Ensemble Based Diagnosis of the Track Errors of Super Typhoon Mangkhut (2018)

Ling HUANG*, Qilin WAN, Chunxia LIU, and Huijun HUANG

Institute of Tropical and Marine Meteorology of China Meteorological Administration, Guangzhou 510000

(Received July 8, 2019; in final form January 2, 2020)

ABSTRACT

Super Typhoon Mangkhut (2018) was the most high-impact typhoon in 2018 because of its long lifespan and significant intensity. The operational track forecasts in the short-to-medium range (deterministic and probabilistic forecast) showed a great uncertainty and the forecast landing points varied with different lead times. This study applied ensembles of high-resolution ECMWF forecasts to investigate the major factors and mechanisms of the bias production of the Mangkhut forecast track. The ensembles with the largest track bias were analyzed to examine the possible bias associated factors. The results suggested that environmental steering flows were the main cause for the erroneous southward track error with a variance contribution of 72%. The tropical cyclone (TC) size difference and the interaction of the TC with the subtropical high (SH) were other two key factors that contributed to the track error. Particularly, larger TCs may have led to a stronger erosion of the southern part of the SH, and thus induced significant changes in the large-scale environment and eventually resulted in an additional northward movement of TC.

Key words: Super Typhoon Mangkhut, ECMWF ensemble, tropical cyclone track

Citation: Huang, L., Q. L. Wan, C. X. Liu, et al., 2020: Ensemble based diagnosis of the track errors of Super Typhoon Mangkhut (2018). J. Meteor. Res., 34(2), 353–367, doi: 10.1007/s13351-020-9086-x.

1. Introduction

Short-range (less than 3 days) operational forecasts of tropical cyclone (TC) tracks have made significant progress over the past four decades owing to the substantial improvement of numerical weather prediction (NWP) models and the operational application of multi-model consensus techniques (Sheets, 1990; Goerss and Jeffries, 1994; Heming and Radford, 1998; Williford et al., 1998; Rappaport et al., 2009; Brown et al., 2010; Peng et al., 2017). However, significant errors still exist in short-to-medium range (approximately 5 days) forecasts (Hamill et al., 2013). For instance, some 5-day forecasts of Super Typhoon Megi (2010), which made a sharp turn to the north, had a forecast track error greater than 1500 km (Qian et al., 2013). In the requirement for extending the lead time (out to 5–7 days ahead) mentioned in the last International TC Workshop (ITCW) in 2014, issuing TC positions 5 days ahead has become a global standard. In addition to TC-related strong winds, storm surges, and flash flooding, the TC tornado risk to coastal regions is also closely associated with TC tracks (e.g., Bai et al., 2017, 2020). Providing accurate TC track forecast days prior to the landfall is urgent and essential for disaster preparedness.

Many previous studies have highlighted the complexity of TC movements associated with storm structure errors (Wang and Holland, 1996; Carr III and Elsberry, 2000a), environmental wind errors (e.g., Carr III and Elsberry, 2000b; Majumdar et al., 2006; McTaggart-Cowan et al., 2006; Peng and Reynolds, 2006; Kehoe et al., 2007; Wu et al., 2007, 2009; Chen et al., 2009; Galarneau and Davis, 2013), and multiscale interactions of TCs with the background environment (e.g., Lander and Holland, 1993; Wu and Emanuel, 1993; Ritchie and Holland, 1997; Simpson et al., 1997). A deterministic forecast alone may not be adequate to capture the most likely TC track. To avoid the limitation of the deterministic forecast, a number of operational centers, such as the ECMWF, US NCEP, China Meteorological Adminis-
tradition (CMA), Korean Meteorological Administration (KMA), and Japan Metrological Agency (JMA), have built ensemble prediction systems for the operational probability forecasts. The Observing System Research and Predictability Experiment (THORPEX) of the World Meteorological Organization established a data archive of operational ensemble predictions, named the THORPEX International Grand Global Ensemble (TIGGE; Bougeault et al., 2010), with the aim of improving the accuracy of 1-day to 2-week high-impact weather forecasts and advancing the development of ensemble forecasting techniques (Richardson et al., 2005). These ensemble prediction systems have opened up opportunities for researchers to use ensemble data for a wide range of studies, particularly on predictability and dynamical processes.

Previous studies have applied ensemble forecasts to diagnose forecast errors of TC track. Munsell and Zhang (2014) demonstrated that the erroneous eastern track forecasts of one of the costliest natural disasters in US history, Hurricane Sandy (2012), were related to the meridional component of steering flows. In addition, the ensemble-based analysis also suggested that the track errors were closely related to the selections of dynamics cores, physical parameterizations, spatial resolutions of the model, stochastic physics algorithms, initial condition uncertainties (Bassill, 2014; Magnusson et al., 2014; Melhauser et al., 2017), the amplitude of the ridge on the poleward side of Sandy (Magnusson et al., 2014), and the interaction of Sandy with midlatitude circulations (Elsberry et al., 2014). Torn et al. (2015) further found that the eastward track of Sandy was characterized by a lower-amplitude upper-tropospheric anticyclone on the poleward side of Sandy, which in turn was associated with westerly perturbation steering flows. In addition, the TC timing and location that involved mesoscale processes and land–ocean processes (Elsberry et al., 2014) and the structure and amplitude of the initial wind perturbations (Yamaguchi and Majumdar, 2010) were found to be associated with the TC track errors. These above mentioned findings suggested that the complex factors that contributed to the TC track errors can be revealed from the perspective of ensemble-based analysis.

Super Typhoon Mangkhut (2018) was the most severe TC in 2018, which affected tens of millions of lives and caused billions of dollars (US $) in economic costs during its landfall in the Philippines and China. It is of great importance to improve the forecast of TC landing points of such major typhoons to avoid unnecessary warnings and inefficient allocations of emergency management resources. This study attempts to evaluate the possible error sources that contributed to the uncertainties of the short-to-medium range TC track forecasts based on the ensemble-based analysis, including the impacts of environmental flows, TC structures, and the interaction of TC and background environment. To avoid the influence of different initial conditions and model formulations in diagnosing the source of track variability by comparing forecasts from two or more models, this study employs a large ensemble of forecasts from the ECMWF provided by the TIGGE dataset. The dynamical processes leading to the errors of TC tracks were investigated to address the following issues: 1) how the error of environmental flows contributed to the TC track error and 2) how the TC regulated the environment, which then fed back onto its motion. Ensemble members with better TC tracks were compared with the members with track bias (compared with the observations). Statistical analysis was also implemented to quantitatively discuss the correlation between TC track errors and the abovementioned possible factors.

