Experiments with the Improved Dynamical–Statistical–Analog Ensemble Forecast Model for Landfalling Typhoon Precipitation over South China

MA Yun-qi (马蕴琦), REN Fu-min (任福民), JIA Li (贾 莉), DING Chen-chen (丁晨晨)
(State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing 100081 China)

Abstract: In recent work, three physical factors of the Dynamical-Statistical-Analog Ensemble Forecast Model for Landfalling Typhoon Precipitation (DSAEF_LTP model) have been introduced, namely, tropical cyclone (TC) track, TC landfall season, and TC intensity. In the present study, we set out to test the forecasting performance of the improved model with new similarity regions and ensemble forecast schemes added. Four experiments associated with the prediction of accumulated precipitation were conducted based on 47 landfalling TCs that occurred over South China during 2004–2018. The first experiment was designed as the DSAEF_LTP model with TC track, TC landfall season, and intensity (DSAEF_LTP-1). The other three experiments were based on the first experiment, but with new ensemble forecast schemes added (DSAEF_LTP-2), new similarity regions added (DSAEF_LTP-3), and both added (DSAEF_LTP-4), respectively. Results showed that, after new similarity regions added into the model (DSAEF_LTP-3), the forecasting performance of the DSAEF_LTP model for heavy rainfall (accumulated precipitation >250 mm and >100 mm) improved, and the sum of the threat score (TS250 + TS100) increased by 4.44%. Although the forecasting performance of DSAEF_LTP-2 was the same as that of DSAEF_LTP-1, the forecasting performance was significantly improved and better than that of DSAEF_LTP-3 when the new ensemble schemes and similarity regions were added simultaneously (DSAEF_LTP-4), with the TS increasing by 25.36%. Moreover, the forecasting performance of the four experiments was compared with four operational numerical weather prediction models, and the comparison indicated that the DSAEF_LTP model showed advantages in predicting heavy rainfall. Finally, some issues associated with the experimental results and future improvements of the DSAEF_LTP model were discussed.

Key words: landfalling tropical cyclone; heavy rainfall forecast; DSAEF_LTP model; forecasting performance; South China

CLC number: P456.8  Document code: A

1 INTRODUCTION

Tropical cyclones (TCs) are cyclonic atmospheric vortices with warm cores and low-pressure structures that are generated over tropical or subtropical warm oceans, and are often accompanied by heavy rainfall, strong winds and storm surges. Among these accompanying phenomena, the heavy rainfall associated with TCs occurs frequently and often causes severe disasters (Chen et al. [1]; Chen and Xu [2]; Pant and Cha [3]), and a plethora of studies have shown that frequency of extreme TC precipitation has increased under the context of global warming (Kim et al. [4]; Emanuel [5]; Touma et al. [6]). In China, nearly three million people were evacuated due to the continuous heavy rain produced by severe tropical storm Bilis (0604), and the direct economic losses reached RMB 35 billion (approximately USD 5 billion) (Gao et al. [7]; Liu and Cui [8]). In the USA in 2017, Hurricane Harvey stalled over southeastern Texas for several days, bringing about a rainfall over 1500 mm in Nederland and Texas, and causing 106 deaths and USD 125 billion economic losses (Oldenborgh et al. [9]; Bosma et al. [10]; Xi et al. [11]). Therefore, improving the forecasting ability for TC precipitation is essential for the prevention and mitigation of typhoon-related disasters and the protection of human life and property.

At present, the forecasting of TC precipitation relies mainly on numerical weather prediction (NWP) by dynamical models, wherein there are two main ways that the forecasting skill can be improved. The first is to develop and improve NWP directly. For instance, some researchers have focused on the development of data assimilation to improve the initial field of models (Zhu et al. [12]; Kuvark and Bhishma [13]; Wang et al. [14]; Shen et al. [15]), while others have sought to improve the parameterization schemes for different physical processes (Yu et al. [16]; Xu et al. [17]; Castorina et al. [18];...
Zittis et al. [19]; Shi et al. [20]; Reddy et al. [21]) or advance the technology around the method of ensemble forecasting (Hsiao et al. [22]; Hong et al. [23]). The other way is to combine dynamical models with statistical methods (referred to as the dynamical-statistical approach), which can be generally divided into three categories (Ren and Xiang [24]). The first category involves using TC tracks predicted by using dynamical models and historical rainfall observational data, and the TC precipitation forecast is obtained from the perspective of the climatic average (Marks et al. [25]; Lee et al. [26]; Lonfat et al. [27]); the second way is to predict TC rainfall by adopting TC track forecasts and the rainfall integration from the initial rainfall rates (Kidder et al. [28]; Liu [29]; Ebert et al. [30]), and the third type works by constructing a dynamical-statistical scheme that consists of various internal TC variables and its environmental fields (Li and Zhao [31]; Zhong et al. [32]).

It is worthwhile to mention some recent studies on the hybrid analog-ensemble-based forecasts. Based on the mean intensity changes from the 10 best historical track analogs, Elsberry and Tsai [33] developed a situation-dependent intensity prediction (SDIP) technique for western North Pacific TCs, and successfully applied the SDIP technique to other TC cases (Tsai and Elsberry [34,35]).

