Unified Ensemble Mean Forecasting of Tropical Cyclones Based on the Feature-Oriented Mean Method

JING ZHANG,[a,b,c,d] JIE FENG,[c,d,e] HONG LI,[a,b] YUEJIAN ZHU,[f] XIEFEI ZHI,[g,h] AND FENG ZHANG[c,d]

[a] *Shanghai Typhoon Institute, China Meteorological Administration, Shanghai, China*
[b] *Key Laboratory of Numerical Modeling for Tropical Cyclones, China Meteorological Administration, Shanghai, China*
[c] *Department of Atmospheric and Oceanic Sciences, Fudan University, Shanghai, China*
[d] *Institute of Atmospheric Sciences, Fudan University, Shanghai, China*
[e] *Innovation Center of Ocean and Atmosphere System, Zhuhai Fudan Innovation Research Institute, Zhuhai, China*
[f] *NOAA/NWS/NCEP/EMC, College Park, Maryland*
[g] *Key Laboratory of Meteorological Disasters, Ministry of Education/Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters, Nanjing University of Information Science and Technology, Nanjing, China*
[h] *Nanjing Joint Center for Atmospheric Research, Nanjing, China*

ABSTRACT: Operational and research applications generally use the consensus approach for forecasting the track and intensity of tropical cyclones (TCs) due to the spatial displacement of the TC location and structure in ensemble member forecasts. This approach simply averages the location and intensity information for TCs in individual ensemble members, which is distinct from the traditional pointwise arithmetic mean (AM) method for ensemble forecast fields. The consensus approach, despite having improved skills relative to the AM in predicting the TC intensity, cannot provide forecasts of the TC spatial structure. We introduced a unified TC ensemble mean forecast based on the feature-oriented mean (FM) method to overcome the inconsistency between the AM and consensus forecasts. FM spatially aligns the TC-related features in each ensemble field to their geographical mean positions before the amplitude of their features is averaged. We select 219 TC forecast samples during the summer of 2017 for an overall evaluation of the FM performance. The results show that the TC track consensus forecasts can differ from AM track forecasts by hundreds of kilometers at long lead times. AM also gives a systematic and statistically significant underestimation of the TC intensity compared with the consensus forecast. By contrast, FM has a very similar TC track and intensity forecast skill to the consensus approach. FM can also provide the corresponding ensemble mean forecasts of the TC spatial structure that are significantly more accurate than AM for the low- and upper-level circulation in TCs. The FM method has the potential to serve as a valuable unified ensemble mean approach for the TC prediction.

SIGNIFICANCE STATEMENT: The ensemble mean forecast of tropical cyclones (TC) remains challenging. The crucial problem is the traditional arithmetic mean (AM) as a pointwise statistic disregards the geographical displacement of TC position and structure in individual ensemble members. This results in over smoothing and distortion of TC structure, particularly degrading the TC intensity forecast skill. Although the use of a consensus approach can improve TC track and intensity forecasts, it cannot provide the ensemble mean of TC spatial structure. Our study introduces a unified ensemble mean scheme for TC forecasts based on the feature-oriented mean (FM) technique. FM can provide the TC track and intensity forecasts and the corresponding TC spatial structure. FM makes a spatial-dependent alignment of TC-related features in each ensemble field to their mean position before averaging the amplitude of TC features. FM has very similar forecast skill of TC track and intensity as the consensus approach. Meanwhile, it significantly outperforms AM in predicting the TC spatial structure of variables such as pressure and wind.

KEYWORDS: Tropical cyclones; Ensembles; Postprocessing

1. Introduction

Tropical cyclones are severe weather events and have huge impacts on human activities and property. Tropical cyclone activities involve complex dynamic and thermodynamic processes across multiple spatiotemporal scales (Yanai 1964; Rotunno and Emanuel 1987; Bryan and Rotunno 2009).

Although numerical weather prediction models have improved significantly in recent years, the simulation and prediction of tropical cyclones are still constrained by deficiencies in the physics and dynamics of the models. It is therefore still challenging to predict the track, intensity and structure of tropical cyclones using a single deterministic forecast.

The ensemble forecast technique has led to significant progress in weather (Toth and Kalnay 1993, 1997; Molteni et al. 1996; Zhu et al. 2012; Feng et al. 2014; Yuan et al. 2018) and climate (Goddard et al. 2001; Doblas-Reyes et al. 2013; Zhang and Zhi 2015; J. Zhang et al. 2015; Hou et al. 2018; Yuan et al. 2018) predictions in recent decades. Ensemble forecast systems for tropical cyclones have been developed and advanced in various operational centers based on both global (Puri et al. 2001;

Corresponding authors: Jie Feng, fengjiefj@fudan.edu.cn; Hong Li, lih@typhoon.org.cn

Hamill et al. 2011; Li et al. 2016) and regional (Zhang and Krishnamurti 1999; Zhang et al. 2014; H. B. Zhang et al. 2015; Lu et al. 2016; Li et al. 2019) forecast models.

A notable advantage of ensemble forecasts over deterministic forecasts is their ability to extract the possible true states from the subspace of ensemble members and to estimate the uncertainty with the ensemble spread (Buizza 1997; Toth and Kalnay 1997; Toth et al. 2001; Zhu et al. 2002). A typical example is the prediction of the track and intensity of tropical cyclones. The most widely used approach to derive such forecasts for tropical cyclones from individual ensemble members is the so-called consensus forecast. The consensus forecast explores how to assign the optimum weights on each forecast member, including equal or unequal weights, to average the track and intensity of tropical cyclones from individual forecast members (Goerss 2000; Goerss et al. 2004; Sampson et al. 2005; Burton 2006; Krishnamurti et al. 2010; Qian et al. 2012; Dong and Zhang 2016). The best forecast members can be selected in advance by verifying the 12-h lagged real-time ensemble forecasts and removing any apparent outliers (Qi et al. 2014; Dong and Zhang 2016; Zhang and Yu 2017). A variety of studies have shown that the consensus approach can significantly improve the accuracy of track and intensity forecasts for tropical cyclones relative to deterministic forecasts (Goerss 2000; Sampson et al. 2005; Krishnamurti et al. 2010; Dong and Zhang 2016).

In ensemble forecasting, the most widespread approach to predict the future states of three-dimensional (3D) spatial fields is to take the arithmetic mean (AM) of the ensemble member fields. The AM is a pointwise statistical concept providing the best sample-based estimate of the expected value of any single variable (Leith 1974; Li et al. 2018; Feng et al. 2019). The AM filters out any unpredictable features in the forecast member fields and reduces the forecast errors relative to deterministic forecasts (Toth and Kalnay 1997; Buizza et al. 2005; Feng et al. 2020). However, the application of the AM method to predictions of tropical cyclone structures has a well-known, but as yet unresolved, problem.

The tropical cyclones in individual forecast members have both position and amplitude-related errors (Hoffman et al. 1995) and these errors amplify as the forecast lead time increases. The use of the AM implements a pointwise average that does not account for the deviations in the locations of tropical cyclones in ensemble fields and the spatial coherence of atmospheric variable fields. Such averaging may cause unrealistic smoothing and distortion of the structure of the tropical cyclone and an underestimation of the intensity. As a consequence, operational ensemble predictions of tropical cyclones routinely use the respective consensus mean forecast of the track and intensity of tropical cyclones rather than directly deriving these parameters from the AM field of the tropical cyclone.

A crucial problem in forecasting the tropical cyclone ensemble mean is the inconsistency of the consensus forecast of the track and intensity of the tropical cyclone compared with the AM forecast of the structure of the tropical cyclone. In other words, the track of the tropical cyclone in the AM forecasts may deviate from the consensus track, and the physical structure of the tropical cyclone in the AM method may not reflect the tropical cyclone intensity obtained from the consensus method.

To eliminate the positional deviations of the tropical cyclone among the ensemble members, global and regional forecast systems at the National Centers for Environmental Prediction use a tropical cyclone relocation procedure for ensemble initialization and data assimilation (Liu et al. 2006; Tallapragada et al. 2015). This method removes the tropical cyclone 3D vortex structure from each ensemble member field as a whole and relocates them to the same observed location of the tropical cyclone center for a better estimation of the background error covariance (Lu et al. 2016; Feng and Wang 2019). In addition to the model initialization, this relocation procedure can also be applied to the postprocess and diagnostics for ensemble forecasts. Averaging these relocated tropical cyclone ensemble members may give a more realistic ensemble mean structure of the tropical cyclone vortices. However, an apparent problem is that the relocation procedure results in discontinuities between the structures within and outside a tropical cyclone in individual ensemble members. This problem is generally remedied by spatial smoothing (Liu et al. 2020). The simple relocation procedure also assumes that all the structures related to tropical cyclones, such as the inner core and the environment in each member, are shifted by the same displacement vector, which is unrealistic.

We propose a unified ensemble mean forecast for tropical cyclones based on the feature-oriented mean (FM) method (Feng et al. 2020). In contrast with the traditional AM method, the FM method aligns the features related to the tropical cyclone in each member field—including the inner core, outflow, and large-scale environment of the tropical cyclone—to their respective mean position across all members before taking the mean of the aligned members. Unlike the relocation procedure as an adjustment or shift based on the central location, the FM uses a feature-based 3D vector alignment that takes the spatial coherence of the variable fields into account (Feng et al. 2020). The FM method eliminates the positional deviations of the tropical cyclone across ensemble members and averages the amplitude of features, ensuring realistic smoothing and spatial continuity of the tropical cyclone structure. The track and intensity of the tropical cyclone are then derived directly from the FM forecast field, preserving the consistency between the track and intensity of the tropical cyclone and its structure.

The paper is organized as follows. Section 2 briefly introduces the simplified and computationally efficient FM method. The experimental setup, including the selection of the case studies, the data for forecasts and verification, and the evaluation metrics, are presented in section 3. Section 4 evaluates the FM for the prediction of tropical cyclones and compares it with the consensus and AM approaches. A summary and discussion are provided in section 5.