Section 2 introduces the data and methods. An overview of Super Typhoon Mangkhut (2018) and its forecasts are presented in Section 3. The major results for the factors governing the track forecasts revealed by the ECMWF ensemble are shown in Section 4. Finally, a summary and discussion are provided in Section 5.

2. Data and methods

2.1 Data

The 5-day ensemble forecasts and high-resolution deterministic forecasts were provided by the ECMWF and downloaded from the TIGGE data archive. The ECMWF originally provided a 50-member ensemble forecast with the T639 spectral truncation and a deterministic forecast with a resolution of approximately 9 km. The ensemble forecast was then commonly interpolated (by TIGGE) onto $0.5^\circ \times 0.5^\circ$ horizontal grids for the ensemble forecast and $0.1^\circ \times 0.1^\circ$ for the high-resolution forecast at intervals of 6 h. The initial perturbations of the ECMWF ensemble were constructed from a combination of initial singular vectors (as before) and ensemble of data assimilation-based perturbations, which replace the evolved singular vectors, since 2010 (Bonavita et al., 2011; Lang et al., 2012). The TC best track data, including the TC positions and central minimum pressure at 6-h intervals, were obtained from the CMA (http://tcdata.typhoon.org/en/index.html; Ying et al., 2014). The observations of maximum wind speed at surface level were
obtained from the quality-controlled automated weather stations (AWSs) provided by the CMA. There were approximately 10,000 AWSs in South China with an average distance of approximately 10 km (Huang and Luo, 2017).

2.2 TC track forecast diagnosed by ECMWF

The TC tracks of the ensemble forecast used in this study were diagnosed by the ECMWF as TC-related NWP products, as recommended by the sixth WMO International Workshop on TCs (available in ftp://tigge@tigge-ldm.ecmwf.int/cxml; Hewson and Titley, 2010). The products provided the diagnosed tracks for the deterministic forecast, the ensemble members, and the analyses, with a precision of 0.1°. The cyclones are detected by using a hybrid of geopotential minimum/vorticity maximum techniques, following the method of König et al. (1993) who allowed the early part of a cyclone’s evolution to sometimes be represented as a low level vorticity maximum at 850 hPa, and the later part of the cyclone be represented as the geopotential minimum at 1000 hPa. This hybrid technique was considered more accurate and reasonable since there could be no geopotential minimum found at the early stage of TC. Additionally, an upper-tropospheric steering wind is employed to estimate future and past positions, using the “half-time tracking” method. Please refer to Hewson and Titley (2010) for more details.

2.3 Absolute angular momentum (AAM) flux

Possible reasons for the difference in the change of TC sizes of the ensemble members, which will be discussed in Section 4, were investigated through the role of lower-level AAM flux (AAMF; Chan and Chan, 2013). It has been indicated that the change in TC size is positively proportional to the change in the lower-tropospheric AAM import (Chan and Chan, 2013). The AAM for a unit mass is given by

\[
\text{AAM} = v_\theta r + \frac{1}{2} f r^2, \tag{1}
\]

where \(v_\theta\) is the tangential wind, \(r\) is the radius from the TC eye, and \(f\) is the Coriolis parameter. The AAMF across a circle centred at the TC eye at a radius \(r\) is given by

\[
\text{AAMF}(r) = \frac{\int_0^{2\pi} \left(v_\theta r + \frac{f r^3}{2}\right) v_r d\theta}{\int_0^{2\pi} r d\theta}, \tag{2}
\]

where \(\theta\) is the azimuth and \(v_r\) is the radial wind. To minimize the uncertainty in understanding the relationship of changes in AAM and TC size, the composite change of AAMF \((\Delta\text{AAMF}_{\text{com}})\) was created by averaging all the changes every 6 h for the focus period:

\[
\Delta\text{AAMF}_{\text{com}} = \frac{\sum_{i=1}^{n-1} [\text{AAMF}(T_{i+1}) - \text{AAMF}(T_i)]}{n-1}, \tag{3}
\]

where \(n\) is the times calculated for the focus period. Such a \(\Delta\text{AAMF}_{\text{com}}\) represents the averaged AAM transport for the entire focus period.

3. Overview of Super Typhoon Mangkhut (2018)

3.1 Case review

Super Typhoon Mangkhut was the most intense TC with a long lifespan over the western North Pacific and South China Sea in 2018. It originated as a tropical depression at approximately 45° longitude east of the Philippines at 0600 UTC 7 September 2018. This tropical depression sufficiently strengthened to officially be named a typhoon within 6 h. Mangkhut then strengthened as it moved northwest in the next 8 days, with a minimum central pressure that dropped to 910 hPa (Fig. 1a). It made landfall in northeastern Philippines with strong winds at 1800 UTC 14 September, causing 65 deaths and affecting millions of residents. Mangkhut landed on the coast of Guangdong Province in China at 0900 UTC 16 September. The observed coastal wind speed reached 35 m s\(^{-1}\) just before it rapidly weakened to a tropical depression by 0600 UTC 17 September (Fig. 1b).

3.2 Southern erroneous track in ECMWF high-resolution deterministic forecasts

The deterministic forecasts of ECMWF showed that the track uncertainty varies with the leading time (Fig. 2). Interestingly, nearly all of the forecast tracks at different initialized time (up to 5 days) have a southward bias compared to the best track (Fig. 2a). The forecast landing points that were initialized earlier (i.e., 0000 UTC 12, 1200 UTC 12, 0000 UTC 13, and 1200 UTC 13 September) were more southward than those from the later initialized forecasts, suggesting that longer-range forecasts were less reliable than shorter-range forecasts (Fig. 2a). The position error can also be presented by the averaged latitude error (the forecast latitude minus the latitude of the best track) every 6 h from 1200 UTC 15 to 0000 UTC 17 September 2018 (Fig. 2b). Note that the averaged latitude error showed an increase from the forecast initialized at 0000 UTC 12 to 1200 UTC 13 September. The forecast initialized at 1200 UTC 12 had a larger error than that at 0000 UTC 12, and the error initialized at 0000 UTC 13 decreased compared with that of the
former forecast (Fig. 2b). The error increase suggested that the forecast track experienced a southward to northward modulation within a time period of 24 h. Overall, the forecast track showed a great uncertainty, especially for longer-range forecasts. The southward erroneous track of nearly all of the forecasts increased the difficulty to predict the exact landing location.

