SCIENCE CHINA
Earth Sciences

• RESEARCH PAPER •

May 2013 Vol.56 No.5: 878–889
doi: 10.1007/s11430-012-4513-1

Forecast scheme and strategy for extended-range predictable components

ZHENG ZhiHai[1*], HUANG JianPing[2], FENG GuoLin[1] & CHOU JiFan[2]

[1] *National Climate Center, China Meteorological Administration, Beijing 100081, China;*
[2] *College of Atmospheric Sciences, Lanzhou University, Lanzhou 730000, China*

Received May 18, 2012; accepted August 29, 2012; published online September 30, 2012

Although extended-range forecasting has exceeded the limit of daily predictability of weather, there are still partially predictable characteristics of meteorological fields in such forecasts. A targeted forecast scheme and strategy for extended-range predictable components is proposed. Based on chaotic characteristics of the atmosphere, predictable components and unpredictable random components are separated by using the standpoint of error growth in a numerical model. The predictable components are defined as those with slow error growth at a given range, which are not sensitive to small errors in initial conditions. A numerical model for predictable components (NMPC) is established, by filtering random components with poor predictability. The aim is to maintain predictable components and avoid the influence of rapidly growing forecast errors on small scales. Meanwhile, the analogue-dynamical approach (ADA) is used to correct forecast errors of predictable components, to decrease model error and statistically take into account the influence of random components. The scheme is applied to operational dynamical extended-range forecast (DERF) model of the National Climate Center of China Meteorological Administration (NCC/CMA). Prediction results show that the scheme can improve forecast skill of predictable components to some extent, especially in high predictability regions. Forecast skill at zonal wave zero is improved more than for ultra-long waves and synoptic-scale waves. Results show good agreement with predictability of spatial scale. As a result, the scheme can reduce forecast errors and improve forecast skill, which favors operational use.

extended-range forecast, predictability, predictable components, analogue-dynamical approach

Citation: Zheng Z H, Huang J P, Feng G L, et al. Forecast scheme and strategy for extended-range predictable components. Science China: Earth Sciences, 2013, 56: 878–889, doi: 10.1007/s11430-012-4513-1

Since establishment of the first atmospheric general circulation model, major numerical prediction centers around the world have performed many dynamical extended-range forecast (DERF) studies [1, 2]. However, operational application was rarely done, because of poor predictive skill. With growing awareness of climate systems, especially deepening understanding of extended-range predictability such as the Madden-Julian Oscillation (MJO) and stratosphere-troposphere interactions [3, 4], a foundation was laid for extended-range forecast (ERF) development. In the last decade, data assimilation techniques and model performance have greatly improved, and ensemble prediction techniques have become widely used. Together, these have provided favorable conditions for improving ERF skill. Thus, the world's leading numerical prediction centers have once again taken great interest in ERF, and have gradually developed their own operational or quasi-operational ERF systems [5–8]. However, the forecast skill of these systems needs further improvement.

At present, DERF error remains large. Its growth mainly results from two factors: First, atmospheric instabilities amplify uncertainties in initial conditions, causing indistin-

*Corresponding author (email: zhengzh@cma.gov.cn)

guishable states of the atmosphere to diverge rapidly on small scales (known as internal error growth); second, numerical model imperfection leads to external error growth, i.e. model error [9, 10]. The first aspect is an atmospheric predictability problem, which is unavoidable because the time scale of extended-range has exceeded the limit of daily predictability. The second aspect is related to errors in the approximate description of the actual atmosphere by a numerical model. Reducing prediction errors associated with these two aspects is key to improving the skill of DERF.

Atmospheric predictability is closely linked to spatial and temporal scales. As stated earlier, the lead time of extended-range forecasts exceeds the limit of daily weather forecast predictability. However, there are still predictable meteorological field characteristics. Both observational and dynamical theoretical studies suggest the existence of predictable components of meteorological fields on the extended-range time scale [11–13]. Monin [14] proposed that to constructively solve the problem of long-range predictability, it is necessary to describe characteristics of meteorological fields predicted in this period. Significant differences exist among the predictabilities of extended-range, short-range and seasonal forecasts. The means of extended-range forecasts using extended-range predictable characteristics provides the basis for improving predictive skill, but there are relatively few works in this area. In recent years, based on atmospheric low-frequency oscillation characteristics, encouraging results have been obtained using certain statistical methods, in an attempt to forecast low-frequency information with characteristic time scale longer than extended-range [15]. However, further research is needed on predictable characteristics of separated low-frequency information. Based on chaotic characteristics of the atmosphere, estimating the growth rate of prediction errors is a major focus in predictability research [16–20]. Given these inherent chaotic characteristics, Chou et al. [21] recently proposed a ERF forecast scheme based on current numerical models, in which numerical model state variables are divided into two parts: Predictable components (PCs) on the extended-range time scale, and unpredictable stochastic components. They stated that predictable components are insensitive to small errors in initial conditions within the prediction period (i.e. extended-range), and can thereby be predicted using numerical models. However, how to predict these components needs to be further investigated.

With a view to the above issues, and based on the current operational DERF model (T63L16) of the NCC/CMA, we established a numerical model targeted for extended-range predictable components (NMPC). We retained predictable components with higher predictability, and filtered out unpredictable stochastic components in the computation, to avoid impacts of rapid growth of the prediction error of small-scale components. The NMPC model can perform deterministic prediction of PCs. Similar to the DERF model, errors in the NMPC are inevitable, and it does not account for the statistical effect of stochastic components on PCs because of direct removal of these stochastic components from the model. Studies have shown that the analogue-dynamical approach (ADA) can effectively combine the advantages of dynamic and statistical methods to improve numerical model prediction [22–27]. Therefore, targeted at the above two problems, the ADA was used in combination with historical data in deterministic prediction of PCs, to correct prediction error in the NMPC (i.e. reduce model error) and to statistically consider the effect of the random components on the predicted PCs. ERF experiments of 6–15 days were performed with the T63L16 model according the above reasoning, and results show that the ADA can improve the ERF skill of the model.

1 Calculation of PCs and establishment of NMPC model

There are objective PCs of the extended-range time scale in meteorological variable fields, and the means to effectively separate them in numerical models is the basis of deterministic forecast. Predictability is an inherent property of the atmosphere, and the numerical model is the primary means of prediction. Thus, separating PCs is necessary to consider the predictability of the atmosphere itself and the prediction ability of the model. That is, the separated PCs are predictable in the real atmosphere, and the numerical model is also potentially able to predict them.

To reduce computation, it is necessary to reduce the degree of freedom of meteorological variable fields since the large degree of freedom in numerical models. Leading empirical orthogonal function (EOF) components are often associated with low-frequency atmospheric circulation patterns, so one can use the EOF method to carry out this reduction, and to identify independent variables supporting the attractor from historical data. Among these, EOFs explaining a larger variance reflect most information on variations of the original variable field, and mainly delineate the slowly varying portion of the real atmosphere, which often has higher predictability [28, 29]. Therefore, whether a numerical model has the potential capability to predict those EOF components is key to separate the PCs.

