A COMBINED VERIFICATION METHOD FOR PREDICTABILITY OF PERSISTENT HEAVY RAINFALL EVENTS OVER EAST ASIA BASED ON ENSEMBLE FORECAST

WU Zhi-peng (吴志鹏), 1, CHEN Jing (陈 静), 2, ZHANG Han-bin (张涵斌), 1, CHEN Fa-jing (陈法敬), 2, ZHUANG Xiao-ran (庄萧然) 1
(1. Chongqing Meteorological Observatory, Chongqing 401147 China; 2. Center of Numerical Weather Prediction of CMA, Beijing 100081 China; 3. College of Atmospheric Science, Nanjing University of Information &Science Technology, Nanjing 210044 China)

Abstract: Persistent Heavy Rainfall (PHR) is the most influential extreme weather event in Asia in summer, and thus it has attracted intensive interests of many scientists. In this study, operational global ensemble forecasts from China Meteorological Administration(CMA) are used, and a new verification method applied to evaluate the predictability of PHR is investigated. A metrics called Index of Composite Predictability (ICP) established on basic verification indicators, i.e., Equitable Threat Score(ETS) of 24h accumulated precipitation and Root Mean Square Error(RMSE) of Height at 500hPa, are selected in this study to distinguish "good" and "poor" prediction from all ensemble members. With the use of the metrics of ICP, the predictability of two typical PHR events in June 2010 and June 2011 is estimated. The results show that the “good member” and "poor member" can be identified by ICP and there is an obvious discrepancy in their ability to predict the key weather system that affects PHR. “Good member” shows a higher predictability both in synoptic scale and mesoscale weather system in their location, duration and the movement. The growth errors for "poor" members is mainly due to errors of initial conditions in northern polar region. The growth of perturbation errors and the reason for better or worse performance of ensemble member also have great value for future model improvement and further research.

Key words: persistent heavy rainfall; verification method; predictability; ensemble prediction; error analysis

CLC number: P456.7 Document code: A
https://doi.org/10.16555/j.1006-8775.2020.004

1 INTRODUCTION

Heavy rainfall is one of the common extreme weather events in Asia in summer monsoon season. The definition for a heavy rainfall is described by National Meteorological Center of China as a rainfall process with 24h accumulated rainfall being equal to or more than 50 mm. A Persistent Heavy Rainfall (PHR) is described as a rainfall process with 24h accumulated rainfall being equal to or more than 50 mm and a duration of at least three days by Chen et al. [1]. If there is a consecutive two-day rainfall which is less than 50mm during a persistent heavy rainfall process, it means the persistent heavy rainfall process has ended. The persistent heavy rainfall is characterized by long duration, very strong precipitation and severe impact (Chen and Zhai [2]). Bao et al. [3] and Ding et al. [4] studied the performance of models and explored the shortcomings of model forecast.

A common method to quantitatively evaluate the predictability of weather is to use the Numerical Weather Prediction (NWP). The method’s basic idea is that initial condition errors would gradually accumulate and the NWP model would lose its predictability. Guo et al. [5] and Fan et al. [6] applied the cell-to-cell mapping method to analyze the predictability. Currently, forecast verification metrics are often used to evaluate the performance of the NWP models. For example, Root Mean Square Errors (RMSE) of height field is often used to evaluate the predictability of the atmosphere circulation. The verification metrics of Threat Score (TS) and Equitable Threat Score (ETS) are often applied to evaluate the predictability of category precipitation. Chen et al. [7] applied TS, bias score, miss rate and false alarm to validate Mesoscale Model version5 (MM5) on the different category precipitation. Gong et al. [8] applied the SAL quantitative verification method to evaluate the structure(S), amplitude(A) and location(L) of the rainfall forecast from the models of T213 of CMA, T639 of CMA and Japan. According to researches of Chen et al. [9], Zhou et al. [10], Wang et al. [11], Casati et al. [12], Atger [13], Ebert et al. [14], and Wang et al. [15], the quantitative estimation of the predictability of PHR...
require not only the evaluation of daily forecast performance, but also the evaluation of its overall performance. It becomes a hot issue in the field of prediction of disastrous weather. In general, there are some achievements on the quantitative estimation of predictability of weather events; however, there are no mature methods currently available for the prediction of PHR and much work still needs to be done.