Recently, Ren et al. [36] developed the Dynamical-Statistical-Analog Ensemble Forecast Model for Landfalling Typhoon Precipitation (hereafter referred to as the DSAEF_LTP model) and preliminarily applied it to rainfall forecast experiments for 21 landfalling TCs (LTCs) from 2012 to 2016 over South China. The results showed that the model has good forecasting performance at magnitudes $\geq 100$ mm and $\geq 250$ mm with two initial physical factors: TC track and TC landfalling season. Subsequently, Ding et al. [37] introduced TC intensity into the model and conducted rainfall forecast experiments for the same 21 LTCs over South China, which demonstrated that introducing TC intensity greatly improved the model’s forecasting ability. Furthermore, Jia et al. [38] added new similarity regions into the DSAEF_LTP model and conducted simulation experiments on super typhoon Lekima (2019). In addition, Jia et al. [39] further added five ensemble forecast schemes based on the original two ensemble forecast schemes (mean and maximum) of the DSAEF_LTP model, and corresponding simulation experiments were conducted on 10 LTCs in 2018.

Based on the above review, it is clear that only a small number of samples have thus far been adopted to conduct experiments with the new similarity regions and ensemble forecast schemes added into the DSAEF_LTP model-namely, an individual TC case (Lekima) for similarity regions, and TC samples from one year for ensemble forecast schemes. In other words, there is currently a deficiency of large-sample experiments to examine the improvement and applicability of the DSAEF_LTP model after introducing new parameters (similarity regions and ensemble forecast schemes). This is the motivation of this study. Accordingly, many TC samples seriously affecting South China spanning 15 years (from 2004 to 2018) were adopted to examine the forecasting performance of the DSAEF_LTP model with new parameters added. Section 2 describes the data and main methods used in this study, Section 3 illustrates the experimental design for introducing new similarity regions and ensemble forecast schemes into the DSAEF_LTP model and the forecast procedures of the DSAEF_LTP model. The results are analyzed in Section 4 and the summary and discussion are provided in Sections 5.

2 DATA AND METHODS

2.1 Data

Table 1 lists the names of the 47 TCs that were selected as experimental samples. They all occurred during 2004–2018 and their tracks (both the best tracks and operational NWP model forecast tracks) were obtained from the China Meteorological Administration (CMA)/ National Meteorological Information Center (NMIC). The historical best tracks at 6-h intervals during 1960–2016, including the position and strength of TCs, were obtained from the Shanghai Typhoon Institute (Ying et al. [40]; Lu et al. [41]).

For each target TC among the 47 TCs, the observed precipitation fields of similar historical TCs which occurred before the target TC since 1960 were integrated to make the forecast. The observed precipitation data during 1960–2018, which were archived at 24-h intervals, commencing at 1200 UTC on each day and including 190 stations in South China, were also obtained from the CMA / NMIC. Furthermore, to compare the forecast performance of the DSAEF_LTP model with that of NWP models, corresponding TC rainfall forecast data were obtained from three global forecast systems-namely, the European Centre for Medium-Range Weather Forecasts (ECMWF) model, the National Centers for Environmental Prediction GFS (Global Forecast System) model and the CMA-GFS (Global / Regional Assimilation Prediction System) model run by the CMA-as well as one regional model-the Shanghai Meteorological Service WRF (Weather Research and Forecasting) ADAS Real-Time Modeling System (CMA-SH9). The ECMWF, GFS, CMA-GFS and CMA-SH9 models have the horizontal resolution of 0.125° $\times$ 0.125°, 0.25° $\times$ 0.25°, 0.25° $\times$ 0.25° and 0.09° $\times$ 0.09°, respectively. And in order to process the data uniformly, the horizontal resolution of the four models have been converted to 0.1° $\times$ 0.1° by using bilinear interpolation.

2.2 Methods

The TC rainfall forecast procedure of the DSAEF_LTP model includes the following four steps:
<table>
<thead>
<tr>
<th>Sample classification (39 tropical cyclones from 2004 to 2016)</th>
<th>Names of tropical cyclones from 2004 to 2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training samples</td>
<td>2004: Rananim (0414)</td>
</tr>
<tr>
<td></td>
<td>2005: Sanvu (0510), Vicente (0516), Damrey (0518)</td>
</tr>
<tr>
<td></td>
<td>2006: Bilis (0604), Kaemi (0605)</td>
</tr>
<tr>
<td></td>
<td>2007: Pabuk (0707), Sepat (0709), Lekima (0715)</td>
</tr>
<tr>
<td></td>
<td>2008: Fung-Wong (0808), Nuri (0812), Hagupit (0814), Higos (0817)</td>
</tr>
<tr>
<td></td>
<td>2009: Nangka (0904), Molave (0906), Ketsana (0916)</td>
</tr>
<tr>
<td></td>
<td>2010: Chanthu (1003), Fanapi (1011)</td>
</tr>
<tr>
<td></td>
<td>2011: Nock-Ten (1108), Nesat (1117), Nalgae (1119)</td>
</tr>
<tr>
<td></td>
<td>2012: Doksuri (1206), Kai-Tak (1213), Son-Tinh (1223)</td>
</tr>
<tr>
<td></td>
<td>2013: Rumbia (1306), Soulik (1307), Jebi (1309), Utor (1311), Usagi (1319), Wutip (1321), Haiyan (1330)</td>
</tr>
<tr>
<td></td>
<td>2014: Rammasun (1409), Kalmaegi (1415)</td>
</tr>
<tr>
<td></td>
<td>2015: Linfa (1510), Mujigae (1522)</td>
</tr>
<tr>
<td></td>
<td>2016: Mirinae (1603), Nida (1604), Sarika (1621), Haima (1622)</td>
</tr>
<tr>
<td>Independent samples (eight tropical cyclones from 2017 and 2018)</td>
<td>2017: Nesat (1709), Hato (1713), Pakhar (1714), Mawar (1716), Doksuri (1719)</td>
</tr>
<tr>
<td></td>
<td>2018: Ewiniar (1804), Son-Tinh (1809), Mangkhut (1822)</td>
</tr>
</tbody>
</table>