Variance analysis is one of the methods used to gauge predictability. It can effectively distinguish the relative magnitudes of climate signal and noise [30, 31], and was therefore used to gauge the predictability of EOF components. In this analysis, internal variance between ensemble numbers represents the sensitivity of the numerical model to initial conditions; its ratio over the climatic external variance reflects the magnitude of predictability. When the variance ratio is very large, the external forcing signal is almost completely hidden by uncertainties of the chaotic system and useful information is even less than that provided by the climate state; thereby it is unpredictable [32, 33]. The detailed procedure for determining PCs is as follows. After reducing the degree of freedom, climate variance of each EOF component and internal variance between ensemble numbers are calculated. The ratio of the internal over external variance is defined as R; the predictable EOF components whose R<1 are then defined as PCs. Determining PCs in such a way not only considers predictability of those components in the real atmosphere, but also potential predictability of the model for those components.

After separating the PCs whose error grow slowly during prediction, it is set that

$$F=\begin{bmatrix} u_1(1) & u_1(2) & \cdots & u_1(p) \\ u_2(1) & u_2(2) & \cdots & u_2(p) \\ \vdots & \vdots & \vdots & \vdots \\ u_n(1) & u_n(2) & \cdots & u_n(p) \end{bmatrix}_{n\times p},$$
$$G=\begin{bmatrix} u_1(1) & u_2(1) & \cdots & u_n(1) \\ u_1(2) & u_2(2) & \cdots & u_n(2) \\ \vdots & \vdots & \vdots & \vdots \\ u_1(p) & u_2(p) & \cdots & u_n(p) \end{bmatrix}_{p\times n}, \tag{1}$$

where F is the standardized orthogonal basis, G is its transpose, and p is the number of predictable EOF components. The transform relationship between model variable φ and PCs ψ is

$$\varphi = F\psi \ , \tag{2}$$

and its inverse transform is

$$\psi = G\varphi \ . \tag{3}$$

The above two transform processes are mutually reversible. Then, the prediction of PCs from time step i to i+1 can be expressed as

$$\varphi_i \Rightarrow G\varphi_i = \psi_i \Rightarrow F\psi_i = \varphi_i^* \Rightarrow M(\varphi_i^*) = \varphi_{i+1} \Rightarrow G\varphi_{i+1} = \psi_{i+1}, \tag{4}$$

where M is the numerical model used, and φ_i^* represents the initial value fields after filtering random components of small scale. The prediction equation for PCs is $\psi_{i+1} = GM(FG\varphi_i)$; thereby, the model state is projected onto the PCs. If the information of random components is eliminated in the prediction process, then the original model becomes the NMPC model.

Atmospheric changes on various time scales possess a fractal structure. There are complicated nonlinear interactions between components of different time scales. After determining the time length of prediction, the object of the forecast should be determined according to this length. The principle for determining the prediction object is that the predictability limit of the object must be longer than the time length of prediction, whereas weather phenomena whose predictability limits are shorter than that length are unpredictable. The NMPC model was constructed based on this idea, by filtering stochastic components whose predictability limit is shorter than the time length of prediction, while reserving PCs whose predictability limit is longer than the that length during computation.

2 Principles for ADA

Numerical weather prediction is essentially an initial and boundary value problem, and the NMPC model can be expressed as $\psi_T = M_s(\psi_0)$, where $\psi_0 \in R^n$ indicates the nearly true observed initial conditions at time t_0, and $\psi_T \in R^n$ indicates the prediction fields of PCs at time T. M_s is the nonlinear numerical model of PCs. It is only an approximation of the evolution of PCs in the real atmosphere (denoted as H_s). Model error (denoted as E) results from the difference between M_s and H_s, which is also dependent on initial and boundary conditions, i.e.

$$E(\psi_0) = H_s(\psi_0) - M_s(\psi_0) \ . \tag{5}$$

The error term in the NMPC includes error caused by model defects and by the statistical effect of filtered random components on the reserved PCs. From the point of view of the inverse problem, several studies have converted the prediction problem of model variables to one of numerical model error, for reducing prediction error. These studies developed an ADA for complex numerical prediction models [22–27], in which prediction error of the historical analogue was used to estimate the unknown error term E. The principle is that the initial value ψ_0 is regarded as a small disturbance superimposed on a historical analogue field, $\psi_0 = \tilde{\psi}_i + \psi_i'$; meanwhile $\|\psi_i'\| << \|\psi_0\|$. Substituting the historical analogue reference state $\tilde{\psi}_i$ into eq. (5), one obtains

$$E(\tilde{\psi}_i) = H_s(\tilde{\psi}_i) - M_s(\tilde{\psi}_i) \ , \tag{6}$$

where $H_s(\tilde{\psi}_i)$ is the real evolution of the ith analogue reference state (i.e. the observation is known); $M_s(\tilde{\psi}_i)$ is the model prediction, which is also known. Therefore, the prediction error of the analogue reference state, $E(\tilde{\psi}_i)$, can be calculated. Since ψ_0 is very close to $\tilde{\psi}_i$, the first order Taylor expansion of $E(\tilde{\psi}_i)$ can be performed in the vicinity of ψ_0, and when $\|\psi_i'\|$ is sufficiently small, one obtains

$$E(\tilde{\psi}_i) = E(\psi_0 - \psi_i') \approx E(\psi_0) - \psi_i' \frac{\partial E}{\partial \psi}\Big|_{\psi=\psi_0}. \tag{7}$$

If one can find k closest historical analogues from a large number of historical records, then

$$E(\psi_0) \approx \frac{1}{A}\sum_{i=1}^{k} a_i E(\tilde{\psi}_i) + \frac{1}{A}\frac{\partial E}{\partial \psi}\Big|_{\psi=\psi_0} \sum_{i=1}^{k} a_i \psi_i', \tag{8}$$

where a_i is the weight determined by the extent of similarity, and $A = \sum_{i=1}^{k} a_i$. In the case where $\sum_{i=1}^{k} a_i \psi_i'$ is as small as possible,

$$E(\psi_0) \approx \frac{1}{A}\sum_{i=1}^{k} a_i E(\tilde{\psi}_i). \tag{9}$$

Eq. (9) is in fact equivalent to using the prediction error information of historical analogues to estimate the current one. The advantages of the ADA are as follows. It does not require reestablishment of the numerical model by using historical data. The ADA considers not only model systematic error, but also contains information on evolution of model error of the analogue manifold with a close initial field, without involving the detailed function forms of prediction errors and model variables. This avoids the difficulties of directly establishing a statistical connection between prediction errors and model variables. Therefore, the ADA is easily transplanted.

The ADA has been greatly developed in theory and application [22–27], but two fundamental unresolved problems remain. First is how to ensure consistency in the evolution of prediction errors of similar initial value fields, since the numerical model is sensitively dependent on the initial values. The second problem is how to select analogue fields, because the degree of freedom of numerical models is very large. Under the condition of existing historical data, it is too difficult to find very close analogues of variables with such a large degree of freedom.