2. Methodology

The atmosphere presents organized structures with spatiotemporal coherence. It is widely recognized that atmospheric forecast errors consist of both position- and amplitude-related

components (Hoffman et al. 1995; Ravela et al. 2007; Beezley and Mandel 2008; Feng et al. 2020; Ji et al. 2020). Most traditional operational and research applications still use the AM forecast, which calculates the pointwise average of ensemble fields, disregarding the spatiotemporal coherence of variable structures. It unrealistically smooths the features in ensemble fields in both space and time. The concept of the FM method is to adjust the features in each ensemble field to their mean position before taking the mean of the amplitude (Feng et al. 2020). The adjustment of the features in each forecast field is carried out using the field alignment technique.

a. Field alignment

The field alignment technique was pioneered by Ravela et al. (2007). Its basic concept is to establish a two-dimensional (2D) mapping between two fields with similar, but spatially misaligned, features. We use $\mathbf{X}$ and $\mathbf{Y}$ to represent the original and target 2D variable fields, respectively, with similar, but displaced, features. Assuming that the features in $\mathbf{X}$ can be aligned along a 2D vector field (i.e., the displacement vector), forming a new 2D variable field $\mathbf{X}'$, then field alignment variationally solves the optimum displacement vector field to minimize the root-mean-square difference between the aligned field $\mathbf{X}'$ and the target field $\mathbf{Y}$. The optimally aligned field $\mathbf{X}'$ is also derived as an output solution of the field alignment technique.

Unlike other techniques primarily applied to the adjustment of the spatial structure between two fields (e.g., Hoffman et al. 1995; Du et al. 2000; Nehrkorn et al. 2003, 2014), the field alignment technique does not rely on the estimated posterior covariance of the field after alignment and has a unique tunable parameter (i.e., the wavenumber parameter l). The wavenumber parameter l in the field alignment technique controls the smoothness of the displacement vector. The alignment of features above the scale of l in the original field is sought, while the smaller scale features are moved with the larger scales. An optimum value of l generally distinguishes the large-scale well-organized features and the small-scale random structures.

More details about the field alignment algorithm can be found in Ravela et al. (2007) and Ravela (2014). The field alignment technique has been applied to a wide variety of areas, including data assimilation (Ravela et al. 2007), forecast verification (Ravela et al. 2007; Ravela 2014; Jankov et al. 2020), nowcasting (Ravela 2012), and spatiotemporal error propagation (Feng and Toth 2017).

b. Feature-oriented mean method

The key concept behind the FM method is to align the features in each ensemble field to their mean position (Feng et al. 2020). The traditional AM may not be able to characterize the positional mean of features accurately due to the distorted structure. The FM method adopts a simple and computationally efficient vector operation to estimate the displacement of features in individual ensemble fields from the mean position and makes appropriate adjustments.

Following the schematic diagram in Feng et al. (2020, Fig. 2), we assume $\mathbf{x}_j$ is a 2D variable field of a randomly selected member from an N-member ensemble ($j = 1, 2, 3, \ldots, N$). The FM method consists of the following five steps.

1) Compute the displacement vector $\mathbf{D}_{ji}$ between the given ensemble field $\mathbf{x}_j$ and all other fields $\mathbf{x}_i$ ($i = 1, 2, 3, \ldots, N$) using the field alignment technique with a specified wavenumber parameter l.

2) Calculate the average of the derived displacement vector field $\mathbf{D}_{ji}$, where $\overline{\mathbf{D}_j} = (1/N)\sum_{i=1}^{N}\mathbf{D}_{ji}$. The term $\overline{\mathbf{D}_j}$ is an estimate of the translation operator that indicates how the given ensemble field $\mathbf{x}_j$ is aligned to the mean position.

3) Align the given ensemble field $\mathbf{x}_j$ to the mean position along the mean displacement vector $\overline{\mathbf{D}_j}$, deriving the aligned member $\mathbf{x}'_j$.

4) Repeat procedures 1–3 for each ensemble field $\mathbf{x}_j$ ($j = 1, 2, 3, \ldots, N$).

5) Calculate the arithmetic mean of all the aligned members $\mathbf{x}'_j$ to derive the FM field $\overline{\mathbf{x}} = (1/N)\sum_{j=1}^{N}\mathbf{x}'_j$.

The FM method is conceptually very similar to the field coalescence method proposed by Ravela (2012). Both the field coalescence and the FM method align the features in each ensemble field to the mean of their positions using the field alignment technique as an alignment tool. However, as discussed in Feng et al. (2020), the field coalescence method solves a more complex variational minimization problem. The cost function to be minimized measures the average distance of the mean field from the individual aligned ensemble fields standardized by the posterior ensemble covariance. By contrast, the FM method solves a set of much simpler field alignment minimization problems and then makes a simplified vector calculation to estimate the mean position regardless of the posterior ensemble covariance.

Tropical cyclones have complex dynamic and thermodynamic structures that involve multivariate and multilayer interactions. The minimum sea level pressure (SLP) is a crucial variable that indicates the location, intensity and primary circulation of tropical cyclones. The SLP is also a continuous smooth field, making it appropriate for the field alignment and the FM method. Therefore a set of field alignment calculations in the FM method is implemented for the variable SLP to derive the displacement vectors of each ensemble member. The displacement vectors for all other variable fields at different levels are not calculated, but the vectors from the variable SLP are used directly for each corresponding ensemble member. This FM scheme maintains the integrity of the spatial structure of tropical cyclones and the cross-variable covariance relationship.

3. Experimental setup

a. Selection of case studies

To evaluate the performance of the FM method as applied to predictions of tropical cyclones, we selected 25 tropical cyclones over the North Atlantic Ocean and the western and eastern North Pacific Ocean from July to September 2017. This time period was selected because there were numerous very intense and long-lasting tropical cyclone activities over the tropics, such as Hurricanes Harvey and Jose, and Super Typhoon Noru, which caused catastrophic losses of both property and life. Figure 1 shows the 6-hourly observed tropical cyclone tracks of
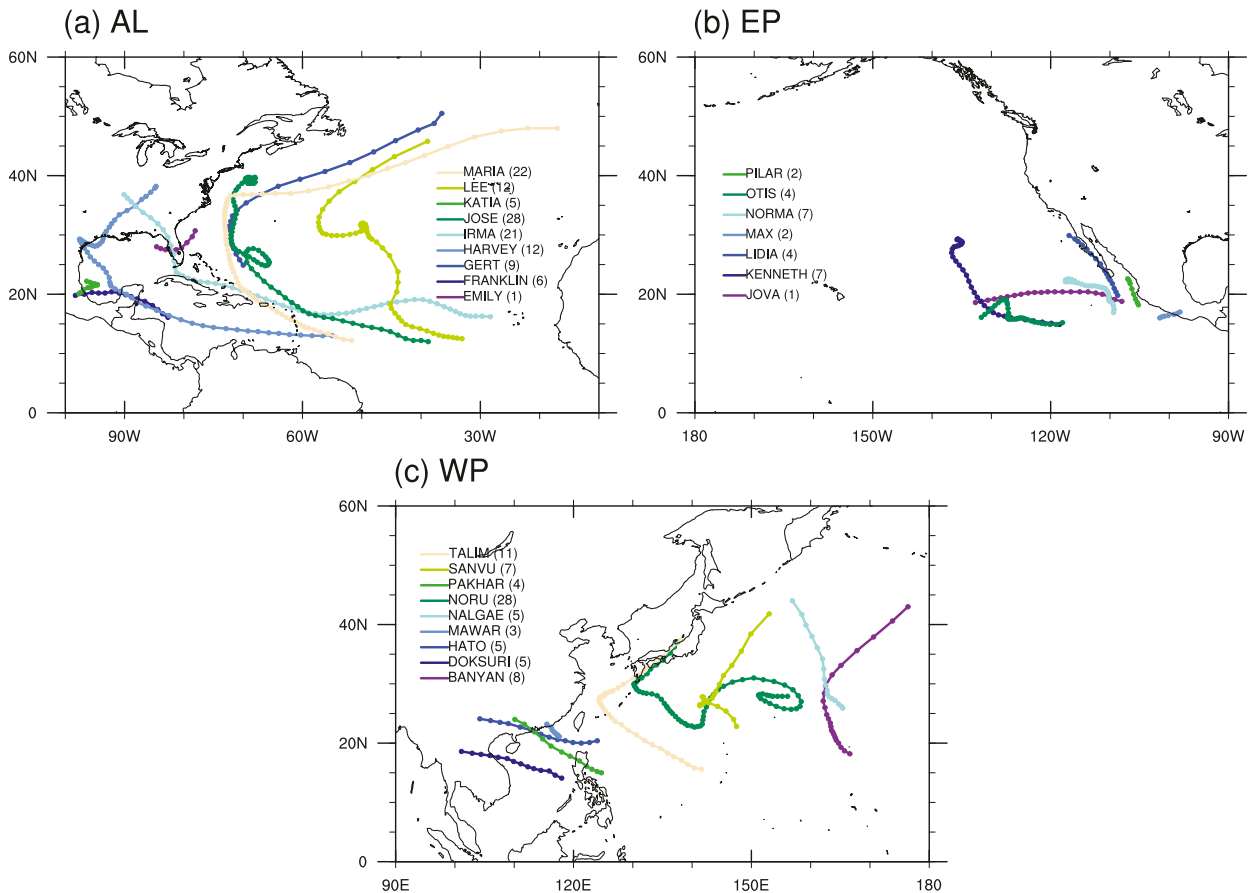
FIG. 1. The 6-hourly observed tropical cyclone tracks from the best track dataset. Atlantic Ocean (AL), eastern Pacific Ocean (EP), and western Pacific Ocean (WP). The numbers in the parentheses indicate the number of forecast samples for each tropical cyclone.

all the selected cases. A total of 219 forecast samples are selected for the 25 tropical cyclone case studies, all of which were initialized at 0000 or 1200 UTC between 25 July and 20 September 2017. The forecast samples have initial observed tropical cyclone intensities reaching at least tropical storm strength. When the observed tropical cyclone is weaker than the tropical storm strength, the forecasts at the same valid time would be omitted. Therefore, the valid number of forecast samples decreases for longer lead times.

b. Forecast and verification data

The ensemble forecast data used in this study are from the Global Ensemble Forecast System of the National Centers for Environmental Prediction (Toth and Kalnay 1993, 1997; Zhu et al. 2012; Zhou et al. 2016, 2017). The real-time operational Global Ensemble Forecast System consists of 20 perturbed ensemble members and has a semi-Lagrangian horizontal resolution of T574 (~34 km) and 64 vertical levels (Zhou et al. 2017). The forecast data used here have homogeneous horizontal grids with (0.5° × 0.5°) resolution and a 5-day forecast lead time at an interval of 12 h.

