The prediction of Northern Hemisphere tropical cyclone extended life cycles by the ECMWF ensemble and deterministic prediction systems. Part I: Tropical cyclone stage. *Mon. Wea. Rev.*, **143**, 5091–5114, https://doi.org/10.1175/MWR-D-13-00385.1.

Hong, S.-Y., and J.-O. J. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc.*, **42**, 129–151.

——, Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134**, 2318–2341, https://doi.org/10.1175/MWR3199.1.

Jin, H., M. S. Peng, Y. Jin, and J. D. Doyle, 2014: An evaluation of the impact of horizontal resolution on tropical cyclone predictions using COAMPS-TC. *Wea. Forecasting*, **29**, 252–270, https://doi.org/10.1175/WAF-D-13-00054.1.

Kain, J. S., 2004: The Kain–Fritsch convective parameterization: An update. *J. Appl. Meteor.*, **43**, 170–181, https://doi.org/10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2.

Kehoe, R. M., M. A. Boothe, and R. L. Elsberry, 2007: Dynamical tropical cyclone 96-and 120-h track forecast errors in the western North Pacific. *Wea. Forecasting*, **22**, 520–538, https://doi.org/10.1175/WAF1002.1.

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. *J. Climate*, **24**, 488–508, https://doi.org/10.1175/2010JCLI3751.1.

Krzanowski, W. J., and Y. Lai, 1988: A criterion for determining the number of groups in a data set using sum-of-squares clustering. *Biometrics*, **44**, 23–34, https://doi.org/10.2307/2531893.

Lander, M. A., 1996: Specific tropical cyclone track types and unusual tropical cyclone motions associated with a reverse-oriented monsoon trough in the western North Pacific. *Wea. Forecasting*, **11**, 170–186, https://doi.org/10.1175/1520-0434(1996)011<0170:STCTTA>2.0.CO;2.

Liu, B., and L. Xie, 2012: A scale-selective data assimilation approach to improving tropical cyclone track and intensity forecasts in a limited-area model: A case study of Hurricane Felix (2007). *Wea. Forecasting*, **27**, 124–140, https://doi.org/10.1175/WAF-D-10-05033.1.

MacQueen, J., 1967: Some methods for classification and analysis of multivariate observations. *Proc. Fifth Berkeley Symp. on Mathematical Statistics and Probability*, Oakland, CA, University of California, 281–297.

Magnusson, L., J. D. Doyle, W. A. Komaromi, R. D. Torn, C. K. Tang, J.C.L. Chan, M. Yamaguchi, and F. Zhang, 2019: Advances in understanding difficult cases of tropical cyclone track forecasts. *Trop. Cyclone Res. Rev.*, **8**, 109–122, https://doi.org/10.1016/j.tcrr.2019.10.001.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.*, **102**, 16 663–16 682, https://doi.org/10.1029/97JD00237.

Moon, J., D. H. Cha, M. Lee, and J. Kim, 2018: Impact of spectral nudging on real-time tropical cyclone forecast. *J. Geophys. Res. Atmos.*, **123**, 12 647–12 660, https://doi.org/10.1029/2018JD028550.

Nakamura, J., U. Lall, Y. Kushnir, and S. J. Camargo, 2009: Classifying North Atlantic tropical cyclone tracks by mass moments. *J. Climate*, **22**, 5481–5494, https://doi.org/10.1175/2009JCLI2828.1.

Neumann, C. J., and J. M. Pelissier, 1981: Models for the prediction of tropical cyclone motion over the North Atlantic: An operational evaluation. *Mon. Wea. Rev.*, **109**, 522–538, https://doi.org/10.1175/1520-0493(1981)109<0522:MFTPOT>2.0.CO;2.

Noh, Y., W. Cheon, S. Hong, and S. Raasch, 2003: Improvement of the K-profile model for the planetary boundary layer based on large eddy simulation data. *Bound.-Layer Meteor.*, **107**, 401–427, https://doi.org/10.1023/A:1022146015946.

Nystrom, R. G., F. Zhang, E. B. Munsell, S. A. Braun, J. A. Sippel, Y. Weng, and K. Emanuel, 2018: Predictability and dynamics of Hurricane Joaquin (2015) explored through convection-permitting ensemble sensitivity experiments. *J. Atmos. Sci.*, **75**, 401–424, https://doi.org/10.1175/JAS-D-17-0137.1.

Pielke, R. A., Jr., J. Gratz, C. W. Landsea, D. Collins, M. A. Saunders, and R. Musulin, 2008: Normalized hurricane damage in the United States: 1900–2005. *Nat. Hazards Rev.*, **9**, 29–42, https://doi.org/10.1061/(ASCE)1527-6988(2008)9:1(29).

Powell, M. D., and S. D. Aberson, 2001: Accuracy of United States tropical cyclone landfall forecasts in the Atlantic basin (1976–2000). *Bull. Amer. Meteor. Soc.*, **82**, 2749–2768, https://doi.org/10.1175/1520-0477(2001)082<2749:AOUSTC>2.3.CO;2.

Raju, P., J. Potty, and U. Mohanty, 2011: Sensitivity of physical parameterizations on prediction of tropical cyclone Nargis over the Bay of Bengal using WRF model. *Meteor. Atmos. Phys.*, **113**, 125–137, https://doi.org/10.1007/s00703-011-0151-y.

Short, C. J., and J. Petch, 2018: How well can the Met Office Unified Model forecast tropical cyclones in the western North Pacific? *Wea. Forecasting*, **33**, 185–201, https://doi.org/10.1175/WAF-D-17-0069.1.

Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., https://doi.org/10.5065/D68S4MVH.

Sun, Y., and Coauthors, 2017: Impact of ocean warming on tropical cyclone track over the western North Pacific: A numerical investigation based on two case studies. *J. Geophys. Res. Atmos.*, **122**, 8617–8630, https://doi.org/10.1002/2017JD026959.

Torn, R. D., T. J. Elless, P. P. Papin, and C. A. Davis, 2018: Tropical cyclone track sensitivity in deformation steering flow. *Mon. Wea. Rev.*, **146**, 3183–3201, https://doi.org/10.1175/MWR-D-18-0153.1.

Velden, C. S., and L. M. Leslie, 1991: The basic relationship between tropical cyclone intensity and the depth of the environmental steering layer in the Australian region. *Wea. Forecasting*, **6**, 244–253, https://doi.org/10.1175/1520-0434(1991)006<0244:TBRBTC>2.0.CO;2.

Waldron, K. M., J. Paegle, and J. D. Horel, 1996: Sensitivity of a spectrally filtered and nudged limited-area model to outer model options. *Mon. Wea. Rev.*, **124**, 529–547, https://doi.org/10.1175/1520-0493(1996)124<0529:SOASFA>2.0.CO;2.

Yamada, H., T. Nasuno, W. Yanase, and M. Satoh, 2016: Role of the vertical structure of a simulated tropical cyclone in its motion: A case study of Typhoon Fengshen (2008). *SOLA*, **12**, 203–208, https://doi.org/10.2151/sola.2016-041.

Zhang, Q., L. Wu, and Q. Liu, 2009: Tropical cyclone damages in China 1983–2006. *Bull. Amer. Meteor. Soc.*, **90**, 489–496, https://doi.org/10.1175/2008BAMS2631.1.