(1) Obtain the forecast track of the target LTC. (2) Determine the generalized initial values (GIVs) that include variables that may influence LTC precipitation—namely, GIVs are determined by using physical factors, including both TC internal variables (e.g., intensity, and size) and environmental variables (e.g., vertical wind shear, subtropical high, and monsoon). TC track, landfall season and intensity are included in the model. (3) Identify the similarity of the GIVs determined in the second step between the target TC and the historical TCs and thus select N top analogs that are closest to the target TC. Considering that TC track can roughly determine the distribution of TC precipitation, TC track similarity is identified first, and the similarity of other factors in the GIVs are identified on the basis of TC track similarity to further screen out similar historical TCs. (4) Conduct an ensemble prediction of LTC rainfall. More detailed forecast procedures are described in Section 3.2.

In the DSAEF_LTP model, the objective TC track Similarity Area Index (TSAI) is used to determine the similarity of TC tracks (Ren et al. [42]) in step 3, whose principle is to calculate the area enclosed by the historical TCs and the target TC; the smaller the area, the higher the similarity.

In step 4, the Objective Synoptic Analysis Technique (OSAT) is used to partition precipitation generated by TCs from the historical accumulated precipitation (Ren et al. [43–44]; Wang et al. [45]).

Threat Score (TS) is a traditional and effective method for meteorologists to evaluate the forecast accuracy of precipitation forecasts. This study used this method to evaluate the forecasting ability of the DSAEF LTP model and dynamical model for accumulated LTC precipitation. The calculation formula is:

\[
TS = \frac{\text{hits}}{\text{hits} + \text{misses} + \text{false alarms}}
\]

The value of TS is between 0 and 1; the closer the value to 1, the higher the forecasting ability. In the formula, “hits” indicates the number of stations with correct forecasts in which both the observed (rain gauge data) and forecasted precipitation reach a certain magnitude; “misses” refers to the number of stations where the observed precipitation reaches a certain magnitude but the forecasted precipitation does not reach this magnitude; and “false alarms” is the number of stations where the forecasted precipitation reaches a certain magnitude but the observed precipitation does not reach this magnitude. In addition, the DSAEF_LTP model is more prone to false alarms in previous studies (Ding et al. [38]; Jia et al. [39]), so the False Alarm Rate (FAR) was also calculated to evaluate the efforts to reduce false alarms of the modified DSAEF_LTP model for heavy rainfall prediction. The calculation formula is as follows:

\[
\text{FAR} = \frac{\text{false alarms}}{\text{hits} + \text{false alarms}}
\]

The precipitation evaluation index mainly focuses on false alarms and ignores misses. The value is between 0 and 1, and the closer it is to 0, the lower false alarms are.

3 EXPERIMENTAL DESIGN AND FORECAST PROCEDURES

3.1 Experimental design

The DSAEF_LTP model consists of characteristic
parameters. Table 2 lists the values and their corresponding physical significances. In this study, this model contains three physical factors: TC track (determined by P2, P3 and P4), TC landfall season (P5), and TC intensity (P6) (Ding et al. [38]). Moreover, the similarity regions (P2) and ensemble forecast schemes (P8) are important parameters that may influence the forecasting performance of the model. And the first 15 values of P2 and mean, maximum values of P8 in Table 2 are contained in the model at the initial time when the model has been developed. Five similarity regions were added-namely, schemes 16–20 of P2 (Jia et al. [39]), which are shown in the Fig. 1 in this study, and five new ensemble forecast schemes (Jia et al. [39]) of P8 were added (Table 2) into the model, including the optimal percentile (90th percentile in this study), fuse of station-based ensemble methods, probability matching mean, equal difference-weighted mean, and TSAI-weighted mean of field-based ensemble methods. The algorithm of each ensemble scheme is listed in Table 3. In order to examine the forecasting performance after introducing the new parameters into the DSAEF_LTP model, accumulated precipitation forecast experiments were conducted using four different configurations of the DSAEF_LTP model (DSAEF_LTP-1 to 4). These configurations represent the DSAEF_LTP model with three physical factors, namely, TC track, TC landfall season and intensity (DSAEF_LTP-1), and, based on DSAEF_LTP-1, with the five new ensemble forecast schemes added (DSAEF_LTP-2), the five new similarity regions added (DSAEF_LTP-3), and with both the ensemble forecast schemes and similarity regions added (DSAEF_LTP-4).

Figure 2 shows the distribution of the 190 rain gauge stations in South China, the experimental area. In the experiments, the time period 2004–2018 was selected and TCs with daily precipitation (at least one rain gauge station) ≥ 100 mm in South China were selected as the target TCs (names listed in Table 1). Among them, 39 TCs from 2004 to 2016 were selected as the training samples (Fig. 3a) to perform simulation experiments to identify the best forecast scheme of the DSAEF_LTP model, and eight TCs from 2017 to 2018 were the independent samples (Fig. 3b) to test the forecasting performance of the best scheme. Their tracks are shown in Figs. 3a and b, respectively.