3.3 Southern erroneous track in ECMWF ensemble forecasts

The high-resolution forecasts from ECMWF have been highly respected and widely used by forecasters, as is the ECMWF ensemble forecast. Such high-resolution forecasts were used to help understand the forecast track uncertainty. Figure 3 shows that the ensemble and deterministic high-resolution forecasts initialized from 0000 UTC 12 to 1200 UTC 15 September extending to the final forecast time at 0000 UTC 17 September. The track uncertainty from 1200 UTC 15 to 0000 UTC 17 September gradually decreased as the lead time shortened, as shown by the ensemble spreads of the forecasts initialized at different time (marked by “std” in Fig. 3). The position error in latitude of the forecast track (marked by “err” in Fig. 3) also generally decreased as the lead time shortened. The fact that both of the track error and spread decreased with the decreasing of lead time is consistent with the findings in previous studies (e.g., Qian et al., 2013). Note that the position error of these ensembles experienced an increase among the former three ensemble forecasts (i.e., those initialized at 0000 UTC 12, 1200 UTC 12, and 0000 UTC 13 September), as did their high-resolution deterministic forecasts (Fig. 2). More importantly, a large number of the ensemble members still presented a great southward bias, which caused a negative averaged position error for all of the ensemble forecasts, especially in the ensemble forecasts with a longer lead time (Figs. 3a–d).

The southern erroneous track forecast from the deterministic and ensemble forecasts showed that the ECMWF forecast mostly failed to predict the track of Mangkhut beyond a lead time of 2 days. The fact that track errors increase with the lengthening of lead time gives forecasters little confidence in the track prediction for several days in advance. Many scientific questions arise from the failure in accurately forecasting the track of Mangkhut, especially regarding how to correct the southern erroneous track in the ensemble scenario. To obtain a better forecast for the landing position for days in advance, it is of great value to investigate the reason for the southern erroneous track bias in the longer-range lead time. In this study, we used the ensemble forecast
initialized at 0000 UTC 12 September to discuss the key factors that contributed to the track forecast.

4. Results

4.1 Differences in ensemble TC tracks

The ensemble tracks initialized at 0000 UTC 12 September showed the largest uncertainty and southward bias (Fig. 3a). The averaged position errors in latitude coordinates for the ensemble members varied from –1.5° to 1° and those in longitude coordinates varied from –1.4° to 2° (Fig. 4a). The fact that the computed ensemble mean (–0.32°) of the latitude error is one magnitude greater than that (0.02°) of the longitude error is the main reason that we primarily focused on the errors in latitude coordinates in the present study. Most of the ensemble members had a negative latitude error, as did the analyses and the deterministic forecast, suggesting the south-
ward bias of the forecast positions (Fig. 4a). The top 20% of the ensemble members with the smallest (“good members”) and largest southern position bias (“poor members”) in latitude, respectively, were selected to further investigate the possible reasons for such a southward bias (Fig. 4a). The tracks of the good members were the closest to the observations for the 5-day forecast in terms of the latitude error (Figs. 4a, b). They nearly overlaid the observed best-track estimate for the whole forecast period. The poor members deviated from the observation and moved more southward shortly after the initial time. The southward bias caused its crossing the mainland of the Philippines (Fig. 4c). Based on the differences among the two groups, the possible factors for the track bias will be discussed in the following two subsections.

4.2 Impacts of steering flows on the TC track

The environmental flows is the primary factor in driving TC motion (e.g., Chan and Gray, 1982; Holland, 1983). Previous studies have suggested that the definition of steering flow varies significantly case by case (e.g., George and Gray, 1976; Dong and Neumann, 1986; Velden and Leslie, 1991; Aberson and DeMaria, 1994). For instance, Chan and Gray (1982) found that the midtropospheric flow (700–500 hPa) in a 5°–7° annulus from the TC center has the best correlation with cyclone movement. Sanders et al. (1980), Dong and Neumann (1986), and Velden and Leslie (1991) found that the depth of the steering layer increases with the TC intensity because a stronger TC normally has a deeper convection that must be steered by a higher-layer environmental flow. Therefore, it is necessary to first determine what vertical level(s) constituted Mangkhut’s steering wind.

To take the case-to-case variability of the depth and radius of the steering flow into account, the steering wind for Mangkhut was determined by using the algorithm in Galarneau and Davis (2013), which involves matching...
the TC motion with a mean vector wind averaged over the horizontal area and vertical layer. This method computes the divergence and vorticity associated with the TC vortex within a given distance of the TC center. The Poisson equation is solved for the stream function and velocity potential, and then the nondivergent and irrotational wind associated with the TC vortex is obtained at any pressure level. The environmental wind vector at any location is obtained by the difference between the total wind vector and the TC vortex irrotational and nondivergent vector. A domain-averaged environmental wind is then computed by averaging the environmental wind within a certain radius and vertical depths starting at 850 hPa. In the present study, this process was repeated for different radii and vertical depths to determine the steering level that most closely matches the TC motion within ±12 h.

To determine the steering flow that best matches the actual TC motion for the entire forecast period, the optimal steering flow determined by the mean-absolute difference between Mangkhut’s forecast motion and the area-average environmental wind was computed with various radii ranging from 1°–8° from the TC center and vertical depths ranging from the shallowest layer of 850–700 hPa to the deepest layer of 850–100 hPa. This process was repeated every 6 h between 12 and 108 h. Figure 5 shows that the lowest mean-absolute difference was obtained for a TC removal radius of 4° and the 850–200-hPa layer wind for not only the whole ensemble (Fig. 5a) but also for the good (Fig. 5b) and poor (Fig. 5c) members. As a consequence, the steering wind was defined as the 850–200-hPa layer-average environmental wind within 4° to the TC center, which is similar to that in Galarneau and Davis (2013) for Hurricane Earl (2010) and Torn et al. (2015) for Hurricane Sandy (2012).