Ensemble prediction is currently a common technology of numerical prediction. Rather than a single forecast from with an NWP model, multiple forecasts are produced by making small alterations either to the initialization conditions or to the forecast model itself, or both, as introduced by Lorenz [10], Du et al. [17-18] and Li et al. [19]. Due to the slight difference of each ensemble member, some have good forecasting performance and some do not. It provides a new perspective for the study of weather and climate predictabilities.

Wen et al. [20] used the dataset of ensemble seasonal forecasts of European Centre of Medium Range Weather Forecast (ECMWF) to assess the potential predictability of the seasonal climate anomaly. For the predictability of PHR, two main highlights should be noticed. One is the model's ability to forecast atmosphere circulation, the other is the ability to forecast precipitation. In this paper, a composite index which assesses the forecast abilities of the atmosphere circulation and precipitation is needed to quantitatively evaluate the model's ability to predict PHR. In addition, ensemble members provide a mix of good and bad predictions. If the best and worst behaving members could be chosen and the differences between their initial conditions and their changes be analyzed, the growth error rate could be certified and the model’s forecast ability would improve (Thompson [21]; Jiang et al. [22]; Zhong et al. [23]).

Based on the ensemble prediction data and verification index, this paper investigates the method to verify the predictability of PHR. A verification metric, which combines the precipitation statistical verification method and verification index of atmosphere circulation forecast, is developed. An Index of Composite Predictability (ICP) is also established. By using ICP, two PHR events that occurred in Yangtze River region were analyzed. By identifying two members with the best and worst forecast performance and comparing the differences between initial condition errors, their growth, and the impact of key weather systems on PHR, we analyzed the rationality and reliability of ICP. This research would lay a foundation for the general assessment of predictability of PHR and mechanism of error growth in NWP model. Furthermore, with the increasing variety of numerical products and ensemble members today, there is a huge amount of data produced, complicating forecaster’s decision-making process. It is hoped that this paper will provide possible solutions to help synthesize more individual ISVs to form a more reasonable and comprehensive ICP in order to better predict a specific target.

2 DATA AND METHOD

2.1 Data

The outputs from CMA T213 ensemble global forecasting system are adopted in this paper. This T213-based ensemble prediction system includes 15 members with the horizontal resolution of 0.5625°. The observations are the precipitation observed from 2412 stations of the National Meteorological Information Center of China. The distribution of the stations can be seen in Fig. 1. There are 524 stations in the area between 26°-35° N, 112° E to the east coast of the Huai River Basin (encircled in the red box). The model outputs are interpolated to the stations in a bilinear way to calculate

Figure 1. Distribution of 2412 stations in China. The east of the Huai River Basin (the region in the red rectangle) is studied in this paper.
the index of verification for persistent heavy rainfall. The National Centers for Environmental Prediction (NCEP) reanalysis data with the horizontal resolution of 1°×1° are used to diagnose analysis of the atmospheric background.

2.2 The two PHR events

Two of the longest events are selected from the PHR events and they are not affected by typhoon systems in the Huai River Basin between 2009 and 2011. The first event started on June 17 and ended on June 25 in 2010 (referred to as Event I hereinafter). And the second one appeared on June 4, 2011 and continued for nine days (hereafter referred to as Event II). Event I ranks third in terms of integrated intensity among the PHR rainfall events in the Huai River Basin from 1950 to 2010, which makes it an excellent example of PHR events.

The total number of the stations with 24-hour accumulated precipitation being more than or equal to 50 mm and the daily precipitation of those two events are shown in Table 1. Obviously, Event I has three significant rainfall peak values. They are on June 17, June 19 and June 24. The rainfall process on June 19 is the strongest and the most extensive. And there are 12 stations where rainfall reaches the value of 250mm. Event II has two strong rainfall peak values. The first is from the June 4 to June 6, 2011, and the second is from June 9 to June 12. June 8 is a temporary stagnation of Event II. The total number of the stations where rainfall exceed 50mm has a good positive correlation with daily precipitation (not shown here).

Table 1. Precipitation of the two persistent heavy rainfall events in the Huai River Basin from 2010 to 2011. Here, the amount of daily precipitation refers to the total rainfall of those 524 stations.

