ADVANCES IN ATMOSPHERIC SCIENCES, VOL. 30, NO. 5, 2013, 1249–1259

A Forecast Error Correction Method in Numerical Weather Prediction by Using Recent Multiple-time Evolution Data

XUE Hai-Le[1,2] (薛海乐), SHEN Xue-Shun*[2,3] (沈学顺), and CHOU Ji-Fan[1] (丑纪范)

[1] *School of Atmospheric Sciences, Lanzhou University, Lanzhou* 730000

[2] *State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing* 100081

[3] *Center for Numerical Prediction, China Meteorological Administration, Beijing* 100081

(Received 1 November 2012; revised 12 December 2012; accepted 10 January 2013)

ABSTRACT

The initial value error and the imperfect numerical model are usually considered as error sources of numerical weather prediction (NWP). By using past multi-time observations and model output, this study proposes a method to estimate imperfect numerical model error. This method can be inversely estimated through expressing the model error as a Lagrange interpolation polynomial, while the coefficients of polynomial are determined by past model performance. However, for practical application in the full NWP model, it is necessary to determine the following criteria: (1) the length of past data sufficient for estimation of the model errors, (2) a proper method of estimating the term "model integration with the exact solution" when solving the inverse problem, and (3) the extent to which this scheme is sensitive to the observational errors. In this study, such issues are resolved using a simple linear model, and an advection–diffusion model is applied to discuss the sensitivity of the method to an artificial error source. The results indicate that the forecast errors can be largely reduced using the proposed method if the proper length of past data is chosen. To address the three problems, it is determined that (1) a few data limited by the order of the corrector can be used, (2) trapezoidal approximation can be employed to estimate the "term" in this study; however, a more accurate method should be explored for an operational NWP model, and (3) the correction is sensitive to observational error.

Key words: numerical weather prediction, past data, model error, inverse problem

Citation: Xue, H.-L., X.-S. Shen, and J.-F. Chou, 2013: Forecast error correction method in numerical weather prediction by using recent multiple-time evolution data. *Adv. Atmos. Sci.*, **30**(5), 1249–1259, doi: 10.1007/s00376-013-2274-1.

1. Introduction

Numerical weather prediction (NWP) has become the key of weather forecasting since it was initiated by Richardson (1922) and was accurately determined with a simplified model by Charney et al. (1950). The development of computer technology has increased the capability of weather simulation, and forecasting by the NWP model has been improved dramatically. However, total avoidance of errors in numerical weather forecasting appears to be impossible due to errors in the initial conditions because of the observational error and imperfect data assimilation, in addition to model deficiencies due to model discretization and approximate physical processes. Efforts to improve assimilation, model dynamics, and sub-grid scale processes have lessened these two types of error. However, unresolved phenomena and model errors persist regardless of parameterization accuracy and grid resolution. Although it is reasonable to consider the NWP as an initial-value problem from the perspectives of mathematics and physics, huge amounts of past data may only be applied if the NWP is considered as an inverse problem. Therefore, a method of using such past observations in the NWP system is worth exploring. This study develops an online algorithm to correct model forecast errors by reasonable application of past data.

For the purpose of enhancing the availability of model forecasts, various approaches of forecast error

*Corresponding author: SHEN Xueshun, shenxs@cma.gov.cn

correction have been offered and can be classified as offline and online bias correction. With model output statistics (MOS), known as one popular offline correction method (Glahn and Lowry, 1972; Carter et al., 1989), statistical relationships between the model output and observations are calculated over a multiyear historical period. Because this method requires a lengthy training period for obtaining stable statistics, expensive computer resources are needed. Moreover, frequent updating of the operational models creates invalid statistics. To overcome these challenges, two additional methods have been presented on the basis of a time series shorter than one month for model output and observation. The first is the running-mean correction method (Eckel and Mass, 2005; Hacker and Rife, 2007), which assumes that the forecast error within one day can be linearly estimated using a time series of model output and observation with lags of 24 h. The other is based on the Kalman filter (KF), which is similar to the first method although recent errors are weighted more than past errors (Delle Monache et al., 2006; McCollor and Stull, 2008; Delle Monache et al., 2011). Limitations of this method are that sudden changes in forecast error caused by rapid transitions from one weather regime to another not likely predicted (Delle Monache et al., 2011), and it is an offline rather than online corrector. Offline bias correction has no dynamic effect on the forecast. The online correction is conducted at every time step of the model integration to retard error growth, and internal and external errors are permitted to interact nonlinearly throughout the integration. Nudging, or Newtonian relaxation, is a simple method used to reduce the systematic errors by adding artificial sources and sinks (24-h error divided by one day) as constant nudging terms to the model tendencies at each time step (Hoke and Anthes, 1976; Saha, 1992; DelSole et al., 2008). Leith (1978) proposed a statistical method in which model bias and systematic errors are assumed to be linearly dependent on flow anomalies. Although this method is simple and adjustable, it is valid only for linear models and is subject to sampling errors and expensive computation. By using a nonlinear quasi-geostrophic model, DelSole and Hou (1999) modified Leith's scheme and examined its availability for practical applications. They determined that the modified Leith scheme improved the forecasting skill to a large extent. In addition, Danforth et al. (2007) divided the 6-h total model error into model bias, periodic, and nonperiodic components. They further corrected these three types of errors by using various methods including nudging, obtained by time averaging the errors over several years; diurnal correction, based on the leading empirical orthogonal functions (EOFs) of the analysis; and the state-dependent correction, which was simplified from Leith's scheme, respectively. The effectiveness of this approach has been tested by two different types of models including quasi-geostrophic and primitive equation models.