Existing studies have barely addressed the first difficulty. In fact, the hypothesis that prediction error is not extremely sensitive to initial values was used in deriving eqs. (7)–(9), which require that the error function be smooth. Only under such a premise can small quantities of high-order and $\sum_{i=1}^{k} a_i \psi_i'$ be neglected. A basic feature of the real atmosphere and existing numerical models is sensitivity to initial values; thus, direct use of the ADA in existing numerical models has great limitations. The basic characteristic of separated PCs here is that in the prediction period, initial error of the dynamical equations does not significantly increase, i.e. their numerical solutions are insensitively dependent on small errors in the initial conditions. This provides a theoretical basis for avoiding the first difficulty above. Meanwhile, hindcast error of the NMPC not only includes model error caused by its own defects, but also contains the statistical effect of filtered stochastic components on the reserved PCs. Therefore, use of the ADA in the prediction of PCs not only corrects the model error, but also considers the statistical effect of the stochastic components on the PCs.

Regarding the second difficulty, it is common to use a key variable field, such as the 500 hPa geopotential height, to determine historical analogues. However, using only one layer information cannot usually be a good representative of the initial state. The degree of freedom of separated PCs is greatly reduced relative to the original variable field, which avoids degree of freedom difficulties and effectively eliminates high frequency components of the small scale. The selected analogues mainly contain large-scale information, and thus their persistence is relatively good. Therefore, the multivariate integrated similarity index was used to assess the degree of similarity between two initial fields. Preliminary tests show that this method was able to effectively find the historical analogues [29].

3 Numerical experiments using operational model

3.1 Experimental scheme

The prediction scheme designed was as follows:

(1) Establishment of the PC prediction model. It was assumed that the atmosphere is in the same climatic state in the 30-day period from 7 days before to 23 days after the initial time. For example, for an initial time in the numerical experiments of 12:00 UTC on January 16, 2008, the atmosphere from January 9 to February 7 (and during the prediction period January 17–31) was considered to be in the same climatic state. The EOF method was first used to compress the degrees of freedom of the daily spectral coefficient anomaly fields of model variables, at each level in the 30-day periods from 1971–2000. This filtered out components caused by random errors. Then the lagged average forecast was used to produce the 20-member ensemble prediction results, and the variance analysis method was used to calculate the ratio of internal over external variance for each component, after filtering random components. Those components with variance ratio less than 1 were determined as predictable. Finally, in the T63L16 model computation, random components were filtered out by doing a transform/inverse transform of model variables and predictable components once every 24 h, as in eqs. (2)–(4), until the end of prediction.

(2) Historical analogue selection. Using spectral coefficient data of initial fields, the projection of the variable at each level on the PCs was the variable used to select historical analogues. Euclidean distance was used to evaluate similarity of two fields, wherein the explained variance of

each PC was taken as the weight. A best analogue was selected from each year of the 30-year period, and the four best analogues were selected from these as the reference state, to ensure that the historical analogues and current initial value had similar climate backgrounds.

(3) Error diagnosis and correction. After determining historical analogues, each historical analogue $\tilde{\psi}(t_0')$ was integrated forward to δt (6 h), and its prediction error from time t_0' to $t_0' + \delta t$ was obtained from historical data. Then, the prediction error of current initial value $\psi(t_0)$ from time t_0 to $t_0 + \delta t$ was estimated from prediction errors of the historical analogues with an equal weight, and added to the model prediction value $\tilde{\psi}(t_0 + \delta t)$. In the second sub-period, the time integration began directly from $\psi(t_0' + \delta t)$ to save computation, and the prediction error correction of the initial value at time $t_0 + 2\delta t$ was simply repeated with the above procedure, until the end of prediction. That is to say, the prediction error correction of the initial value was performed at a time interval of 6 h, using the historical analogue information.

It is worth highlighting that the prediction scheme using historical data was targeted for the NMPC model, rather than simple direct application of the T63L16 model. We focused on the 6–15-day forecast; the first five-day forecast was not considered. As to the latter forecast, each scale component can be considered predictable, especially since the numerical resolution of the T63 mode is relatively coarse. Therefore, all components were maintained in the first five-day prediction computation. Subsequently, the degree of freedom of variables was first compressed, and then unpredictable components filtered out. To save computation without loss of generality, five independent cases at 12:00 UTC on January 16, 2004–2008 were taken as initial fields in the numerical experiments. Since the forecast object is PCs, projections of the original observation data on the PCs were taken as the "real" observations. Prediction results using the ADA were examined against those of the NMPC model, with the anomaly correlation coefficient (ACC) and root-mean-square error (RMSE) for global (20°–90°S, 0°–360°E), Northern Hemispheric extratropics (20°–90°N, 0°–360°E), Southern Hemispheric extratropics (20°–90°S, 0°–360°E) and the tropics (20°N–20°S, 0°–360°E).

3.2 Daily circulation forecast

3.2.1 500 hPa geopotential height

Figure 1 illustrates the ACC of the daily forecast and the real observation of regional 500 hPa geopotential height in the five cases. ACC of the ADA using historical data are clearly superior to those of the NMPC model. There are significant regional differences in the amount of improvement in ACC. In the Northern Hemispheric extratropics (Figure 1(b)), the ACC of the ADA is slightly higher mainly after the 10th day, but slightly lower than that of the NMPC before then. In the Southern Hemispheric extratropics and the tropics (Figure 1(c), (d)), atmospheric predictability is greater. The improvement in both is obvious, especially in the Southern Hemispheric extratropics, the improvement is

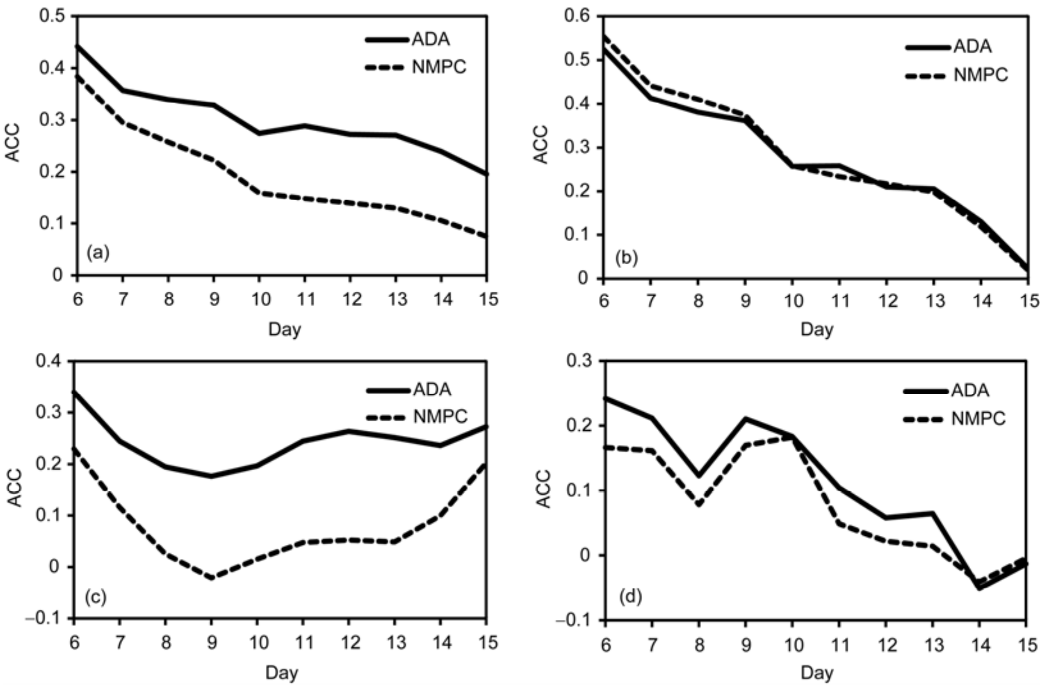

Figure 1 ACCs of the daily forecast and the real observation of regional 500 hPa height field for globe (a), Northern Hemispheric extratropics (b), Southern Hemispheric extratropics (c), and tropics (d).

particularly significant. Globally (Figure 1(a)), the reduction of the ADA ACC score with increasing prediction time is smaller than that of the NMPC, which is important for extending the prediction time.