<table>
<thead>
<tr>
<th>Events (June 17-25, 2010)</th>
<th>Date</th>
<th>Total number of stations</th>
<th>Daily rainfall amount (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event I</td>
<td>June 17</td>
<td>54</td>
<td>8522</td>
</tr>
<tr>
<td></td>
<td>June 18</td>
<td>39</td>
<td>5391</td>
</tr>
<tr>
<td></td>
<td>June 19</td>
<td>91</td>
<td>13211</td>
</tr>
<tr>
<td></td>
<td>June 20</td>
<td>35</td>
<td>4778</td>
</tr>
<tr>
<td></td>
<td>June 21</td>
<td>11</td>
<td>2622</td>
</tr>
<tr>
<td></td>
<td>June 22</td>
<td>18</td>
<td>3022</td>
</tr>
<tr>
<td></td>
<td>June 23</td>
<td>31</td>
<td>4967</td>
</tr>
<tr>
<td></td>
<td>June 24</td>
<td>50</td>
<td>7161</td>
</tr>
<tr>
<td></td>
<td>June 25</td>
<td>6</td>
<td>2201</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Events (June 4-12, 2011)</th>
<th>Date</th>
<th>Total number of stations</th>
<th>Daily rainfall amount (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event II</td>
<td>June 04</td>
<td>77</td>
<td>9064</td>
</tr>
<tr>
<td></td>
<td>June 05</td>
<td>34</td>
<td>5908</td>
</tr>
<tr>
<td></td>
<td>June 06</td>
<td>33</td>
<td>6442</td>
</tr>
<tr>
<td></td>
<td>June 07</td>
<td>10</td>
<td>1989</td>
</tr>
<tr>
<td></td>
<td>June 08</td>
<td>0</td>
<td>104</td>
</tr>
<tr>
<td></td>
<td>June 09</td>
<td>54</td>
<td>8874</td>
</tr>
<tr>
<td></td>
<td>June 10</td>
<td>35</td>
<td>5689</td>
</tr>
<tr>
<td></td>
<td>June 11</td>
<td>20</td>
<td>3260</td>
</tr>
<tr>
<td></td>
<td>June 12</td>
<td>19</td>
<td>4114</td>
</tr>
</tbody>
</table>

2.3 Methodology of ICP

Verification methods for NWP model usually adopt a set of prediction and observation data and calculate an Index of Verification (IV) for individual variable so as to quantitatively present the forecast ability of the model. The evaluation methods and indexes vary from one variable to another. For example, threat score (TS), equitable threat score (ETS) and bias score are often used to verify categorical precipitation forecast, while anomaly correlation coefficient (ACC) and root mean square error (RMSE) are more often used to evaluate the forecast of geopotential height at high layer. Usually, the relationship between the model performance and the verification score can be divided into two scenarios: 1. The bigger the score, the better the performance, for example both ETS score and ACC are positive relation score; 2. The negative correlation, such as RMSE. An index, called unified index of verification (UIV) is defined in formula (1). UIV is defined as IV when the relationship is positive. The greater the UIV value, the better the forecast.

\[
\text{UIV} = \begin{cases} 
\text{IV} & \text{score is positive} \\
-\text{IV} & \text{score is negative}
\end{cases}
\]  

(1)

\[
\text{AV}_{m}^{i} = \frac{\sum_{n=1}^{w} \text{UIV}_{n} \text{U}_{m}}{\sum_{n=1}^{w} \text{UIV}_{n} / M}
\]  

(2)

A new dimensionless index which shows the differences of performance in predicting individual variable between the ensemble members, Anomaly of Verification (AV), is defined in formula (2), where \( w \) is the number of ensemble prediction members, \( M \) is the total number of ensemble members and \( i \) is the forecast lead time. At the forecast lead time of \( i \), ensemble member of \( m \) giving a positive maximum value of \( \text{AV}_{m}^{i} \) indicates the best forecast performance, and vice versa. If \( \text{AV}_{m}^{i} \) is equal to 0, the performance of the ensemble member of \( m \) is precisely at the average skill level. In order to evaluate the forecast skill from forecast lead time of 1 to lead time (LT) for the ensemble member of \( m \), Index of Single Variable (AV) is defined as the mean of all forecast LT.

\[
\text{ISV}(\text{AV}_{m}) = \sum_{i=1}^{LT} \text{AV}_{m}^{i} / LT,
\]  