The above methods are based on the traditional concept of NWP, which is considered as an initial-value and forward problem. That is, the forecasts of the traditional NWP depend on the initial value and the model. If the NWP is considered as an inverse problem, however, the past data including observations and previously determined model performance can be utilized. Hence, the NWP as an inverse problem can treat numerical weather forecasting both dynamically and statistically. Gu (1958) offered this concept and successfully applied it for reconstructing global three-dimensional data by using a quasi-geostrophic model and observations over land. At that time, no observations over oceans had been reported, and no inhabitants were known to be in the area. Chou (1974) expanded on Gu's (1958) concept, and further proposed that NWP could be considered as an inverse problem also because the past data should be utilized in conducting the NWP. Since that time, several studies have been conducted to establish the theoretical foundation NWP and to explore possible applications (Qiu and Chou, 1989; Huang and Wang, 1992; Cao, 1993; Gu, 1998). However, these studies focused on the development of a dispersion model (Huang and Wang, 1992) or a self-memorization model (Cao, 1993; Gu, 1998) to utilize past data; no research explored the application to operational NWP models. The basic concept of treating the NWP as the inverse problem is to estimate forecast error due to model imperfections by solving the inverse problem. Chou and his colleagues proposed two approaches for such estimation that include an analog scheme (Bao et al., 2004; Ren and Chou, 2005, 2006, 2007), and an extrapolation scheme (Da, 2011). The former utilizes analog information diagnosed from past observations and model outputs, while the latter is an optimization problem for constructing an object function by using the recent observations and model outputs as constraints to estimate forecast error. Compared with the forward NWP model, the extrapolation scheme could have a more precise solution because it continuously observes model trajectory, which was demonstrated by Chou (1974). Da (2011) proposed an algorithm that considered that the model errors due to model deficiencies can be expressed as a Lagrange interpolation polynomial, while the coefficients of this polynomial can be obtained by solving the above-mentioned optimization problem. Da's (2011) approach could have high potential in practical NWP because utilization of the short length of past data is possible in this algorithm. However, his study presented only a theoretical framework and did not investigate the details of practical application in operational NWP. Thus, three main problems remain. The first is that Da's (2011) work did not explain how to choose the length of past data according to various scales of errors, although this point is very important for practical application. The second is that Da (2011) gave an approximate approach to obtain the solution of the so-called "model integration with exact solution" in the inverse problem and did not discuss the possibility for full NWP model application. The third is that Da (2011) did not consider the possible effect of errors existent in the past data, which is inevitable for observation and assimilation.

Therefore, the following three factors must be determined for practical application in an operational NWP model: (1) a sufficient length of time series of past data for proper estimation of the model errors, (2) a method by which the term "model integration with the exact solution" is estimated when solving the inverse problem, and (3) the extent to which this approach is sensitive to the observational errors.

These issues are examined in the present study. In this paper, section 2 describes the approach, and section 3 gives idealized experiments for analyzing the three factors. On the basis of the results, we further examine the approach in section 4 by using a one-dimensional nonlinear advection–diffusion model. Section 5 presents our conclusions.

2. The approach

Generally, the NWP model can be written as the following initial problem:

$$\frac{\partial \boldsymbol{\psi}}{\partial t} = \boldsymbol{M}(\boldsymbol{\psi}) , \tag{1a}$$

$$\boldsymbol{\psi}(t)|_{t=0} = \boldsymbol{\psi}_0 , \tag{1b}$$

where $\boldsymbol{\psi}(t)$ is the vector form of state variables including wind, temperature, moisture, and pressure; $\boldsymbol{M}$ is the model operator; and Eq. (1b) is the initial condition. For brevity, the following discussion will focus on one state variable of the model equation at the space-discretized point:

$$\frac{\partial \boldsymbol{\psi}}{\partial t} = \boldsymbol{M}(\boldsymbol{\psi}) , \tag{2a}$$

$$\boldsymbol{\psi}(t)|_{t=0} = \boldsymbol{\psi}_0 . \tag{2b}$$

Usually, Eq. (2a) is an approximation to the actual atmosphere due to the discretization and model parameterizations. If we denote the error of model as $\varsigma(t)$ and assume n times of data ψ_{-1}, $\psi_{-2}, \ldots \psi_{-n}$ as exact solutions in the past with time interval δ, we have the following equations:

$$\frac{\partial \psi}{\partial t} = M(\psi) + \varsigma(t) , \tag{3a}$$

$$\psi(0) = \psi_0 , \tag{3b0}$$

$$\psi(-\delta) = \psi_{-1} , \tag{3b1}$$

$$\psi(-2\delta) = \psi_{-2} , \tag{3b2}$$

$$\cdots\cdots$$

$$\psi(-n\delta) = \psi_{-n} . \tag{3bn}$$

Here, ψ becomes the exact solution of the atmosphere rather than the prognostic variable of Eq. (2a).

Suppose that $H_{-\left(\frac{2k-1}{2}\right)}$ $(k = 0, 1, \ldots, n\ )$ is an error factor at the midpoint interval between $-k\delta$ and $-(k-1)\delta$. $\varsigma(t)$ can be expressed as a Lagrange interpolation polynomial from $-n\delta$ to δ according to Da's (2011):

$$\varsigma(t) = \sum_{k=0}^{n} l_{-\frac{2k-1}{2}}(t) H_{-\frac{2k-1}{2}} , \tag{4}$$

$$l_{-\frac{2k-1}{2}} = \prod_{i=0, i\neq k}^{n} \frac{t + \dfrac{2i-1}{2}\delta}{-\dfrac{2k-1}{2}\delta + \dfrac{2i-1}{2}\delta} \qquad k = 0, 1, 2, \ldots, n . \tag{5}$$

It is noted that $H_{\frac{1}{2}}$ is the error factor at the midpoint of the model forward integration period δ; therefore, the Lagrange interpolation polynomial can be used to correct forecast error at half or one "future" time interval. From Eqs. (4) and (5), $n+1$ are unknown factors $H_{\frac{1}{2}}, H_{-\frac{1}{2}}, \ldots, H_{-\left(\frac{2n-1}{2}\right)}$ in determining the coefficients of the Lagrange interpolation polynomial. The inverse problem constructed by Eq. (3a) and Eqs. (3b0)–(3bn) can be solved by integrating Eq. (3a) for each interval δ. Thus, n equations correspond to n intervals but $n+1$ unknown factors. Therefore, the inverse problem is underdetermined. An additional hypothesis is necessary for closing the problem and will be discussed subsequently.