3.2 Forecast procedures

The DSAEF_LTP model consisted of eight characteristic parameters, each with several different values in this study. These values produced a numerical combination, and one combination was one forecast scheme. The purpose of the simulation tests on the 39

<table>
<thead>
<tr>
<th>Parameters (1–8)</th>
<th>Tested values</th>
<th>Number of values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial time (P1)</td>
<td>1–3 for 12:00, 00:00 UTC on the day of LTC precipitation falling on land and 12:00 UTC on the day before.</td>
<td>3</td>
</tr>
<tr>
<td>Similarity region (P2)</td>
<td>A parameter of TSAI defined as a rectangle with the diagonal points $M$ and $N$. $M$ is the TC locations at 0, 12, 24, 36, or 48 h prior to the initial time, and $N$ is the TC locations at 0, 6, or 12 h prior to the maximum lead time (i.e., the end of the TC track predicted by NWP). 15 similarity regions are combined by $M$ and $N$. The 16th to 20th newly added similarity regions are shown in Fig. 1.</td>
<td>20</td>
</tr>
<tr>
<td>Threshold of the segmentation ratio of a latitude extreme point (P3)</td>
<td>A TSAI parameter: 1–6 for 0.9, 0.8, 0.7, 0.6, 0.5 and 0.4, respectively.</td>
<td>6</td>
</tr>
<tr>
<td>Seasonal similarity (P5)</td>
<td>1–5 for the whole year, May to Nov, July to Sept, the same landfall month with the target TC, and within 15 days of the target TC landfall time, respectively.</td>
<td>5</td>
</tr>
<tr>
<td>Intensity similarity (P6)</td>
<td>Four categories: average and maximum intensity on first rainy day, average and maximum intensity on all rainy days. Five levels: all grades (grade 1 tropical depression to grade 6 super typhoon), same grade and above.</td>
<td>4×5</td>
</tr>
<tr>
<td>Number of TCs with the top $N$ closest similarity (P7)</td>
<td>1–10 for 1, 2, …, and 10, respectively.</td>
<td>10</td>
</tr>
<tr>
<td>Ensemble forecast scheme (P8)</td>
<td>Mean, maximum, optimal percentile, fuse, probability matching mean (PM), equal difference-weighted mean (ED-WM), and TSAI-weighted mean (TSAI-WM).</td>
<td>7</td>
</tr>
<tr>
<td>Total number of schemes</td>
<td>$3 \times 20 \times 3 \times 6 \times 5 \times 4 \times 5 \times 10 \times 7 = 7,560,000$</td>
<td></td>
</tr>
</tbody>
</table>
Table 3. Improved ensemble schemes in the DSAEF_LTP model.

<table>
<thead>
<tr>
<th>Type</th>
<th>Name</th>
<th>Computational procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Station-based ensemble methods</td>
<td>Mean</td>
<td>1. The precipitation forecast of each station is calculated separately.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. For a station, the average precipitation of n selected analogs at this station [Pr(i), i=1, 2, ..., n] as the final forecast result of the station, Prep = ( \frac{\sum_{i=1}^{n} Pr(i)}{n} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. The forecast results of all stations form the forecast precipitation field.</td>
</tr>
<tr>
<td></td>
<td>Optimal percentile (90th percentile in this study)</td>
<td>1. For each station, pre(r), i=1, 2, ..., n is sorted from small to large. pre(r) is the precipitation ranked r.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. ( d = 1 + (n - 1) \times 0.9 )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. The integer part of d is r and the decimal part is f</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Prep = ( Pre(r) + \frac{[Pre(r+1) - Pre(r)] \times f}{d} )</td>
</tr>
<tr>
<td></td>
<td>Fuse</td>
<td>Calculation rules of forecast precipitation at each station:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1. If Max (pre(i)) ( \geq 100 ) mm, Prep = Max(pre(i))</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. If the 90th percentile value of Pre(i) ( \geq 50 ) mm, the Prep equals the 90th percentile value of Pre(i)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. If the 75th percentile value of Pre(i) ( \geq 50 ) mm, the Prep equals the 75th percentile value of Pre(i)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. If the median value of Pre(i) ( \geq 10 ) mm, the Prep equals the median value of Pre(i)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. If none of the above conditions can be met, the Prep equals the 10th percentile value.</td>
</tr>
</tbody>
</table>
1. Arrange all precipitation data for the $n$ members of the 190 stations in ascending order (containing $190 \times n$ stations’ rainfall data). Divide the data from $190 \times n$ largest to smallest into 190 equal parts, retaining the median of each part and recording them as $\text{prem}(k)$, $k=1, 2, ..., 190$.

2. For a station, the average precipitation of $n$ selected analogs at this station is $\text{Prea} = \frac{\sum_{i=1}^{n} \text{Pre}(i)}{n}$; the Prea of 190 stations is ranked from largest to smallest; the ranking of each station’s prea is recorded as $k$.

3. Corresponding to the $\text{prem}(k)$ of each station based on the $k$ of each station, and $\text{prem}(k)$ is the predicted precipitation for this station, $\text{Prep} = \text{prem}(k)$.

**Field-based ensemble methods**

**Probability matching mean (PM)**

The weight of precipitation for the selected similar TC whose similarity rank $i$ is $W(i) = \frac{2 \times n - i \times 2}{(3 \times n - 1) \times n}$ ($i=1, 2, ..., n$), the forecast precipitation is $\text{Prep} = \sum_{i=1}^{n} W(i) \times \text{Pre}(i)$.