(3)

The maximum and minimum value of ISV(AV) corresponds to the best and the worst performance of members respectively; if a given \( m \) gives a positive maximum value of \( \text{ISV}(\text{AV}_{m}) \), that \( m \) gives the best forecast performance. But \( m \) varies from the index of verification. For example, when ISV (ETS) is defined as ETS score of category precipitation, the ensemble member \( m \) may give the positive maximum value. But when ISV (RMSE) is defined as RMSE score of 500hPa height, the ensemble member \( m \) may give 10th ISV (RMSE) value, which is below the average skill.
PHR events, Index of Composite Predictability (ICP) is defined as:

$$\text{ICP}_m = \begin{cases} \text{ISV}(\text{ETS}_m) + \text{ISV}(\text{RMSE}_m), & \text{if } \text{ISV}(\text{ETS}_m) \times \text{ISV}(\text{RMSE}_m) \geq 0 \\ 0, & \text{else} \end{cases}$$  \hspace{1cm} (4)$$

The ETS scores decrease with the increase of forecast lead times and have an obvious fluctuation. The RMSE scores are approximately linear in its growth and the error growth rates change greatly with each ensemble member and forecast lead times (Fig. 2b and Fig. 3b). Taking the RMSE scores of member 2 for example in PHR Event I, the member 2 has minimum RMSE score at the lead time of 168 hour and the maximum RMSE score at the lead time of 240 h. For PHR Event I, both the ISV(ETS) and ISV(RMSE) of member 10 is clearly larger than that of member 1 (Fig. 2c). The ICP of PHR Event I (showed in Fig. 2d) indicates that member 10 is the “good member” and member 1 is “poor member”. For PHR Event II, the ETS value of member 2 is the largest while that of member 11 is the smallest (Fig. 3c). The maximum and the minimum of ICP in Fig. 3d refer to member 10 and member 1 respectively. This is because the ETS of moderate precipitation and the RMSE of 500hPa geopotential height have opposite signs in the ISV for member 2. According to formula (4), the value of ICP of member 2 is equal to 0. Thus, the maximum ICP of member 10 indicates that it is a “good member”, while the minimum ICP of member 1 indicates that it is a “poor member”.

3 RESULTS AND DISCUSSION

3.1 “Good member” and “poor member” from ICP scores

The “good member” and “poor member” calculated from ICP scores of two PHR events mentioned above are analyzed in this section. Fig. 2 and Fig. 3 give the ETS scores of moderate rainfall, the RMSE scores of 500hPa geopotential height field in East Asia, the ISV(ETS) scores, ISV(RMSE) scores and the ICP scores for PHR Event I and PHR Event II. The model initial time is at 1200 UTC on 14 June 2010 and 1200 UTC on 1 June, 2011.

Both of Fig. 2a and Fig. 3a show the ETS scores of moderate rain for each individual ensemble member.

**Figure 2.** The verification scores for the PHR Event I. (a) The ETS of precipitation at threshold value ≥10mm. (b) The RMSE of 500hPa geopotential height in East Asia. (c) The ISV_{ETS} (blue bars), and the ISV_{RMSE} (red bars) of all ensemble member. (d) The ICP of each individual member (after filtering conflict ISV behavior members).
It is worth noting the sign of the ISV(ETS) and ISV (RMSE) for each individual member. For PHR Event I, 12 of 15 ensemble members have the same signs between ISV(ETS) and ISV(RMSE) at a forecast lead time. But for Event II, only 5 of 15 ensemble members have the same signs between ISV(ETS) and ISV(RMSE) and 10 of 15 ensemble members have opposite sign. This indicates that a better prediction in atmospheric general circulation, a much more satisfied prediction in precipitation, and vice versa.

3.2 The predictability for PHR events

In this section, by using the “good member” and “poor member” identified from ICP scores for the two PHR events mentioned before, we analyze the predictability of large scale general circulation, mesoscale weather systems, precipitation intensities and locations.

We first analyze the predictability of large scale atmospheric circulation of the PHR Event I. Fig. 4 shows the analysis and predictions of the mean geopotential height, temperature and wind speed at 500hPa of the “good member” and “poor member” for both the PHR Event I and the PHR Event II.