To integrate Eq. (3a) from $-k\delta$ to $-(k-1)\delta (k = 0, 1, \ldots, n)$, the left-hand-side (lhs) of Eq. (3a) becomes $\psi_{-(k-1)} - \psi_{-k}$, the second term of right-hand-side (rhs) of Eq. (3a) is a linear combination of $H_{\frac{1}{2}}$, $H_{-\frac{1}{2}}$, $\ldots$, $H_{-\left(\frac{2n-1}{2}\right)}$, and the first term of rhs of Eq. (3a) can be written as $\int_{-k\delta}^{-(k-1)\delta} M(\psi)dt$. To recall the above assumption of the given past exact solutions in Eqs. (3b0)–(3bn), the term $\int_{-k\delta}^{-(k-1)\delta} M(\psi)dt$ is assumed to be obtained by integration of the model with the exact solution. By algebraic manipulation,

$H_{-\frac{1}{2}},\ldots,H_{-\left(\frac{2n-1}{2}\right)}$ can be expressed by the following vector formulation:

$$\boldsymbol{H}=\boldsymbol{\Lambda}^{-1}\boldsymbol{A}-H_{\frac{1}{2}}\boldsymbol{\Lambda}^{-1}\boldsymbol{\alpha}\ , \qquad (6)$$

where

$$\boldsymbol{\alpha}=\left(\int_{-\delta}^{0}l_{\frac{1}{2}}(t)dt\int_{-2\delta}^{-\delta}l_{\frac{1}{2}}(t)dt\ldots\int_{-n\delta}^{-(n-1)\delta}l_{\frac{1}{2}}(t)dt\right)^{\mathrm{T}},$$

$$\boldsymbol{H}=\left(H_{-\frac{1}{2}}H_{-\frac{3}{2}}\ldots H_{-\frac{2n-1}{2}}\right)^{\mathrm{T}},\boldsymbol{A}=(A_1A_2\ldots A_n)^{\mathrm{T}}\ ,$$

$$A_k=\psi_{-(k-1)}-\psi_{-k}+\int_{-k\delta}^{-(k-1)\delta}M(\psi)dt, k=1,\ldots,n\ ,$$

$$\boldsymbol{\Lambda}=\begin{pmatrix}\int_{-\delta}^{0}l_{-\frac{1}{2}}(t)dt & \int_{-\delta}^{0}l_{-\frac{3}{2}}(t)dt & \cdots & \int_{-\delta}^{0}l_{-\frac{2n-1}{2}}(t)dt\\ \int_{-2\delta}^{-\delta}l_{-\frac{1}{2}}(t)dt & \int_{-2\delta}^{-\delta}l_{-\frac{3}{2}}(t)dt & \cdots & \int_{-2\delta}^{-\delta}l_{-\frac{2n-1}{2}}(t)dt\\ \cdots & \cdots & \cdots & \cdots\\ \int_{-n\delta}^{-(n-1)\delta}l_{-\frac{1}{2}}(t)dt & \int_{-n\delta}^{-(n-1)\delta}l_{-\frac{3}{2}}(t)dt & \cdots & \int_{-n\delta}^{-(n-1)\delta}l_{-\frac{2n-1}{2}}(t)dt\end{pmatrix}.$$

There is only one unknown factor, $H_{\frac{1}{2}}$, in Eq. (6). Da (2011) defined two types of norms of $\varsigma(t)$ for the purpose of obtaining $H_{\frac{1}{2}}$. Here, we adopt one of the definitions:

$$J(H_{\frac{1}{2}})=\sqrt{\int_{-n\delta}^{\delta}[\varsigma(t)]^2}.$$

It is easy to solve $H_{\frac{1}{2}}$ by minimizing the norm. The above approach indicates that by giving the minimized modification of the model by introducing a correction term, the model forecast can be kept accurate for the past given accurate solutions, and the forward forecast can be improved by minimizing the norm $J(H_{\frac{1}{2}})$. This approach is equivalent to the original scheme proposed by Chou, which has been theoretically proved as effective for improving the model forecast (Chou, 1974).

3. Discussions on practical application

For the purpose of implementing the above approach into an operational NWP model, three problems should be investigated and resolved, as mentioned in the Introduction. Moreover, the computing cost must be evaluated when implementing the scheme into operation. In this section, a linear model is designed to investigate these issues.

3.1 *Linear model and experiments*

A simple model is first designed to analyze these questions:

$$\frac{d\psi}{dt}=C_1\sin(\omega_1t+\varphi_1)+C_2\cos(\omega_2t+\varphi_2)\ , \qquad (7)$$

where ψ is a function of time; C_1 and C_2 are amplitudes; ω_1 and ω_2 are frequencies; and φ_1 and φ_2 are initial phases. Eq. (7) is assumed to be an accurate model, while the following is assumed to be an imperfect model due to the disregarding of the second term of Eq. (7):

$$\frac{d\psi}{dt}=C_1\sin(\omega_1t+\varphi_1)\ . \qquad (8)$$

Generally, the error source term should be very small. Here, the amplitude of this error term is set as $C_1\gg C_2$. The following equation integrates Eq. (7):

$$\psi(t)-\psi(t_{00})=-\frac{C_1}{\omega_1}\cos(\omega_1t+\varphi_1)+\frac{C_2}{\omega_2}\sin(\omega_2t+\varphi_2)\ . \qquad (9)$$

With initial time $t_{00}=-n\delta$ and initial condition $\psi(-n\delta)=\psi_{-n}$, we can obtain the exact values ψ_0, $\psi_{-1},\ldots\psi_{-n}$ for correction and the "future" data for verification according Eq. (9) . Suppose $A=C_1/\omega_1$, $B=C_2/\omega_2$, and $A\gg B$. Integrating Eq. (7) from $-k\delta$ to $-(k-1)\delta$ results in

$$\begin{aligned}\psi_{-k\delta}-\psi_{-(k-1)\delta}=&-A\{\cos[\omega_1(-k\delta)+\varphi_1]-\\&\cos[\omega_1(-k+1)\delta+\varphi_1]\}+\\&B\{\sin[\omega_2(-k\delta)+\varphi_2]-\\&\sin[\omega_2(-k+1)\delta+\varphi_2]\}\ , \qquad (10)\end{aligned}$$

where the first term of rhs, the so-called "integration of the model with exact solution" , is denoted as $\boldsymbol{\Psi}_k=-A\{\cos[\omega_1(-k\delta)+\varphi_1]-\cos[\omega_1(-k+1)\delta+\varphi_1]\}$. Because this model is simple and analytical, the exact solution of $\boldsymbol{\Psi}_k$ is easily obtained. By using this idealized model, the approach given in section 2 can be easily tested to discuss the issues.