The meridional steering flow was further analyzed to examine the impact of steering flows on the southward errors in the TC tracks. As shown in Fig. 6a, the position errors in latitude coordinates for the poor members directly showed the southern biases. The TC position of these members overlaid with the southern biases. The TC position of the time when Mangkhut landed at the Philippines.

The meridional movement of the two groups can be partly explained by the steering flow. Figure 6b shows the meridional steering wind as a function of forecast time. The meridional steering winds of the two groups were mostly southerlies, suggesting the overall southward movement of TC was indeed associated with steer-
ing flows. To quantitatively investigate the relationship between the steering flow and the southward movement, the correlation between the steering flows and the position error was examined. The meridional steering flows averaged every 12 h from 1200 to 1400 UTC 12 September positively correlated with the forecast position errors (latitude coordinates in Figs. 6c–f). Such a positive correlation indicates that the stronger the northward meridional steering flows, the more northward movement of TC. We also noticed that even though the good members generally moved more northward than the poor members, the meridional steering flows of some good members were similar to those of some poor members, especially during the early stages (Figs. 6c–e). This phenomenon suggests that in addition to the influence of steering flows, there probably were other processes that could contribute to the track divergence between the good and poor groups. In the present study, we also examined the TC structure and the TC–SH interaction to discuss how these two factors contributed to the TC track difference between the good and poor groups in the next two subsections.

4.3 Differences in TC tracks associated with TC size

Previous studies implied the change in TC structure might affect the track. By analyzing the tracks of 145 TCs over the western North Pacific, Lee et al. (2010) found that TC tracks varied with TC sizes, which suggests that the change in TC size possibly affects the TC track. Based on numerical experiments, Sun et al. (2014) found that a TC could feedback to the large-scale environmental conditions through thermal effects and thus modify the background circulation and change the TC track. In this study, we examined the relationship of the TC size and track, which may provide some insights for better understanding the mechanisms of the TC track change.

To further quantitatively analyze the impact of the TC track difference, we examined the correlation between the steering flows and the position error. The scatter plots in Figs. 6c–f show the relationship between the averaged meridional steering flow and the mean position error at 72-, 96-, and 120-h forecast times. The red, magenta, and grey dots in (c–f) indicate the good members, poor members, and the other ensemble members. The black line in (c–f) indicates the linear regression.

Fig. 6. (a) Position errors (forecast minus observation) in latitude coordinates for good (red) and poor (blue) members. (b) Meridional steering flow (m s\(^{-1}\)) for good (red) and poor (blue) members, as a function of the forecast hour. The pink and light blue shadings indicate the spread of the members in each group. Scatter plots of the averaged meridional steering flow from forecast time of (c) 12–24, (d) 24–36, (e) 36–48, and (f) 48–60 h (x-axis) vs. mean position error in latitude coordinates averaged by the position error at 72, 96, and 120-h forecast times (y-axis). The red, magenta, and grey dots in (c–f) indicate the good members, poor members, and the other ensemble members. The black line in (c–f) indicates the linear regression.
size on the TC track, the TC size in this work was first defined by the mean azimuth radius of 34 knots of the surface winds, which was widely used to define the TC size (e.g., Wu et al., 2015). Figure 7 shows the evolution of TC size every 12 h. At the initial time (0000 UTC 12 September), the TC sizes in the two groups were nearly the same (Fig. 7). After that, the TCs with more northward component became larger than those with more southward component. The relative size difference between the poor and good members became the largest at 0000 UTC 15 September, which was associated with the difference of the zonal position errors for these two groups after 0000 UTC 13 September (Fig. 6a). The TC sizes of the poor members were larger than those of the good member after 0000 UTC 16 September, which may be associated with the TC approaching to the coasts of China. The friction of the land surface may destruct the TC structure, hence affected the influence of TC size on TC track. Therefore, the TC size comparison after 0000 UTC 16 September will not be discussed. The quantitative relationship between the TC sizes and tracks were also investigated. A positive correlation (approximately 0.6) between a TC’s size and its future latitude error indicated that the larger TC at earlier stages tends to move more northward than the smaller TCs (not shown). Such results suggest that the TC size could influence the TC movement. The larger TC size, the more northward component of the TC.

The initial sizes of the good and poor members were comparable, but they evolved in different sizes along the forecast time. Since the TC track bias was associated with the TC size, it is necessary to investigate the possible reasons for the difference-size enlargement among the two groups. To examine how the TC sizes in the two groups evolved, the role of lower-level AAMF on the TC size change was investigated (e.g., Chan and Chan, 2013). Figure 7 suggests that the difference between the poor and good members became visible at 1200 UTC 12 September and decreased after 0000 UTC 15 September. The composite ΔAAMF was then calculated for both of the good and poor groups (Figs. 8a, b) from 0600 UTC 12 to 0000 UTC 15 September every 6 h. As discussed in Chan and Chan (2013), the change in TC size is positively related to the change in the lower-tropospheric AMM import. As shown in Fig. 8a, the lower-level (1000–850 hPa) AAM import of the good members was strong enough to reach the region within 150 km from the TC center. In contrast, the AAM import of the poor was resisted by the AAM export at approximately 150 km from the TC vortex center (Fig. 8b), which suggests the lower efficiency of the AAM import in enlarging the TC size. Wang and Wang (2013) have also suggested that the radial transport of AAM was the major contributor to the azimuthal mean tangential wind tendency, which can be used to estimate the TC size change. Therefore, with stronger and deeper AAM import, the good members tended to enlarge more than the poor members.

4.4 Interaction of TC and subtropical high (SH)

The location and intensity of the SH are already known to have a major impact on TC motion (Elsberry, 1987; Elsner et al., 2000; Kossin et al., 2010; Colbert and Soden, 2012); when a strong SH extended to the west and south, the TCs remained at low latitudes, whereas when the SH contracted to the east and north, the TCs recurved into the western Atlantic. That is, the straight-moving TCs were associated with a westward extension and strengthening of the SH, whereas the recurving TCs were associated with a weakening of the high, which indicates that the stronger SH may have prevented the TCs from heading north. Meanwhile, TCs also have great impacts on the SH. Ren et al. (2007) found that TCs with different tracks could lead to the changes in the western Pacific SH (WPSH) pattern. Zhong and Hu (2007) pointed out that TC activities in the WNP could weaken the WPSH intensity and significantly influence regional climate over East Asia.