The verification for the tropical cyclone structure uses an independent dataset which is the initial analysis of the operational control forecasts from the European Centre for

Medium-Range Weather Forecasts (ECMWF). The data have a horizontal resolution of spectral triangular truncation T639 (~32 km; Buizza 2014) and are interpolated to the same spatial resolution (0.5° × 0.5°) as the forecast verification. Both the forecast and verification data can be downloaded from the THORPEX Interactive Grand Global Ensemble (TIGGE) website (http://apps.ecmwf.int/datasets/data/tigge). The tropical cyclone track and intensity forecasts are verified against the best track dataset over the western North Pacific Ocean and the North Atlantic and eastern Pacific oceans. The former is issued by the Joint Typhoon Warning Center (www.metoc.navy.mil/jtwc/jtwc.html?western-pacific) and the latter two are from the National Hurricane Center (www.nhc.noaa.gov/data/#hurdat). The best track dataset contains comprehensive information for each storm, including the location, maximum winds and central pressure. The position of the tropical cyclone center in forecasts is defined using a simple local extreme method (Ryglicki and Hart 2015). First, the tropical cyclone center in the best track dataset at the same valid time is used as a first guess and the minimum SLP is searched over a circle of radius (~500 km). If the minimum SLP is lower than that of each grid point along the 100 km radius, the corresponding position is adopted as the tropical cyclone center. Otherwise, the resulting latitude and longitude

TABLE 1. Sample-mean global and regional RMSE (hPa) of the 2.5-day original ensemble member forecasts and the corresponding aligned forecasts as a function of the wavenumber parameter l for the SLP. The regional RMSE is for a 600 km $\times$ 600 km domain surrounding the tropical cyclones in the global forecasts. The initial time of the forecast cases ranges from 0000 UTC 1 Aug 2017 to 0000 UTC 5 Aug 2017 at an interval of 24 h. Each case has 20 ensemble member forecasts.

	Original	$l = 64$	$l = 128$	$l = 192$	$l = 256$
Global RMSE (hPa)	3.10	2.66	1.67	1.60	1.58
Tropical cyclone RMSE (hPa)	2.86	2.12	1.28	1.19	1.14

were used as the next guess. The use of the 100-km radius is to confirm that the searched minimum SLP is a low pressure center. This step also ensures that the tropical cyclone center can be searched outward if the storm is not found within the 500-km radius.

c. Evaluation metrics

We compare the FM method against the consensus approach in tropical cyclone track and intensity forecasts. The consensus forecast uses the faster, equal-weight mean of all ensemble members, as preferred by the operational centers (Goerss 2000; Goerss et al. 2004; Sampson et al. 2005; Burton 2006; Krishnamurti et al. 2010). The tropical cyclone track errors are calculated using the great circle distance. The tropical cyclone intensity is represented by the central minimum SLP. We compared the FM method against the traditional AM in tropical cyclone structure forecasts using the root-mean-square errors (RMSEs) of the SLP and the wind fields at 10 m and 500 hPa as metrics.

4. Results

a. Optimizing the wavenumber parameter l

The unique wavenumber parameter l is tuned in the field alignment algorithm. To estimate the optimum parameter for the alignment of the member fields in the FM algorithm, we select five forecast cases of the SLP at the 60-h[1] lead time initialized from 0000 UTC 1 August 2017 to 0000 UTC 5 August 2017 at a 24-h interval. Each case has 20 ensemble forecasts. These global ensemble forecasts are aligned to the ECMWF analysis fields at the same valid time using the field alignment method with different values of l. The key to the performance of the FM method is the accuracy of the estimated mean position for ensemble members and the effectiveness of the alignment of each ensemble member to their mean position. The expected mean positions of features among a group of ideal ensemble members would be very close to those presented in the analysis field. Therefore, we use the analysis field as an alternative target field to test the optimal value of l for the alignment of ensemble forecasts. The global RMSE between the aligned forecast fields and the analysis fields averaged over all the samples (i.e., 20 $\times$ 5) are calculated

as a function of l (see the second row of Table 1). Because this study is focused on tropical cyclone forecasts, the RMSE averaged over a 600 km $\times$ 600 km domain surrounding the Nalgae and Noru tropical cyclones in these global ensemble forecasts of the SLP is also assessed to optimize the value of l. The RMSE variation as a function of l is also evaluated for each case, which produces qualitatively similar results as the sample mean result (see Table 2).

Both results show that the RMSE is substantially reduced even after aligning the SLP fields using $l = 64$. The RMSE decreases further with increasing values of l, but the decrease in RMSE with increasing l becomes much more gradual for $l >$ 192. This value of 192 roughly corresponds to 200-km scale motion, which means that the finer ($<$200 km) scales are almost unpredictable at 2.5 days (60 h) and their alignments are nearly random and useless. Alignment should therefore be carried out for only larger scale ($>$200 km) features that are fully or partially predictable. The following results use the wavenumber parameter $l = 192$ for the field alignment algorithm taking into consideration the computational expense. A larger optimum value of l may be used in a higher resolution model in which the finer scale features are better resolved and more predictable.

Figure 2 shows an example of the alignment of an arbitrarily selected 2.5-day ensemble forecast valid at 1200 UTC 3 August 2017 against the analysis. It shows that the original forecast (red contours) and the analysis (black contours) are similar, but show some spatial displacement. For example, Typhoon Noru over the western North Pacific presents an apparent eastward positional error in the forecast relative to the analysis (green box in Fig. 2a and enlargement in Fig. 2c). Such a displacement throughout the globe is well described by the displacement vectors (blue arrows), with the direction and length of the arrows indicating the displaced direction and distance, respectively. The aligned field (blue contours) obtained using the field alignment technique shows a much better positional match of features to the analysis field (black contours; Fig. 2b). The SLP structure associated with Typhoon Noru in the forecast is aligned to a similar location as in the analysis, including the vortex and the large-scale circulation (green box in Fig. 2b and enlargement in Fig. 2d).

b. Alignment of ensemble members

Before a quantitative evaluation of the FM is carried out, we provide a virtual demonstration of the alignment of the ensemble members to the mean position. Figure 3 shows the SLP in the original and aligned ensemble members at lead times of 1 (solid lines), 2.5 (dashed lines), and 5 (dotted lines) days for

---

[1] The optimal parameter l may slightly vary with different lead times, but such variation within the first 5 days is not considered in this study.

TABLE 2. As in Table 1, but for individual cases.

		Original	$l = 64$	$l = 128$	$l = 192$	$l = 256$
Global RMSE (hPa)	Case 1	2.96	1.87	1.67	1.60	1.57
	Case 2	3.06	1.84	1.59	1.56	1.56
	Case 3	3.37	2.27	1.93	1.83	1.81
	Case 4	3.20	4.47	1.62	1.54	1.51
	Case 5	2.91	1.85	1.52	1.45	1.41
Tropical cyclone RMSE (hPa)	Case 1	3.14	2.33	1.20	1.0	1.05
	Case 2	2.73	2.48	1.27	1.38	1.18
	Case 3	2.73	2.09	1.56	1.26	1.23
	Case 4	3.03	2.07	1.21	1.17	1.15
	Case 5	2.65	1.53	1.13	1.11	1.11

Hurricane Maria (Fig. 3a) and Typhoon Noru (Fig. 3b). In both cases, the location, intensity and structure of the tropical cyclone in the original members tend to be more diverse as the lead time increases. Averaging these original members (i.e.,

the arithmetic mean method) inevitably results in a spurious smoothing of the structure of the tropical cyclone, causing suboptimum mean forecasts of the track, intensity and structure. By contrast, the tropical cyclones in the processed
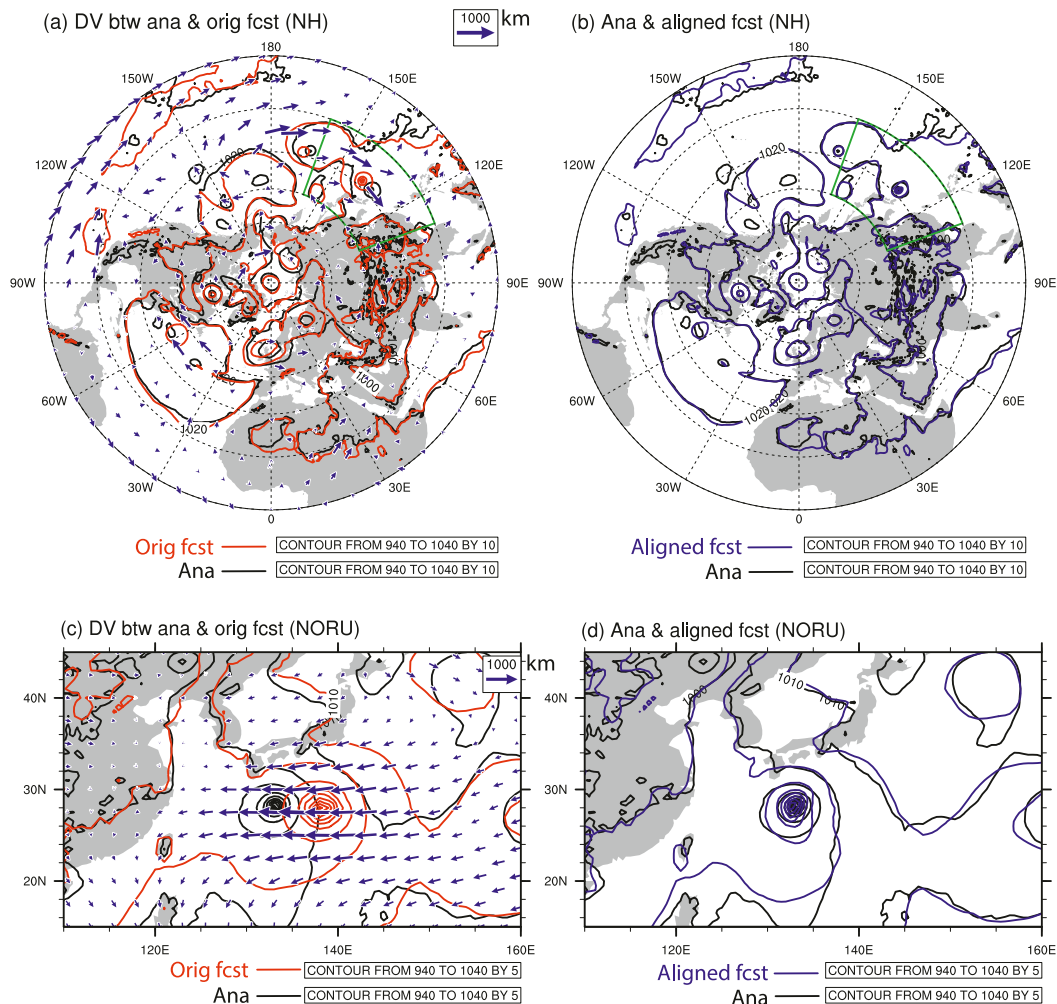