The most time-consuming part of the scheme is the computation of the Lagrange interpolation polynomial at each time step and model grid point. Hence, an

Table 1. Test descriptions. ($n = 2, 5, 10, 20$; $m = 1, 2, 4, 8, 16$).

Test set	Test name	Tests description
1	NO_m	Period of error source $T = m\delta$ for past data without observation error
	Tn_m	Period of error source $T = m\delta$ for n-time past data without observation error and series of $\boldsymbol{\Psi}_k$ of true value
2	Kn_m	Period of error source $T = m\delta$ for n-time past data without observation error and series of $\boldsymbol{\Psi}_k$ computed by trapezoidal integration method
3	KEn_m	Period of error source $T = m\delta$ for n-time past data with observation error and series of $\boldsymbol{\Psi}_k$ computed by trapezoidal integration method

effective algorithm is crucial for controlling expenses. Wu and He (2007) expanded a product of n-linear factors as a simple n-order polynomial:

$$\prod_{i=1}^{n}(x+a_i) = D_0x^n + D_1x^{n-1} + \cdots + D_nx^0 , \quad (11)$$

where

$$D_0 = 1,$$

$$D_1 = \sum_{j=1}^{n} a_j,$$

$$D_k = \sum_{j=1}^{n-k+1} a_j \left(\sum_{i=j+1}^{n-i-2} D_{k-1,i}\right), k = 2, 3, \ldots n,$$

and $D_{k-1,i}$ is the ith term of D_{k-1}. Let

$$x = t, a_1 = -\frac{1}{2}\delta, \ldots a_j = \frac{2(j-1)-1}{2}\delta,$$

$$a_{j+1} = \frac{2(j+1)-1}{2}\delta, \ldots a_n = \frac{2n-1}{2}\delta,$$

and

$$K_k = \prod_{i=0, i\neq k}^{n} \frac{1}{-\frac{2k-1}{2}\delta + \frac{2i-1}{2}\delta};$$

Eq. (5) becomes

$$l_{-\frac{2k-1}{2}} = K_k(D_{k,0}t^n + D_{k,1}t^{n-1} + \cdots + D_{k,n}t^0) . \quad (12)$$

With Eq. (12), Eq. (4) becomes,

$$\varsigma(t) = \sum_{k=0}^{n}\left[\sum_{i=0}^{n} K_i D_{i,0} H_{-\frac{2(i-1)}{2}}\right] t^{n-k} . \quad (13)$$

Compared with (4) and (5), it is obvious that (12) and (13) are easier for computing Lagrange operators and the Lagrange polynomial. Once the polynomial coefficients are determined, its integration can be easily obtained. To add error term (13) at rhs of (8), the imperfect model with correction is

$$\frac{d\psi}{dt} = C_1 \sin(\omega_1 t + \varphi_1) + \sum_{k=0}^{n}\left[\sum_{i=0}^{n} K_i D_{i,0} H_{-\frac{2(i-1)}{2}}\right] t^{n-k} . \quad (14)$$

To investigate the effects of various time scale model errors and various lengths of time series of past data, three idealized experiments are designed as shown in Table 1. Regarding to the length of past data sufficient for estimation of the model errors, the first experiment examines the ideal or optimal effect of correction by using all exact past data and series of $\boldsymbol{\Psi}_k$. In the second experiment, the term $\boldsymbol{\Psi}_k$, i.e., the so-called "integration of the model with exact solution" is estimated by trapezoidal rather than analytical integration for the purpose of investigating the proper method of computing this term for practical application. The third experiment focuses on the impacts of observational errors on the effectiveness of the method. Here, the parameters are given as δ=21 600 s, A=20, B=0.5, $\varphi_1 = \varphi_2 = 0$, $\omega_1 = 2\pi/(20\times24\times3600)$. With the initial condition $\psi(-20\delta) = 2$ and related ω_2 in Table 1, the past "observations" and "future" exact solution used for verification can be obtained from Eq. (9). With the initial condition $\psi(0) = \psi_0$, the forecast, i.e., "future" data, by the imperfect model can be obtained from (8), and the corrected forecast by using the approach presented in this paper can be obtained from Eq. (14). Because it is impossible to determine the analytical expression of the NWP model error sources, an approximation of $\boldsymbol{\Psi}_k$ is used. The tendencies of points $-k\delta$ and $-(k-1)\delta$ can be obtained from Eq. (8) by integrating one step, then the approximate $\boldsymbol{\Psi}_k$ can be computed with the tendencies of the two points by the trapezoidal integration method. In the second and third experiments, the trapezoidal integration method is explored to examine the application for future implementation into an NWP model.

3.2 *Results and discussions*

In experiment 1, we assumed the error as having different periods $T = m\delta$ to mimic the various types of NWP model errors. Then, different lengths of past data were applied to investigate the proper lengths of past data in correcting the forecast errors with various periods. Figure 1 gives non-dimensional analytical and simulated errors with different periods. As shown in Figs. 1a and b, errors with a time scale of 1δ or 2δ

