The DOI for this manuscript is
DOI:10.2151/jmsj.2022-029
J-STAGE Advance published date: March 10th, 2022

Improvement of the Ensemble Methods in the Dynamical–Statistical–Analog Ensemble Forecast Model for Landfalling Typhoon Precipitation

Li JIA[1, 2], Fumin REN[2], Chenchen DING[1,2], Zuo JIA[3], Mingyang WANG[1,2]

Yuxu CHEN[4], and Tian FENG[5]

1 Key Laboratory of Meteorological Disaster, Ministry of Education (KLME), Nanjing University of Information Science and Technology, Nanjing, China

2 State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing 100081, China;

3 CSSC Marine Technology Co., Ltd., Beijing 100070, China;

4 Shantou Meteorological Bureau, Guangdong 515000, China;

5 Haikou Meteorological Station, Haikou Meteorological Bureau, Hainan 570100, China

Submitted: December 20, 2020

First Revised: March 29, 2021; Second Revised: August 1, 2021

Response to comments from JMSJ Editorial Committee: 29-Oct-2021

Response to Editor: 19-Dec-2021

----

1) Corresponding author: Dr. Fumin REN, State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing 100081 China.

E-mail: fmren@163.com

Tel: +86 139-1032-4105

Abstract

The Dynamical–Statistical–Analog Ensemble Forecast model for landfalling typhoon precipitation (the DSAEF_LTP model) identifies tropical cyclones (TCs) from history data that are similar to a target TC, and then assembles the precipitation amounts and distributions of those identified to obtain those of the target TC. Two original ensemble methods in the DSAEF_LTP model, mean and maximum, tend to under- and over-forecast TC precipitation, respectively. In addition, these two methods are unable to forecast precipitation at stations beyond their maxima. To overcome the shortcomings and improve the forecast performance of the DSAEF_LTP model, the following five new ensemble methods are incorporated: optimal percentile, fuse, probability matching mean, equal difference-weighted mean, and TSAI (Tropical cyclone track Similarity Area Index)-weighted mean. Then, model experiments for landfalling TCs over China in 2018 are conducted to evaluate the forecast performance of the DSAEF_LTP model with the new ensemble methods. Results show that the overall performance of the optimal percentile (the 90th percentile) ensemble method is superior, with the false alarm rate lower than that of the original ensemble methods. As compared to five operational numerical weather prediction models, the improved DSAEF_LTP model shows advantages in predicting accumulated rainfall, especially with the rainfall of over 250 mm. When implementing the experiments, above results, however, it is found that the model forecast performance varies, depending on the type of TC tracks.

That is, the accumulated rainfall forecast for westbound TCs is significantly better than that of northbound TCs. To address this issue, different schemes are used to forecast the accumulated rainfall of TCs with the two different track types. The precipitation forecast performance for westbound and northbound TCs, using the 90th percentile and the probability-matched ensemble mean ensemble method, respectively, is much better than that using a single ensemble method for all the TCs.

Keywords: landfalling tropical cyclone, heavy precipitation forecast, dynamical statistical model, ensemble forecast

1. Introduction

China is the country with the world's most frequent landfalling tropical cyclones (TCs, also known as typhoons in the western North Pacific) and TC-related disasters (Chen and Meng, 2001; Zhang et al., 2009) that include strong winds, storm surges, and heavy rainfall. TC-related strong winds and surges primarily occur in coastal areas near the landfall sites of TCs, while the rainfall of TCs can cause widespread and significant damage, even affecting the hinterland (Chao et al., 2005; Chen et al., 2010; Luitel, 2016). Besides, many studies have shown that, although the number of landfalling TCs (LTCs) in China has decreased (Ren et al., 2011; Gu et al., 2016; Knutson et al., 2020), the number of disasters caused by the LTCs has increased (Emanuel et al., 2005; Chan et al., 2008; Barthel et al., 2012; Weinkle et al., 2012). The mechanisms and forecasts of TC precipitation have attracted much attention (Chen et al., 2006; Woo et al., 2014; Rogers, 2018).

Numerical weather prediction (NWP) models are the main tools for LTC precipitation forecast. The continuous development of key techniques in NWP has improved NWP-based precipitation prediction for LTCs significantly. There are two main categories of these studies. The first one focuses on the improvement of the initial fields of the NWP models by assimilation technology. Many studies（Xiao et al., 2007; Zhao et al., 2012; Zhang and Pu, 2014; Zhu et al., 2016）showed that the forecast performance could be improved by using assimilation techniques. The second one focuses on the improvement of the parameterization of different physical processes. Ma and Tan (2009) and Yu et al. (2013) improved the forecast performance for the typhoon precipitation by Kain-Fritsch convective parameterization scheme; Xue et al. (2007) improved the parameterization scheme suitable for the forecast of the typhoon precipitation in Zhejiang and Fujian provinces. However, overall, the ability of NWP models to forecast LTC precipitation remains limited (Marchok et al., 2007; Wang et al., 2012; Ma, 2014). Thus, some researchers have explored alternative methods other than NWP models for forecasting LTC precipitation. In this regard, the dynamical–statistical method has received considerable attention (Zhong et al., 2009; Wei, 2012a, b; Li et al., 2015). Recently, Ren et al. (2020) proposed the Dynamical–Statistical–Analog Ensemble Forecast (DSAEF) model and then applied it to LTC precipitation forecasts (DSAEF_LTP). This model searches for TCs that are similar to a target TC in accordance with the similarity of the generalized initial value (GIV) that contains the value of some factors affecting TC precipitation. TC track and landing season are considered as the two major factors in the first version of the DSAEF_LTP model. The word of "generalized" means that both the observed value before the time to forecast (initial time) and the forecasted value after the initial time are included. Then, the accumulated precipitation data of TCs that are similar to the target TC are treated as an ensemble precipitation forecast for the target TC. The model has been further improved on its forecasting performance by considering the GIV of a new variable (i.e., TC intensity) and modify parameter ranges of the existing parameters (Ding et al., 2020; Jia et al., 2020).

In recent years, quantitative precipitation forecast (QPF) techniques based on ensemble techniques have been developed rapidly (Ebert, 2001; Clark et al., 2017;

Sofiati and Nurlatifah, 2019), and have also been applied to LTC precipitation forecasts (Cheung et al., 1999; Zhang et al., 2007; Chen et al., 2016). Among the various ensemble prediction methods, an important class is the integration of ensemble members or multi-model predictions, including the probability matching mean (PM) (Clark et al., 2012; Fang et al., 2013; Surcel et al., 2014), multi-model similar integration (Chen et al., 2005; Lin et al., 2013), optimal percentile (Dai et al., 2016), and ensemble pseudo-bias-corrected QPF (Novak et al., 2014; Alexander et al., 2019; Binh et al., 2020) methods, which yield the most possible single-value forecast by extracting or overlaying valid information.

Ensemble forecast is a key technology of the DSAEF_LTP model because it determines the forecast performance when similar TCs are selected. However, this model only contains mean and maximum ensemble methods, which have their disadvantages in terms of the high rates of misses and false alarms, respectively. Besides, the largest predicted rainfall in a given station may reach is the maximum historical TC precipitation of the station. Thus, there remains considerable room for the ensemble methods of the DSAEF_LTP model to improve. Applying new ensemble methods to the DSAEF_LTP model is likely to further improve its forecast performance. Therefore, the goal of this study is to develop new ensemble methods in the DSAEF_LTP model and evaluate whether its forecast performance can be further improved.

The paper is structured as follows: The next section describes the data and methods. Section 3 presents experiment design. Section 4 analyzes the results. A summary and discussion are given in the final section.

2. Data and methods

2.1 Data

The data used in this paper include historical observed precipitation data during 1960–2018 that were archived at 24-h intervals at 1200 UTC by the China Meteorological Administration (CMA). The data are from 2027 rain gauge stations covering most of China (2006 on mainland China and 21 on Taiwan Island).

To compare the forecast performance of the DSAEF_LTP model with NWP models, we employ three global models and two regional models—namely, the European Centre for Medium-Range Weather Forecasts (ECMWF) model; the Global Forecast System (GFS) of the National Centers for Environmental Prediction; the Global/Regional Assimilation and Prediction System (GRAPES) model run by the CMA (cma.gov, 2011); Shanghai Meteorological Service WRF ADAS Real-Time Modeling System (SMS-WARMS) (Xu et al., 2016); and Rapid-refresh Multi-scale Analysis and Prediction System (RMAPS) developed by Institute of Urban Meteorology, CMA (Tao et al., 2019). The corresponding rainfall forecast data of these models are obtained with the grid spacing of $0.1° \times 0.1°$.

The historical best-track data at 6-h intervals during 1960–2018, including the position and strength of TCs, are from the Shanghai Typhoon Institute (Ying et al., 2014). Additionally, the operational NWP model-forecast tracks of 10 TCs, whose precipitation amounts are to be forecast, are obtained from the CMA.

2.2 Methods

The Objective Synoptic Analysis Technique for partitioning TC precipitation (Ren et al., 2001 and 2007; Wang et al., 2006) is used in this paper. This method can identify the precipitation generated by TCs from daily observed precipitation data based on the distance between the stations and the precipitation centers. There are 1041 TCs from 1960 to 2018 being identified through this method. The onshore precipitation period in China during one single TC, which is named after influence period, is obtained as well. The precipitation discussed in this paper, whether observed or predicted by the DSAEF_LTP model or NWP models, is the total process precipitation during the influence period.

The DSEAF_LTP model is used to identify historical TCs that are similar to a target TC. These identified TCs occurred before the target TC are named as analogs. Then, the DSEAF_LTP model uses these analogs' precipitation to obtain ensemble forecast. The specific steps of the DSAEF_LTP model are given in section 3.2.

To identify TCs whose tracks are similar to the target TC, the objective TC track Similarity Area Index (TSAI) (Ren et al., 2018) is used. The principle of the TSAI is to calculate the area enclosed by the track of the historical TCs and the target TC over a certain region. The smaller the TSAI value is, the higher is the similarity.