FIG. 2. (a) The 2.5-day SLP forecast of a randomly selected ensemble member (red contours) and the analysis (black contours) at the same valid time of 1200 UTC 3 Aug 2017 and their displacement vectors (blue arrows). (b) The aligned field (blue contours) of the forecast in (a) and the analysis (black contours). (c) As in (a), and (d) as in (b), but for the area of the green box shown in (a) and (b).
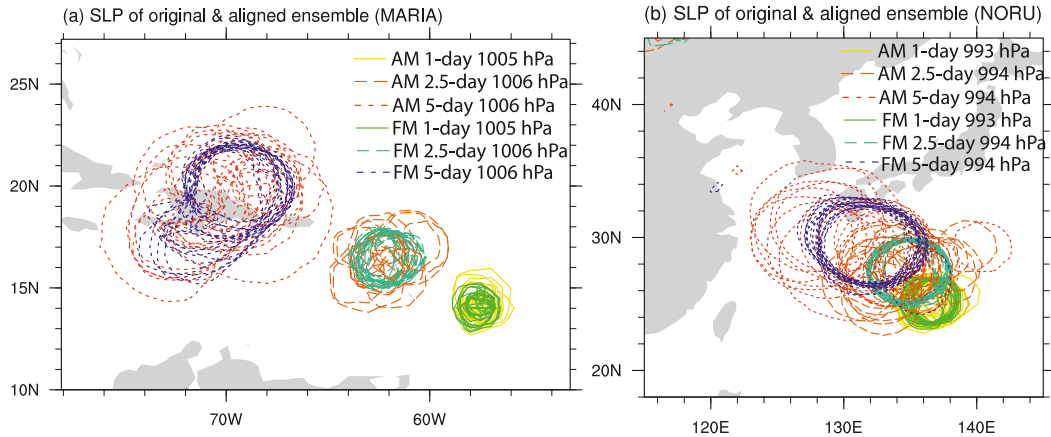
FIG. 3. (a) SLP of the original and aligned forecast members of Hurricane Maria at the 1-day (solid lines), 2.5-day (dashed lines), and 5-day (dotted lines) lead times initialized at 0000 UTC 17 Sep 2017. (b) SLP of the original and aligned forecast members of Typhoon Noru at the 1-day (solid lines), 2.5-day (dashed lines), and 5-day (dotted lines) lead times initialized at 0000 UTC 1 Aug 2017.

ensemble members of FM method are aligned to the mean of their positions, avoiding the spurious smoothing caused by the displacement of the tropical cyclone in individual ensemble members. The aligned ensemble members greatly reduce the position-related spread among the original members, facilitating an accurate calculation of the mean amplitude of features.

c. Errors in the tropical cyclone track and intensity forecasts

Tropical cyclone track and intensity forecasts are crucial metrics with respect to the prediction of tropical cyclones. Figure 4 compares the sample-mean tropical cyclone track (Fig. 4a) and the minimum SLP (Fig. 4b) errors of the consensus (red), FM (blue), and AM (green) forecasts. The

uncertainty bars show the 95% confidence interval for the mean error. The three approaches have slight, statistically insignificant differences in the tropical cyclone track forecasts (Fig. 4a).

In terms of the tropical cyclone intensity forecast (Fig. 4b), the AM method shows a poorer forecast skill than the other two methods after 2 days. The differences are statistically significant beyond 3.5 days; namely, the intensity forecast errors of the consensus and FM approaches are out of the uncertainty ranges of the AM forecast errors. The differences between the consensus and FM intensity forecast errors are statistically insignificant, although the FM errors are slightly larger ($\sim$3 hPa) than the consensus forecast errors. The marginally lower errors of the consensus forecasts could be possibly attributed to the
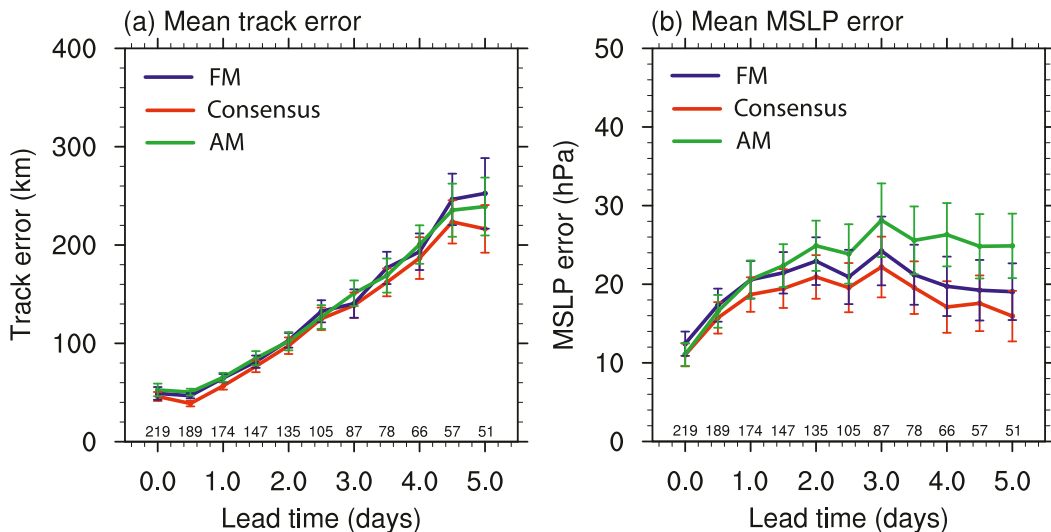

FIG. 4. Sample-mean tropical cyclone (a) track and (b) minimum SLP (MSLP) errors of the consensus (red curve), feature-oriented mean (FM; blue curve), and arithmetic mean (AM; green curve) forecasts. The uncertainty bars show the 95% confidence interval for the mean error. The numbers at the bottom show the number of forecast samples at different lead times.
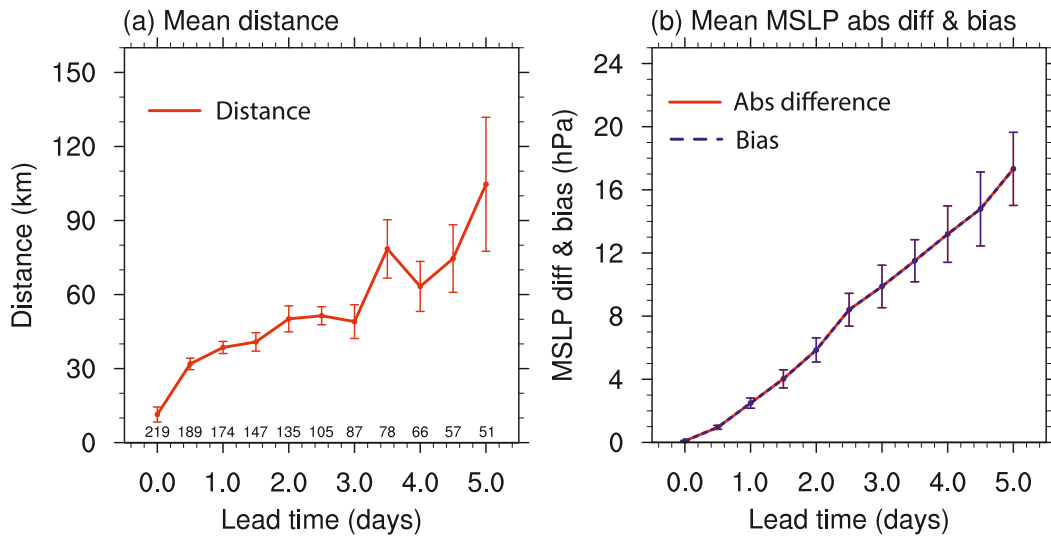
FIG. 5. (a) Sample-mean distance of the tropical cyclone centers (red curve) and (b) the sample-mean absolute difference (red solid curve) and bias (blue dashed curve) of the tropical cyclone MSLP for the arithmetic mean relative to the consensus forecasts as a function of the lead time. The uncertainty bars show the 95% confidence interval for the sample-mean values. The absolute difference and bias in Fig. 5b have exactly the same uncertainty bars. The numbers at the bottom show the number of forecast samples at different lead times.

averaging offset of the random errors contained in the minimum SLP of individual ensemble forecasts. By contrast, the tropical cyclone intensity in the FM method is estimated after the aligned forecasts are averaged and thus may contain more random errors.