**Equal difference-weighted mean (ED-WM)**

The weight of precipitation for the selected similar TC whose similarity rank $i$ is $W(i) = \frac{A(i)}{\sum_{i=1}^{n} A(i)}$, and the forecast precipitation is $\text{Prep} = \sum_{i=1}^{n} W(i) \times \text{Pre}(i)$.

---

**Figure 2.** Distribution of the 190 rain gauge stations in South China.

**Figure 3.** Tracks of (a) 39 TCs training samples from 2004 to 2016 and (b) eight TCs independent samples from 2017 to 2018.
training samples was to identify the best scheme—that is, to identify the optimal parameter values for predicting heavy rainfall. The process of determining the best forecast scheme was as follows: First, the forecast track after the initial time (P1) obtained from operational TC forecast data and the observed track prior to the initial time were merged into a complete track as the target TC track. Next, in the given similarity region (P2), the TSAI was calculated between the target TC track and all historical TC tracks (before the target TC since 1960), and organized from low to high. The lower the TSAI, the higher the track similarity. In this step, P2, P3 and P4 jointly determined the track similarity. Then, on the basis of TC track similarity, the seasonal (P5) and TC intensity (P6) similarities were identified in order to eliminate the historical TCs with large differences in landfall time and intensity from the target TC. Finally, an optimized ensemble forecast scheme (P8) was adopted to assemble the remaining precipitation fields corresponding to the N top-ranked historical TCs (P7 determined the number N), and the predicted precipitation for the target TC was obtained.

Under ideal conditions, 7, 560, 000 forecast schemes could be combined from eight parameters. However, some TCs cannot be fully valued on certain parameters, such as the initial time (P1) or the similarity region (P2); therefore, the number of common schemes suitable for all 39 TCs should be equal to or less than 7, 560, 000. For operational forecasting of typhoons in China, heavy rainfall in accumulated precipitation associated with TCs was evaluated at the thresholds of 100 mm and 250 mm. Considering the advantage the DSAEF_LTP model showed in forecasting heavy rainfall (Ren et al. [7]), these two thresholds were applied in this study. And the TS was used to evaluate the forecasting ability of the model. The largest TS250+TS100 of the 39 TCs’ common forecast schemes was selected as the best forecast scheme, where TS250 and TS100 represent the average TS value of 39 TCs for predicting TC accumulated rainfall that is ≥ 250 mm and ≥ 100 mm, respectively. Finally, the best forecast scheme was applied to the independent sample forecasts, which were compared with the precipitation fields forecasted by the four dynamical models [three global models (ECMWF, GFS, and CMA-GFS) and one mesoscale model (CMA-SH9)].

4 RESULTS

This section illustrates the improvement and applicability of the DSAEF_LTP models (i.e., DSAEF_LTP-1 to 4) in predicting the accumulated rainfall associated with LTCs through simulation experiments, forecast experiments, and analysis of representative typical cases, which are discussed in Section 4.1, 4.2 and 4.3, respectively.

4.1 DSAEF_LTP model simulation experiments

Based on the forecast procedures of the DSAEF_LTP model described in Section 3.2, four groups of simulation experiments were conducted on the 39 TCs of the training samples to identify the best forecast schemes. The TSs (i.e., TS100 and TS250) from the 23, 988 forecast schemes in DSAEF_LTP-1, 83, 958 schemes in DSAEF_LTP-2, 118, 530 schemes in DSAEF_LTP-3 and 527, 408 schemes in DSAEF_LTP 4 are shown by black dots in Figs. 4a-d, respectively. Each dot in the figures represents a forecast scheme, and the

![Figure 4](image_url)

Figure 4. Scatterplots of average threat scores (TS100-TS250) from (a) 23, 988 forecast schemes of DSAEF_LTP-1, (b) 83, 958 forecast schemes of DSAEF_LTP-2, (c) 118, 530 forecast schemes of DSAEF_LTP-3, and (d) 527, 408 forecast schemes of DSAEF_LTP-4. TS250 and TS100 represent the threat scores for predicting accumulated rainfall that is ≥250 mm and ≥100 mm, respectively. The red dots are the best forecast scheme of the DSAEF_LTP model.
red dot is the best forecast scheme in the experiment, which is determined by the maximum TS250+TS100. It is clear from Figs. 4a and b that the maximum TS250+TS100 of DSAEF_LTP-2 are the same as those of DSAEF_LTP-1, which is 0.3124 (0.0749 + 0.2375), indicating that the model adopting the five new ensemble forecasting schemes has the same forecast performance as the initial two ensemble forecasting schemes, while DSAEF_LTP-3 and DSAEF_LTP-4 have been improved both in TS100 and TS250. The value of TS250+TS100 increases to 0.3645 (0.955+0.2690) and 0.4127 (0.1162+0.2965), respectively, which shows that when the new similarity regions are added, or the new similarity regions and ensemble forecast schemes are added simultaneously, the DSAEF_LTP model’s forecasting performance improves. Table 4 lists the parameter values of the best forecast scheme for each experiment. The optimized parameters of DSAEF_LTP-1 and DSAEF_LTP-2 are the same, and DSAEF_LTP-3 and DSAEF_LTP-4 adopt the newly added parameters, which also indicates that the improved parameters can improve the model’s performance to forecast LTC heavy rainfall.