The threat score (TS) and bias score ($BIAS$), which are widely used in the operational weather prediction, are the two basic criteria for determining the forecast performance in this study. TS is defined as $TS = \frac{hits}{hits + misses + false\ alarms}$, indicating the fraction of correctly predicted forecast events. It varies from 0 to 1. The closer it is to 1, the higher is the hit rate. $BIAS$ is defined as $BIAS = \frac{hits + false\ alarms}{hits + misses}$, indicating whether the forecast system has a tendency to underestimate ($BIAS < 1$) or overestimate ($BIAS > 1$). $Hits$ denotes the number of stations which the event is forecast to occur, and does occur; $misses$ is the number of stations which the event is forecast not to occur, but does occur; $false\ alarms$ is the number of stations which the event is forecast to occur, but does not occur.

Since 100 and 250 mm are important thresholds used in the operational forecasts of extreme precipitation for LTCs in China, and since the DSAEF_LTP model shows advantages in predicting extreme precipitation (Ren et al., 2020), the two values are used for the precipitation thresholds of interest for this study. $TS100(BIAS100)$ and $TS250\ (BIAS250)$ are $TS(BIAS)$ defined as the two thresholds above 100 and 250 mm, respectively. To evaluate the forecast performances at the two thresholds, we apply $TSsum = TS100 + TS250$; $BSsum = \pm(|BIAS100 - 1| + |BIAS250 - 1|)$, where the symbol depends on whether $(BIAS100 + BIAS250 - 2)$ is positive or negative; namely, positive values indicate overprediction while negative values indicate underprediction. Accordingly, a larger *TSsum* or a smaller absolute value of *BSsum* indicates a better forecast performance of the DSAEF_LTP model at these two thresholds.

3. Experiment design

3.1 Experiment samples

Ten LTCs that occurred in 2018 over China from June to September are selected as samples. Usually, seven or eight LTCs occur during this period; however, in 2018, there were ten LTCs. These LTCs caused widespread heavy precipitation over the coastal areas of China, which posed a great challenge in terms of precipitation forecast. Figure 1 shows the observed tracks of the 10 LTCs selected for the experiment and their TC numbers. The intensities of these 10 TCs range from tropical storm (wind speed ≥ 17.2m/s) to super typhoon (wind speed ≥ 51.0 m/s). The single-station maximum precipitation during one TC varies greatly from 116.4 to 618.9 mm. They made landfall in South or East China, and moved westward or northward afterward.

3.2 Steps in applying the DSAEF_LTP model

The DSAEF_LTP model used to perform accumulated precipitation simulation experiments involves four steps (Ren et al., 2020) as shown in Figure 2. Table 1 lists the parameters (i.e., P1 to P8) of the DSAEF_LTP model. Specific steps are given as follows.

(1) Obtaining the forecast TC track. As shown in Table 1, the initial time (P1) is determined by the landfall day of a target TC. The first step is to combine the observed track of the target TC before the initial time and the forecast track after the initial time into its complete track. The observed track is the historical best track data of the Shanghai Typhoon Institute, as mentioned in section 2.1. The forecast TC track can be obtained by the NWP model.

(2) Constructing the GIV. The second step involves constructing the GIV for variables that have impacts on LTC precipitation, which includes TC track, landfall

season, and intensity. For example, both the observed and predicted tracks for the target TC are treated as the GIV.

(3) Identifying *m* analogs. The third step is to discriminate the similarity of the GIV constructed in the second step between the target TC and the historical TCs, and then select *m* top analogs that resemble most the target TC. Parameters P2 to P6 are used in this step. P2 limits the region where similar tracks are found; and P3 and P4 are used to determine the bend and degree of overlap of two tracks respectively. TSAI can be calculated only if the values of P3 and P4 meet certain conditions. Thus, P1 to P4 determine the track similarity. The similarity between TC landfall seasons and intensities can be divided into different types, as defined by P5 and P6 in Table 1 respectively.

For example, if P1 is 1, the initial time is 1200 UTC on the day of TC precipitation occurring on land. If P2 is 2, P3 is 3 and P4 is 4, the TSAI are calculated in the second similarity region when the bending degree of TC tracks is less than 0.3 and the degree of longitude (latitude) overlap of TC tracks is greater than 0.6. Then, historical TCs are ranked according to the TSAI. If P5 is 5 and P6 is (1,4), the ranked TCs, whose landfall times are 15 days different from the target TC and average intensity on the first rainy day are the same grades as the target TC, can be seen as analogs. Ultimately, *m* (depending on P7) analogs with the GIVs that are most similar to the GIV of the target TC could be selected, and their accumulated precipitation amounts are the ensemble members of the DSAEF_LTP model.

(4) Finding the ensemble LTP of the analogs. The final step is to derive the target TC

accumulated precipitation by assembling the ensemble members with the ensemble methods decided by P8 in Table 1 and Table 2, as described in detail in section 3.3.

3.3 Ensemble methods in the DSAEF_LTP model

This study uses the DSAEF_LTP model with new ensemble methods to perform simulation experiments. The previous version of the DSAEF_LTP model only had two ensemble methods (i.e., mean and maximum). The forecast rainfall at a station can be the maximum or mean value of rainfall of the *m* analogs at that station. In this study, five new ensemble methods have been added, namely, optimal percentile, fuse, PM, equal difference-weighted mean (ED-WM), and TSAI-weighted mean (TSAI-WM).

The specific calculation steps of the seven ensemble methods are given in Table 2. The mean and maximum ensemble methods forecast the precipitation at each station by calculating the average and max precipitation, respectively, at each station of the selected analogs. Since these two methods always tend to underestimate and overestimate precipitation, respectively, percentiles were introduced. To get the optimal percentile of the best forecast performance, the 60th to 95th percentiles, at 5 percentile intervals, are applied to simulating the precipitation of the 10 LTCs. Results show that the 90th percentile is the optimal one. Thus, the 90th percentile is adopted in this study.

The fuse ensemble method is also adopted to obtain the target TC's precipitation by employing different percentile ensemble methods determined by the precipitation of *m* analogs in order to achieve better forecast performance. This method can be implemented by following the calculation rules shown in Table 2. The criteria in the fuse are checked in order. If one criterion is met, the rest will not be checked.

Because the forecasted precipitation at a station by using these four methods (i.e., mean, maximum, 90th percentile and fuse) only ensemble m analogs' precipitation at the station, the forecast precipitation of a certain station cannot be affected by data from other stations. These methods are called station-based ensemble methods. However, they have two drawbacks: First, they are unable to forecast precipitation at a certain station beyond the historical maximum of itself. Second, they greatly reduce the amount of historical data that can be used in the precipitation forecast at a certain station. Thus, three field-based ensemble methods were added to take advantage of information from all stations.

Historical precipitation data from the remaining stations are directly used when using PM to forecast the precipitation at a station. By using this method, the higher the average precipitation of the selected analogs at a certain station, the higher is the forecasted precipitation. The forecast values, whose algorithm is given in Table 2, depend on the precipitation of the similar TCs selected at all stations.

The ED-WM ensemble method can be achieved by assigning equal differential weights to the precipitation amounts of the selected m analogs in order of similarity. That is, the higher the similarity is, the more weight will be given to the precipitation of that analog is. Thus, the weight of precipitation for each similar TC selected is $W(i) = \frac{(2\times m-i)\times 2}{(3\times m-1)\times m}(i = 1,2,..,m)$.

TSAI-WM takes an important indicator of TSAI as the similarity between TCs into account. Thus, it may be more valid than simply considering the similarity rank. Since the smaller the TSAI is, the higher is the degree of similarity, taking the reciprocal of the TSAI for each selected m analog to obtain $A(i) = \frac{1}{TSAI(i)} (i = 1,2,..,m)$ and further obtain the precipitation weight of these analogs, $W(i) = \frac{A(i)}{\sum_{i=1}^{m} A(i)}$. The sum of the weights of m analogs of the ED-WM and TSAI-WM ensemble methods are both 1. The ensemble forecast precipitation is $Prep$, $Prep = \sum_{i=1}^{m} W(i) \times Pre(i)$. The weights of ED-WM and TSAI-WM depend on the rank of analogs, which is determined from the data of all stations, and thus affect the forecast results.

See section 4 for the performances of these seven ensemble methods.

3.4 Steps for selecting the best scheme

As each parameter in the DSAEF_LTP model has several different options, thousands of combinations are possible. Each combination is referred to as a forecast scheme. The purpose of the experiment is to determine the best scheme with the highest *TSsum* when an ensemble method was chosen, and then compare the highest *TSsum* under seven ensemble methods. Thus, seven experiments are designed in this study by applying different ensemble methods.

Steps for selecting the best scheme in an experiment were as follows: First, the *TS100* and *TS250* of every scheme are calculated when simulating a single TC. Due to the short impact period of some TCs, some options of the initial time (P1) and similarity region (P2) could not be chosen. Thus, the number of valid schemes for a TC is always less than or equal to the total number of the schemes given in Table 1. The second step is to select the schemes that could yield forecast for all the 10 LTCs. These schemes are called common schemes. The third step was to calculate the *TS100*, *TS250* and *TSsum* of each common scheme, i.e., the mean *TS100*, *TS250* and *TSsum* of each common scheme for the 10 LTCs. The common scheme with the maximum *TSsum* in each experiment could then be regarded as the best scheme in that experiment.

It should be mentioned that, with the different ensemble methods, the values of the remaining parameters of the best scheme can be different. Since the ensemble methods in every experiment are different, we represent an experiment by the name of the ensemble method used in that experiment. The performance of an ensemble method refers to the performance of the best scheme in the experiment with this ensemble method.

4. Results

4.1 Comparison of results in seven experiments

Seven experiments are conducted and the best scheme was selected for each experiment. The best scheme of an experiment was determined by their maximum *TSsum*. Table 3 and Fig. 3 show the choice of parameters and *TS* (including *TSsum*, *TS100*, and *TS250*) for the best schemes of the seven experiments respectively. It is evident from Table 3 that the parameter values of the best scheme with different ensemble methods are similar. The criteria used by the model to select similar TCs are similar. This means that there is always a criterion for selecting similar TCs that makes the DSAEF_LTP model better forecast performance. In other words, the stability of the model is satisfactory. Especially, these values appear to be the same between the maximum and fuse methods, as well as the ED-WM and TSAI-WM methods. However, the *TS* values in Fig.3 are different. That is, if parameters P1 to P7 are assigned values, the forecast performance is determined by the ensemble method (P8). This indicates that the ensemble method plays an important role in determining the forecast performance of the DSAEF_LTP model.