The results in Fig. 4 indicate that the consensus and FM approaches are comparable in the tropical cyclone track and intensity forecasts because they similarly extract tropical cyclone position and intensity information from individual ensemble forecasts. The FM method is superior because its tropical cyclone position and intensity are directly estimated from the mean forecast field, which is the so-called unified ensemble mean. However, the consensus approach offers a scalar mean prediction of the position and intensity of tropical cyclones that are inconsistent with those of AM. The spurious smoothing of the displaced tropical cyclone structures by the AM method has little statistical impact on the tropical cyclone track forecast, but degrades the intensity forecast skill compared with the consensus and FM methods.

d. Inconsistency between consensus and arithmetic mean forecasts

The consensus method is widely used to predict the mean track and intensity of tropical cyclones. Its algorithm is essentially different from the direct estimation of the track and intensity of tropical cyclones based on the AM forecast field. To quantify the differences between the two approaches, we provide the sample-mean distance of tropical cyclone centers (Fig. 5a) and the sample-mean absolute difference (red solid curve) and bias (blue dashed curve) of the minimum SLP of tropical cyclones (Fig. 5b) for the AM relative to the consensus forecasts. Their respective 95% confidence intervals are given by the uncertainty bars.

Figure 5a shows that the mean distance of the tropical cyclone centers between the two methods increases with the lead time and is statistically significant at all lead times. A similar variation occurs for the mean absolute difference in the tropical cyclone central SLP (Fig. 5b), reaching nearly 15 hPa at day 5. The mean bias of the AM relative to the consensus forecasts has a consistent variation compared to their mean absolute difference (cf. blue and red curves). It indicates that the traditional AM method systematically underestimates the tropical cyclone intensity compared with the consensus forecast due to the smoothing of the displaced tropical cyclone structures in individual members.

e. Comparing tropical cyclone structures in the arithmetic mean, feature-oriented mean, and consensus forecasts

Following the comparison of the track and intensity of tropical cyclones, this section compares the structure of tropical cyclones in the AM, FM and consensus forecasts. Figures 6a, 6b, 6e, and 6f show the tropical cyclone track consensus forecast overlain on the 5-day AM fields of the SLP (black contours) and the circulation for Hurricane Maria and Typhoon Noru. The same variables of the unified FM forecasts are also shown for comparison (Figs. 6c,d,g and h).

Figure 6 shows that the minimum SLP (maximum wind amplitude) in the AM method is apparently higher (weaker) than that in the FM method as a result of the oversmoothing of the displaced tropical cyclone vortices in the AM method (see Fig. 3). This smoothing also results in a poorly resolved eyewall of the tropical cyclone in the AM method for both Hurricane Maria and Typhoon Noru. The centers of the tropical cyclone in the consensus forecast (red dots) deviate from those in the AM forecast (red crosses) at both the sea surface and 500 hPa in both cases (Figs. 6a,b,e and f) and the distance
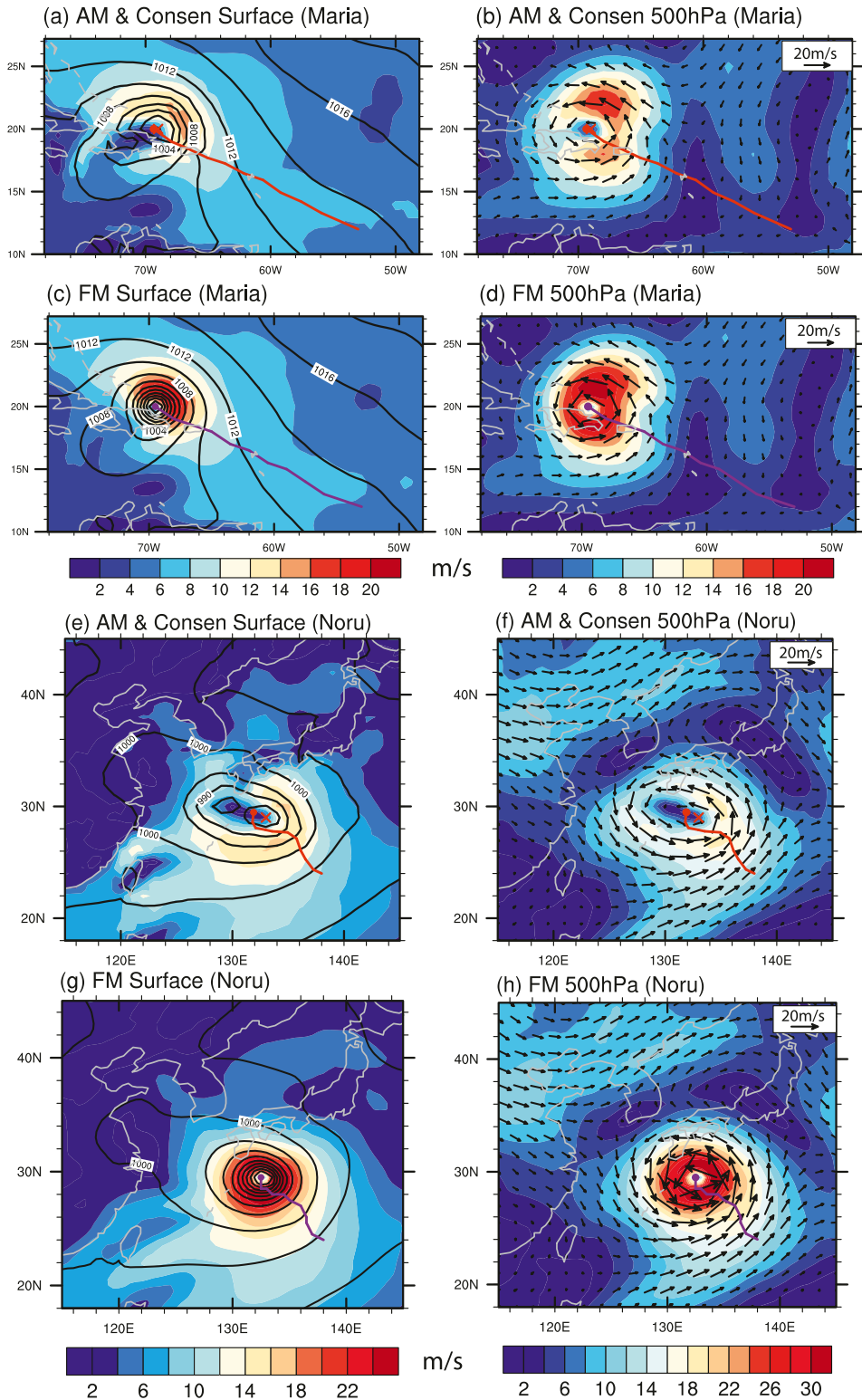
FIG. 6. Tropical cyclone track consensus forecast (red lines) overlain on the 5-day arithmetic mean (AM) fields of (a) the SLP (black contours) and 10-m wind amplitude (shaded) and (b) the 500-hPa wind vectors (black arrows) and amplitude (shaded) for Hurricane Maria initialized at 0000 UTC 17 Sep 2017. The red dots and crosses denote the tropical cyclone centers for the 5-day consensus and arithmetic mean forecasts, respectively. (c),(d) As in (a) and (b), respectively, but for the feature-oriented mean forecast. The tropical cyclone track forecasts of the feature-oriented mean are shown in purple. (e)–(h) As in (a)–(d), but for Typhoon Noru initialized at 0000 UTC 1 Aug 2017.
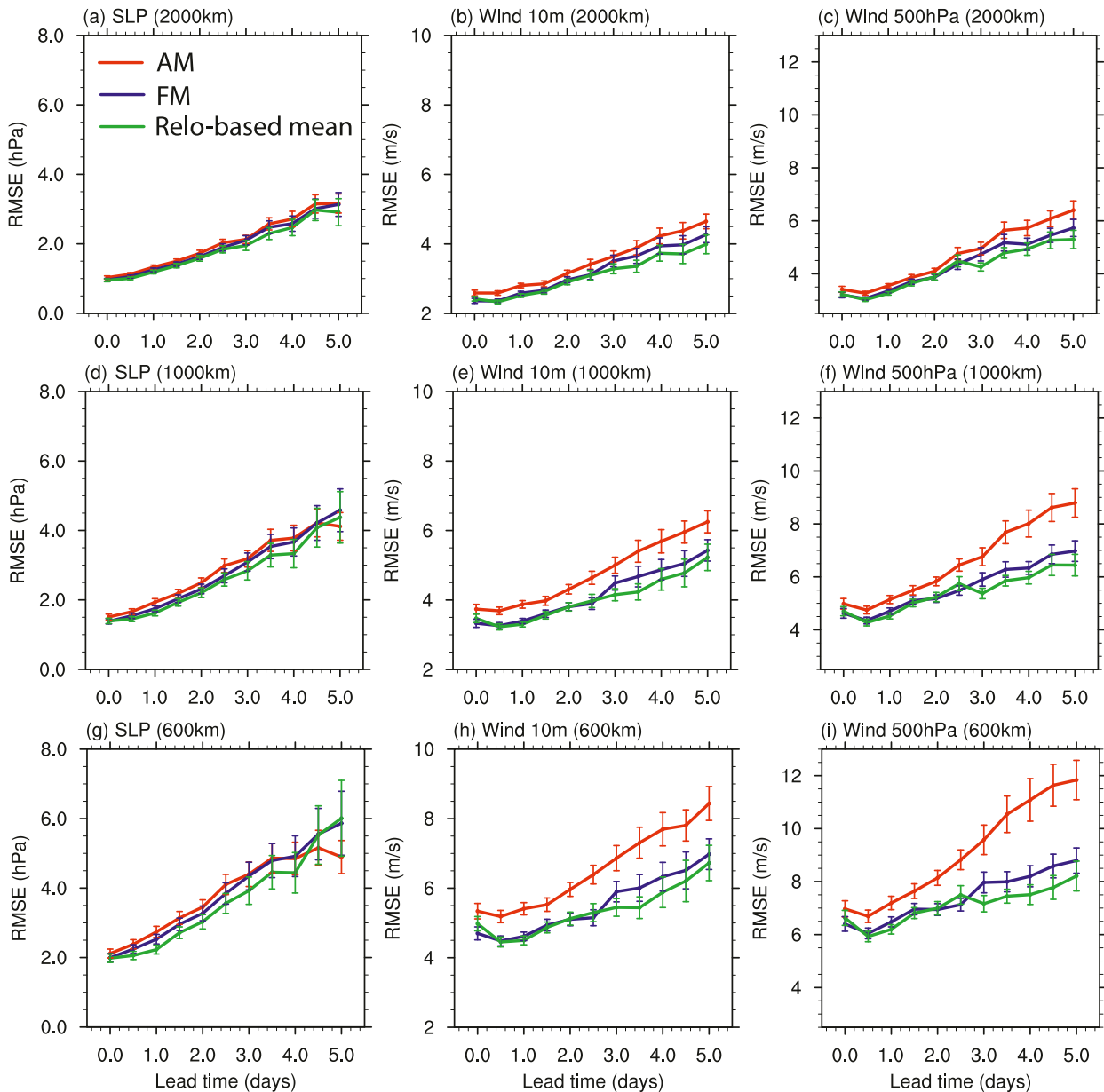
FIG. 7. Sample-mean RMSE of the arithmetic mean (red curve), feature-oriented mean (blue curve), and relocation-based mean (green curve) forecasts averaged over a 2000 km × 2000 km square domain in predicting the structures related to tropical cyclones of the (a) SLP, (b) 10-m winds, and (c) 500-hPa winds. (d)–(f) As in (a)–(c), but for 1000 km × 1000 km. (g)–(i) As in (a)–(c), but for 600 km × 600 km. The square domains in the arithmetic mean, feature-oriented mean, and analysis fields at the same valid time are centered at their own tropical cyclone centers for verification. The uncertainty bars give the 95% confidence interval for the sample-mean error.