To further illustrate the effectiveness of the historical data use, Figure 2 shows the RMSE of prediction results. It is observed in the figure that in various regions, the ADA effectively reduced prediction errors compared to the NMPC. Furthermore, the reduction becomes more obvious with increasing prediction time. However, the case is slightly different in the Southern Hemispheric extratropics, where the improvement decline with prediction time. Overall, with increased prediction time, prediction error of the ADA increases very slowly, and is lower than that of the NMPC. Thus, the prediction of the ADA using historical data is improved relative to the NMPC, in terms of both ACC or RMSE, especially for longer forecast times.

3.2.2 850 hPa geopotential height

Improvements of the 850 hPa geopotential height, more closely connected with daily surface weather variables, are shown in Figures 3 and 4. From the viewpoint of global ACC scores, prediction skill of the 850 hPa geopotential height is more clearly improved than that of 500 hPa, but the improvement shows great regional variation. Improvement is particularly obvious in the Southern Hemispheric extratropics. Nevertheless, in the Northern Hemispheric extratropics and tropical region, ADA prediction skill is even lower than that of the NMPC. In contrast to ACC, RMSEs are definitely improved by the ADA over nearly all regions. The improvement in the Northern Hemispheric extratropics is less than in other regions. Regional distributive characteristics of the RMSE scores indicate that improvement of RMSE was dependent on regional predictability of the variable at the same level. By comparing 500 hPa with 850 hPa, it is easy to see that improvement of ADA prediction is greater at 850 hPa. This is opposite the vertical distribution of the ACC of the NMPC. This indicates that improvement of numerical model prediction using historical data is closely related to the model prediction capability for PCs. That is to say, when the numerical prediction model is very accurate, the ADA prediction results degenerate to the forecast results of the dynamical method. However, when the historical analogue is very accurate, the ADA prediction results degenerate to the forecast results of the historical analogue. The current situation for both numerical model and historical analogue is still far from these two limiting circumstances. Therefore, there is still much room for development and potential application of the ADA, which takes full advantage of both dynamical models and historical data information.

3.3 Average circulation forecast

Figures 5–8 show the prediction skill of the regional pentad and 6th–15th day mean forecasts of 500 and 850 hPa geo-

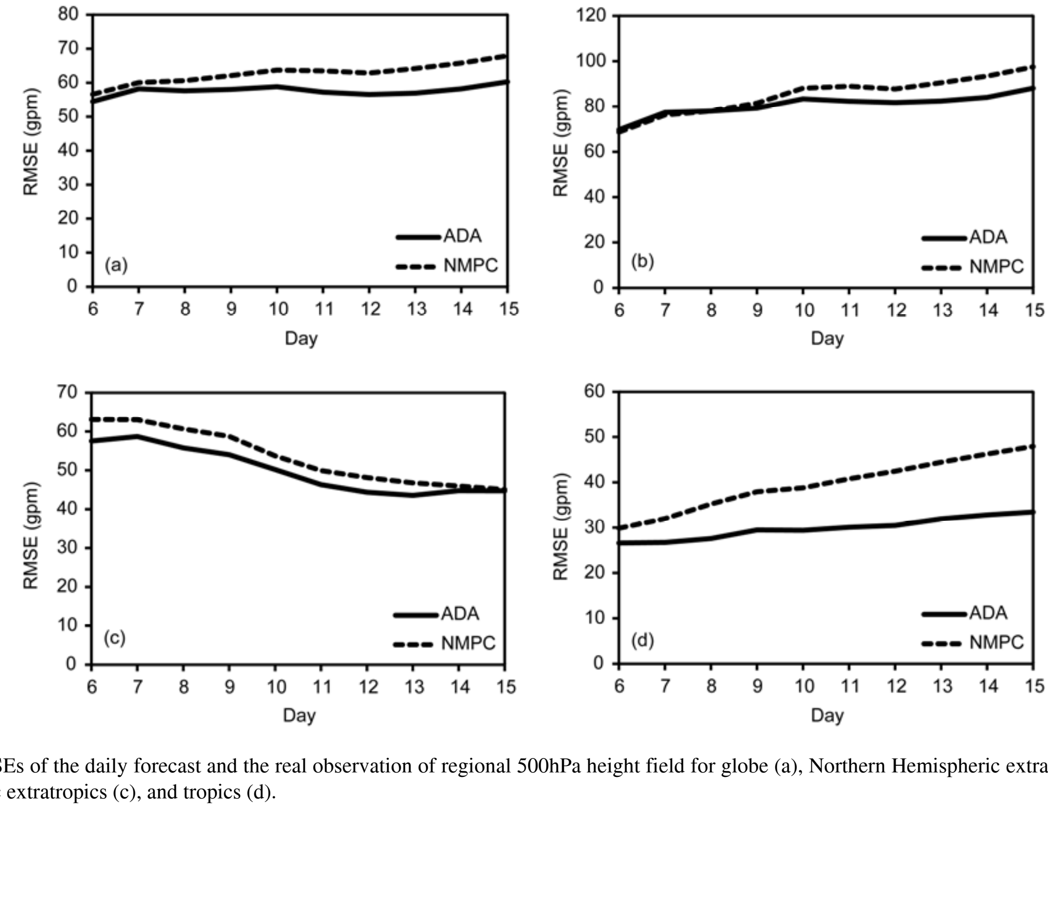

Figure 2 RMSEs of the daily forecast and the real observation of regional 500hPa height field for globe (a), Northern Hemispheric extratropics (b), Southern Hemispheric extratropics (c), and tropics (d).