As can be seen from Fig.3, the station-based ensemble methods (the first four ensemble methods in Table 2) show better forecast performance than the field-based ensemble methods; the overall forecast performance of the 90th percentile is the best, i.e., the *TSsum* of the best scheme with the 90th percentile ensemble method is the highest. This may be due to the fact that the precipitation distribution of the selected analogs by the DSAEF_LTP model is very similar to that of the target TC. Therefore, obtaining the ensemble forecast using the precipitation of the station itself performs better. The fuse and maximum ensemble schemes rank the second. They have the same *TSsum* value because they obtain the same forecast of precipitation of more than 100 mm. A difference between the two methods is that the fuse scheme reduces the rate of misses for less than 100 mm precipitation. The *TS250* is maximized when the 90th percentile is adopted, while the *TS100* is the highest when the ensemble method is fuse or maximum. This is consistent with the conclusion of some previous studies (e.g., Chen et al., 2015; Li et al., 2018). This shows that for different levels of precipitation forecast, using different percentile of precipitation of selected analogs might improve forecast performance. Besides, the different *TSsum* values of the first four ensemble methods from those of the last three ensemble methods are mainly reflected in predicting the precipitation of over 100 mm. The advantage of using the station-based ensemble methods in terms of the forecast performance of over 250 mm is small, which may be due to the fact that over 250 mm rainfall of analogs is relatively scattered in distribution. Besides, the forecast performance of PM is better than that of the other two field-based ensemble methods. This is because only this method directly uses the precipitation data of all stations to obtain forecast at a certain station of concern.

The *TS* for individual TCs by the best schemes in the seven experiments is given in Fig.4, showing that generally, the station-based ensemble methods outperform the field-based ensemble methods. This is most evident in TC1823, in which the *TSsum* of the 90th percentile is 0.369 higher than that of the TSAI-WM ensemble method, followed by TC1816, in which the *TSsum* of fuse is 0.348 higher than that of the PM value. Figure 4 also shows that the forecast performance of each ensemble method for TC1808 is significantly different from that for the other LTCs. The 90th percentile and fuse predictions, which perform better than the other ensemble methods, are less effective, while the field-based ensemble method performs better. This is because precipitation at the other stations can be used as ensembles in the field-based ensemble methods, which leads to the forecasted precipitation exceeding the historical extreme value. The field-based ensemble method makes up for the fact that the station-based ensemble method cannot predict extreme precipitation that exceeds the historical record at certain stations.

4.2 Forecast comparison between the DSAEF_LTP model and five NWP models

As shown in Fig. 3, the *TSsum* and *TS250* values of the best schemes in the seven experiments exceed those of the three NWP and two regional NWP models. However, for the prediction of precipitation exceeding 100 mm, only the fuse, 90th percentile, and maximum method outperform the performances of all the NWP models except for the GFS.

Figure 4 compares the forecast performance of the DSAEF_LTP model in the seven experiments to that of the NWP models for 10 LTCs. The *TSsum* of the 90th percentile ensemble method ranks top three while simulating most of the LTCs' accumulated precipitation. Three LTCs (i.e., 1810, 1812, and 1814) are poorly predicted by the 90th percentile ensemble method. The single-station observed maximum total rainfall amounts of these three LTCs are the three smallest among the 10 LTCs, with values of 182.7, 224.8, and 295.7 mm, respectively, as indicated by dotted lines in Fig. 4. The advantage of the forecasts by the DSAEF_LTP model in this experiment is mainly in the prediction of precipitation for LTCs with the large amounts of accumulated precipitation.

4.3 DSAEF_LTP model track-type experiments and results

Figure 4 demonstrates that even the best scheme for each experiment poorly simulates the precipitation of TC1810, TC1812, TC1814, and TC1818, which are all northbound TCs (Fig. 1). Since the best scheme for the current experiments produces relatively poor simulations of the northbound TCs compared to the westbound ones, different schemes for TCs with different track types are considered for the simulation of the accumulated precipitation. Thus, the track-type experiments are conducted, in which the 10 LTCs are grouped into two based on their tracks. Namely, westbound TCs (i.e., TC1804, TC1809, TC1816, TC1822, and TC1823) and northbound TCs (i.e., TC1808, TC1810, TC1812, TC1814, and TC1818). The common schemes of the five TCs in each experiment are first selected, and then the *TS100*, *TS250*, and *TSsum* values for the common schemes of the two experiments are calculated, separately. The scheme with the largest *TSsum* is considered as the best scheme.

The LTCs with the two different track types are simulated with the different best-performing schemes. Results show that the selected best scheme for the westbound TCs is the same as that for the 10 TCs. That is, the parameters of P1-P7 take values of 1, 20, 1, 6, 3, 2, 5, and 3 (Table 3), respectively, with the 90th percentile ensemble method used. By comparison, for the northbound TCs, the parameters of P1-P7 in the best scheme take values of 2, 20, 1,5, 3, 4, 3, and 5, respectively, with the PM ensemble method applied. The precipitation forecasts for the westbound TCs are better when a station-based ensemble method is selected, whereas there is little advantage of the station-based ensemble method for the northbound TCs. Besides, the average *TSsum* of the field-based ensemble method is 0.014 higher than that of the station-based method. The stations with maximum precipitation associated with LTCs during 1960–2018 are given in Fig. 5. The map shows the stations with a maximum precipitation, along with the times that a station being the maximum total rainfall station. The better forecast performances of the station-based ensemble approach for the westbound TCs and the field-based ensemble method for the northbound TCs may be attributed to the large precipitation centers of the westbound TCs that are located in Southern China (i.e., Hainan, Guangdong, Fujian provinces and Taiwan Island). These precipitation centers are usually concentrated on some stations, while the precipitation levels vary widely between stations. Thus, for one particular meteorological station, obtaining the ensemble forecast result by assembling the precipitation of the station itself is reasonable. By comparison, large-value centers of the northbound TC are less frequent and more scattered. Thus, using the TC precipitation information of a single station itself may smooth out the large values or overestimate the precipitation of this station. However, PM can combine the accurate precipitation location of the ensemble-averaged forecast and the good precipitation magnitude evaluated through selected ensemble members to obtain a better forecast.

By comparing Figs. 3 and 6, it can be seen that the *TSsum* of each ensemble method has risen for the two track types of LTCs. The new ensemble methods increase the *TSsum* of the westbound TCs. The *TSsum* with the 90th percentile method for the westbound TCs increased 0.191 more than that of the northbound TCs. By comparing Figs. 4 and 7, the most obvious improvement of *TSsum* after classifying the TC track types occurs for TC1808. Compared to the five NWP models, the superiority of the DSAEF_LTP model is obvious in the case of TC1816. Besides, the forecast performance of different ensemble methods varies greatly for TC1823. Therefore, in the next subsection, the precipitation forecasts of these three representative LTCs are compared in the context of the relative advantages and disadvantages of applying the various ensemble methods.

4.4 Analysis of three representative LTC cases

a. TC1808

As indicated in the preceding subsection, TC1808 is a northbound TC, which is best forecasted by the 90th percentile (Fig.3) and PM ensemble method (Fig.6) before and after considering the track type, respectively. The *TSsum* increases from 0.53 to 1.471 and the *BSsum* changes from +2.125 to −0.529 after considering its track type. Figure 8 compare the predicted precipitation of TC1808 by these two schemes to the observed. TC1808 has a station with accumulated precipitation exceeding 250 mm, and the track type experiment reproduces it successfully. For the precipitation of more than 100 mm, there are five large-valued centers. If the selected best scheme does not consider the track type of this TC, only one large-valued center in northern Taiwan Island can be simulated, but with the precipitation in southern Taiwan Island overpredicted. After considering the track type, three of the five large-valued centers can be simulated. The simulated precipitation for Taiwan Island and Zhejiang Province is much improved, with little evidence of overprediction. Thus, better forecasts can be obtained by using different schemes for TCs of different track types. However, both experiments produce poor precipitation simulations for inland areas. After classifying the TC tracks, underprediction still exists inland.

b. TC1816

TC1816 is a westbound typhoon, and best forecasted when the 90th percentile ensemble method is used (Fig. 6). The simulated precipitation of has a *TSsum* of 0.701 and a *BSsum* of −0.36. In contrast, the *TSsum* values of this case for the five NWP models (i.e., GFS, GRAPES, ECMWF, SMS-WARMS, and RMAPS) are 0.698, 0.424, 0.487, 0.520, and 0.260, respectively, while their *BSsum* values are +0.643, −1.413, +0.712, −1.218, and −0.524. In addition, Fig. 9 shows that for precipitation above 250 mm, the GFS forecast is better than that of the other models, but there is severe overprediction in precipitation ranging from 100 to 250 mm. The DSAEF_LTP model can simulate the precipitation over 250 mm. However, there is a bias in simulating a large-valued zone in Hainan Province. That is, the simulated heavy precipitation occurs generally over northeastern Hainan whereas the large-valued region predicted by the DSAEF LTP model appears in southwestern Hainan. Also, the DSAEF_LTP model does not perform well in terms of the large-valued precipitation region in Guangdong Province. For the predicted precipitation of greater than 100 mm, the DSAEF_LTP model exhibits clearly some advantages compared to the NWP models. The predicted precipitation distributions, especially over coastal areas, are very similar to those observed, allowing areas of high precipitation values to be forecasted without overprediction. In short, the forecast result of the DSAEF_ LTP model has the highest hit rate with minimum range deviations.

c. TC1823

Figure 10 compares the forecast precipitation of TC1823 with the best scheme of each ensemble method from the track-type experiments to the observed precipitation.

It is evident that the forecast performance for TC1823 is the best with the 90th percentile ensemble method, followed by the maximum and fuse ensemble methods, whereas precipitation of more than 100 mm cannot be simulated by the other ensemble methods (cf. Figs. 10 and 7). This TC produced more than 100 mm accumulated precipitation at 12 stations, with four of them recording more than 250 mm. The DSAEF_LTP model using the 90th percentile ensemble method predicts more than 100 mm precipitation at 7 stations, and more than 250 mm precipitation at 3 stations. However, this method setup underestimates the precipitation of above 250 mm and overestimates the precipitation over 100 mm (Fig. 10d). As compared to the original ensemble methods in the DSAEF_LTP model, the 90th percentile outperforms the mean (Fig. 10b) and maximum (Fig. 10c) ensemble methods. The precipitation distribution predicted by the 90th percentile is similar to that predicted by the maximum ensemble method, but the false alarm rate of the former drops significantly. The latter point can be seen from the *BIAS250* and *BIAS100* of the 90th percentile and maximum ensemble methods: they are 0.750 and 1.750, and 1.250 and 4.583, respectively. Besides, Fig. 10e looks similar to Fig. 10c, and Figs. 10f–10h look similar to Fig.10b, because only two analogs are selected as the ensemble members.