reaches nearly 200 km for Typhoon Noru. This shows the inconsistency in the information about the position of tropical cyclones in the AM and consensus forecasts.

By contrast, the track of the tropical cyclones in the FM (purple line) traverses the location of the minimum wind amplitude at the sea surface and 500 hPa in both cases (Figs. 6c,d,g and h) because the position of the tropical cyclone is derived directly from the FM forecast field. This confirms the validity of using the displacement vectors of the SLP ensemble fields for the alignment of other variables, such as the wind fields at 500 hPa.

The AM is currently the most popular approach to provide 3D spatial field forecasts because it reduces the overall errors relative to a deterministic forecast (e.g., Buizza et al. 2005; Yuan et al. 2018). It is therefore of interest to compare the forecast skill of the AM and the FM in predicting the structure of tropical cyclones. Figure 7 shows the RMSE of the AM (red curve) and the FM (blue curve) averaged over a specified

square domain in predicting the structures of the SLP and the winds at 10 m and 500 hPa related to tropical cyclones. The square domains of the AM, the FM and the analysis fields at the same valid time are centered at their own centers of the tropical cyclone for verification. Different domain sizes are used to demonstrate the sensitivity of the comparison results, including 2000 km × 2000 km, 1000 km × 1000 km, and 600 km × 600 km.

The results show that the FM only has slightly smaller RMSE than the AM in predicting the SLP field for all the domain sizes that are statistically insignificant. By contrast, the superiority of the FM over the AM is much more remarkable for the wind field, where it is statistically significant for all domain sizes and both the surface and upper levels. Reducing the domain size for verification further expands the advantage of the FM. This suggests that the improvement of the FM relative to the AM is mainly in the prediction of the vortex-related primary circulation. These results imply that AM has limited skill in predicting extreme events, although it is intended to minimize forecast errors by filtering out unpredictable noise. The FM recognizes and eliminates displacements among features and simply averages their amplitude. The FM therefore has the potential to act as a better and more skillful ensemble mean algorithm for the prediction of extreme events.

In view of the shortcomings of the AM method, the tropical cyclone ensemble mean field can also be derived using a similar concept to the consensus forecast. Within a predefined domain size, the structures related to the tropical cyclone in individual ensemble fields are extracted as a whole and relocated to the consensus mean of the position of the tropical cyclone. The relocation-based ensemble mean field is then calculated by simply averaging the relocated structure within the overlapping domain. The relocation-based mean (green curve) is also compared with the AM and FM in terms of the regional RMSE as the metric used in Fig. 7. The relocation-based mean forecast performs much better than the AM and shows slightly lower regional RMSE than the FM in the wind fields beyond 2.5 days. Nonetheless, a suboptimal assumption for the relocation-based ensemble mean is all its structures related to the tropical cyclone, such as the vortex and the outflow environment, in a single ensemble field have a homogeneous displacement from their respective mean position. Figure 8 gives an example to demonstrate the flaw in the relocation-based mean forecasts of the structure of tropical cyclones. Figure 8 shows the original, relocated and aligned ensemble fields (Figs. 8a–c) and their corresponding mean fields (Figs. 8d–f)—that is, the AM, relocation-based mean, and FM forecasts.

The original ensemble fields (Fig. 8a) present remarkable positional displacements across their members, including the inner vortex (e.g., the 1000-hPa contour) and the outer large-scale circulation (e.g., the 1015-hPa contour). By contrast, the inner vortices in the relocated (Fig. 8b) and aligned (Fig. 8c) ensemble fields are adjusted to their mean position with a significant reduction in the positional diversity to below 200 km. However, the outer circulation of the tropical cyclone (i.e., 1015 hPa) in the relocated ensemble fields still maintains a considerable positional deviation, especially for the eastern and northern edges of Hurricane Maria with a positional range > 400 km (highlighted by the red arrow). By comparison,

the members aligned by the FM only have a positional range within 200 km. These results suggest that the adjustment of the outer circulation of the tropical cyclone by the relocation scheme is less effective than that of the FM scheme based on the spatial-dependent field alignment technique. As a result, the mean field of the FM is closer to the reference than the other two mean fields, especially the more realistic narrow trough toward the east to the north of Hurricane Maria.

5. Summary and discussion

Ensemble member forecasts of tropical cyclones become more dispersed in terms of the location, intensity and structure of the tropical cyclone as the predictability gradually degrades. The traditional pointwise arithmetic mean method for ensemble forecasts of tropical cyclones disregards the displacement of the structures related to tropical cyclones in individual ensemble fields, leading to an unrealistic smoothing of the features of the tropical cyclone and suboptimal ensemble mean forecasts of the track and intensity. The consensus approach has therefore been adopted to give a weighted or unweighted average of the track and intensity information for tropical cyclones in individual forecast members.

The consensus forecast has become a routine approach to provide guidance forecasts of the track and intensity of tropical cyclones in operational forecasts. However, ensemble mean forecasts of the spatial structure of tropical cyclones corresponding to the track and intensity consensus forecasts are not yet available. This study introduces a unified ensemble mean for the prediction of tropical cyclones by applying the feature-oriented mean method. Unlike the simple relocation procedure, the feature-oriented mean method implements a spatial-dependent alignment of the structure related to the tropical cyclone in each ensemble member toward the mean of their positions before the amplitude is averaged. The track and intensity forecasts for tropical cyclones can be directly estimated from the feature-oriented mean field.

To evaluate the performance of the feature-oriented mean method in predicting the track, intensity and structure of tropical cyclones, we selected 219 forecast samples from 25 tropical cyclones during the summer of 2017. We found that the track and intensity forecasts for tropical cyclones from the arithmetic mean and the consensus methods are remarkably inconsistent. Specifically, the track forecasts of the two approaches have statistically significant differences in distances throughout all lead times. The arithmetic mean forecasts systematically underestimate the intensity of tropical cyclones relative to the consensus forecast. By contrast, the feature-oriented mean method provides ensemble mean forecasts of the track and intensity of tropical cyclones and their corresponding spatial structure.

In terms of the track forecasts for tropical cyclones, the sample-mean evaluation shows that the arithmetic mean, consensus, and feature-oriented mean methods do not present statistically significant differences. The feature-oriented mean and consensus forecasts have a similar forecast skill for the intensity of tropical cyclones and both methods perform better than arithmetic mean beyond 2 days. With respect to the
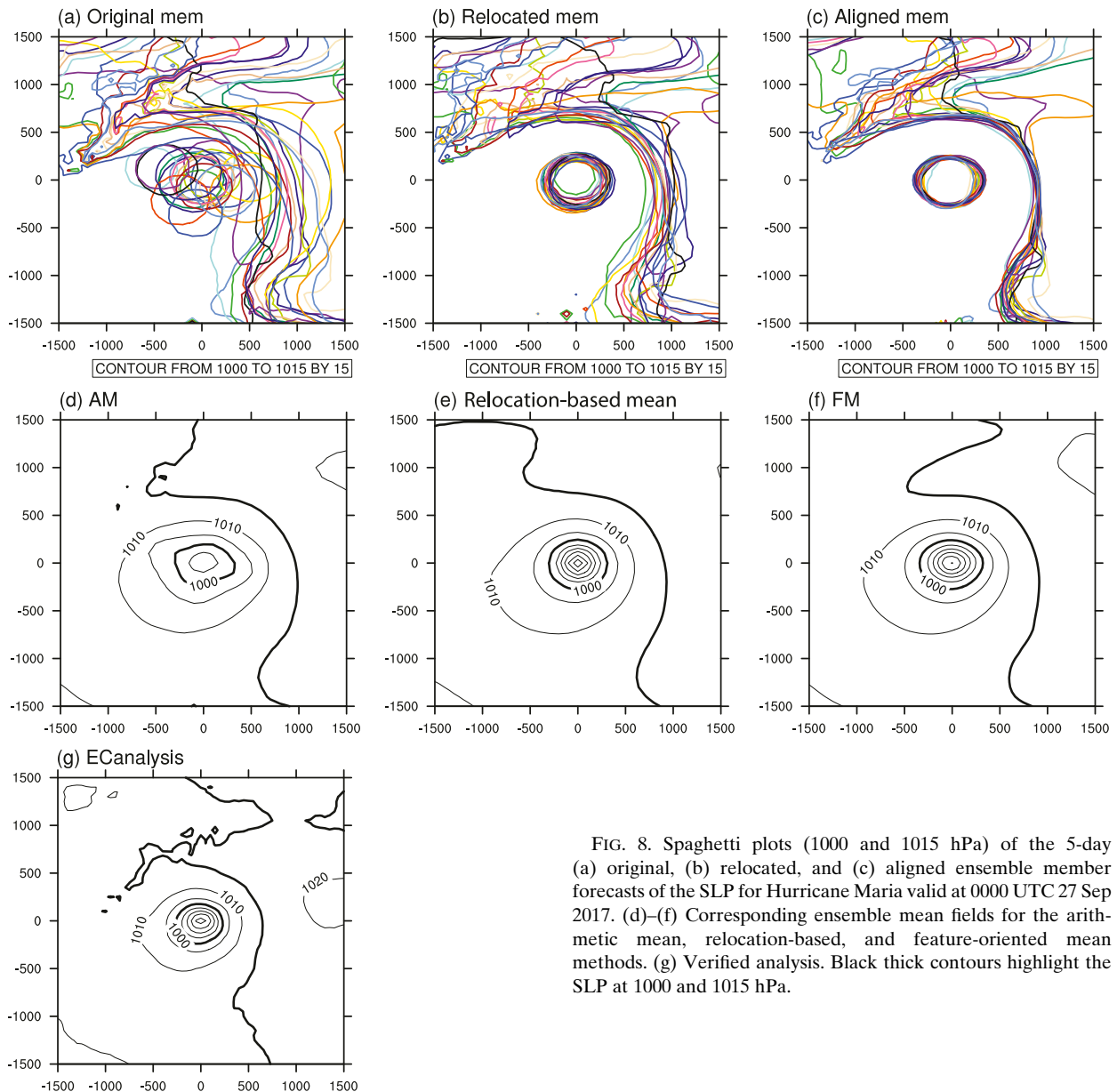
FIG. 8. Spaghetti plots (1000 and 1015 hPa) of the 5-day (a) original, (b) relocated, and (c) aligned ensemble member forecasts of the SLP for Hurricane Maria valid at 0000 UTC 27 Sep 2017. (d)–(f) Corresponding ensemble mean fields for the arithmetic mean, relocation-based, and feature-oriented mean methods. (g) Verified analysis. Black thick contours highlight the SLP at 1000 and 1015 hPa.