The interaction between TC and SH was discussed in this study to examine another possible factor for the track divergences among these three groups of members. The impact of TC size on the SH extent was examined quantitatively by the distance of the TC center to the SH centroid. The SH centroid was defined as the centroid of the 5880-gpm isopleth. The spatial deviation of the SH centroid is associated with the TC structure. Given that the large-scale SH is relatively quasi-stationary within a
short period, the interaction of a TC and SH would “erode” the 5880-gpm isopleth toward northeast, which made the computed SH centroid be located farther northeast to the TC center. In Fig. 9, the computed distances of SH centroids and TC centers for good members are always longer than those for poor members, even though the TC centers in good members are located more northeast. The results suggest that TCs with larger sizes had stronger effects on the shrinking of the SH extent.

Analyses within the framework of a moving vortex were performed to further study the TC–SH interaction. We calculated the correlation between the 500-hPa geopotential height and the future average forecast position error in latitude coordination (forecast – observation; Fig. 10) based on the moving-vortex framework. The negative correlation region in the north of the TC suggested that the stronger the SH to the north of the TC, the smaller the value of the position error (i.e., the more southward movement; Fig. 10), which was consistent with the idea of the stronger SH blocking the TC from moving more northward. The interaction between the TC and SH may have occurred via “erosion”, which was affected by the TC size and the strength of the SH (Qian et al., 2013). Consistent with Qian et al. (2013), larger TCs have broader areas of lowered heights that overlap more with the SH, which would naturally weaken both the outer edges of the SH. A positive feedback of this “erosion” may have led to more northward motion; the stronger and larger TCs pushed the south edge of the SH and generated northward track bias, the TCs were then blocked less from the northward-drifted SH so that they moved more northward, and the northward moving TCs then pushed the SH further northward in return. Instead of

Fig. 8. Radius–height cross-section composite of the change of in AAMF ($10^6$ m$^3$ s$^{-2}$) for (a) good and (b) poor members. Solid black contours indicate the zero change. The dots indicate the composite analysis that passed the 90% confident test. Note that a positive change can imply an increase in AMM export or a decrease in AM import, while a negative change can imply an increase in AMM import or a decrease in AMM export.

Fig. 9. Evolution of the distance of the TC center to the SH centroid for good (red) and poor members (blue) from 0000 UTC 12 to 17 September 2018 at 24-h intervals.
preventing the TC Sandy from recurving in Qian et al. (2013), the SH prevented the TC Mangkhut from pushing northward in this study. The stronger TCs could thus have pushed the SH towards the north until they reached a balance. Such an erosion mechanism could be one of the main reasons that the good members moved more northward than the poor members.

4.5 Contribution of steering flow and TC size on TC track

The roles of steering flows and TC sizes on the forecast southward bias were discussed in this subsection. The relative contribution of the meridional steering flows and the TC sizes to the forecast track errors were obtained through the explained variance (the square of the correlation; Meehl et al., 2016). The explained variance between the 5-day averaged TC track error in latitude coordinates and the 5-day averaged meridional steering flows reached 0.72, while the explained variance between averaged TC track and the averaged TC size was 0.07 (all passed 90% confident test), indicating a 72% and 7% relative contribution of these two variables on the TC track errors. The results suggested that the meridional steering flows made the contribution to the TC track variance nearly 10 times as large as that from the TC size. This provided further evidence for the dominant role steering flows played in determining the southward TC track errors.

5. Summary and discussion

This study explored the source of the erroneous southward track of Supper Typhoon Mangkhut (2018) that was present within a 50-member ECMWF ensemble and high-resolution deterministic forecast initialized five days prior to the TC landfall. Two groups of ensemble members were designed by comparing the position errors with respect to the TC best track. A subset of 20%
ensemble members that forecast tracks were located to the south of the best track were selected as the poor members, while the 20% ensemble members whose tracks were close to the best track were grouped as the good members. The forecast differences were then compared to determine the dynamical processes responsible for the southern bias of the TC track. The major conclusions are as follows:

The steering flows defined by the mean absolute vector wind difference between the motion of Mangkhut and the environmental flow was closely associated with the northward movement of the entire ensemble. The correlation between the TC track errors and the meridional steering flows suggested that the stronger the northward meridional steering flows, the more northward movement of the TC.

The TC size was another factor that has impacts on the TC track. Larger-size TCs tended to move more northward than smaller-size TCs. The mechanism of the TC size on influencing the TC track may be associated with the interaction of TC and SH.

The divergence of the TC track forecast was likely a result of the interaction between the TC and its surrounding environment. As a TC approached the SH, it would lead to the shrinking of the SH extent as well as the change in steering flows and eventually influence the TC track. Larger TCs may “erode” more to the SH, weakening the SH accordingly and thus reducing the resistance from the SH that prevents the TC from moving northward. The northward-moving TCs may further weaken the SH until the TCs and SH are force-balanced. Therefore, the interaction between a TC and the SH is crucial to forecast TC tracks.

The relative contribution of steering flows and TC sizes to the forecast track error was further examined by the explained variance. The results suggested that the steering flow was the main factor that contributed 72% variance in influencing the TC track, while the TC size only contributed 7% variance.

To better represent the interaction of the TC and SH, accurate TC size forecasts are needed. Unfortunately, numerical models still have difficulty in accurately representing TC sizes. Ways to improve these representations include advanced data assimilation algorithms to ingest real-time targeted observations (Jung et al., 2012), advanced tropical vortex initialization techniques for better initialization (McTaggaet-Cowan et al., 2006), and advanced numerical models (improved dynamic core, parameterization scheme, etc.) to reduce the model error during forward integration (e.g., Zhang, 1994; Leung et al., 1999; Lee and Suh, 2000; Sun et al., 2014, 2015). Further studies on the mechanism of the change of TC size (e.g., due to the AAM transport) are needed to provide a scientific foundation for developing the numerical models.