5. Summary and data

In this study, five new ensemble methods are added to the original DSAEF_LTP model proposed by Ren et al. (2020), and then 7 experiments with different ensemble methods are carried out for 10 LTCs over China June-September of 2018. The best scheme for each experiment is selected and compared with five NWP models (i.e., ECMWF, GRAPES, GFS, SMS-WARMS, and RMAPS). To achieve better forecast performance, the track-type experiments are also carried out. Major results can be summarized as follows:

- The 90th percentile ensemble method performs best in LTC precipitation forecasts of the new ensemble methods tested. With this method, the *TS250* and *TS100* values for the best scheme of the DSAEF_LTP model are 0.184 and 0.209, respectively. The *TSsum* of the 90th percentile ensemble method (i.e., *TS250* = 0.158, and *TS100* = 0.215) higher than that of the maximum ensemble method, ranking the former as the first before the new ensemble methods are added. The *TSsum* of the mean ensemble method, which is the intrinsic ensemble method in the DSAEF_LTP model, ranks the fifth. In general, the *TSsum* of the best scheme with the station-based ensemble method is higher than that of the field-based ensemble method. The difference in the *TS* of these two kinds of ensemble methods is mainly reflected in forecasting the precipitation of over 100 mm.
- As compared with the *TSsum* to the five NWP models, the *TSsum* of the best schemes of the DSAEF_LTP model with the new ensemble methods are higher. The main advantage of the DSAEF_LTP model lies in predicting the precipitation of over 250 mm.
- To address the relatively poor precipitation forecast of northbound TCs by the best schemes of the DSAEF_LTP model, ten TCs are divided into two groups according to their tracks and then track-type experiments are conducted. Results show that the *TSsum*

of the best schemes with the seven ensemble methods exhibits significant improvements for the northbound TCs. When the 90th percentile method is adopted for the westbound TCs, and PM for the northbound TCs, the *TSsum* of the best schemes are the highest. This may be due to the fact that the accumulated precipitation centers over southern China are frequently concentrated at some stations, whereas those over northern China are scattered and the total precipitation at many stations varies greatly in magnitude.

• The above results are further demonstrated from an analysis of three representative TC cases (i.e., TC1808, TC1816, and TC1823), confirming that the forecast performance of the DSAEF_LTP model can be improved by adopting a new ensemble method. The hit rate can be further increased, and with reduced false alarm rates after considering different track types.

Since the early publication of the DSAEF_LTP model, we have made some improvements. Previous studies (i.e., Ding et al., 2020; Jia et al., 2020) focused mainly on how to select more reasonably similar TCs, and the problem of high false alarm rates has been less researched. The current study focuses on the improvement of the ensemble methods in the DSAEF_LTP model. Based on the results shown herein, we may conclude that applying different ensemble methods under different situations will help improve the forecast performance of the DSAEF_LTP model, which might then be applied to the other ensemble forecast studies. However, only 10 TCs are chosen as the objects of the experiments in this study. Thus, the applicability of the best schemes needs further tests. In the future, large-sample experiments with the DSAEF_LTP model should be carried out to determine the most suitable scheme for LTC precipitation over China or other regions through training and independent forecast experiments, before being used for operational TC precipitation forecasting. Moreover, more variables that influence TC precipitation, especially background environment variables, such as vertical wind shear, relative humidity should be considered in the DSAEF_LTP model. When the GIV in the DSAEF_LTP model contains enough variables influencing TC precipitation, the forecast performance can be further improved. The analogs selected by the GIV similarity can even include global environment changes because different global environments mean different GIVs.

Data Availability Statement

The historical observed precipitation data used during this study are available from http://data.cma.cn/data/cdcdetail/dataCode/A.0012.0001.html. The precipitation forecast data from ECMWF, GFS and T639 model are available from https://www.ecmwf.int/en/forecasts/datasets; https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/global-forcast-system-gfs; http://data.cma.cn/data/cdcdetail/dataCode/F.0003.0001.html. The historical best-track data are from https://tcdata.typhoon.org.cn/zjljsjj_zlhq.html. The operational forecast tracks of TCs are obtained from the CMA.

Acknowledgements. The authors would like to express their sincere thanks to two anonymous reviewers; the Editor of JMSJ, Dr. Shunji Kotsuki; and Prof. Da-Lin Zhang of the University of Maryland for their helpful suggestions and comments. This work was supported by the National Key R&D Program of China (Grant No. 2019YFC1510205), the Hainan Provincial Key R&D Program of China (SQ2019KJHZ0028), the National Natural Science Foundation of China (Grant No. 41675042), and the Jiangsu Collaborative Innovation Center for Climate Change.

List of abbreviations

BIAS	Bias Score
CMA	China Meteorological Administration
ECMWF	European Centre for Medium-range Weather Forecasts
ED-WM	Equal Difference-Weighted Mean
GFS	Global Forecast System
GIV	Generalized Initial Value
GRAPES	Global/Regional Assimilation and Prediction System
LTC	Landfalling Tropical Cyclone
NWP	Numerical Weather Prediction
PM	Probability-matching Mean
QPF	Quantitative Precipitation Forecasting
RMAPS	Rapid-refresh Multi-scale Analysis and Prediction System
SMS-WARMS	Shanghai Meteorological Service WRF ADAS Real-Time Modeling System
TC	Tropical Cyclone
the DSAEF_LTP model	the Dynamical–Statistical–Analog Ensemble Forecast model for Landfalling Typhoon Precipitation
TS	Threat Score
TSAI	Tropical cyclone track Similarity Area Index
TSAI-WM	TSAI-weighted mean

References

Alexander, S. and A. Smith, 2019: Quantifying the predictability of a predictand: Demonstrating the diverse roles of serial dependence in the estimation of forecast skill. Quarterly Journal of the Royal Meteorological Society, **145**:718, 40-52.

Balaguru, K., G. R. Foltz, L. R. Leung, and K. A. Emanuel, 2016: Global warming-induced upper-ocean freshening and the intensification of super typhoons. Nat. Commun., **7**, 13670, https://doi.org/10.1038/ncomms13670.

Barthel, F., and E. Neumayer. 2012. A Trend Analysis of Normalized Insured Damage from Natural Disasters. Climatic Change, **113**: 215–237. doi:10.1007/s10584-011-0331-2.

Binh T., L. Le, T. Le, K. Bui, V. Le, H. Ly, and I. Prakash, 2020: Development of advanced artificial intelligence models for daily rainfall prediction. Atmospheric Research, **237**, 104845.

Chan, J. C. L., 2008: Decadal variations of intense typhoon occurrence in the western North Pacific. Proc. Roy. Soc. London, **464**, 249–272.

Chao, Y.Y., J.G. Alves, and H.L. Tolman, 2005: An Operational System for Predicting Hurricane-Generated Wind Waves in the North Atlantic Ocean. Wea. Forecasting, **20**, 652–671.

Chen, B., Y. Guo, K. Dai, Q. Qian, 2016: Research for the Ensemble Member Optimization Correction Technique on Typhoon Rainstorm Forecast and Its Application Experiment. Meteor Mon,**42(12)**:1465-1475. (in Chinese)

Chen B., Dai K., Guo Y. Precipitation verification and analysis of ECMWF ensemble statistic products in 2013 flooding season [J]. Torrential Rain and Disasters, 2015,34(1):64-73

Chen, L. and Z. Meng, 2001: Researches on tropical cyclones in China in the past decade. Chinese Journal of Atmospheric Sciences, **03**:420-432. (in Chinese)

Cheung, K., J. Chan, 1999: Ensemble forecasting of tropical cyclone motion using a barotropic model, Part I. Mon. Wea. Rev., **127(6)**:1229-1243.

Cheung, K., J. Chan, 1999: Ensemble forecasting of tropical cyclone motion using a barotropic model, Part II. Mon. Wea. Rev., **127(11)**:2617-2640.

Clark, A., Weiss S., Kain J., et al., 2012: An Overview of the 2010 Hazardous Weather Testbed Experimental Forecast Program Spring Experiment. Bull Amer Meteor Soc, **93(1)**:55-74.

Clark, A., 2017: Generation of Ensemble Mean Precipitation Forecasts from Convection-Allowing Ensembles. Wea. Forecasting, **32**, 1569–1583. https://doi.org/10.1175/WAF-D-16-0199.1

cma.gov (homepage on the internet). (2011) China Meteorological Administration. Available at: http://www.cma.gov.cn/en2014/20150311/20170728/2017072803/201708/t20170822446943.html; accessed 01 September 2019.

Dai, K., Y. Cao, Q. Qian, S. Gao, S. Zhao, Y. Chen, C. Qian, 2016: Situation and Tendency of Operational Technologies in Short and Medium Range Weather Forecast. Meteor Mon, **42(12)**:1445-1455. (in Chinese)

Ding, C., F. Ren, Y. Liu, J. L. McBride, and T. Feng, 2020: Improvement in the Forecasting of Heavy Rainfall over South China in the DSAEF_LTP Model by Introducing the Intensity of the Tropical Cyclone. Wea. Forecasting, **35**, 1967–1980.

Ebert, E., 2001: Ability of a Poor Man's Ensemble to Predict the Probability and Distribution of Precipitation. Mon. Wea. Rev., **129**, 2461–2480.

Emanuel, K. A., 2005b: Increasing destructiveness of tropical cyclones over the past 30 years. Nature, **436**, 686–688.

Fang, X., Kuo Y., 2013. Improving ensemble-based quantitative precipitation forecasts for topography-enhanced typhoon heavy rainfall over Taiwan with a modified probability-matching technique. Mon Wea Rev, **141(11)**:3908-3932.

Gu, C., J. Kang, G. Yan, Z. Chen, 2016: Spatial and Temporal Variability of Northwest Pacific Tropical Cyclone Activity in a Global Warming Scenario. Journal of Tropical Meteorology, **22**.