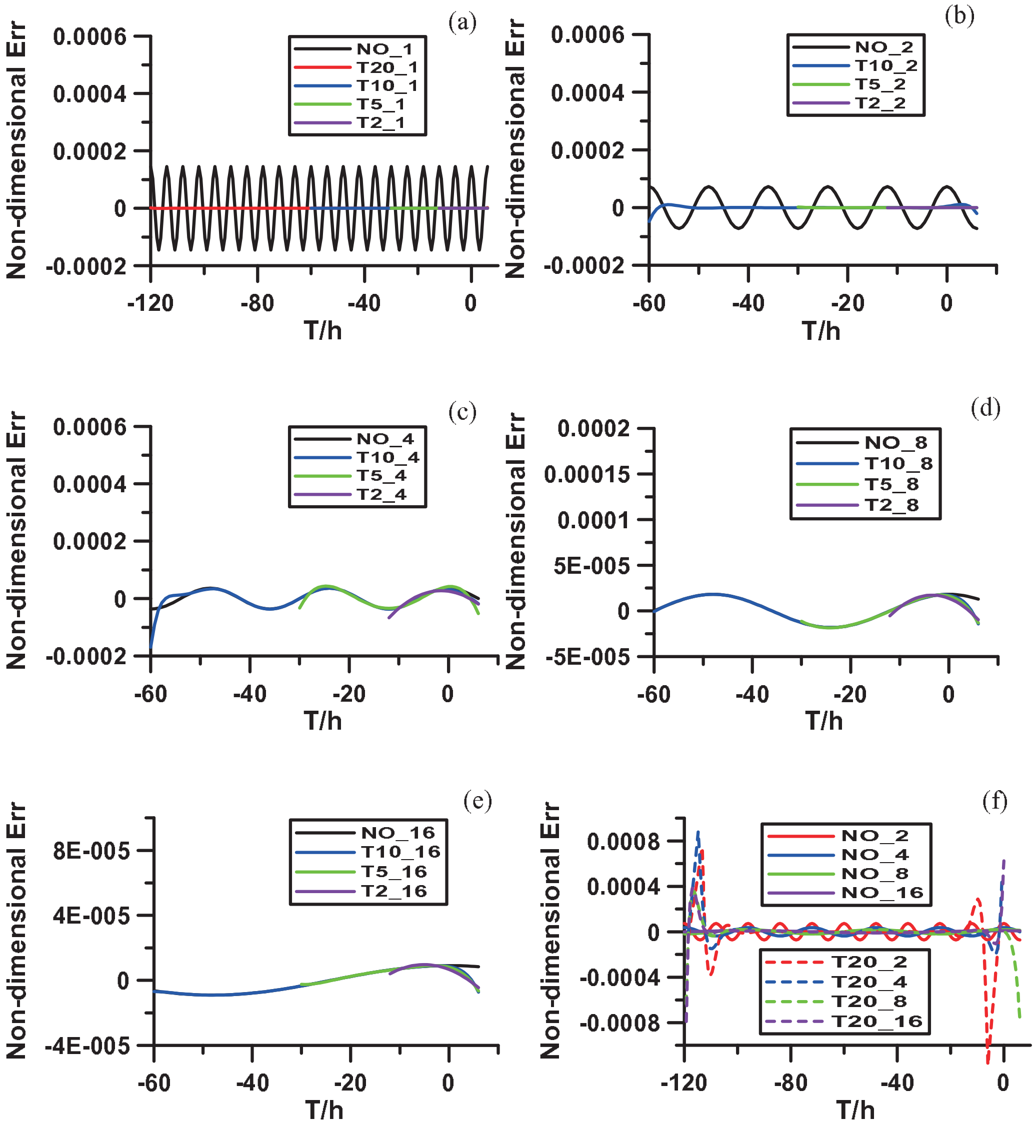

Fig. 1. Performance of error source simulation by various ranks of correctors for the following error source scales: (a) $T = 1\delta$, (b) $T = 2\delta$, (c) $T = 4\delta$, (d) $T = 8\delta$, (e) $T = 16\delta$, and (f) $T = 20\delta$. The legends are described in Table 1.

cannot be simulated using the polynomial regardless of how many times past data were used because the corrector we used could not resolve the high-frequency error factors. This result indicates that the high-frequency errors in NWP may not be corrected using the method presented in this paper. Moreover, Fig. 1f indicates that a 20-order corrector creates unstable results simulated both in the beginning and last intervals of entire simulation period, which implies that a high-order corrector or excessively long past data may not be appropriate for correcting the model forecast errors when this approach is used. Table 2 summarizes the results with invalid applications, which are denoted as $\times$. Except for the above two cases, as shown in Figs. 1c, d, and e, 2-order, 5-order or 10-order polynomial correctors can effectively simulate the error functions along nearly the entire period with error time scales of 4δ, 8δ, or 16δ. Deviations can also be found at the end points corresponding to the interval between the last point of past data and the first point of "future" data. Nevertheless, it can be concluded that these correctors are highly effective at least at half intervals. Table 2 summarizes the valid applications, which are denoted as $\surd$. Overall, the above results indicate that low-order correctors are suitable for low-frequency error sources and are unreliable for high-frequency error sources. This assumption is reasonable because the interval of the past data sampling is limited, and this

Table 2. Validity of correctors for various time scales of error.

T	2-oder	5-order	10-order	20-order
1δ	×	×	×	×
2δ	×	×	×	×
4δ	√	√	√	×
8δ	√	√	√	×
16δ	√	√	√	×

period determines the temporal resolution for resolving the high-frequency errors. The smallest period of error that can be easily corrected is approximately 4δ. Moreover, the vast amount of past data may not improve results because the order of the polynomial depends on the sampling number of past data. Thus, an excessively high-order corrector, such as an excessive amount of past data, would cause overshooting or undershooting of error estimation by using polynomials. Figure 2 presents the 6-h averaged error corresponding to various periods of errors. It is obvious that 2-order, 5-order, and 10-order correctors effectively reduce forecast errors if the periods of error sources are larger than 4δ.

In experiment 1, the term $\boldsymbol{\Psi}_k(k = 1, \cdots, n)$, "model integration with exact solution", was analytically determined. However, this term does not practically apply to the operational NWP model. In this paper, we recommend the trapezoidal integration method for estimating this term. The second experiment was designed to test the validity of this estimation, which is similar to experiment 1 except for the $\boldsymbol{\Psi}_k$ computation. Assuming errors with periods of 4δ, 8δ, and 16δ, Fig. 3 gives the 0–8 h forecast errors. It was determined that the 0–7 h forecast errors can be

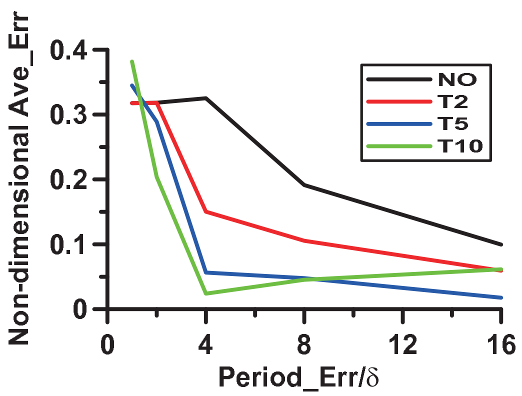

Fig. 2. 6-h average forecast error. Low-frequency error sources ($T \geqslant 4\delta$) can be corrected by low-rank correctors. The legends are described in Table 1.

Table 3. Corrector coefficients of sensitivity to trapezoidal integration and observation error.