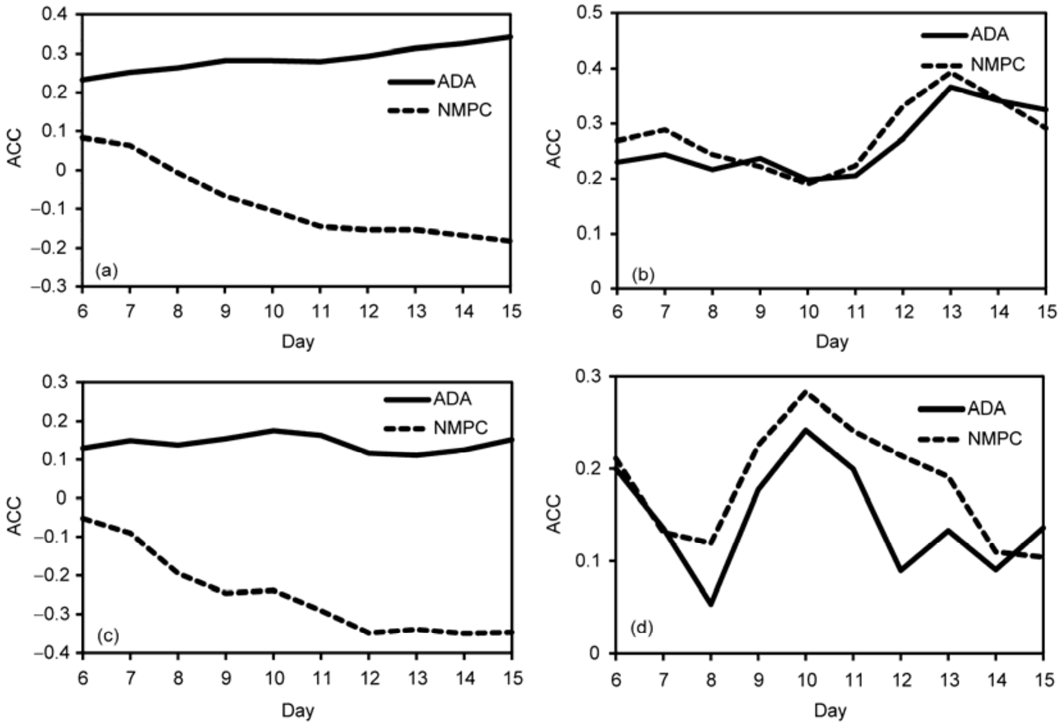

Figure 3 Same as in Figure 1, but for 850 hPa.

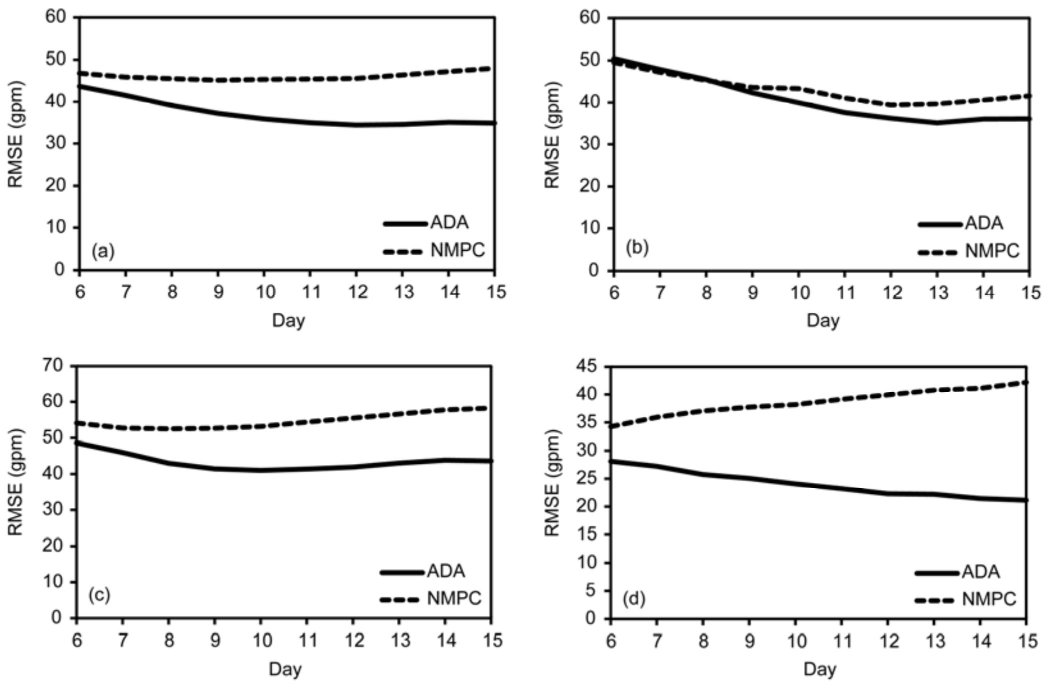

Figure 4 Same as in Figure 2, but for 850 hPa.

potential heights. It is seen from Figures 5–8 that average circulation at the two levels are improved, with improvement decreasing from the lower to the upper level, consistent with the circumstances of daily circulation forecasting. In particular, whether in terms of ACC or RMSE, improvement of the circulation forecast of the third pentad using the ADA is significantly greater than that of the second pentad. This is encouraging and has an important meaning, especially for forecasting on longer time scales (for example, 10–30 days). In the future, the authors will

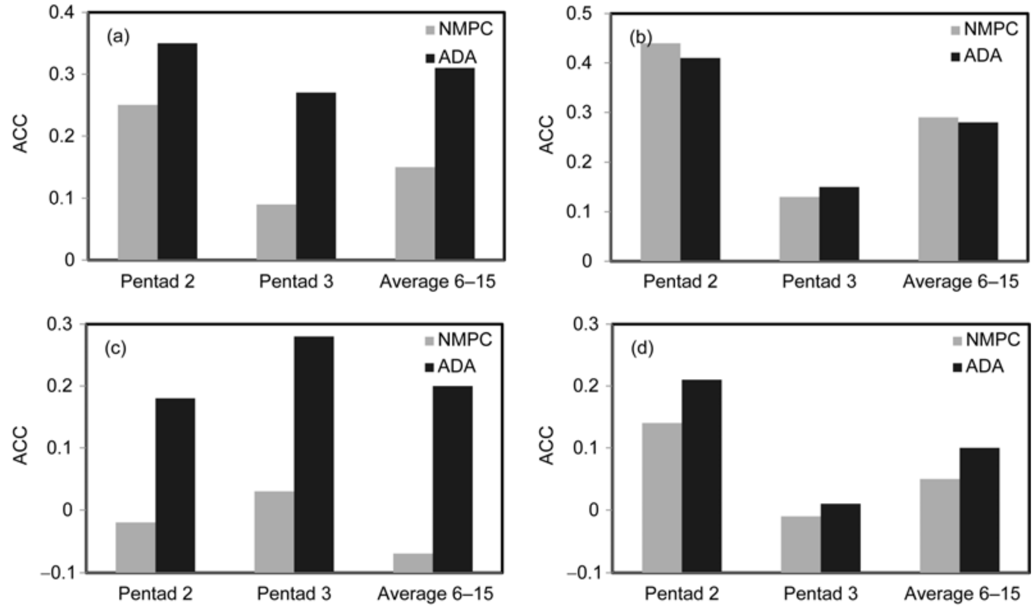

Figure 5 ACCs of the pentadly forecast and the real observation of regional 500 hPa height field for globe (a), Northern Hemispheric extratropics (b), Southern Hemispheric extratropics (c), and tropics (d).