forecast of the structure of tropical cyclones, the feature-oriented mean presents a significantly smaller RMSE than the arithmetic mean in predicting both the vortex and environmental circulation of tropical cyclones for all lead times. A case study of Hurricane Maria suggests that the feature-oriented mean method performs better than the relocation-based ensemble mean in predicting the environmental circulation of the tropical cyclone. This is because the relocation procedure does not account for the displacement of the large-scale environment of the tropical cyclone, although it moves the tropical cyclone vortices to the consensus mean position.

With respect to the computational efficiency, the feature-oriented mean method is easy to implement in a parallel computing framework because it adjusts each ensemble forecast independently. The process of each ensemble member can be submitted to a single central processing unit, which makes dramatic savings in the cost of computing. The field alignment step is executed for a global or regional domain, not simply for tropical cyclones. This means that the feature-oriented mean method can simultaneously provide ensemble mean forecasts of multiple variables over the entire model domain in addition to the activities of tropical cyclone. This greatly expands the time-efficient application of the feature-oriented mean method to operational forecasts for different regions and different weather and climate events.

The results of this study, despite using a limited number of examples, provide an initial evaluation of the application of the feature-oriented mean method to ensemble predictions of

tropical cyclones. We plan to use more examples of tropical cyclones to give a more comprehensive assessment of the feature-oriented mean method in the future. Some consensus approaches have been developed to optimally select high-skill tropical cyclone forecast members from single-model or multimodel ensembles to give an unequally weighted average of forecast members. This idea can also be extended to the application of the feature-oriented mean method to further improve the forecast skill for tropical cyclones. It is worth emphasizing that the superior forecasts of the 3D structure of tropical cyclones using the feature-oriented mean method compared with the arithmetic mean and relocation-based mean methods may have more applications—for example, improving the prediction of the position and intensity of tropical cyclone–related wind gusts and heavy precipitation, which could provide useful guidance for tropical cyclone forecasters. Convection-scale ensemble forecasting is another area for the potential application of the feature-oriented mean method.

Data availability statement. Both the forecast and verification product can be downloaded from the THORPEX Interactive Grand Global Ensemble (TIGGE) website (http://apps.ecmwf.int/datasets/data/tigge). The best track dataset over the western North Pacific Ocean and the Atlantic and eastern Pacific Oceans are issued by the Joint Typhoon Warning Center (JTWC; https://www.metoc.navy.mil/jtwc/jtwc.html?western-pacific) and the National Hurricane Center (NHC; https://www.nhc.noaa.gov/data/#hurdat), respectively.

REFERENCES

Beezley, J. D., and J. Mandel, 2008: Morphing ensemble Kalman filters. *Tellus*, **60A**, 131–140, https://doi.org/10.1111/j.1600-0870.2007.00275.x.

Bryan, G. H., and R. Rotunno, 2009: The maximum intensity of tropical cyclones in axisymmetric numerical model simulations. *Mon. Wea. Rev.*, **137**, 1770–1789, https://doi.org/10.1175/2008MWR2709.1.

Buizza, R., 1997: Potential forecast skill of ensemble prediction, and spread and skill distributions of the ECMWF Ensemble Prediction System. *Mon. Wea. Rev.*, **125**, 99–119, https://doi.org/10.1175/1520-0493(1997)125<0099:PFSOEP>2.0.CO;2.

——, 2014: The TIGGE global, medium-range ensembles. ECMWF Tech. Memo. 739, 53 pp., https://www.ecmwf.int/sites/default/files/elibrary/2014/7529-tigge-global-medium-range-ensembles.pdf.

——, P. L. Houtekamer, Z. Toth, G. Pellerin, M. Wei, and Y. Zhu, 2005: A comparison of the ECMWF, MSC, and NCEP global ensemble prediction systems. *Mon. Wea. Rev.*, **133**, 1076–1097, https://doi.org/10.1175/MWR2905.1.

Burton, A., 2006: Sharing experiences in operational consensus forecasting. *Proc. Sixth Int. Workshop on Tropical Cyclones*, San Jose, Costa Rica, WMO/CAS/WWW, Topic 3a., http://severe.worldweather.org/iwtc/document/Topic_3a_Andrew_Burton.pdf.

Doblas-Reyes, F., and Coauthors, 2013: Initialized near-term regional climate change prediction. *Nat. Commun.*, **4**, 1715, https://doi.org/10.1038/ncomms2704.

Dong, L., and F. Zhang, 2016: OBEST: An observation-based ensemble subsetting technique for tropical cyclone track prediction. *Wea. Forecasting*, **31**, 57–70, https://doi.org/10.1175/WAF-D-15-0056.1.

Du, J., S. L. Mullen, and F. Sanders, 2000: Removal of distortion error from an ensemble forecast. *Mon. Wea. Rev.*, **128**, 3347–3351, https://doi.org/10.1175/1520-0493(2000)128<3347:RODEFA>2.0.CO;2.

Feng, J., and Z. Toth, 2017: Spatial extended estimates of analysis and short-range forecast error variances. *Tellus*, **69A**, 1325301, https://doi.org/10.1080/16000870.2017.1325301.

——, and X. G. Wang, 2019: Impact of assimilating upper-level dropsonde observations collected during the TCI field campaign on the prediction of intensity and structure of Hurricane Patricia (2015). *Mon. Wea. Rev.*, **147**, 3069–3089, https://doi.org/10.1175/MWR-D-18-0305.1.

——, R. Q. Ding, D. Q. Liu, and J. P. Li, 2014: The application of nonlinear local Lyapunov vectors to ensemble predictions in Lorenz systems. *J. Atmos. Sci.*, **71**, 3554–3567, https://doi.org/10.1175/JAS-D-13-0270.1.

——, and Coauthors, 2019: The relationship between deterministic and ensemble mean forecast errors revealed by global and local attractor radii. *Adv. Atmos. Sci.*, **36**, 271–278, https://doi.org/10.1007/s00376-018-8123-5.

——, J. Zhang, Z. Toth, M. Peña, and S. Ravela, 2020: A new measure of ensemble central tendency. *Wea. Forecasting*, **35**, 879–889, https://doi.org/10.1175/WAF-D-19-0213.1.

Goddard, L., S. J. Mason, S. E. Zebiak, C. F. Ropelewski, R. Basher, and M. A. Cane, 2001: Current approaches to seasonal to interannual climate predictions. *Int. J. Climatol.*, **21**, 1111–1152, https://doi.org/10.1002/joc.636.

Goerss, J. S., 2000: Tropical cyclone track forecasts using an ensemble of dynamical models. *Mon. Wea. Rev.*, **128**, 1187–1193, https://doi.org/10.1175/1520-0493(2000)128<1187:TCTFUA>2.0.CO;2.

——, C. R. Sampson, and J. M. Gross, 2004: A history of western North Pacific tropical cyclone track forecast skill. *Wea. Forecasting*, **19**, 633–638, https://doi.org/10.1175/1520-0434(2004)019,0633:AHOWNP.2.0.CO;2.

Hamill, T. M., J. S. Whitaker, M. Fiorino, and S. G. Benjamin, 2011: Global ensemble predictions of 2009's tropical cyclones initialized with an ensemble Kalman filter. *Mon. Wea. Rev.*, **139**, 668–688, https://doi.org/10.1175/2010MWR3456.1.

Hoffman, R. N., Z. Liu, J. Louis, and C. Grassotti, 1995: Distortion representation of forecast errors. *Mon. Wea. Rev.*, **123**, 2758–2770, https://doi.org/10.1175/1520-0493(1995)123<2758:DROFE>2.0.CO;2.

Hou, Z. L., J. P. Li, R. Q. Ding, J. Feng, and W. S. Duan, 2018: The application of nonlinear local Lyapunov vectors to the Zebiak–Cane model and their performance in ensemble prediction. *Climate Dyn.*, **51**, 283–304, https://doi.org/10.1007/s00382-017-3920-6.