Coef.	2-order	5-order	10-order
$C2$	0.4	10.8	172.0
$C1$	11.9	48.0	1329.1
$C0$	6.8	1.4	95.2

reduced using the 2-order, 5-order, and 10-order correctors for these three types of errors. Moreover, higher order correctors gave better results for long-period error sources. The results of this experiment indicate that the use of trapezoidal integration to estimate $\boldsymbol{\Psi}_k$ does not significantly influence the error correction.

The above experiments assumed that no errors were in the past data. In practical applications, the past data can only be given using data assimilation or re-analysis; errors in the past data are inevitable. To test the influence of such errors on the correction, random errors were added to the past data. Here, the amplitudes of random errors were set to approximately 0.1%–1.0% of the past data. This experiment is similar to experiment 2 except for the addition of these random errors. As is evident in Fig. 4, the low-order correctors such as 2-order and 5-order, gave nearly the same results as those without random errors shown in Fig 3. However, the conditions change with the application of high-order correctors. Here, the 10-order corrector appeared to be highly sensitive to the errors existent in the past data. Table 3 shows the most sensitive changes of polynomial coefficients after the random errors were added to the past data. Obviously, higher orders correspond to increased changes in polynomial coefficients; those of the 10-order corrector showed the sharpest changes in this experiment. These results suggest that attention should be paid when using higher-order correctors in practical applications.

4. Application to a nonlinear model

On the basis of the linear model, a forecast error correction method and the issues related to its practical applications in NWP are discussed in above section. It was determined that low-order correctors such as 2-order or 5-order effectively reduced the forecast errors with low frequency. However, operational NWP models are the highly nonlinear systems. Therefore, it is necessary to test the method and confirm the conclusions drawn in section 3 by using a nonlinear model before implementation into the full NWP model. A nonlinear advection–diffusion equation is employed here

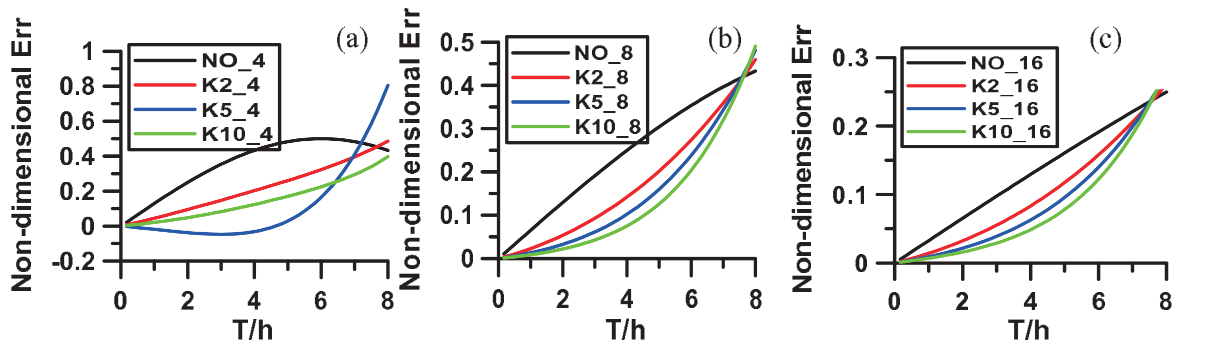

Fig. 3. Forecast error corrected by using trapezoidal integration for the following time scales: (a) $T = 4\delta$, (b) $T = 8\delta$, and (c) $T = 16\delta$. The legends are described in Table 1.

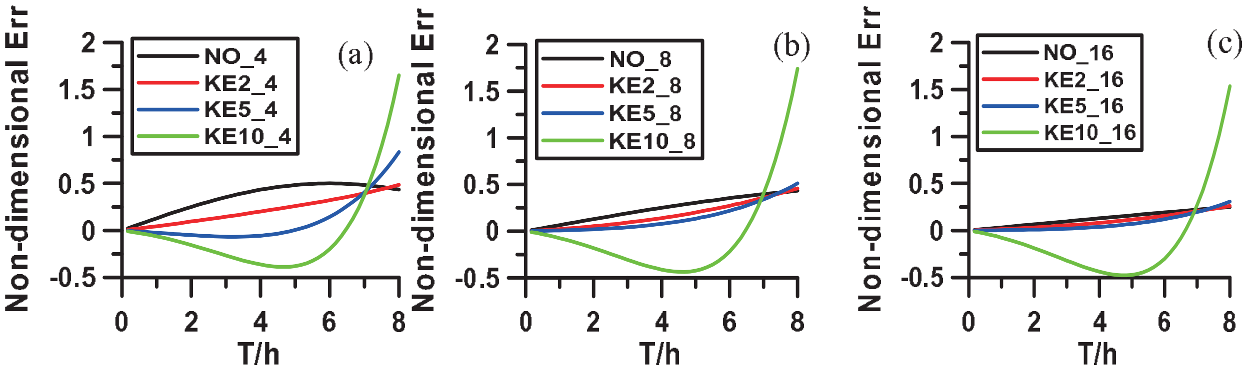

Fig. 4. Forecast error corrected by with observation error for the following time scales: (a) $T = 4\delta$, (b) $T = 8\delta$, and (c) $T = 16\delta$. The legends are described in Table 1.

as

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} - \sigma\frac{\partial^2 u}{\partial x^2} = E(x,t)\,, \tag{15}$$

where $E(x,t)$ is the error term; $E(x,t) = 0$, Eq.(15) is assumed to be an accurate model. $E(x,t)$ is defined as $E(x,t) = A\cos(\omega t - 2\pi x/L)$ to consider the errors with various frequencies. The diffusion coefficient, error frequency, and amplitude are given as $\sigma = 600\ 000$, $\omega = 2\pi/(16\times 21\ 600)$, and $A = 2.0E-5$, respectively. The initial value is given as $u(x,0) = 1+\cos(2\pi x/L)$(units: m s^{-1}), and the periodic boundary condition is utilized. Here, the time step is taken as $dt = 600$ s, the grid interval $dx = 60\ 000$ m, the number of grid points $nx = 21$, and $L = (nx-1)dx$.