### Table 4. Optimized parameters of the best schemes of the DSAEF_LTP models.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>DSAEF_LTP-1</th>
<th>DSAEF_LTP-2</th>
<th>DSAEF_LTP-3</th>
<th>DSAEF_LTP-4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial time (P1)</td>
<td>1 (1200 UTC on the day of LTC precipitation falling on land)</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Similarity region (P2)</td>
<td>2 (TC locations at the initial time and 12 h prior to the maximum lead time as a rectangle)</td>
<td>2</td>
<td>18 (listed in Fig. 1)</td>
<td>18</td>
</tr>
<tr>
<td>Threshold of the segmentation ratio of a latitude extreme point (P3)</td>
<td>1 (0.1)</td>
<td>1 (0.1)</td>
<td>2 (0.2)</td>
<td>3 (0.3)</td>
</tr>
<tr>
<td>The overlapping percentage threshold of two TC tracks (P4)</td>
<td>6 (0.4)</td>
<td>6 (0.4)</td>
<td>2 (0.8)</td>
<td>4 (0.6)</td>
</tr>
<tr>
<td>Seasonal similarity (P5)</td>
<td>2 (May-Nov)</td>
<td>2</td>
<td>2</td>
<td>1 (Whole year)</td>
</tr>
<tr>
<td>Intensity similarity (P6)</td>
<td>(4, 2)</td>
<td>(4, 2)</td>
<td>(4, 1)</td>
<td>(4, 1)</td>
</tr>
<tr>
<td>Number (N) of TCs with the top N closest similarity (P7)</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>Ensemble forecast scheme (P8)</td>
<td>2 (Maximum)</td>
<td>2</td>
<td>2</td>
<td>3 (90th percentile)</td>
</tr>
</tbody>
</table>

### 4.2 DSAEF_LTP model forecast experiments

With the best forecast schemes identified, the DSAEF_LTP model can be applied to predict the accumulated rainfall associated with the eight independent samples from 2017 to 2018. Among the four configurations of DSAEF_LTP, the best forecast schemes of DSAEF_LTP-1 and DSAEF_LTP-2 are the same, and obviously, the same forecast results will therefore be obtained. Accordingly, these two models are hereafter collectively referred to as DSAEF_LTP-1(2) for later discussion.

Figure 5 compares the average TSs of the eight TCs’ rainfall produced by the best forecast schemes of the DSAEF_LTP models (DSAEF_LTP-1 to 4) to those produced by the four dynamical models (ECMWF, GFS, CMA-GFS and CMA-SH9) mentioned in Section 2.1. For the DSAEF_LTP models, the sum of TS values (TS250 + TS100) of DSAEF_LTP-1(2) is 0.3222 (0.0452 + 0.2770), and that of DSAEF_LTP-3 is 0.3365 (0.0640+0.2725), which represents an increase by 4.44% after adding the new similarity regions. Although the TS250 of DSAEF_LTP-3 drops slightly, the TS100 increases by 0.0188, improving the overall forecasting performance. In contrast, DSAEF_LTP-4 shows significant improvements both in TS250 and TS100, which performs the best and ranks the first in all forecast models, with a total TS of 0.4039 (0.0774 + 0.3265), increasing by 25.36% relative to DSAEF_LTP-1(2). Among the four dynamical models, ECMWF performs the best both in TS250 and TS100, and the value of TS250+TS100 is 0.2474 (0.0188 + 0.2286). Obviously, the DSAEF_LTP models perform better than the dynamical models for predicting accumulated rainfall that is ≥250 mm and ≥100 mm; and DSAEF_LTP-1, the configuration without the addition of new parameters, already has advantages compared to the dynamical models. In conclusion, adding new similarity regions to the DSAEF_LTP model, or adding both similarity regions and ensemble forecast schemes, helps to improve the model’s forecasting performance for LTC heavy rainfall. That is, DSAEF_LTP-3 and DSAEF_LTP-4 offer improvement, with the latter improving more significantly than the former.

Figures 6a and b compare the TSs associated with
individual LTCs from the best scheme of the DSAEF_LTP model to those from the four dynamical models. Fig. 6a only shows five LTCs (numbered TC1714, TC1716, TC1804, TC1809 and TC1822) with single-station maximum total rainfall exceeding 250 mm, because the accumulated rainfall of the other three LTCs was less than 250 mm, for which it would have been meaningless to analyze their TS250. For LTCs with precipitation ≥250 mm, Fig. 6a indicates that none of the forecast models (i.e., the eight models in this study) provides a value larger than null TSs for TC1716 and TC1822. This might be because the number of stations with precipitation ≥250 mm is only one and four, which is more difficult to predict accurately. For the remaining LTCs, only DSAEF_LTP-3 and DSAEF_LTP-4 have greater than null TSs for TC1714, and the latter provides higher scores. None of the four dynamical models gives TS values higher than null TSs for TC1809, but the DSAEF_LTP model provides better forecast results; and, compared with DSAEF_LTP-1(2), the total TS values of

Figure 5. Average threat scores (TS250, TS100 and TS250+TS100) of eight TCs in independent samples produced by the best scheme of the DSAEF_LTP models, compared with those of the four dynamical models (i.e., ECMWF, CMA-GFS, GFS and CMA-SH9).