Knutson, T., and Coauthors, 2020: Tropical Cyclones and Climate Change Assessment: Part II: Projected Response to Anthropogenic Warming. Bull. Amer. Meteor. Soc., **101**, E303–E322, https://doi.org/10.1175/BAMS-D-18-0194.1.

Jia, L., Z. Jia, F. Ren, C. Ding, M. Wang and T. Feng, 2020: Introducing TC intensity into the DSAEF_LTP model and simulating precipitation of super-typhoon Lekima (2019). Q J R Meteorol Soc, **146**: 3965– 3979. https://doi.org/10.1002/qj.3882

Lin, J., Z. Zong, X. Jiang, 2013: The verification report of multi-model integrated QPF products from 2010-2011. Wea Forecast Rev, **5(1)**: 67-74. (in Chinese)

Li Q., Lan H., Chan J., et al., 2015: An operational statistical scheme for tropical cyclone includes rainfall forecast. Journal of Tropical Meteorology, **21(2)**:101-110.

Li W., Che Q., Jiang J., et al. Applications of ensemble statistic-integrating method to short-range torrential rain forecast in Hubei province[J]. Torrential Rain and Disasters,2018,37(5):455-461.

Luitel B, Villarini G, Vecchi GA, 2018: Verification of the skill of numerical weather prediction models in forecasting rainfall from U.S. landfalling tropical cyclones. J Hydrol, **556**:1026–1037.

I Sofiati and A Nurlatifah, 2019: The prediction of rainfall events using WRF (weather research and forecasting) model with ensemble technique. IOP Conference Series: Earth and Environmental Science, **374**, 012036.

Ma, L., 2014: Research progress on China typhoon numerical prediction models and associated major techniques. Progress in Geophysics, **29(3)**: 1013-1022.

Ma L M and Tan Z M. 2009: Improving the behavior of the cumulus parameterization for tropical cyclone prediction: Convection trigger. Atmospheric Research, **92(2)**: 190-211.

Marchok, T., R. Rogers, and R. Tuleya, 2007: Validation Schemes for Tropical Cyclone Quantitative Precipitation Forecasts: Evaluation of Operational Models for U.S. Landfalling. Cases. Wea. Forecasting, **22**, 726–746, https://doi.org/10.1175/WAF1024.1

Novak, D. R., C. Bailey, K.F. Brill, P. Burke, W.A. Hogsett, R. Rausch, and M. Schichtel, 2014: Precipitation and Temperature Forecast Performance at the Weather Prediction Center. Wea. Forecasting, **29**, 489–504, https://doi.org/10.1175/WAF-D-13-00066.1

Ren, F., G. Wu, X. Wang et al., 2011: Tropical cyclones affecting China in the last 60 years [M]. Beijing: China Meteorological Press, 43-74.

Ren F., Qiu W., Jiang X., Wu L., Xu Y., Duan Y., 2018: An objective track similarity index and its preliminary application to predicting precipitation of landfalling tropical cyclones. Weather Forecasting, **33(6)**: 1725–1742

Ren, F., C. Ding, D. Zhang, D. Chen, H. Ren, and W. Qiu, 2020: A Dynamical-Statistical-Analog Ensemble Forecast Model: Theory and an Application to Heavy Rainfall Forecasts of Landfalling Tropical Cyclones. Mon. Wea. Rev., **148(4)**, 1503-1517.

Rogers R., 2018: Summary of Fourth International Workshop on Tropical Cyclone Landfall Processes (IWTCLP-4). WMO/IWTC-9, Honolulu, Hawaii, December 3-7.

Surcel, Madalina,Zawadzki, Isztar,Yau, M K. On the Filtering Properties of Ensemble Averaging for Storm-Scale Precipitation Forecasts. Monthly Weather Review,2014,**142(3)**.

Tao J., Zhao H., Yi X., Liu Y., Jia H.,2019: Causal Analysis of a Short-Time Strong Rainfall Based on RMAPS and Observation Data. Meteorological Science and Technology,2019,47(02):299-311.

Wang Y., X. Shen, D. Chen, 2012: Verification of tropical cyclone rainfall predictions from CMA and JMA global models. Journal of Tropical Meteorology, **18(4)**: 537-542.

Wei C., 2012a: RBF neural networks combined with principal component analysis applied to quantitative precipitation forecast for a reservoir watershed during typhoon periods. Journal of Hydrometeorology, **13(2)**: 722-734

Wei C., 2012b: Wavelet support vector machines for forecasting precipitation in tropical cyclones: Comparisons with GSVM, regression, and MM5.Weather & Forecasting, **27(2)**: 438-450

Weinkle, J., R. Maue, and R. Pielke, 2012: Historical Global Tropical Cyclone Landfalls. J. Climate, **25**, 4729–4735, https://doi.org/10.1175/JCLI-D-11-00719.1

Woo W.C., W. Hogsett, M. Mohapatra, K. Nagata, P. Otto, et al., 2014: Challenges and Advances related to TC Rainfall Forecast. The Third International Workshop on Tropical Cyclone Landfall Processes (IWTCLP-III), Jeju, **8-10 Nov 2014**.

Xiao Q., Y.H. Kou, J. Sun, et al., 2007: An approach of Radar reflectivity data assimilation and its assessment with the inland QPF of typhoon Rusa(2002) at landfall[J]. Journal of Appl. Meteor. Climatol.**45**:14-22.

Xu T., Li J., Yang Y., Wang X., Chen B.,2016: Verification of SMS-WARMS V2.0 Model Forecast Results. Meteorological Monthly,2016,**42(10)**:1176-1183.

Ying, M., W. Zhang, H. Yu, X. Lu, J. Feng, Y. Fan, Y. Zhu, and D. Chen, 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

Yu, X., Park S., Lee Y., et al., 2013. Quantitative precipitation forecast of a tropical cyclone through optimal parameter estimation in a convective parameterization. Scientific Online Letters on the Atmosphere Sola, **9(4)**: 36-39.

Zhang, H., Pu Z., 2014. Influence of Assimilating Surface Observation on Numerical Prediction of Landfalls of Hurricane Katrina (2005) with an Ensemble Kalman Filter. Monthly Weather Review, **142(8)**: 2915-2934.

Zhang, Q., C. Zhang, Z. Zhang et al., 2007: Study on the uncertainty of ensemble forecasting of tropical cyclone. Chinese Journal of Geophysics, 2007, **50(3)**:701-706. (in Chinese)

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

Zhao, K. X. Li M. Xue, B.J.Jou and W.C. Lee. 2012: Short-term forecasting through intermittent assimilation of data from Taiwan and mainland China coastal radars for Typhoon Meranti (2010) at landfall. Journal of Geophysical Research, 117, D06108.

Zhong, Y., H. Yu, W. Teng, P. Chen, 2009: A Dynamic Simulated Scheme for Tropical Cyclone Quantitative Precipitation forecast. Journal of Applied Meteorological Science, 2009,**20(01)**:17-27. (in Chinese)

Zhu L., Q. Wan, X. Shen, Z. Meng, et al., 2016: Prediction and Predictability of High-Impact Western Pacific Landfalling Tropical Cyclone Vicente (2012) through Convection-Permitting Ensemble Assimilation of Doppler Radar Velocity. Mon. Wea. Rev., **144**: 21-43.

List of Figures

Fig.1 Track distribution of 10 landfalling TCs over China in 2018.

Fig.2 Flowchart of the DSAEF_LTP model.

Fig.3 Threat scores (*TSsum*, *TS250* and *TS100*) for accumulated LTC precipitation forecasts of the best schemes of the DSAEF_LTP model in the seven experiments and five NWP models (i.e., ECMWF, GRAPES, GFS, SMS-WARMS and RMAPS).

Fig.4 Threat scores (vertical color bars) of the best schemes of the DSAEF_LTP model in the seven experiments and the five NWP models (i.e., ECMWF, GRAPES, GFS, SMS-WARMS and RMAPS) for the rainfall forecast of 10 LTCs. Dashed lines represent the observed maximum accumulated precipitation (mm) associated with LTCs.

Fig.5 The maximum accumulated precipitation distribution of TCs during 1960–2018. The colored solid points indicate the stations with single-station observed maximum total rainfall (mm) of each TC during 1960–2018. The colors of these solid points show the precipitation amount. Other markers indicate the frequency of a station with the maximum total rainfall.

Fig.6 *TSsum* of the best schemes with seven ensemble methods in the track-type experiments.

Fig.7 As in Fig.4, but for the track-type experiments.

Fig.8 Distribution of accumulated precipitation (mm) for TC1808: (a) observation; (b) prediction of the 90th percentile ensemble method of the DSAEF_LTP model with the

highest $TSsum$; (c) prediction of the PM ensemble method of the DSAEF_LTP model with the highest $TSsum$ for the westbound classification.

Fig.9 Accumulated precipitation (mm) for TC1816: (a) observed; (b) the scheme of the DSAEF_LTP model with the 90th percentile ensemble method in the track-type westbound TC experiment; (c) GFS; (d) GRAPES; (e) ECMWF; (f) SMS-WARMS; (g) RMAPS.

Fig.10 Accumulated precipitation (mm) for TC1823: (a) observed; (b–g) the best scheme of the DSAEF_LTP model with mean, maximum, 90th percentile, fuse, PM, ED-WM and TSAI-WM ensemble methods, respectively, in the track-type westbound TC experiment.