Forecasters have paid attention to the high-resolution deterministic model, which neglects the various possible scenarios in the forecast. It is important to recognize that the ensemble forecast technique is a good way to represent the uncertainty of the NWP system, especially for longer-range forecasts. It is believed that ensemble forecasts provide probabilistic information that is of more value for assessing the limitations of inevitable errors in the initial state than deterministic forecasts alone (e.g., Tracton and Kalnay, 1993; Du et al., 1997; Speer and Leslie, 1997; Zhu, 2005; Bartholmes et al., 2009). Super ensembles based on multiple ensembles are an effective way to reduce the uncertainty. With the increasing number of ensemble members, more scenarios will be produced. Forecasters need to be able to determine a set of a prior best forecasts among the large number of ensemble members. Developing a more objective tool that can help analyze large ensembles is thus of great importance.

Acknowledgments. The TIGGE data were downloaded from http://apps.ecmwf.int/datasets. The observed data can be achieved by contacting the corresponding author. Additional gratitude is extended to the three anonymous reviewers who have greatly aided in the improvement of this paper.

REFERENCES
Aberson, S. D., and M. DeMaria, 1994: Verification of a nested barotropic hurricane track forecast model (VICBAR). Mon. Wea. Rev., 122, 2804–2815, doi: 10.1175/1520-0493(1994)122<2804:VOANBH>2.0.CO;2.
Bai, L. Q., Z. Y. Meng, L. Yan, et al., 2017: An integrated damage, visual, and radar analysis of the 2015 Foshan, Guangdong, EF3 tornado in China produced by the landfalling Typhoon Mujigae (2015). Bull. Amer. Meteor. Soc., 98, 2619–2640, doi: 10.1175/BAMS-D-16-0015.1.
Bai, L. Q., Z. Y. Meng, K. Sueki, et al., 2020: Climatology of tropical cyclone tornadoes in China from 2006 to 2018. Sci. China Earth Sci., 62, 37–51, doi: 10.1007/s11430-019-9391-1.
Bartholmes, J. C., J. Thielen, M. H. Ramos, et al., 2009: The European flood alert system EFAS—Part 2: Statistical skill assessment of probabilistic and deterministic operational forecasts. Hydrol. Earth Syst. Sci., 13, 141–153, doi: 10.5194/hess-13-141-2009.
Bassill, N. P., 2014: Accuracy of early GFS and ECMWF Sandy (2012) track forecasts: Evidence for a dependence on cumulus parameterization. Geophys. Res. Lett., 41, 3274–3281, doi: 10.1002/2014GL059839.
Bonavita, M., L. Raynaud, and L. Isaksen, 2011: Estimating back-
ground-error variances with the ECMWF Ensemble of Data Assimilations system: Some effects of ensemble size and day-to-day variability. *Quart. J. Roy. Meteor. Soc.*, 137, 423–434, doi: 10.1002/qj.756.

Bougeault, P., Z. Toth, C. Bishop, et al., 2010: The TORPEX interactive grand global ensemble. *Bull. Amer. Meteor. Soc.*, 91, 1059–1072, doi: 10.1175/2010BAMS2853.1.

Brown, D. P., J. L. Beven, J. L. Franklin, et al., 2010: Atlantic hurricane season of 2008. *Mon. Wea. Rev.*, 138, 1975–2001, doi: 10.1175/2009MWR3174.1.

Carr, III L. E., and R. L. Elsberry, 2000a: Dynamical tropical cyclone track forecast errors. Part I: Tropical region error sources. *Wea. Forecasting*, 15, 641–661, doi: 10.1175/1520-0434(2000)015<0661:DTCFTE>2.0.CO;2.

Carr, III L. E., and R. L. Elsberry, 2000b: Dynamical tropical cyclone track forecast errors. Part II: Midlatitude circulation influences. *Wea. Forecasting*, 15, 662–681, doi: 10.1175/1520-0434(2000)015<0662:DTCFTE>2.0.CO;2.

Chan, J. C. L., and W. M. Gray, 1982: Tropical cyclone motion and environmental geostrophic flows. *Mon. Wea. Rev.*, 110, 1354–1374, doi: 10.1175/1520-0493(1982)110<1354:TCMASP>2.0.CO;2.

Chan, K. T. F., and J. C. L. Chan, 2013: Angular momentum transport and synoptic flow patterns associated with tropical cyclone size change. *Mon. Wea. Rev.*, 141, 3985–4007, doi: 10.1175/MWR-D-12-00204.1.

Chen, J. H., M. S. Peng, C. A. Reynolds, et al., 2009: Interpretation of tropical cyclone forecast sensitivity from the singular vector perspective. *J. Atmos. Sci.*, 66, 3383–3400, doi: 10.1175/2009JAS3063.1.

Colbert, A. J., and B. J. Soden, 2012: Climatological variations in North Atlantic tropical cyclone tracks. *J. Climate*, 25, 657–673, doi: 10.1175/JCLI-D-11-00341.1.

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

Du, J., S. L. Mullen, and F. Sanders, 1997: Short-range ensemble forecasting of quantitative precipitation. *Mon. Wea. Rev.*, 125, 2427–2459, doi: 10.1175/1520-0493(1997)125<2427:SREFOQ>2.0.CO;2.

Elsberry, R. L., Ed., 1987: *A Global View of Tropical Cyclones*. Office of Naval Research, Arlington, 185 pp.

Elsberry, R. L., H. C. Tsai, and M. S. Jordan, 2014: Extended-range forecasts of Atlantic tropical cyclone events during 2012 using the ECMWF 32-day ensemble predictions. *Wea. Forecasting*, 29, 271–288, doi: 10.1175/WAF-D-13-00104.1.

Elsner, J. B., K. B. Liu, and B. Kocher, 2000: Spatial variations in major U.S. hurricane activity: Statistics and a physical mechanism. *J. Climate*, 13, 2293–2305, doi: 10.1175/1520-0442(2000)013<2293:SVHUAC>2.0.CO;2.

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

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

Goerss, J. S., and R. A. Jeffries, 1994: Assimilation of synthetic tropical cyclone observations into the Navy Operational Global Atmospheric Prediction System. *Wea. Forecasting*, 9, 557–576, doi: 10.1175/1520-0434(1994)009<0557:ASTCO>2.0.CO;2.