Figure 6. Individual threat scores (vertical colored bars) for predicting the accumulated rainfall using the DSAEF_LTP models compared with those predicted by the four dynamical models (ECMWF, CMA-GFS, GFS and CMA-SH9): (a) TS250 for accumulated rainfall that is ≥250 mm associated with five LTCs (TC1714, TC1716, TC1804, TC1809 and TC1822); the other three LTCs are not shown because their single-station accumulated maximum rainfall is <250 mm; (b) TS100 for accumulated rainfall that is ≥100 mm associated with all eight LTCs. Individual false alarm rate for predicting the accumulated rainfall using the DSAEF_LTP models: (c) FAR250 for accumulated rainfall that is ≥250 mm associated with five LTCs (null means the false alarms and hits are both zero); (d) FAR100 for accumulated rainfall that is ≥100 mm associated with eight LTCs.
DSEAFLTP-3 and DSEAFLTP-4 have significantly increased for TC1809, from 0.0909 to 0.3125 and 0.2727 respectively. For TC1804, the forecasting performance of the improved model decreases. In general, the forecasting performance of the DSEAFLTP-3 and DSEAFLTP-4 models have been improved for individual TCs. In particular, they can predict heavy rainfall with valuable TSs when other forecast models are not able to do so.

When the eight TCs with accumulated rainfall ≥100 mm (Fig. 6b) are forecasted, the best forecasters among the DSEAFLTP models and the dynamical models are compared, which reveals that the DSEAFLTP models perform better than the dynamical models for TC1709 (DSEAFLTP-4), TC1713 (DSEAFLTP-4), TC1719 (DSEAFLTP-4) and TC1809 (DSEAFLTP-3). For TC1716 and TC1804, CMA-GFS ranks the top, with 0.2857 and 0.5182 of TS100, and CMA-SH9 has the best forecast performance for TC1714 and TC1822 among all models. Only for the DSEAFLTP models themselves, except for TC1804, does the precipitation forecasting performance of the remaining seven TCs show improvement after the new similarity regions are added (DSEAFLTP-3), or the new similarity regions and the ensemble forecast schemes (DSEAFLTP-4) are added (Fig. 7). DSEAFLTP-4 has the highest TS100 values for TC1709, TC1713, TC1714 and TC1719, and DSEAFLTP-3 performs the best for TC1716, TC1809 and TC1822.

![Figure 7](image.png)

**Figure 7.** Individual threat scores produced by the DSEAFLTP models (DSEAFLTP 1 to 4) for predicting accumulated rainfall that is ≥100 mm.

Figures 6c and d further compare the FAR associated with individual LTCs from the best scheme of the DSEAFLTP models. For LTCs with precipitation ≥250 mm, the DSEAFLTP-3 and DSEAFLTP-4’s FAR250 of TC1714 and TC1809 decrease from 0.9231, 0.5833 to 0.8, 0.4, respectively. And for LTCs with precipitation ≥100mm, the FAR100 of TC1714, TC1716, TC1719, TC1804, TC1809 and TC1822 also reduced to some extent. In general, the FAR of the DSEAFLTP model can be reduced for certain TCs by adding similarity regions and forecast schemes.

4.3 Analysis of representative typical cases

To gain better insight into the improvement of the DSEAFLTP models after adding the new similarity regions and ensemble schemes, two typical representative cases whose TS250 + TS100 increased with the new similarity regions or ensemble schemes added-namely, TC1713 and TC1822 are selected for further analysis in Figs. 8 and 9.

For TC1713, the distribution of heavy rainfall (≥100 mm) predicted by the DSEAFLTP models (DSEAFLTP-1 to 4) (Figs. 8b-d), ECWMF and CMA-GFS (Figs. 8e and f) are consistent with rain gauge observations, while GFS and CMA-SH9 (Figs. 8g and h) have large range regions with heavy rainfall missed to forecast (i.e., where the observed precipitation reaches the magnitude of ≥100 mm but the forecasted precipitation does not reach this magnitude), especially in Guangxi Autonomous Region, which also leads to lower TS100 values. Overall, the forecasting performance of the DSEAFLTP models is comparable to that of the dynamical models for heavy rainfall ≥100 mm; and among the two groups (four DSEAFLTP models and four dynamical models) of forecast models, the best performer, DSEAFLTP-4, with a TS100 of 0.4096, is slightly higher than the ECMWF model, with 0.4000, which ranks first among the dynamical models. In terms of the DSEAFLTP models themselves, DSEAFLTP-3 and DSEAFLTP-4 have better forecasting performance for heavy rainfall at the border of Guangdong and Guangxi than DSEAFLTP-1(2), and the heavy rainfall regions that missed to forecast have been reduced.

TC1822 was the strongest LTC in China in 2018,
producing precipitation of $\geq 100$ mm in the central and western parts of Guangdong, and the precipitation in some areas exceeded 250 mm (Fig. 9a), where all models fail to predict this amount of precipitation. For heavy rainfall $\geq 100$ mm, the distributions of precipitation predicted by GFS and CMA-SH9 (Figs. 9g and h) are closer to the observations. Although the DSAEF_LTP models also successfully capture the heavy rainfall in Guangdong, there are false-alarm areas in the south of Guangxi, which is the reason why their values of TS100 are slightly inferior to the above two dynamical models. However, in the case of the DSAEF_LTP models, the FAR of DSAEF_LTP-1(2) is 0.6232 for precipitation $\geq 100$ mm, which decreases to 0.5 in DSAEF_LTP-3 and 0.5224 in DSAEF_LTP-4, indicating that the DSAEF_LTP models with the newly added similarity regions and ensemble forecast schemes can to a certain extent reduce the FAR of the prediction.
SUMMARY AND DISCUSSION