Figure 5 presents the time evolution of averaged u as well as the absolute error. It is obvious that when using the 2-order and 5-order correctors, the predicted u was much closer to the analytical solution than that when no correction was used. Moreover, the 2-order corrector gave the better results than the 5-order corrector (Figs. 5b and b). Figure 6 shows a comparison between the true error $E(x,t) = A\cos(\omega t - 2\pi x/L)$ and those estimated by 2-order and 5-order correctors. It is interesting to note that both correctors effectively identified the true error pattern and its evolution prior to the 20 time step ($\delta/2$), as shown in Figs. 6a and b, respectively. This result implies that forecast error can be reduced to a large extent by adding the correctors for a certain time interval, i.e., ($\delta/2$). Moreover, the 2-order corrector gave error amplitudes much close to the true values. The 5-order corrector tended to overestimate the errors, which may be related to the sensitivity of higher-order scheme to the trapezoidal integration approximation in the nonlinear model. This result indicates that the order of corrector may be limited by the nonlinear processes and should thus be considered when applying to operational NWP models. Moreover, the estimated errors after the 20 time step, i.e., ($\delta/2$), changed signs, which indicates that the correctors are valid only for the forward time interval ($\delta/2$).

Although we give only the results concerning low-frequency errors in this paper, various scales of errors have also been tested using this nonlinear model (figures not shown) with similar results to those obtained in the linear model experiment. In summary, the 2-order corrector can reasonably identify errors with low

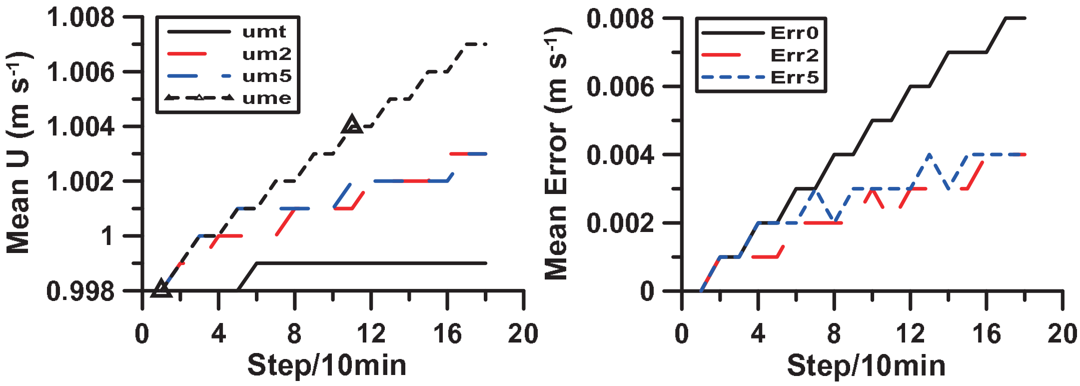

Fig. 5. (a) Mean value of u and (b) its absolute error. umt and ume represent the analytical and predicted solutions of u, respectively. um2 and um5 are mean of predicted u corrected by 2-order and 5-order correctors, respectively. Err0, Err2, and Err5 are mean of absolute error of u uncorrected and that corrected by 2-order and 5-order correctors, respectively.

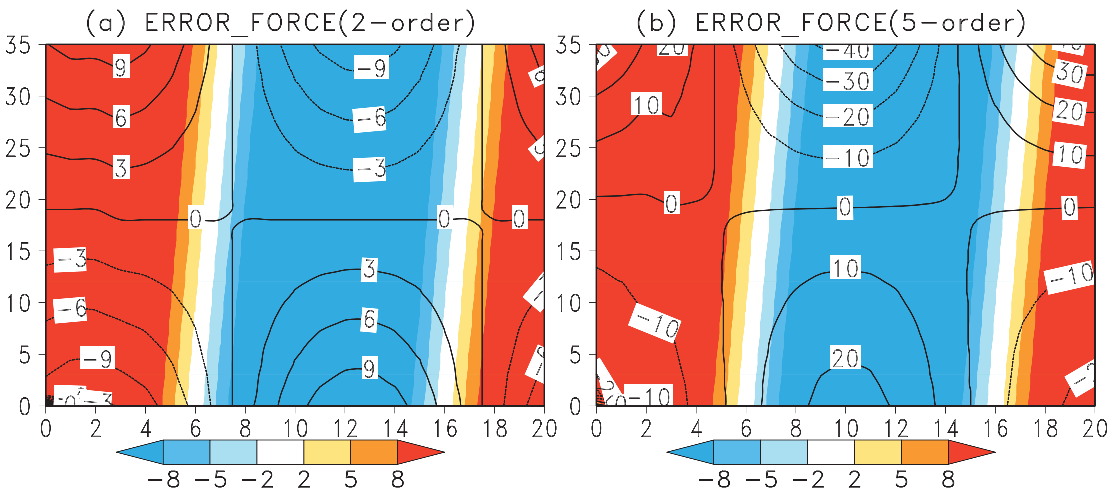

Fig. 6. The true error source ($E\times10^6$) (shading) and simulated error source ($E\times10^6$) (contours). The horizontal and vertical axes represent spatial grids and time step, respectively. (a) Forecast error by 2-order corrector; (b) forecast error by 5-order corrector.

frequency and can effectively reduce the forecast error for a forward time interval of ($\delta/2$). This conclusion has been confirmed by linear and nonlinear model experiments. Generally, this approach can be easily tested using other forms of nonlinear equations; however, it is believed that the same fundamental conclusions would be drawn.

5. Conclusions and discussions

This paper investigated an online approach for correcting NWP forecast errors. The foundation of this approach was to consider the NWP as an inverse problem. By using the past observations and model outputs, the model forecast errors were assumed to be a Lagrange polynomial form in which the coefficients were derived by solving the inverse problem. Theoretically, this approach should be easily implemented into an NWP model with flexible time intervals of past data. For practical application, however, the following key issues required confirmation: (1) the sufficient length of time series of past data for proper estimation of the model errors, (2) a method by which the term “model integration with the exact solution” is estimated when solving the inverse problem, and (3) the extent to which this approach is sensitive to the observational errors.

This paper examined these issues by designing two idealized model experiments that used simple linear and nonlinear models. The main conclusions are summarized in the following points:

(1) High-frequency errors cannot be treated by this approach due to the limited time resolution of past data. The low-frequency errors can be effectively reduced for a certain forward time interval determined by the time resolution of past data.

(2) A few data limited by the order of corrector can be used; an excessive amount of past data may not be valid in correcting the forecast errors. The proper length of past data depends on the order of correctors.