Hamill, T. M., G. T. Bates, J. S. Whitaker, et al., 2013: NOAA’s second-generation global medium-range ensemble reforecast dataset. *Bull. Amer. Meteor. Soc.*, 94, 1553–1565, doi: 10.1175/BAMS-D-12-00014.1.

Heming, J. T., and A. M. Radford, 1998: The performance of the United Kingdom Meteorological Office Global Model in predicting the tracks of Atlantic tropical cyclones in 1995. *Mon. Wea. Rev.*, 126, 1323–1331, doi: 10.1175/1520-0493(1998)126<1323:TPOTUK>2.0.CO;2.

Hewson, T. D., and H. A. Tittley, 2010: Objective identification, typing and tracking of the complete life-cycles of cyclonic features at high spatial resolution. *Meteor. Appl.*, 17, 355–381, doi: 10.1002/met.204.

Holland, G. J., 1983: Tropical cyclone motion: Environmental interaction plus a beta effect. *J. Atmos. Sci.*, 40, 328–342, doi: 10.1175/1520-0469(1983)40<0328:TCMEIP>2.0.CO;2.

Huang, L., and Y. L. Luo, 2017: Evaluation of quantitative precipitation forecasts by TIGGE ensembles for South China during the presummer rainy season. *J. Geophys. Res. Atmos.*, 122, 8494–8516, doi: 10.1002/2017JD026512.

Jung, B.-J., H.-M. Kim, F. Q. Zhang, et al., 2012: Effect of targeted dropsonde observations and best track data on the track forecasts of Typhoon Sinlaku (2008) using an ensemble Kalman filter. *Tellus A*, 64A, 14984, doi: 10.3402/tellusa.v64i4.14984.

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, doi: 10.1175/WAF1002.1.

König, W., R. Sausen, and F. Sielmann, 1993: Objective identification of cyclones in GCM simulations. *J. Climate*, 6, 2217–2231, doi: 10.1175/1520-0442(1993)006<2217:OICCGI>2.0.CO;2.

Lander, M., and G. J. Holland, 1993: On the interaction of tropical-cyclone-scale vortices. I: Observations. *Quart. J. Roy. Meteor. Soc.*, 119, 1347–1361, doi: 10.1002/qj.49711951406.

Lang, S. T. K., M. Leutbecher, and S. C. Jones, 2012: Impact of perturbation methods in the ECMWF ensemble prediction system on tropical cyclone forecasts. *Quart. J. Roy. Meteor. Soc.*, 138, 2030–2046, doi: 10.1002/qj.1942.

Lee, C. S., K. K. W. Cheung, W.-T. Fang, et al., 2010: Initial maintenance of tropical cyclone size in the western North Pacific. *Mon. Wea. Rev.*, 138, 3207–3223, doi: 10.1175/2010MWR3023.1.

Lee, D. K., and M. S. Sut, 2000: Ten-year East Asian summer monsoon simulation using a regional climate model (RegCM2). *J. Geophys. Res. Atmos.*, 105, 29565–29577, doi: 10.1029/2000JD900438.

Leung, L. R., S. J. Ghan, Z. C. Zhao, et al., 1999: Intercomparison of regional climate simulations of the 1991 summer monsoon in eastern Asia. *J. Geophys. Res. Atmos.*, 104, 6425–6454, doi: 10.1029/1998jd200016.

Kossin, J. P., S. J. Camargo, and M. Sitkowski, 2010: Climate modulation of North Atlantic hurricane tracks. *J. Climate*, 23, 3057–3076, doi: 10.1175/2010JCLI3497.1.
Majumdar, S. J., S. D. Aberson, C. H. Bishop, et al., 2006: A comparison of adaptive observing guidance for Atlantic tropical cyclones. Mon. Wea. Rev., 134, 2354–2372, doi: 10.1175/MWR3193.1.

Magnusson, L., J. R. Bidlot, S. T. K. Lang, et al., 2014: Evaluation of medium-range forecasts for Hurricane Sandy. Mon. Wea. Rev., 142, 1962–1981, doi: 10.1175/MWR-D-13-00228.1.

McTaggart-Cowan, R., L. F. Bosart, J. R. Gyakum, et al., 2006: Simulated airborne Doppler radar observations. J. Atmos. Sci., 63, 2508–2528, doi: 10.1175/JAS3777.1.

Meeth, G. A., J. M. Arblaster, C. M. Bitz, et al., 2016: Antarctic sea-ice expansion between 2000 and 2014 driven by tropical Pacific decadal climate variability. Nat. Geosci., 9, 590–595, doi: 10.1038/geo2751.

Melhauser, C., F. Q. Zhang, Y. H. Weng, et al., 2017: A multiple-model convection-permitting ensemble examination of the probabilistic prediction of tropical cyclones: Hurricanes Sandy (2012) and Edouard (2014). Wea. Forecasting, 32, 665–688, doi: 10.1175/WAF-D-16-0082.1.

Munsell, E. B., and F. Q. Zhang, 2014: Prediction and uncertainty of Hurricane Sandy (2012) explored through a real-time cloud-permitting ensemble analysis and forecast system assimilating airborne Doppler radar observations. J. Adv. Model. Earth Syst., 6, 38–58, doi: 10.1002/2013ms000297.

Peng, M. S., and C. A. Reynolds, 2006: Sensitivity of tropical cyclone forecasts as revealed by singular vectors. J. Atmos. Sci., 63, 2508–2528, doi: 10.1175/JAS3777.1.

Peng, X. D., J. F. Fei, X. G. Huang, et al., 2017: Evaluation and error analysis of official forecasts of tropical cyclones during 2005–14 over the western North Pacific. Part I: Storm tracks. Wea. Forecasting, 32, 689–712, doi: 10.1175/WAF-D-16-0043.1.

Qian, C. H., F. Q. Zhang, B. W. Green, et al., 2013: Probabilistic evaluation of the dynamics and prediction of Supertyphoon Megi (2010). Wea. Forecasting, 28, 1562–1577, doi: 10.1175/WAF-D-12-00121.1.

Rappaport, E. N., J. L. Franklin, L. A. Avila, et al., 2009: Advances and challenges at the National Hurricane Center. Wea. Forecasting, 24, 395–419, doi: 10.1175/2008WAF2221281.1.