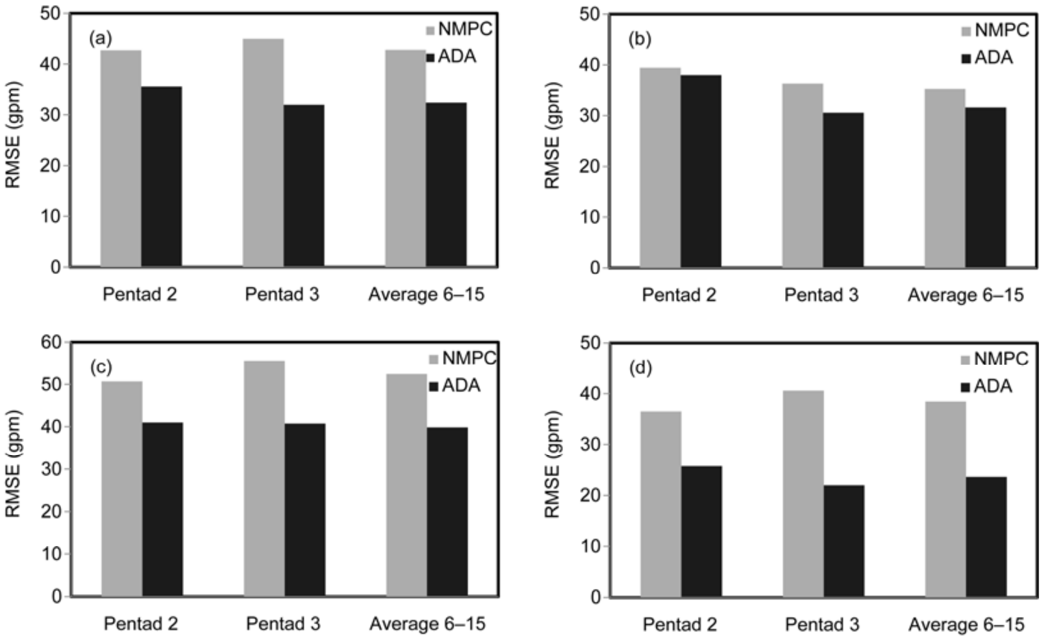

Figure 6 Same as in Figure 5, but for RMSE.

attempt a 10–30-day ERF using the ADA.

It is seen from the regional 6th–15th day average skills that ACC improvement of the NMPC, with use of the ADA, is most evident in the Southern Hemispheric extratropics. Improvement is not clear in the other regions, and it is even smaller than that of the NMPC without the ADA. However, the RMSEs show that skill is improved in nearly all regions, led by the tropics, followed by the Southern Hemispheric extratropics. There is relatively small improvement in the Northern Hemispheric extratropics. This was manifest in the above daily forecast case, thereby indicating that improvement of prediction skill in the NMPC by use of historical data was affected by regional predictability.

PCs are closely linked with spatial and temporal scales; therefore, prediction improvements for waves of different scales in the NMPC using historical data are also worthy of investigation. Figure 9 shows the ACC and RMSE of the 6th–15th day average forecast and real observations of dif-

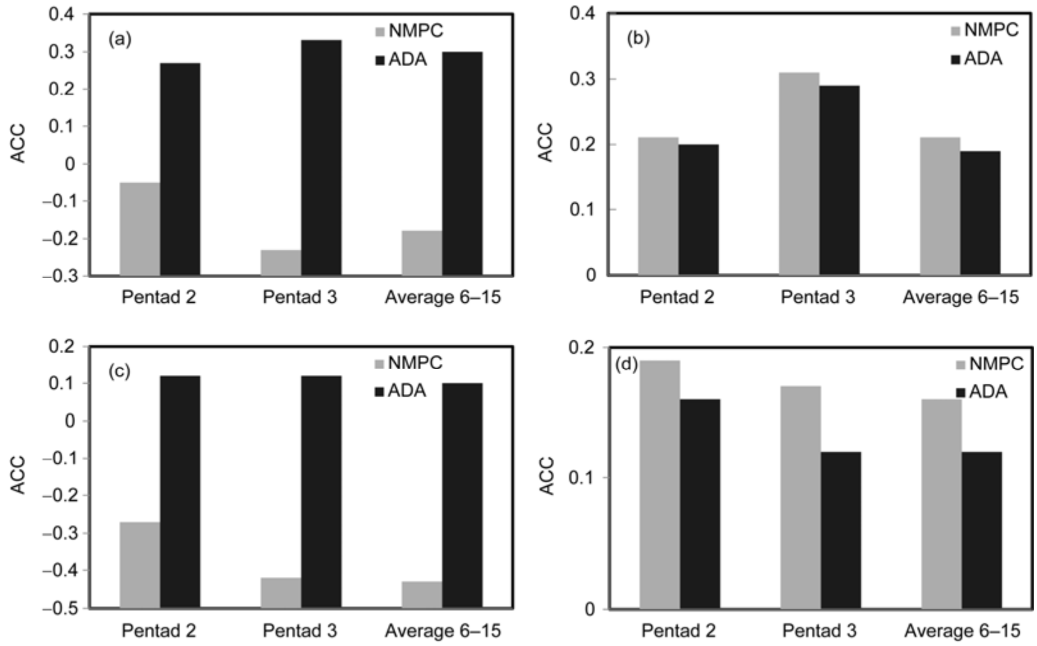

Figure 7 Same as in Figure 5, but for 850 hPa.

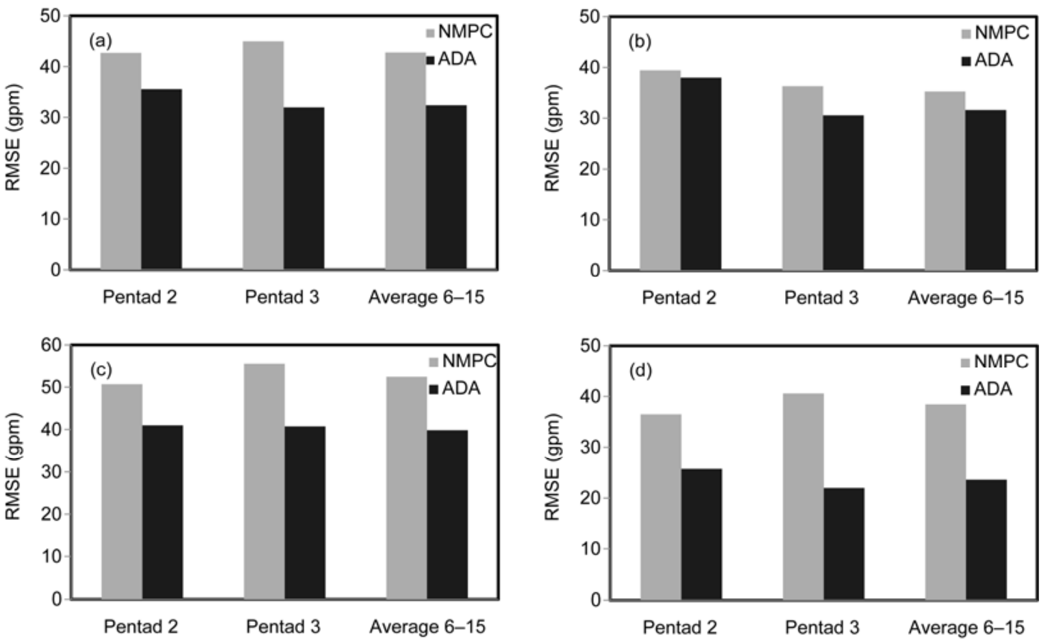

Figure 8 Same as in Figure 6, but for 850 hPa.

ferent scale wave components of 500 and 850 hPa geopotential heights, across various regions. Figure 9 shows that prediction improvement of the zonal wave zero at 500 hPa and 850 hPa using the ADA was most evident, with ACC increasing by 0.40 and 0.82, and RMSE decreasing by 6.70 and 11.52 gpm, respectively. Predictions of ultra-long and long waves are also improved, although to a lesser extent than those of the zonal wave zero.