Let’s have a first look at the analysis of the two PHR events. For the PHR Event I, Fig. 4a shows there are two strong and stable troughs and a ridge in the northern atmospheric circulation in Asia. The troughs are over Lake Baikal and East Asia respectively, and the ridge is over the Mongolian Plateau. This forms the single blocking high pattern during this period. From June 17 to 25, the Mongolian blocking high slowly moves eastward. And the Western Pacific Subtropical High (hereafter referred as subtropical high) is stable. The 5840gpm isohyet line always stays in the Huai River Basin. This makes the warm and wet air from the south and the cold and dry air from the north meet at this region, which stimulates the genesis and the development of mesoscale weather systems. In addition, the Tibet Plateau is controlled by a warm air. And many low pressure disturbance have eastward activities. The combined effect of all of these factors results in the appearance of the persistent heavy rainfall in the Huai River Basin. For the PHR Event II (Fig. 4b), the northern atmospheric circulation contains two troughs and a ridge. The two troughs are in the south of the Ural Mountains and around East Coast of Asia. The subtropical high is stable, and the 5840gpm isohyet line insist in the Huai River Basin which is also the location of the rain belt. The south trough is deep and has a closed low pressure center. Besides, the area of the Huai River Basin has been constantly affected by the fluctuations. On June 8, 2011, the impact of the northwest flow behind the trough leads to a temporary stagnation of the rainstorm.

For both the PHR Event I and the PHR Event II, the “good member” (Fig. 4c, Fig. 4d) has better performance in predicting the Mongolian High than that of “poor member” (Fig. 4e and 4f). The predicted locations and speed of eastward movements of Mongolian High in “good member” are consistent with the analysis (figure omitted). Based on a detailed analysis on the rainstorm process on June 19, 2010, the
“good member” achieves more accurate predictions on the southward movements of the trough which meets the warm and wet air from the south and then leads to the occurrence of the persistent heavy rainfall. The bias of the predicted trough from analysis for the “poor member” is nearly 10 meridional degrees. Moreover, it is found that the prediction on the eastward withdraw process of subtropical high by “good member” is better, which is missing in the “poor member”.

Figure 4. The average atmospheric circulation during the PHR. (a, b) NCEP reanalysis data. (c, d) predictions of “good member”. (e, f) predictions of “poor member” of the mean 500 hPa geopotential height field (black solid lines), the mean temperature field (red dashed lines) and the wind (bars) for the PHR Event I (left column) and PHR Event II (right column).

For the PHR Event I, a key meso-scale weather system was southwest vortex, which often occurs in southwest China at lower layer (famous named as SW vortex) and brings heavy rainfall during it genesis or when it moves eastward. During Event I the SW vortex generated twice, one was from 1200UTC June 18 to 0000UTC June 20, 2010 and the other was from 1200UTC June 23 to 0000UTC June 25, 2010. Fig. 5 shows the observation and predictions on the two scenarios of SW vortex from “good member” and “poor member”. The predictions for the two scenarios of SW vortex on the genesis and eastward movement from “good member” (green line) are both successful although it is slightly biased. But as for the “poor member” (red line), the prediction both on the genesis and the eastward movement of SW vortex failed. Especially for the second scenario of SW vortex both the prediction on its genesis and eastward movement completely failed.

For the PHR Event II, the mesoscale system was the wind shear over the Huai River Basin form June 4 to June 6, 2011 and the eastward movements of Southwest Vortex from June 9 to June 10, 2011. By comparing the mean geopotential height and wind at 700hPa and the
WU Zhi-peng (吴志鹏), CHEN Jing (陈静), et al.

Figure 5. Event I observations (black solid line) and the predictions of the SW Vortex from the “good member” (green solid line) and the “poor member” (red solid line). (a) 1200 UTC on June 18 to 1200 UTC on June 20, 2010, and (b) 1200 UTC on June 23 to 1200 UTC on June 25, 2010.

Predictions from the “good member” and the “poor member” (Fig. 6), we can see that wind shear near 32°N, 107°-120°E persisted for about 3 days. The “good member” defined the wind shear as the sudden change between the north wind and the west wind. However, “poor member” predicted it as the southwest-southeast wind shear.