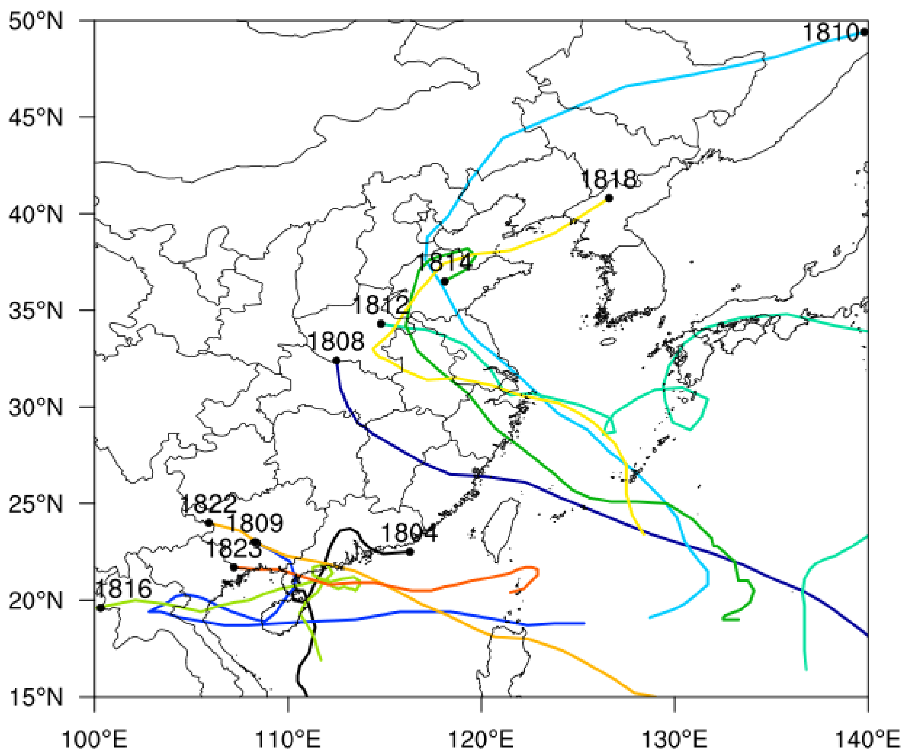

Fig.1 Track distribution of the 10 landfalling TCs over China in 2018.

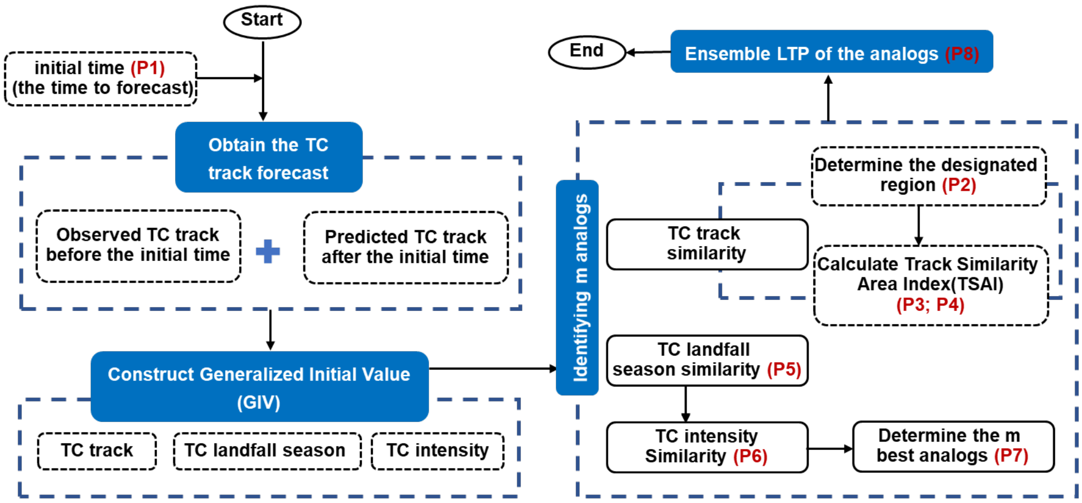

Fig.2 Flowchart of the DSAEF_LTP model.

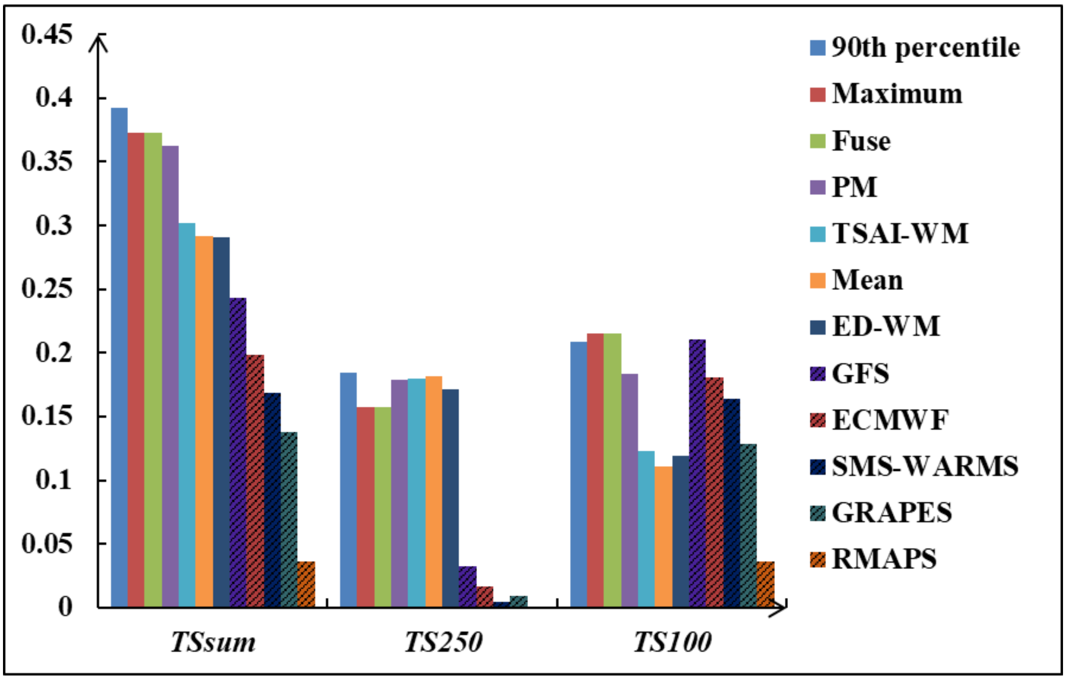

Fig.3 Threat scores (*TSsum*, *TS250* and *TS100*) for accumulated LTC precipitation forecasts of the best schemes of the DSAEF_LTP model in the seven experiments and five NWP models (i.e., ECMWF, GRAPES, GFS, SMA-WARMS and RMAPS).

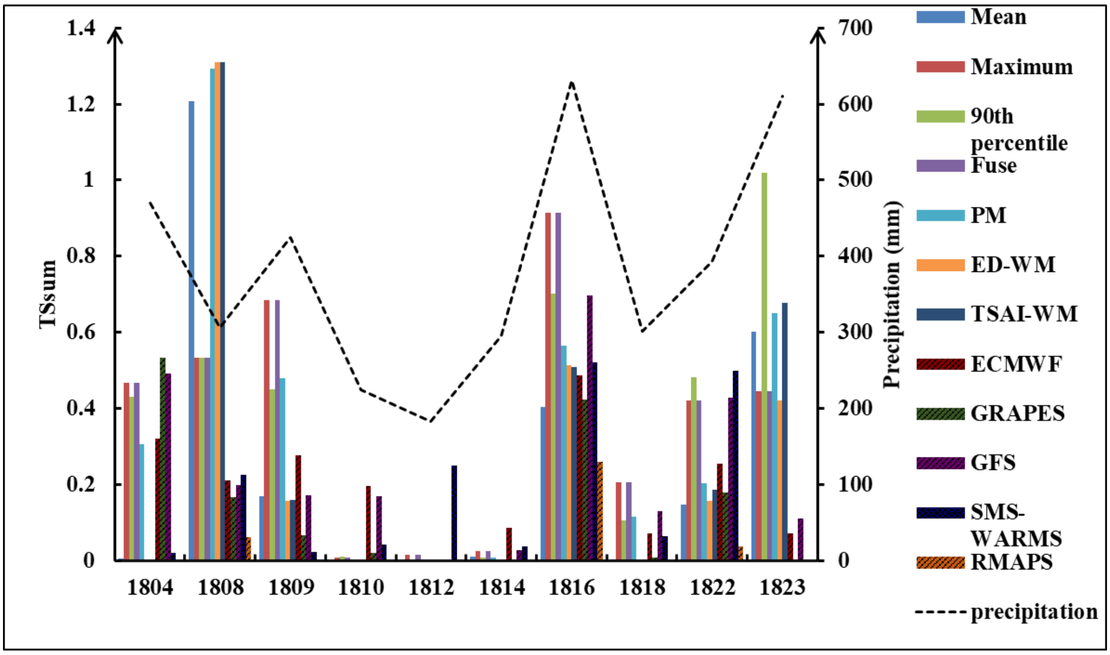

Fig.4 Threat scores (vertical color bars) of the best schemes of the DSAEF_LTP model in the seven experiments and the five NWP models (i.e., ECMWF, GRAPES, GFS, SMS-WARMS and RMAPS) for the rainfall forecast of 10 LTCs. Dashed lines represent the observed maximum accumulated precipitation (mm) associated with LTCs.

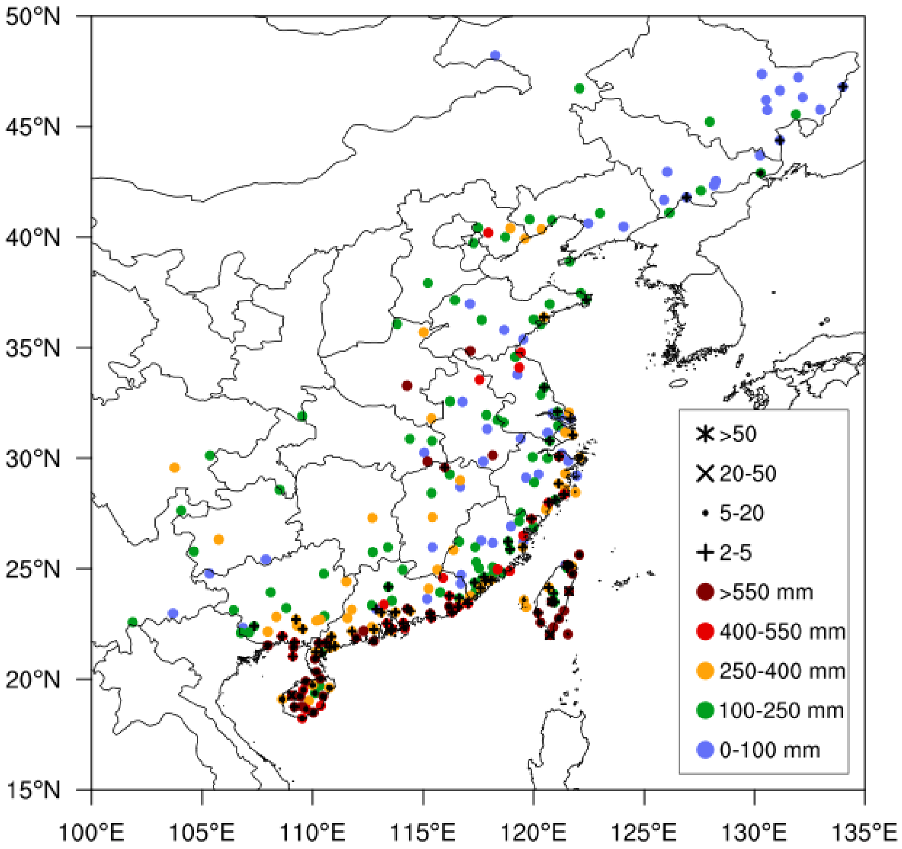

Fig.5 The maximum accumulated precipitation distribution of TCs during 1960–2018. The colored solid points indicate the stations with single-station observed maximum total rainfall (mm) of each TC during 1960–2018. The colors of these solid points show the precipitation amount. Other markers indicate the frequency of a station with the maximum total rainfall.