Ren, S. L., Y. M. Liu, and G. X. Wu, 2007: Interactions between typhoon and subtropical anticyclone over western Pacific revealed by numerical experiments. Acta Meteor. Sinica, 65, 329–340, doi: 10.3321/j.issn:0577-6619.2007.03.003. (in Chinese)

Richardson, D., R. Buizza, and R. Hagedorn, 2005: First Workshop on the THORPEX Interactive Grand Global Ensemble (TIGGE): Final Report. World Meteorological Organization Rep. WMO/TD-No. 1273, 39 pp.

Ritchie, E. A., and G. J. Holland, 1997: Scale interactions during the formation of Typhoon Irving. Mon. Wea. Rev., 125, 1377–1396, doi: 10.1175/125-0493(1997)125<1377:SIDTFO>2.0.CO;2.

Sanders, F., A. L. Adams, N. J. B. Gordon, et al., 1980: Further development of a barotropic operational model for predicting paths of tropical storms. Mon. Wea. Rev., 108, 642–654, doi: 10.1175/1520-0493(1980)108<0642:FDOABO>2.0.CO;2.

Sheets, R. C., 1990: The National Hurricane Center—Past, present, and future. Wea. Forecasting, 5, 185–232, doi: 10.1175/1520-0434(1990)005<0185:TNHCPA>2.0.CO;2.

Simpson, J., E. Ritchie, G. J. Holland, et al., 1997: Mesoscale interactions in tropical cyclone genesis. Mon. Wea. Rev., 125, 2643–2661, doi: 10.1175/1520-0493(1997)125<2643:MITT CG>2.0.CO;2.

Spee, M. S., and L. M. Leslie, 1997: An example of the utility of ensemble rainfall forecasting. Austr. Meteor. Mag., 46, 75–78.

Sun, Y., Z. Zhong, W. Lu, et al., 2014: Why are tropical cyclone tracks over the western North Pacific sensitive to the cumulus parameterization scheme in regional climate modeling? A case study for Megi (2010). Mon. Wea. Rev., 142, 1240–1249, doi: 10.1175/MWR-D-13-00232.1.

Sun, Y., Z. Zhong, and W. Lu, 2015: Sensitivity of tropical cyclone feedback on the intensity of the western Pacific subtropical high to microphysics schemes. J. Atmos. Sci., 72, 1346–1368, doi: 10.1175/JAS-D-14-0051.1.

Torn, R. D., J. S. Whitaker, P. Pegion, et al., 2015: Diagnosis of the source of GFS medium-range track errors in Hurricane Sandy (2012). Mon. Wea. Rev., 143, 132–152, doi: 10.1175/MWR-D-14-00086.1.

Tracton, M. S., and E. Kalnay, 1993: Operational ensemble prediction at the National Meteorological Center: Practical aspects. Wea. Forecasting, 8, 379–398, doi: 10.1175/1520-0434(1993)008<0379:OEPA>2.0.CO;2.

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, doi: 1520-0434(1991)006<0244:TBRC>2.0.CO;2.

Wang, Y. Q., and G. J. Holland, 1996: Tropical cyclone motion and evolution in vertical shear. J. Atmos. Sci., 53, 3313–3332, doi: 10.1175/1520-0469(1996)053<3313:TCMVEV>2.0.CO;2.

Wang, Y. Q., and H. Wang, 2013: The inner-core size increase of Typhoon Megi (2010) during its rapid intensification phase. Trop. Cycl. Res. Rev., 2, 65–80, doi: 10.6057/2013TCRR-02.01.

Williford, C. E., R. J. Correa-Torres, and T. N. Krishnamurti, 1998: Tropical cyclone forecasts made with the FSU Global Spectral Model. Mon. Wea. Rev., 126, 1332–1336, doi: 10.1175/1520-0499(1998)126<1332:TCFMWT>2.0.CO;2.

Wu, C.-C., and K. A. Emanuel, 1993: Interaction of a baroclinic vortex with background shear: Application to hurricane movement. J. Atmos. Sci., 50, 62–76, doi: 10.1175/1520-0469(1993)50<0062:IOABWS>2.0.CO;2.

Wu, C.-C., J.-H. Chen, P.-H. Lin, et al., 2007: Targeted observations of tropical cyclone movement based on the adjoint-derived sensitivity steering vector. J. Atmos. Sci., 64, 2611–2626, doi: 10.1175/1520-0469(2007)64<2611:TOOBSV>2.0.CO;2.

Wu, C.-C., J. H. Chen, S. J. Majumdar, et al., 2009: Intereomparison of targeted observation guidance for tropical cyclones in the northwestern Pacific. Mon. Wea. Rev., 137, 2471–2492, doi: 10.1175/2009MWR2762.1.

Wu, L. G., W. Tian, Q. Y. Liu, et al., 2015: Implications of the observed relationship between tropical cyclone size and intensity over the western North Pacific. J. Climate, 28, 9501–9506, doi: 10.1175/JCLI-D-15-0628.1.

Yamaguchi, M., and S. J. Majumdar, 2010: Using TIGGE data to
diagnose initial perturbations and their growth for tropical cyclone ensemble forecasts. *Mon. Wea. Rev.*, **138**, 3634–3655, doi: 10.1175/2010MWR3176.1.

Ying, M., W. Zhang, H. Yu, et al., 2014: An overview of the China Meteorological Administration tropical cyclone database. *J. Atmos. Oceanic Technol.*, **31**, 287–301, doi: 10.1175/JTECH-D-12-00119.1.

Zhang, G. J., 1994: Effects of cumulus convection on the simulated monsoon circulation in a general circulation model. *Mon. Wea. Rev.*, **122**, 2022–2038, doi: 10.1175/1520-0493(1994)122<2022:EOCCOT>2.0.CO;2.

Zhong, Z., and Y. J. Hu, 2007: Impacts of tropical cyclones on the regional climate: An East Asian summer monsoon case. *Atmos. Sci. Lett.*, **8**, 93–99, doi: 10.1002/asl.158.

Zhu, Y. J., 2005: Ensemble forecast: A new approach to uncertainty and predictability. *Adv. Atmos. Sci.*, **22**, 781–788, doi: 10.1007/BF02918678.