(3) To solve the inverse problem, the term "model integration with exact solution" must be determined. The trapezoidal integration method is proposed and is proved effective in this study. However, further consideration should be given in the operational NWP models because the initialization could lead to problems in obtaining the tendencies of the first model-integrated time steps.

(4) Although the correction is sensitive to observational error, random errors existent in the past data do not significantly affect the efficiency of the low-order correctors in this study.

The above conclusions were determined on the basis of a simple one-point sinusoidal model and an advection–diffusion equation. Because the operational NWP models are more complicated, however, other factors should be considered. For example, although random error and trapezoidal approximation did not significantly affect low-order correctors of approximately 2, the random errors caused by assimilation or by trapezoidal approximation may be quite larger than those presented in this paper. Therefore, such issues should be resolved for the future full NWP model application. Moreover, determining the manner in which high-frequency forecast errors are filtered when using a very low order corrector presented in this paper remains a challenging problem. Its full model application will be studied in future work.

Acknowledgements. This work was funded by the Special Scientific Research Project for Public Interest (GYHY201206009), the National Key Technologies Research and Development Program (Grant No. 2012BAC22B02), the National Natural Science Foundation Science Fund for Creative Research Groups (Grant No.41221064), the Special Scientific Research Project for Public Interest (Grant No. GYHY201006013), and the National Natural Science Foundation of China (Grant No. 41105070). We especially thank DA Chaojiu and YU Haipeng for offering valuable comments to this study.

REFERENCES

Bao, M., Y. Q. Ni, and J. F. Chou, 2004: The experiment of monthly mean circulation prediction using the analogy-dynamical model. *Chinese Science Bulletin*, **49**(12), 1296–1300. (in Chinese)

Carter, R. G., J. Dallavalle, and H. Glahn, 1989: Statistical forecasts based on the National Meteorological Center's numerical weather prediction system. *Wea. Forecasting*, **4**, 401–412.

Cao, H. X., 1993: Self-memorization equation in atmospheric motion. *Science in China* (*B*), **36**(7), 845–855. (in Chinese)

Charney, J. G., R. Fjortoft, and J. von-Neumann, 1950: Numerical integration of the barotropic vorticity equation. *Tellus*, **2**, 237–254.

Chou, J. F., 1974: A problem of using past data in numerical weather forecasting. *Scientia Sinica*, **17**(6), 814–825. (in Chinese)

Da, C. J., 2011: One scheme which maybe improve the forecasting ability of the global (regional) assimilation and prediction system. Ph.D. dissertation, School of Atmospheric Sciences, Lanzhou University, 100pp. (in Chinese)

Danforth, C. M., E. Kalnay, and T. Miyoshi, 2007: Estimating and correcting global weather model error. *Mon. Wea. Rev.*, **135**, 281–299.

Delle Monache, L., T. Nipen, X. Deng, Y. Zhou, and R. B. Stull, 2006: Ozone ensemble forecasts: 2. A Kalman-filter predictor bias correction. *J. Geophys. Res.*, **111**, D05308, 1–15.

Delle Monache, L., T. Nipen, Y. B. Liu, G. Roux, and R. Stull, 2011: Kalman filter and analog schemes to postprocess numerical weather predictions. *Mon. Wea. Rev.*, **139**, 3554–3570.

DelSole, T., and A. Y. Hou, 1999: Empirical correction of a dynamical model. Part I: Fundamental issues. *Mon. Wea. Rev.*, **127**, 2533–2545.

DelSole, T., M. Zhao, P. A. Dirmeyer, and B. P. Kirtman, 2008: Empirical correction of a coupled land–atmosphere model. *Amer. Meteor. Soc.*, **136**, 4063–4076.

Eckel, F. A., and C. F. Mass, 2005: Aspects of effective mesoscale,short-range ensemble forecasting. *Wea. Forecasting*, **20**, 328–350.

Glahn, H., and D. Lowry, 1972: The use of model output statistics in objective weather forecasting. *J. Appl. Meteor.*, **11**, 1203–1211.

Gu, X. Q., 1998: A spectral model based on atmospheric self-memorization principle. *Chinese Science Bulletin*, **43**(20), 1692–1702. (in Chinese)

Gu, Z. C., 1958: The use of past dat a in numerical weather forecast. *Acta Meteorologica Sinica*, **29**(3), 176–184. (in Chinese)

Hacker, J., and D. Rife, 2007: A practical approach to sequential estimation of systematic error on near-surface mesoscale grids. *Wea. Forecasting*, **22**, 1257–1273.

Huang, J. P., and S. W. Wang, 1992: The experiments of seasonal prediction using the analogy-dynamical model. *Acta Meteorologica Sinica* (*B*), **35**(2), 207–216.

Hoke, J. E., and R. A. Anthes, 1976: The initialization of numerical models by a dynamic initialization technique. *Mon. Wea. Rev.*, **104**, 1551–1556.

Richardson, L. F., 1922: *Weather Prediction by Numerical Process*. Cambridge University Press, 250pp.

Leith, C. E., 1978: Objective methods for weather prediction. *Annual Review of Fluid Mechanics*, **10**, 107–128.

McCollor, D., and R. Stull, 2008: Hydrometeorological accuracy enhancement via post-processing of numerical weather forecasts in complex terrain. *Wea. Forecasting*, **23**, 131–144.

Qiu, C. J., and J. F. Chou, 1989: An analogue-dynamical method of weather forecasting. *Scientia Atmospherica Sinica*, **13**(1), 22–28. (in Chinese)

Ren, H. L., and J. F. Chou, 2005: Analogue correction method of errors by combining both statistical and dynamical methods together. *Acta Meteorologica Sinica*, **63**(6), 988–993. (in Chinese)

Ren, H. L., and J. F. Chou, 2006: Introducing the updating of multi-reference states into dynamical analogue prediction. *Acta Meteorologica Sinica*, **63**(3), 315–324. (in Chinese)

Ren, H. L., and J. F. Chou, 2007: Strategy and methodology of dynamical analogue prediction. *Science in China* (*D*), **50**(10), 1589–1599.

Saha, S., 1992: Response of the NMC MRF model to systematic error correction within integration. *Mon. Wea. Rev.*, **120**, 345–360.

Wu, Y. X., and N. He, 2007: The entirey expanding of Lagrange's interpolation formula. *Journal of Tonghua Teachers College*, **28**(2), 10–12. (in Chinese)