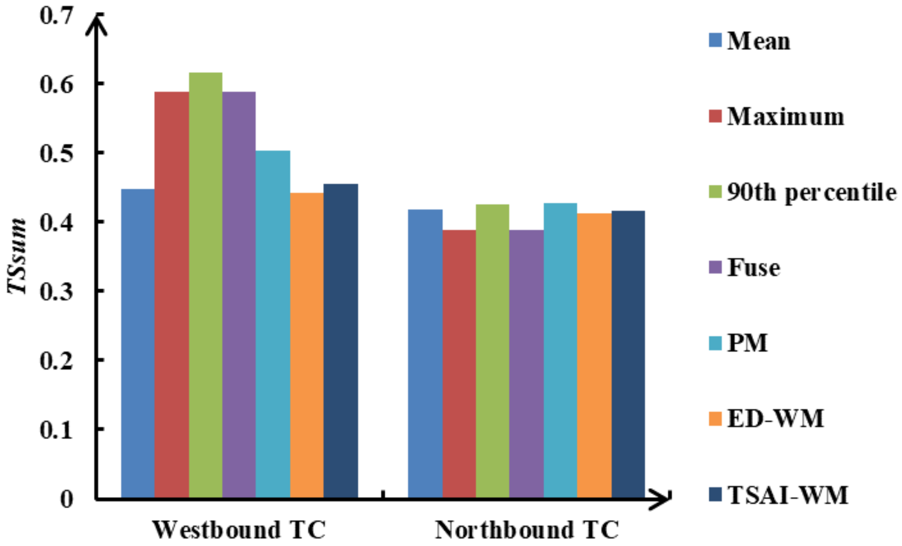

Fig.6 *TSsum* of the best schemes with seven ensemble methods in the track-type experiments.

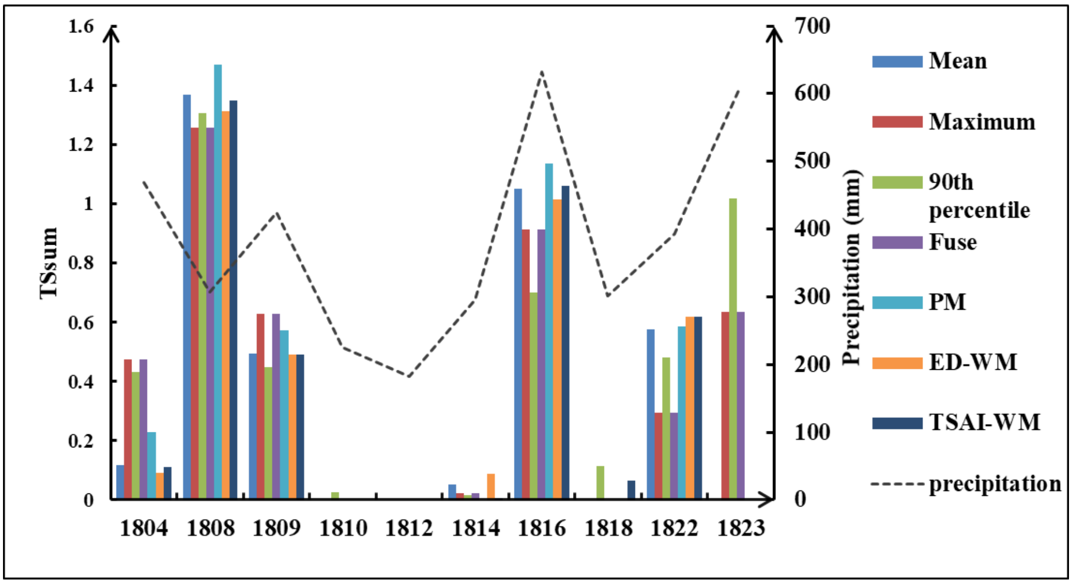

Fig.7 As in Fig.4, but for the track-type experiments.

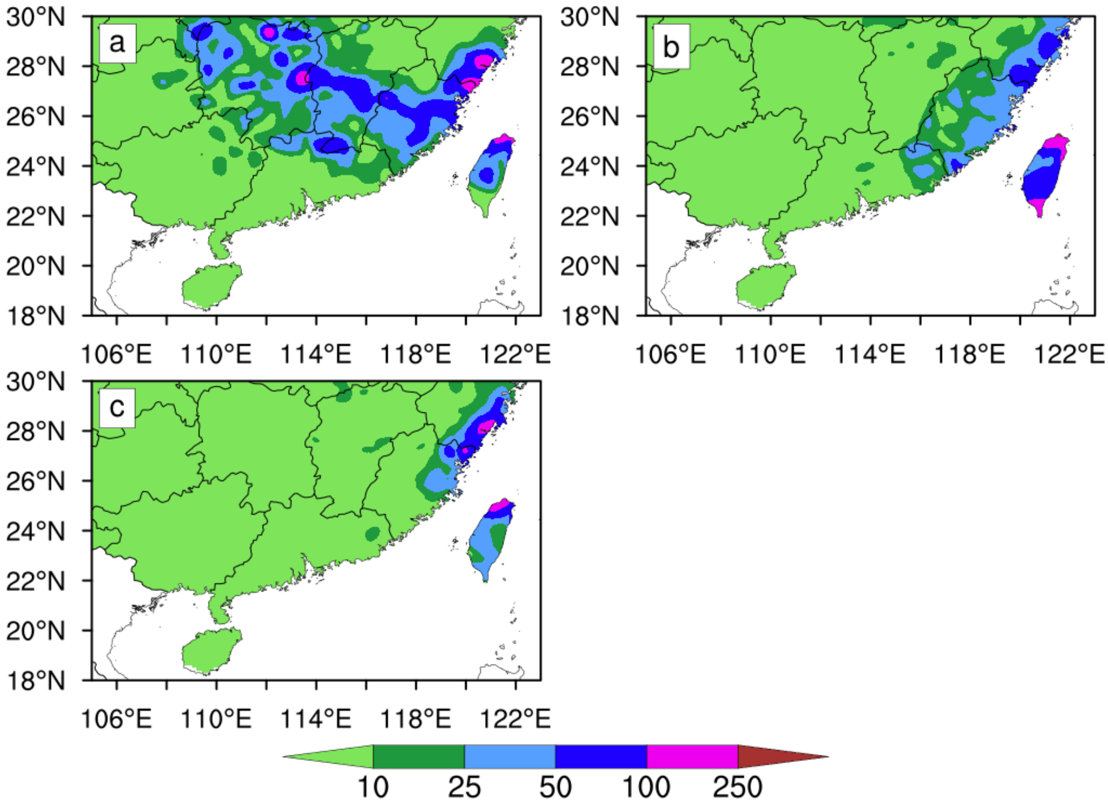

Fig.8 Distribution of accumulated precipitation (mm) for TC1808: (a) observation; (b) prediction of the 90th percentile ensemble method of the DSAEF_LTP model with the highest $TSsum$; (c) prediction of the PM ensemble method of the DSAEF_LTP model with the highest $TSsum$ for the westbound classification.

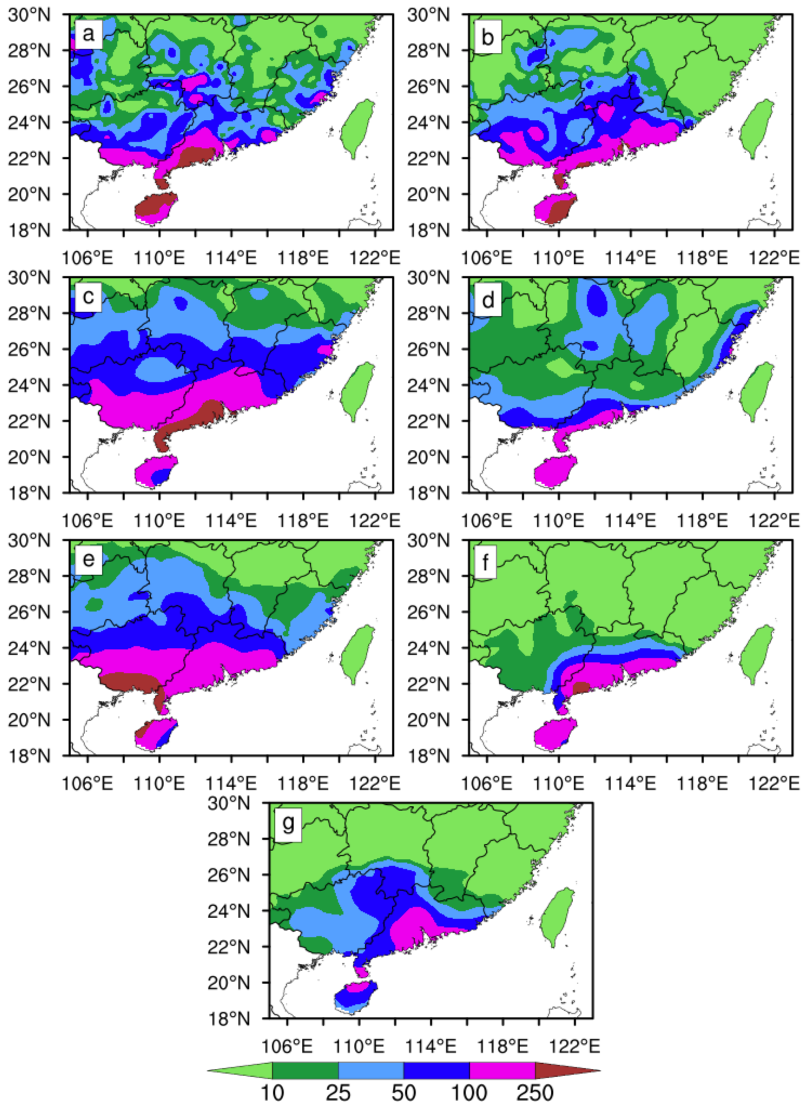

Fig.9 Accumulated precipitation (mm) for TC1816: (a) observed; (b) the scheme of the DSAEF_LTP model with the 90th percentile ensemble method in the track-type westbound TC experiment; (c) GFS; (d) GRAPES; (e) ECMWF; (f) SMS-WARMS; (g) RMAPS.