The forecasting performance of the improved DSAEF_LTP model with new similarity regions and new ensemble schemes added was tested in this study. Four experiments were designed to explore the effects of parameter improvement on the model. Forty-seven LTCs whose rainfall seriously affected South China during 2004–2018 were selected, 39 of which (from 2004–2016) were used as training samples to identify the best forecast scheme (the largest TS250+TS100) and eight of which (from 2017–2018) were independent samples used to test the forecasting performance of the best scheme by comparison with the LTCs’ accumulated rainfall predicted by four dynamical models (ECMWF, GFS, CMA-GFS, and CMA-SH9). The major conclusions are as follows:

1. With the addition of new similarity regions into the model (i.e., DSAEF_LTP-3), the forecasting performance of the DSAEF_LTP model for heavy

Figure 9. Horizontal distribution of the accumulated total rainfall amounts (mm) associated with TC1822 (Mangkhut) from (a) rain gauge observations and predictions from the models including (b) DSAEF_LTP-1, (c) DSAEF_LTP-3, (d) DSAEF_LTP-4, (e) ECMWF, (f) CMA-GFS, (g) GFS, and (h) CMA-SH9. The orange lines are the forecast track of Mangkhut, and the black lines are the tracks of similar historical TCs selected by the best forecast scheme of DSAEF_LTP models.
rainfall (accumulated precipitation \(\geq 250\) mm and \(\geq 100\) mm) improves, and the TS250 + TS100 increases by 4.44\% (from 0.3222 to 0.3365). Although the forecasting performance of DSAEF_LTP-2, which only has the new ensemble schemes added, is the same as that of DSAEF_LTP-1, its forecasting performance is significantly improved and better than that of DSAEF_LTP-3 when new ensemble schemes and similarity regions are added simultaneously (DSAEF_LTP-4), with the TS increasing by 25.36\% (from 0.3222 to 0.4039).

(2) Overall, comparison of the TSs of the four dynamical models shows that the DSAEF_LTP models (DSAEF_LTP 1 to 4) boast advantages in predicting heavy rainfall. For individual LTCs, the DSAEF_LTP models can provide valuable TS250 associated with certain LTCs, especially for those that the four dynamical models find impossible to predict (e.g., TC1714 and TC1809). For accumulated rainfall \(\geq 100\) mm, the forecasting performance of DSAEF_LTP models are superior or slightly inferior to that of the four dynamical models.

(3) The analysis of two representative cases (TC1713 and TC1822) further demonstrated that the improved DSAEF_LTP-3 and DSAEF_LTP-4 can effectively reduce the FAR of the model to a certain extent. And the DSAEF_LTP models can successfully capture most of the areas with heavy rainfall, for which the predicted precipitation patterns are consistent with the observations.

Despite the successful improvement and application of the DSAEF_LTP model, there are several issues that need to be discussed: (i) When the track direction of the target TCs over South China are east-west, the initial similarity regions (i.e., the first 15 values of P2 in Table 2) are narrowed, which leads to that some similar historical TCs cannot be recognized. However, with the square similarity regions newly added, this deficiency has been made up and better forecasting performance has been obtained in DSAEF_LTP-3. (ii) The forecasting performance of the DSAEF_LTP model with the newly added ensemble schemes (DSAEF_LTP-2) is the same as before, while with both the new ensemble schemes and similarity regions added, the performance of DSAEF_LTP-4 significantly improved and became better than that of DSAEF_LTP-3. This suggests that improving forecast skill by introducing new ensemble schemes requires certain conditions. The key to improve the DSAEF_LTP model is the coordination among the various values of each parameter, such as the similarity regions and ensemble schemes. The newly added similarity regions make more similar historical TCs can be selected, and the better prediction result will be gained with optimal ensemble schemes-90\% percentile in this study. (iii) Although the FAR reduced in the improved DSAEF_LTP model, how to further reduce FAR of the model remains an important problem needed to focus on in the future study. The study suggests that (i) conducting simulation experiments with a large number of TC samples is more likely to build a stable DSAEF_LTP model; and (ii) adding appropriate parameters is an effective way to improve the forecasting performance of the DSAEF_LTP model, and the best forecast scheme of the improved model is expected to be further applied to the operational prediction of the heavy rainfall of LTCs over South China. Aside from South China, corresponding experiments also need to be carried out for the whole country, North China, and East China. Moreover, to further improve the model, other verification methods (e.g., MET and contiguous rain area (CRA)) would be considered to comprehensively evaluate the forecasting performance, and physical factors that affect LTC precipitation should also be added into the GIVs of the DSAEF_LTP model, which would contribute to constructing more accurate GIVs to obtain further improvements. These physical factors can be divided into TC characteristics and environmental conditions. The former contained in the current DSAEF_LTP model includes TC track, TC landfall season, and TC intensity. Furthermore, other internal features of TCs (e.g., TC size and TC translation speed) and environmental conditions (e.g., vertical wind shear, subtropical high, low-level jet, etc.) need to be introduced. Therefore, there are many directions for further research to improve the model’s forecasting performance.

REFERENCES


[33] ELSBERRY R L, TSAI H C. Situation-dependent intensity skill metric and intensity spread guidance for


Citation: MA Yun-qi (马蕴琦), REN Fu-min (任福民), et al. Experiments with the Improved Dynamical-Statistical-Analog Ensemble Forecast Model for Landfalling Typhoon Precipitation over South China [J]. Journal of Tropical Meteorology, 2022, 28(2): 139-153, https://doi.org/10.46267/j.1006-8775.2022.011