Figures 5–8 show that prediction skill improvement of the NMPC by using the ADA has obvious regional characteristics. To better display this regional distribution, Figure 10 gives the MSSS of the 6th–15th day prediction of the NMPC using the ADA, compared to the NMPC alone. It is clear from this figure that the greater improvement is at the 850 hPa level, whether in terms of improvement magnitude or regional extent. At 850 hPa, improvement is most distinct in the tropics, followed by Southern Hemispheric mid-high latitudes, except for a low-value band near 30°S. Improve-

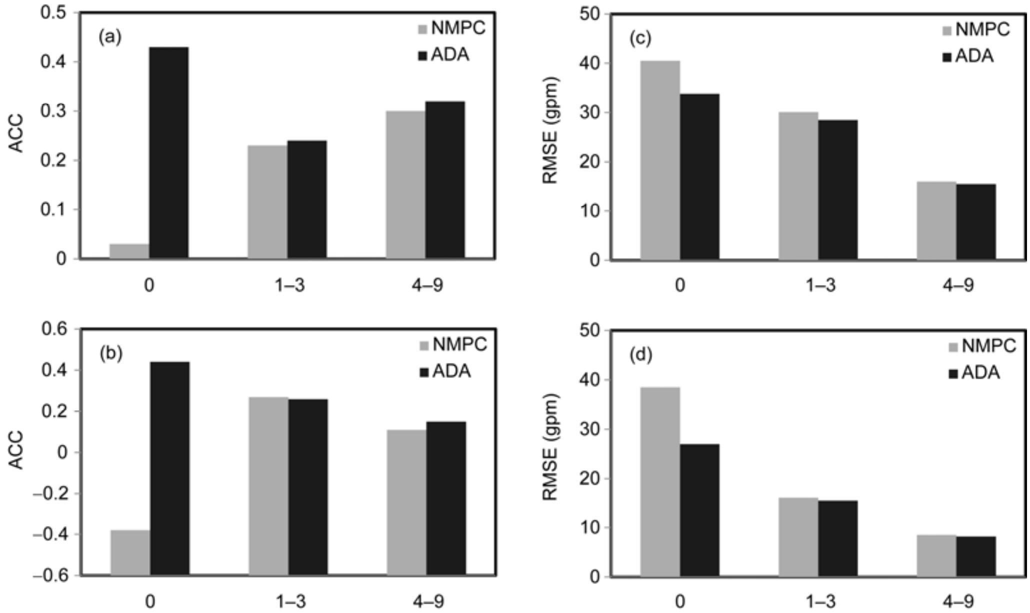

Figure 9 ACCs ((a), (b)) and RMSEs ((c), (d)) of the 6th–15th day mean forecast and the real observation of different scale waves (the abscissa: zonal wave number) of 500 hPa ((a), (c)) and 850 hPa ((b), (d)) height fields.

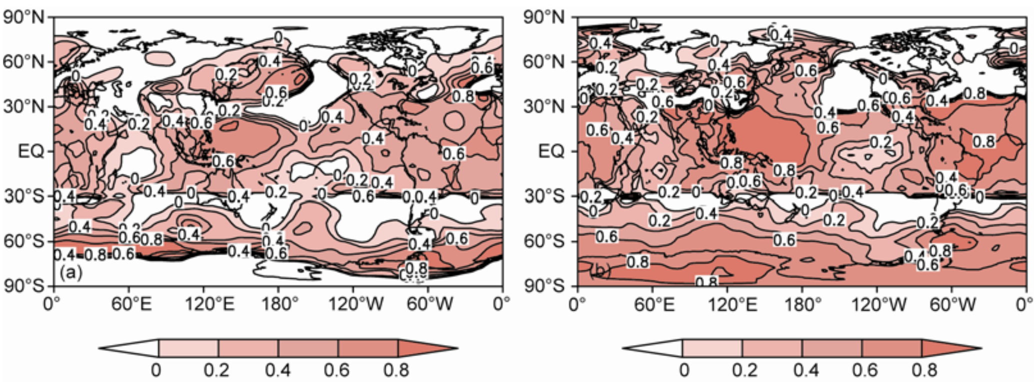

Figure 10 The 6th–15th day averaged MSSS of the 500 hPa (a) and 850 hPa (b) height fields predicted by the ADA compared to by the NMPC.

ment is slight at Northern Hemispheric mid-high latitudes, if viewed meridionally. If viewed latitudinally, it is generally better in the Eastern Hemisphere than in the Western Hemisphere. Compared with the 850 hPa level, skill improvement at 500 hPa is reduced in both improvement magnitude and regional extent, with almost no improvement at polar high latitudes. Likewise, there was a low value band of MSSS near 30°S. However, for Northern Hemispheric mid-high latitudes, the improvement was more evident in the Western Hemisphere than the Eastern Hemisphere.

In summary, prediction experiments of the NMPC using the ADA show that the latter was able to effectively improve model prediction skill. The selection of historical analogues and correction of errors in the experiments were all targeted for parts of model variables excluding small-scale components, thus avoiding the impact on prediction results of the rapid growth of prediction errors in small-scale components. Both procedures were conducted for each model variable at each level. Improvement of prediction ensued for various variables and levels, laying the foundation for operational application of the ADA. Of course, the error correction targeted for deterministic forecasting of PCs is closely related with their number. For the 6th–15th day forecast, since components of the synoptic scale are predictable, prediction error of PCs accounts for a sizeable portion of total prediction error. Reducing this part of the prediction error by use of statistical methods in combination with historical data can improve the forecast performance of the numerical model, as confirmed by the experimental results. The authors will introduce probability prediction of random components into the prediction process, to attain different prediction schemes and strategies targeted to PCs and random components. This will be discussed in future work. In addition, one can use the ADA to make forecasts targeted for a specific prediction variable (such as the 500 hPa geopotential height). That is, the historical analogue selection and error correction are all targeted for this variable. This is intended to increase the predictive skill of the model for that variable. This will be also discussed in future work.