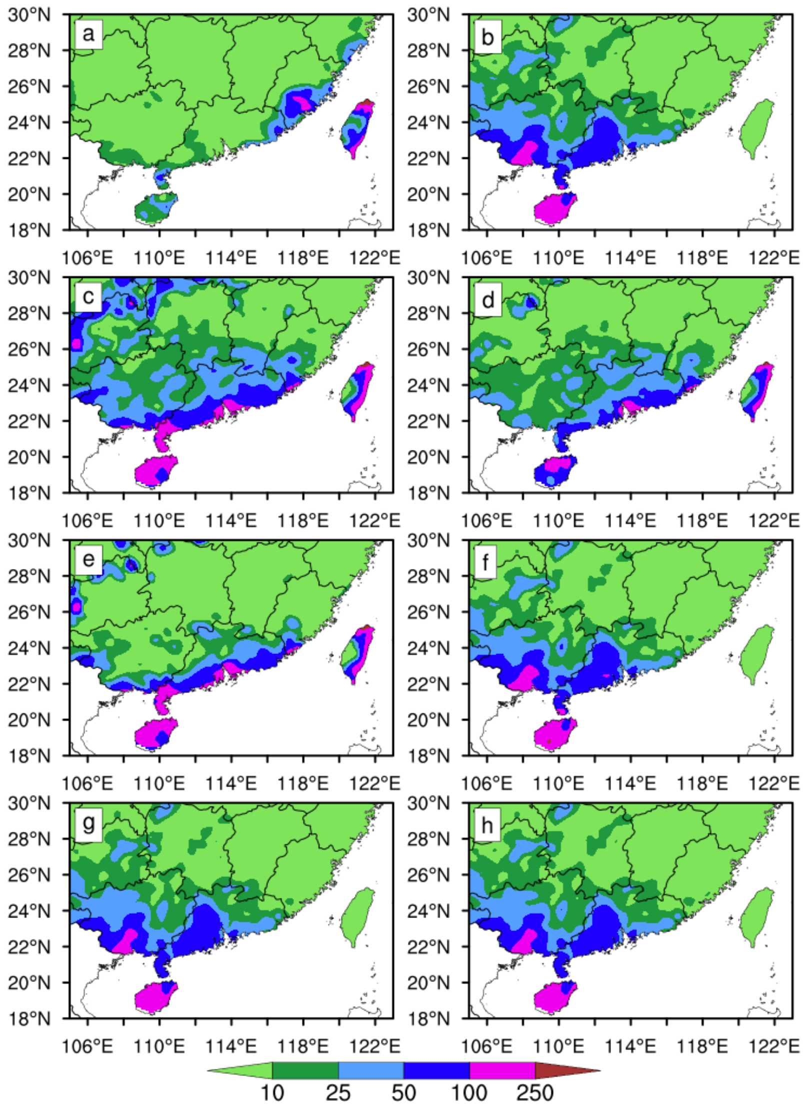

Fig.10 Accumulated precipitation (mm) for TC1823: (a) observed; (b–g) the best scheme of the DSAEF_LTP model with mean, maximum, 90th percentile, fuse, PM, ED-WM and TSAI-WM ensemble methods, respectively, in the track-type westbound TC experiment.

List of Tables

Table 1 Parameters of the DSAEF_LTP model.

Table 2 The improved ensemble methods in the DSAEF_LTP model.

Table 3 Parameter values for the best schemes with seven ensemble methods.

Table 1 Parameters of the DSAEF_LTP model.

Parameter	Description	Experimented values
P1 Initial time	The complete track of the target TC consists of the observed track before the initial time and the forecast track after the initial time.	1: 1200 UTC on Day1 2: 0000 UTC on Day1 3: 1200 UTC on Day0 4: 0000 UTC on Day0 5: 1200 UTC on Day-1 6: 0000 UTC on Day-1 (Day0: the day of TC precipitation occurring on land; Day1: the day after Day0; Day-1: the day before Day0)
P2 Similarity region	A designated region within which the TSAI is calculated. It is a rectangle with diagonal points A and B.	Decided by the predicted TC track, initial time and diameter of the TC. There are 20 experiment values (1-20).
P3 Threshold of the segmentation ratio of a latitudinal extreme point	A parameter of TSAI that represents the bending degree of TC tracks.	1: 0.1 2: 0.2 3: 0.3
P4 Overlapping percentage threshold of two TC tracks	A parameter of TSAI that represents the degree of longitude (latitude) overlaps of TC tracks.	1: 0.9 2: 0.8 3: 0.7 4: 0.6 5: 0.5 6: 0.4
P5 Seasonal similarity	A parameter that indicates the TC landfall time.	1: the whole year 2: May–Nov 3: Jul–Sept 4: the same landfall month as the target TC 5: within 15 days of the target TC landfall time
P6 Intensity similarity	A parameter that indicates the differences between the TC intensity of the target TC and historical TCs. There are four categories of TC intensity that can be chosen. The similarity of TC intensity is divided	Four categories: 1: average intensity on the first rainy day 2: maximum intensity on the first rainy day 3: average intensity on all rainy days 4: maximum intensity on all rainy days Five levels: 1: all grades 2: the target TC intensity is the same grade or above the historical TC 3: the same grade or below 4: only the same grade

	into five levels.	5: the same grade or one grade difference
P7 Number of analog TCs screened for the ensemble forecast	M historical TCs with the first m most similar GIVs to that of the target TC	1–10 for 1, 2… and 10, respectively
P8 Ensemble	Ensemble forecast scheme	1–7 for 7 ensemble methods listed in Table 2
Total number of schemes: 6 × 20 × 3 × 6 × 5 × 4 × 5 × 10 × 7 = 15,120,000		

Table 2 The improved ensemble methods in the DSAEF_LTP model.

Type	Name	Computational procedure
Station-based ensemble methods	Mean	1. The precipitation forecast of each station is calculated separately. 2. For a station, the average precipitation of m selected analogs at this station $[Pre(i), i = 1, 2, \dots, m]$ as the final forecast result of the station, $Prep = \frac{\sum_{i=1}^{m} Pre(i)}{m}$. 3. The forecast results of each station form the forecast precipitation field.
	Maximum	Same as the mean ensemble method, but $Prep = Max\ (pre(i))$
	Optimal percentile (90th percentile in this study)	1. For each station, $pre(i), i = 1, 2, \dots, m$ is sorted from small to large. $pre(r)$ is the precipitation ranked r. 2. $d = 1 + (m - 1) \times 0.9$ 3. The integer part of d is r, and the decimal part is f 4. $Prep = pre(r) + [pre(r + 1) - pre(r)] \times f$
	Fuse	Calculation rules of the forecast precipitation at each station： 1. If $Max\ (pre(i)) \geq$ 100 mm, $Prep = Max\ (pre(i))$; 2. If the 90th percentile value of $Pre(i) \geq 50$ mm，the $Prep$ equals the 90th percentile value of $Pre(i)$; 3. If the 75th percentile value of $Pre(i) \geq 50$ mm，the $Prep$ equals the 75th percentile value of $Pre(i)$; 4. If the median value of $Pre(i) \geq 10$ mm, the $Prep$ equals the median value of $Pre(i)$; 5. If none of the above conditions can be met, the $Prep$ equals the 10th percentile value.
Field-based ensemble methods	Probability matching mean (PM)	1. Arrange all the precipitation data for the m members of 2027 stations in ascending order (containing $2027 \times m$ stations' rainfall data). Divide the $2027 \times m$ data into 2027 equal parts in reverse order, retaining the median of each part and recording them as $prem(k), k =$

<table>
<tr><td></td><td></td><td>$1,2,\ldots,2027$.
2. For a station, the average precipitation of m selected analogs at this station is $Prea = \frac{\sum_{i=1}^{m} Pre(i)}{m}$; the $Prea$ of 2027 stations is ranked in reverse order; the ranking of each station's $prea$ is recorded as k.
3. Corresponding to the $prem(k)$ of each station based on the k of each station, and $prem(k)$ is the predicted precipitation for this station, $Prep = prem(k)$.</td></tr>
<tr><td></td><td>Equal difference-weighted mean (ED-WM)</td><td>The weight of the precipitation for the selected similar TC whose similarity rank i is $W(i) = \frac{(2\times m-i)\times 2}{(3\times m-1)\times m}(i = 1,2,\ldots,m)$, the forecasted precipitation is $Prep = \sum_{i=1}^{m} W(i) \times Pre(i)$</td></tr>
<tr><td></td><td>TSAI-weighted mean (TSAI-WM)</td><td>$A(i) = \frac{1}{TSAI(i)}(i = 1,2,..,m)$; the weight of the precipitation for the selected similar TC whose similarity rank i is
$W(i) = \frac{A(i)}{\sum_{i=1}^{m} A(i)}$, and the forecast precipitation is $Prep = \sum_{i=1}^{m} \mathrm{W}(i) \times Pre(i)$</td></tr>
</table>

Table 3 Parameter values for the best schemes with the seven ensemble methods.

	Mean	Maximum	90th percentile	Fuse	PM	ED-WM	TSAI-WM
P1	2	1	1	1	1	1	1
P2	20	20	20	20	20	20	20
P3	3	1	1	1	2	1	1
P4	5	6	6	6	6	6	6
P5	2	2	3	2	2	2	2
P6	3/3	2/5	2/5	2/5	2/5	1/2	1/2
P7	2	5	3	5	4	5	5