Jankov, I., and Coauthors, 2020: Partition of forecast error into positional and structural components. *Adv. Atmos.*

Sci., **38**, 1012–1019, https://doi.org/10.1007/s00376-021-0251-7.

Ji, L. Y., X. F. Zhi, C. Simmer, S. P. Zhu, and Y. Ji, 2020: Multimodel ensemble forecasts of precipitation based on an object-based diagnostic evaluation. *Mon. Wea. Rev.*, **148**, 2591–2606, https://doi.org/10.1175/MWR-D-19-0266.1.

Krishnamurti, T. N., S. Pattnaik, M. K. Biswas, E. Bensman, M. Kramer, N. Surgi, and T. S. V. V. Kumar, 2010: Hurricane forecasts with a mesoscale suite of models. *Tellus*, **62A**, 633–646, https://doi.org/10.1111/j.1600-0870.2010.00469.x.

Leith, C. E., 1974: Theoretical skill of Monte Carlo forecasts. *Mon. Wea. Rev.*, **102**, 409–418, https://doi.org/10.1175/1520-0493(1974)102<0409:TSOMCF>2.0.CO;2.

Li, H., J. Y. Luo, and M. T. Xu, 2019: Ensemble data assimilation and prediction of typhoon and associated hazards using TEDAPS: Evaluation for 2015–2018 seasons. *Front. Earth Sci.*, **13**, 733–743, https://doi.org/10.1007/s11707-019-0794-4.

Li, J. P., J. Feng, and R. Q. Ding, 2018: Attractor radius and global attractor radius and their application to the quantification of predictability limits. *Climate Dyn.*, **51**, 2359–2374, https://doi.org/10.1007/s00382-017-4017-y.

Li, W. W., Z. Wang, and M. S. Peng, 2016: Evaluating tropical cyclone forecasts from the NCEP Global Ensemble Forecasting System (GEFS) reforecast version 2. *Wea. Forecasting*, **31**, 895–916, https://doi.org/10.1175/WAF-D-15-0176.1.

Liu, Q., S. J. Lord, N. Surgi, Y. Zhu, R. Wobus, Z. Toth, and T. Marchok, 2006: Hurricane relocation in Global Ensemble Forecast System. *27th Conf. on Hurricanes and Tropical Meteorology*, Monterey, CA, Amer. Meteor. Soc., P5.13, https://ams.confex.com/ams/pdfpapers/108503.pdf.

——, and Coauthors, 2020: Vortex initialization in the NCEP operational hurricane models. *Atmosphere*, **11**, 968, https://doi.org/10.3390/atmos11090968.

Lu, X., X. Wang, Y. Li, M. Tong, and X. Ma, 2016: GSI-based ensemble-variational hybrid data assimilation for HWRF for hurricane initialization and prediction: Impact of various error covariances for airborne radar observation assimilation. *Quart. J. Roy. Meteor. Soc.*, **143**, 223–239, https://doi.org/10.1002/qj.2914.

Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliagis, 1996: The new ECMWF Ensemble Prediction System: Methodology and validation. *Quart. J. Roy. Meteor. Soc.*, **122**, 73–119, https://doi.org/10.1002/qj.49712252905.

Nehrkorn, T., R. Hoffman, C. Grassotti, and J.-F. Louis, 2003: Feature calibration and alignment to represent model forecast errors: Empirical regularization. *Quart. J. Roy. Meteor. Soc.*, **129**, 195–218, https://doi.org/10.1256/qj.02.18.

——, B. Woods, T. Auligné, and R. N. Hoffman, 2014: Application of feature calibration and alignment to high-resolution analysis: Examples using observations sensitive to cloud and water vapor. *Mon. Wea. Rev.*, **142**, 686–702, https://doi.org/10.1175/MWR-D-13-00164.1.

Puri, K., J. Barkmeijer, and T. N. Palmer, 2001: Ensemble prediction of tropical cyclones using targeted diabatic singular vectors. *Quart. J. Roy. Meteor. Soc.*, **127**, 709–732, https://doi.org/10.1002/qj.49712757222.

Qi, L., H. Yu, and P. Chen, 2014: Selective ensemble-mean technique for tropical cyclone track forecast by using ensemble prediction systems. *Quart. J. Roy. Meteor. Soc.*, **140**, 805–813, https://doi.org/10.1002/qj.2196.

Qian, C., Y. Duan, S. Ma, and Y. Xu, 2012: The current status and future development of China operational typhoon

forecasting and its key technologies. *Adv. Meteor. Sci. Technol.*, **2**, 36–43.

Ravela, S., 2012: Quantifying uncertainty for coherent structures. *Proc. Comput. Sci.*, **9**, 1187–1196, https://doi.org/10.1016/j.procs.2012.04.128.

——, 2014: Spatial inference for coherent geophysical fluids by appearance and geometry. *IEEE Winter Conf. on Applications of Computer Vision*, Steamboat Springs, CO, IEEE, 925–932, https://doi.org/10.1109/WACV.2014.6836005.

——, K. Emanuel, and D. McLaughlin, 2007: Data assimilation by field alignment. *Physica D*, **230**, 127–145, https://doi.org/10.1016/j.physd.2006.09.035.

Rotunno, R., and K. Emanuel, 1987: An air–sea interaction theory for tropical cyclones. Part II: Evolutionary study using a nonhydrostatic axisymmetric numerical model. *J. Atmos. Sci.*, **44**, 542–561, https://doi.org/10.1175/1520-0469(1987)044<0542:AAITFT>2.0.CO;2.

Ryglicki, D. R., and R. E. Hart, 2015: An investigation of center-finding techniques for tropical cyclones in mesoscale models. *J. Appl. Meteor. Climatol.*, **54**, 825–846, https://doi.org/10.1175/JAMC-D-14-0106.1.

Sampson, C. R., J. S. Goerss, and A. J. Schrader, 2005: A consensus track forecast for Southern Hemisphere tropical cyclones. *Aust. Meteor. Mag.*, **54**, 115–119.

Tallapragada, V., and Coauthors, 2015: Hurricane Weather Research and Forecasting (HWRF) Model: 2015 scientific documentation. NCAR Tech Note NCAR/TN-522+STR, https:doi.org/10.5065/D6ZP44B5.

Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteor. Soc.*, **74**, 2317–2330, https://doi.org/10.1175/1520-0477(1993)074<2317:EFANTG>2.0.CO;2.

——, and ——, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.*, **125**, 3297–3319, https://doi.org/10.1175/1520-0493(1997)125<3297:EFANAT>2.0.CO;2.

——, Y. Zhu, and T. Marchok, 2001: The use of ensembles to identify forecasts with small and large uncertainty. *Wea. Forecasting*, **16**, 436–477, https://doi.org/10.1175/1520-0434(2001)016<0463:TUOETI>2.0.CO;2.

Yanai, M., 1964: Formation of tropical cyclones. *Rev. Geophys.*, **2**, 367–414, https://doi.org/10.1029/RG002i002p00367.

Yuan, H., Z. Toth, M. Peña, and E. Kalnay, 2018: Overview of weather and climate systems. *Handbook of Hydrometeorological Ensemble Forecasting*, Q. Duan et al., Eds., Springer, 1–31, https://doi.org/10.1007/978-3-642-40457-3_10-1.

Zhang, H. B., and Coauthors, 2015: Study of the modification of multi-model ensemble schemes for tropical cyclone forecasts. *J. Trop. Meteor.*, **21**, 389–399.

Zhang, J., W. Duan, and X. Zhi, 2015: Using CMIP5 model outputs to investigate the initial errors that cause the ''spring predictability barrier'' for El Niño events. *Sci. China Earth Sci.*, **58**, 685–696, https://doi.org/10.1007/s11430-014-4994-1.

Zhang, L., and X. F. Zhi, 2015: Multi-model consensus forecasting of low temperature and icy weather over central and southern China in early 2008. *J. Trop. Meteor.*, **21**, 67–75.

Zhang, X., and H. Yu, 2017: A probabilistic tropical cyclone track forecast scheme based on the selective consensus of ensemble prediction systems. *Wea. Forecasting*, **32**, 2143–2157, https://doi.org/10.1175/WAF-D-17-0071.1.

Zhang, Z., and T. N. Krishnamurti, 1999: A perturbation method for hurricane ensemble predictions. *Mon. Wea.*

Rev., **127**, 447–469, https://doi.org/10.1175/1520-0493(1999)127<0447:APMFHE>2.0.CO;2.

——, V. Tallapragada, C. Kieu, S. Trahan, and W. Wang, 2014: HWRF based ensemble prediction system using perturbations from GEFS and stochastic convective trigger function. *Trop. Cyclone Res. Rev.*, **3**, 145–161, https://doi.org/10.6057/2014TCRR03.02.

Zhou, X., Y. Zhu, D. Hou, and D. Kleist, 2016: A comparison of perturbations from an ensemble transform and an ensemble Kalman filter for the NCEP Global Ensemble Forecast System. *Wea. Forecasting*, **31**, 2057–2074, https://doi.org/10.1175/WAF-D-16-0109.1.

——, ——, ——, Y. Luo, J. Peng, and R. Wobus, 2017: Performance of the NCEP Global Ensemble Forecast System in a parallel experiment. *Wea. Forecasting*, **32**, 1989–2004, https://doi.org/10.1175/WAF-D-17-0023.1.

Zhu, Y., Z. Toth, R. Wobus, D. Richardson, and K. Mylne, 2002: The economic value of ensemble-based weather forecasts. *Bull. Amer. Meteor. Soc.*, **83**, 73–83, https://doi.org/10.1175/1520-0477(2002)083<0073:TEVOEB>2.3.CO;2.

——, D. Hou, M. Wei, R. Wobus, J. Ma, B. Cui, and S. Moorthi, 2012: February 2012 upgrade of the NCEP Global Ensemble Forecast System (NAEFS). NOAA, 29 pp., https://www.emc.ncep.noaa.gov/gmb/yzhu/imp/i201109/GEFS_Science_20120208.pdf.