Comparisons between the observations and the

Figure 6. Event II low level mean circulation. (a) T213 reanalysis field of mean geopotential height (black solid lines, unit: gpm) and mean wind (vanes) at 700 hPa, and those forecast fields at 700 hPa by (b) the best member (m10) and (c) the worst member (m1) from 4 to 6 June 2011.
forecasts of the “good member” and the “poor member” (Fig. 7) indicate that both members have a good performance in the forecasts of the eastward developments of Southwest Vortex. The forecast of “poor member” of SW vortex was to the north of the observations. But the two members did not successfully predict the dissipation of SW Vortex. They predicted that the vortex would keep existing and moving eastward, which would have brought significant effects on the precipitation forecasting.

The comparisons between the observations and the predictions of precipitation are shown in Fig. 8 and Fig. 9 for the two PHR events. For the PHR Event I, the performance of the “good member” is better than that of the “poor member”. The precipitation amount and the rain belt predicted by “good member” are much closer to the observations. For example, the “good member” predicted the rain belt and the region with maximum precipitation successfully on June 19 and 20, but the “poor member” predicted the rain belts over the Huai River Basin on June 20 and 21, which was incorrect, making a false prediction of precipitation center in central China.

Figure 9 is the same as Fig. 8 but for PHR Event II. During this event, the first heavy rainfall process is caused by the maintenance of the stable wind shear over the Huai River Basin. Both of the “good member” and the “poor member” predicted the wind shear and the precipitation, but “good member” had a better performance in predicting the location of the wind shear, and its precipitation predictions were closer to the observations.

3.3 The initial error growth of “good member” and “poor member”

In this section, we focus on studying the initial error growth, which has an important impact on predictability of PHR. The PHR Event I will be investigated in terms of its thermal field, geopotential height field, dynamical field, studying its initial error growth and its propagation characteristic by using the “good member” and “poor member”.

3.3.1 Temperature Error Propagation

Figure 10 depicts the initial temperature errors both in the “good member” and “poor member”. The initial temperature errors here are calculated as the difference between the forecast field and the analysis field. Red shaded regions in this figure show the positive errors, which means higher forecasted temperatures than those in analysis. The blue shaded regions, on the other hand, represent lower forecasted temperatures than those of analysis. From this figure, we find that the temperature errors at initial time are negligible in most parts of China except for those of “poor member”, which contain small
negative temperature errors in the middle of China. Errors in “poor member” are mainly found at high latitude as well as Polar Regions. But these errors are not significant at initial time, and the distributions of positive and negative errors at initial time are irregular.

By studying in details, we find that a weak trough centered in $60^\circ$E, $60^\circ$N contained negative temperature errors in the early forecast time, which then grew quickly with time and finally spread to a large area, leading to a failure of forecast.

The temperature error characteristics in “good member” and “poor member” at a series of forecast time are analyzed (Fig. 11). At lead time of 48h, the distributions of temperature errors in “good member” and “poor member” are similar except for trough zones in European regions. The weak trough centered in $60^\circ$E, $60^\circ$N discussed above has moved eastward by $20^\circ$-30$^\circ$ in longitude. Positive errors and negative errors are found in the front and in the back of this trough respectively. At lead time of 96h, the situations in “good member” and “poor member” became different. The “poor member” predicted deeper upper-level troughs and contained larger errors when compared to the “good member”, which contained closer to analysis results with higher forecast temperature in the regions from Northwestern China to Mongolia. At lead time of 144h, the errors increased in both members: In “poor member”, the center of negative temperature errors around troughs moved eastward to Mongolia. Meanwhile, east of the trough, there existed a strong positive error center in the ridge zone. In “good member”, temperature errors were mostly negative around East Asia trough. Such differences between these two members inevitably lead to different precipitation prediction. At lead time of 168h and 214h, temperature error centers further increased. The principle difference in forecast between “good member” and “poor member” at 240h was in the Northeast of China. In
“poor member”, strong negative errors appeared which were derived from the propagation of previous negative error centers in Mongolia; in “good member”, errors were relatively smaller.

Figure 11. Evolution of temperature errors (shaded area) of “poor member” (line one) and “good member” (line 2) at 500 hPa. The black contours are isobars, and the color contours are isothermals. Temperature errors at low levels were influenced by the situations at high levels to some extent. The error propagation direction at low levels had similar characteristics with that at high levels. At 200hPa, temperature errors in “good member” and in “poor member” developed from middle and high latitudes to polar region. They propagated along the westerlies until arriving in East Asia. The magnitude errors in “poor member” were obviously larger than those in “good member”.