4 Discussion and conclusions

The extended-range time scale has exceeded the upper limit of daily weather forecast predictability, but certain predictable meteorological field characteristics remain. How to predict these characters accurately is key to ERF. We developed an NMPC model by retaining PCs with higher predictability and filtering out random components from the model meteorological fields, to avoid the impact of rapid growth of prediction errors of small scale components on prediction results. We analyzed and discussed the two difficulties faced by the ADA, which uses historical data to improve the numerical model. We proposed methods to solve these difficulties, thereby developing the ADA further. We used the ADA to correct prediction errors of the NMPC, and to statistically assess the impact of random components on the predictable components. Results show that the ADA developed is able to effectively raise the prediction skill of PCs in the NMPC model. There is clear skill improvement for highly predictable areas if viewed from the spatial distribution, and for the zonal wave zero followed by ultralong waves and synoptic scale waves, if viewed from the spatiotemporal scale. Both the spatial distribution and spatiotemporal scale of skill improvement suggest that the ADA has the capacity to effectively reduce prediction errors of the NMPC, thereby raising its forecast skill and improving its prospects for operational application.

The success of ERF depends on two important factors: First, on a certain time scale, determining which characteristics of real atmospheric processes are predictable; second, based on an understanding of these predictable processes, how to predict these characteristics in a targeted manner. For the extended-range time scale, there are distinct differences in basic characteristics between PCs and random components. This merits further study, using different prediction schemes and strategies in the predictions. With continuous improvement, the numerical model will be more able to describe PCs on the time scale of the prediction period. Therefore, the more useful information the model can provide, the better the application prospects for the ADA using historical data.

This work was supported by National Natural Science Foundation of China (Grant Nos. 41105070, 40930952 and 41005041), State Key Program of Science and Technology of China (Grant No. 2009BAC51B04), and Meteorological Special Project of China (Grant No. GYHY 201106016).

1 Miyakoda K, Sirutis J, Ploshay J. One-month forecast experiments-without anomaly boundary forcings. Mon Weathe Rev, 1986, 114: 846–869
2 Palmer T N, Brankovic C, Molteni F, et al. The European centre for medium-range weather forecasts (ECMWF) program on extended-range prediction. Bull Amer Meteorol Soc, 1990, 71: 1317–1390
3 Wheeler M, Hendon H. An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. Mon Weathe Rev, 2004, 132: 1917–1932
4 Baldwin M P, Stephenson D B, Thompson D W J, et al. Stratospheric memory and skill of extended-range weather forecasts. Science, 2003, 301: 636–640
5 Li W J, Zhang P Q, Li Q Q, et al. Research and operational application of dynamical climate model prediction system (in Chinese). Chin J Appl Meteorol, 2005, 16: 1–11
6 Chen B M, Ji L R, Yang P C, et al. An approach to improving the dynamical extend-range (monthly) prediction. Chin Sci Bull, 2003, 48: 696–703
7 Vitart F. Monthly forecasting at ECMWF. Mon Weather Rev, 2004, 132: 2761–2779
8 Saha S, Nadiga S, Thiaw C, et al. The NCEP climate forecast system. J Clim, 2006, 19: 3483–3517
9 Chou J. Predictability of the atmosphere. Adv Atmos Sci, 1989, 6: 335–346
10 Danforth C M, Kalnay E, Miyoshi T. Estimating and correcting global weather model error. Mon Weather Rev, 2007, 135: 281–299
11 Chao J P, Ji J J, He J H, et al. Physical basis for a long-range numerical weather prediction method (in Chinese). Sci China Ser B, 1977, (2): 162–172
12 Chou J. Nonlinearity and Complexity of the Atmospheric Sciences (in Chinese). Beijing: China Meteorological Press, 2002, 157
13 Lorenz E N. On the existence of extended range predictability. J Appl Meteorol, 1973, 12: 543–546
14 Monin A S. Weather Forecasting as a Problem in Physics. Cambridge, Mass and London: MIT Press, 1972. 120
15 Sun G W, Xin F, Chen B M, et al. A predicting method on the low-frequency synoptic weather map (in Chinese). Plateau Meteorol, 2008, 27: 64–68
16 Li J P, Wang S H. Some mathematical and numerical issues in geophysical fluid dynamics and climate dynamics. Commun Comput phys, 2008, 3: 759–793
17 Mu M, Li J P, Chou J F, et al. Theoretical study on the predictability of climate system (in Chinese). Clim Environ Res, 2002, 7: 227–235
18 Feng G L, Dai X G, Wang A H, et al. On numerical predictability in the chaos system (in Chinese). Acta Phys Sin, 2001, 50: 606–611
19 Chen B H, Li J P, Ding R Q. Nonlinear local Lyapunov exponent and atmospheric predictability research. Sci China Ser D-Earth Sci, 2006, 49: 1111–1120
20 Li J P, Ding R Q. Temporal-spatial distribution of atmospheric predictability limit by local dynamical analogs. Mon Weather Rev, 2011, 139: 3265–3283
21 Chou J F, Zheng Z H, Sun S P. The think about 109–30 day extended-range numerical weather prediction strategy-facing the atmospheric chaos (in Chinese). Sci Meteorol Sin, 2010, 30: 569–573
22 Huang, J P, Chou J F. Studies on the analogous rhythm phenomenon in coupled ocean-atmosphere system (in Chinese). Sci China Ser B, 1990, 33: 851–860
23 Huang J P, Yi Y H, Wang S W, et al. An analogue-dynamical long-range numerical weather prediction system incorporating historical evolution. Q J R Meteorol Soc, 1993, 119: 547–565
24 Huang J P, Wang S W. Seasonal forecasting experiments using an analogue-dynamic model (in Chinese). Sci China Ser B, 1991, (2): 216–224
25 Bao M, Ni Y Q, Chou J F. The experiment of monthly mean circulation prediction using the analogy-dynamical model. Chin Sci Bull, 2004, 49: 1296–1300

26 Ren H., Chou J F, Huang J P, et al. Theoretical basis and application of an analogue-dynamical model in the lorenz system, Adv Atmos Sci, 2009, 26: 67–77

27 Ren H L, Chou J F. Strategy and methodology of dynamical analogue prediction. Sci China Ser D-Earth Sci, 2007, 50: 1589–1599

28 Li Z J, Ji L R. Analysis of the dependence of predictability on spatial and temporal scales from operational forecasts (in Chinese). Chin J Atmos Sci, 1996, 20: 290–297

29 Zheng Z H, Feng G L, Chou J F, et al. Compression for degrees of freedom in numerical prediction and the error analogy (in Chinese). Chin J Appl Meteor, 2010, 21: 139–148

30 Chen W Y, Van Den Dool H M. Atmospheric predictability of seasonal, annual, and decadal climate means and the role of the ENSO cycle: a model study. J Clim, 1997, 10: 1236–1254

31 Shukla J. Dynamical predictability of monthly means. J Atm Sci, 1981, 38: 2547–2572

32 Xu X G, Li W J, Ren H L, et al. Distributions of prediction capacity of T63L16 model for medium-range forecast at different spatial scales (in Chinese). Acta Meteorol Sin, 2009, 67: 992–1001

33 Zheng Z H, Ren H L, Huang J P. Analogue correction of errors based on seasonal climate predictable components and numerical experiments (in Chinese). Acta Phys Sin, 2009, 58: 7359–7367