3.3.2 KINETIC ENERGY ERRORS GROWTH AND PROPAGATION

Figure 12 is evolution of kinetic energy errors in “poor member”. The propagation of high-level kinetic energy errors is primarily along the movement of long wave troughs and ridges in the westerlies. However, the propagation of kinetic energy errors at low levels is not regular and is influenced by the high level systems. In the view of the north branch of the atmospheric general system, the kinetic errors show some patterns: negative errors in the front of troughs and positive errors in the back of troughs. In other words, regions suffering from low-pressure troughs will have bigger kinetic errors. “Good member” has overall smaller kinetic errors than “poor member”.

Overall, the selection of the statistical evaluation index has a great impact on ICP. The ICP values can be an index of single variable or a composite index of multiple variables. Different ISVs or ICP values can be gained from different statistical assessment indices, which could affect the selection of “good member” and “poor member”. The choice of the statistical evaluation index is the key to successful application of ICP. The capacity of predicting the large scale and the mesoscale systems tend to decrease with the increase of forecast lead time. The genesis and growth of perturbation errors and the reason to cause better or worse performance of ensemble member have great value for future model improvement and further research.

4 SUMMARY AND CONCLUSIONS

In this paper, an evaluation method for the predictability of persistent heavy rainfall event is studied with the use of Global ensemble data and verification metrics. A new index of composite predictability is established to assess the predictability of persistent heavy rainfall events in the Huai River Basin. By using the “good member” and “poor member” identified from ICP scores for two typical PHR events in June 2010 and June 2011, we analyze the predictability of large scale general circulation, meso-scale weather systems, precipitation intensities and locations. The initial error growth and its propagation characteristic which has an important impact on PHR weather predictability was studied by using the “good member” and “poor member” from ICP scores in terms of their thermal field, geopotential height field, and dynamical field. Some primary conclusions can be obtained as follows.

1) A metrics called Index of Composite Predictability (ICP) based on Equitable Threat Score (ETS) of 24h accumulated precipitation and Root Mean Square Error (RMSE) of height at 500hPa is established.
The maximum and minimum value of ICP$_m$ corresponds to the best and the worst performance of members respectively. By using ICP$_m$, the “good member” and “poor member” forecast are recognized based on CMA T213 global ensemble members.

(2) ICP metrics is applied to two persistent heavy rainfall events in the Huai River Basin. The “good member” and “poor member” are easily recognized. The forecasts between the “good member” and “poor member” differ greatly in terms of the large scale circulation, mesoscale systems and precipitation. The performance of the “good member” is much better than that of “poor member”, especially in forecasting PHR events. The results prove its rationality and reliability of ICP metric.

(3) The forecasts of the “good member” on large scale systems in terms of its location, intensity, duration and the evolution of subtropical high and East Asia trough were much better than those of the “poor member”. Furthermore, the accuracy of the predictions on large scale systems is the prerequisite for successful mesoscale predictions.

(4) Accurate forecasting of heavy precipitation is the key to successful prediction of persistent heavy rainfall events, which is closely associated with the prediction on the genesis and development of mesoscale systems. It was found that the good member was also much more accurate than poor member, especially in predicting SW Vortex, such as its occurrence and movement speed.

(5) The comparisons of the error growth between “good member” and “poor member” reveal that most of the temperature errors exist in the high latitudes and polar areas and grow rapidly with time, especially the errors on the trough. Those errors increase quickly with the eastward movement of the trough, and eventually lead to the failure of the forecasts. The kinetic energy errors at upper levels propagate along the planetary wave in the westerly flow, but the lower’s are irregular and correlated with different systems. Generally, the negative errors are in the front of the trough, and the positive errors are behind the trough. Thus, this will lead
to higher energy existing behind the trough.

(6) This research provides a very basic concept that make it possible to synthesize different individual verification indexes together, and form a more comprehensive index in order to pick out the ensemble members with better performance. With the increasing variety of numerical products and ensemble members today, this idea could lay a foundation for the optimal selection to help predict a long-term weather system.

Acknowledgments: The authors would like to thank Dr. Jing Chen and Mr. LI Xin-cheng for translating the manuscript into English.

REFERENCES:


