ADVANCES IN ATMOSPHERIC SCIENCES, VOL. 32, NOVEMBER 2015, 1493–1503

An Online Model Correction Method Based on an Inverse Problem: Part II—Systematic Model Error Correction

XUE Haile*[1,3], SHEN Xueshun[1,2], and CHOU Jifan[3]

[1]*State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing* 100081
[2]*Center for Numerical Prediction, China Meteorological Administration, Beijing* 100081
[3]*School of Atmospheric Sciences, Lanzhou University, Lanzhou* 730000

(Received 6 December 2014; revised 8 April 2015, accepted 5 May 2015)

ABSTRACT

An online systematic error correction is presented and examined as a technique to improve the accuracy of real-time numerical weather prediction, based on the dataset of model errors (MEs) in past intervals. Given the analyses, the ME in each interval (6 h) between two analyses can be iteratively obtained by introducing an unknown tendency term into the prediction equation, shown in Part I of this two-paper series. In this part, after analyzing the 5-year (2001–2005) GRAPES-GFS (Global Forecast System of the Global and Regional Assimilation and Prediction System) error patterns and evolution, a systematic model error correction is given based on the least-squares approach by firstly using the past MEs. To test the correction, we applied the approach in GRAPES-GFS for July 2009 and January 2010. The datasets associated with the initial condition and SST used in this study were based on NCEP (National Centers for Environmental Prediction) FNL (final) data. The results indicated that the Northern Hemispheric systematically underestimated equator-to-pole geopotential gradient and westerly wind of GRAPES-GFS were largely enhanced, and the biases of temperature and wind in the tropics were strongly reduced. Therefore, the correction results in a more skillful forecast with lower mean bias and root-mean-square error and higher anomaly correlation coefficient.

Key words: model error, past data, inverse problem, error estimation, model correction, GRAPES-GFS

Citation: Xue, H. L., X. S. Shen, and J. F. Chou, 2015: An online model correction method based on an inverse problem: Part II—Systematic model error correction. *Adv. Atmos. Sci.*, **32**(11), 1493–1503, doi: 10.1007/s00376-015-4262-0.

1. Introduction

To reduce the forecast errors in numerical weather prediction (NWP) models by improving data assimilation and numerical models is an enduring challenge. As presented in the previous paper (Da, 2011; Xue et al., 2013), an approach to correct the forecast error is proposed through solving an inverse problem of NWP. This approach is of practical value, which is easy to implement as an online model correction into operational NWP systems, regardless of the various sources of forecast error (Chou, 1974; Bao et al., 2004; Ren and Chou, 2007; Da, 2011; Xue et al., 2013). As for operational application, this approach can be an online correction with only a minor modification to the numerical model. As reported in the companion paper (Xue et al., 2015), the model error tendency terms (ME) is obtained by iteratively solving an inverse problem using the past multi-time analyses.

Actually, this approach is a further development of Chou's idea (Chou, 1974). Chou recommended that NWP should use a huge series of past observations rather than consider the NWP model as an initial value problem. Based on Chou's idea, Da (2011) proposed a theoretical framework to predict the forecast error by using the past observations and NWP model. Da suggested that the ME could be expressed as a Lagrangian interpolation polynomial of MEs in past intervals, while the coefficients of polynomial can be determined by solving an inverse problem. In Da's study, observations were assumed to be absolutely accurate. This assumption results in that the order of the polynomial must be equal to the times of past observations, for keeping the problem well-posed. Due to the inevitable errors in observations, Da's method is short of practical value because the high-order polynomial shows strong sensitivity to errors embedded in the past observations (Xue et al., 2013). Moreover, the approach employed in Da (2011) for obtaining MEs is trapezoidal approximation, which is too rough. Hence, in practical application, the order of the polynomial and number of past data are restricted.

In the companion paper (Xue et al., 2015), we propose a more precise and practical method by using an iterative approach to estimate the past ME. By analyzing and comparing with the forecast errors in GRAPES-GFS (Global Forecast System of the Global and Regional Assimilation and Predic-

* Corresponding author: XUE Haile
Email: xuehl@cams.cma.gov.cn

tion System), it is confirmed that the error pattern and temporal evolution, linearly estimated based on past MEs, show the rationality of the iterative method and the potential application of the MEs for online correction. To fulfill this objective, a key question should be considered is how to extrapolate the past MEs to the forecast period so that forecast error can be corrected throughout model integration. Da (2011) suggested using a Lagrangian interpolation polynomial to predict the future ME, but the order of the polynomial shows sensitivity to the errors introduced in the past observations, as discussed above. In reality, the correction should be based on the characteristics of the forecast errors. Danforth et al. (2007) divided the 6-h total forecast error into systematic model error, periodic errors, and non-periodic components. The correction is expected to utilize multiple past MEs more flexibly and offer a more stable influence on model forecasts. In this paper, a least-squares approach is adopted to predict the correction of the systematic errors. In this context, a brief estimation of the systematic errors of GRAPES-GFS is given firstly, in section 2, followed by a description of the least-squares approach and the experimental methods in section 3. Section 4 presents the results of the experiments, followed by discussion and conclusions in section 5.

2. Systematic errors of GRAPES-GFS

In Xue et al. (2015), we used the iterative method to obtain the MEs in the past intervals and verified its validity by comparing the 2-month-mean forecast errors to the estimated error corrections in GRAPES-GFS. For consistency, the model and analyses adopted in the present paper are the same as in Xue et al. (2015). Because we aim to correct the systematic errors of GRAPES-GFS in this part of the study, it is necessary to understand the patterns and evolution of the systematic errors of GRAPES-GFS. Here, systematic ME is defined as the departure of the forecast from the FNL analyses. It is expected that the patterns of systematic errors should be different from each other according to different prediction systems. However, it is surprising that certain aspects of systematic MEs seem to be persistent not only in different versions of the same model, but also across models developed at different climate centers (Berner et al., 2012). Jung and Tompkins (2003) and Jung (2005) delivered studies of the patterns and evolution of the systematic errors of the European Centre for Medium-Range Weather Forecasts (ECMWF) forecasting system. The results indicated that the systematic errors of the ECMWF system maintained a fixed pattern and increased linearly in the short term.

Murphy (1988) decomposed the mean-squared error into systematic and random components; namely, the mean squared error (MSE) and error variance, and Jung and Tompkins (2003) delivered a simpler formula to present systematic error. Here the expression of Jung and Tompkins (2003) is followed as

$$E_{\text{sys}} = \bar{x}_f - \bar{y}_0, \tag{1}$$

where $\bar{x}_f$ and $\bar{y}_0$ are time averages of model forecast and FNL analyses, respectively. Zheng et al. (2006, 2009) employed a similar formula to Eq. (1) to deduce parameters of a linear, first-order Markov stochastic error model in a climate model. If the average period is months of a certain season of multiple years, Eq. (1) can result in climatological systematic errors. We use winter months (December–January–February, DJF) of 5 years (2001–2005) as the average period to obtain the climatological systematic errors in the following analysis.

Climatological systematic errors of 500 hPa geopotential height (marked as $Z500$, where Z means geopotential height) for winters (DJF) of 2001–2005 are shown in Fig. 1. Evidently, two large systematic Z500 errors are apparent in the Northern Hemisphere. The positive one is over the North Pacific to the Arctic and the negative one covers from North Africa to East Asia in the midlatitudes. Climatological systematic errors of $Z850$ for winters (DJF) of 2001–2005 have also been examined (not shown). Compared with $Z500$, the two centers of systematic errors are still existent, which reveals that the strong wintertime bias in the Northern Hemisphere extends throughout the troposphere. Form Fig. 1, we can also see that both the positive bias and negative bias are enhanced as forecast time increases.

The zonally averaged North Hemispheric systematic errors of geopotential height (Z) and zonal wind (u) for winters are shown in Fig. 2a. The positive bias of Z in the extratropics and pole of the Northern Hemisphere extends from the bottom to the top of the atmosphere as an equivalent barotropic structure. In the middle and upper troposphere of the extratropics and subtropics of the Northern Hemisphere, the westerly wind is underestimated, which is consistent with the underestimation of the equator-to-pole pressure gradient. This negative westerly wind bias together with the systematic error structure of Z is governed by geostrophic wind balance in the Northern Hemisphere. For summers (June–July–August, JJA), the results are similar, except a negative geopotential height bias band and easterly wind bias is found at the top of the troposphere (not shown).

For the tropics, the temperature errors and 2D wind vector errors are calculated by the mean difference between model forecasts of GRAPES-GFS and NCEP FNL in the period of the winters (DJF) of 2001–2005. The 2D wind vector errors include v and w errors, where w errors are multiplied by 100. The convergent or divergent winds are dominantly forced by heterogeneous heating. As shown in Fig. 2b, a convergent bias center is located in the lower troposphere between 50°E and 100°E, due to a positive temperature bias center. Another convergent bias center is found in the upper troposphere and stratosphere at around 50°E, due to another positive heating bias in the upper layer. As for meridional wind bias, the negative temperature bias causes a divergent meridional wind bias in the boundary layer in the tropics, which results in a downwelling bias in the troposphere above the equator (not shown). However, the whole tropospheric layer tends to be a convergent bias, due to two positive temperature biases, though there is a layer with negative temperature bias between them (not shown). In the tropics, the temperature errors show a baroclinic structure, as positive bias alternates

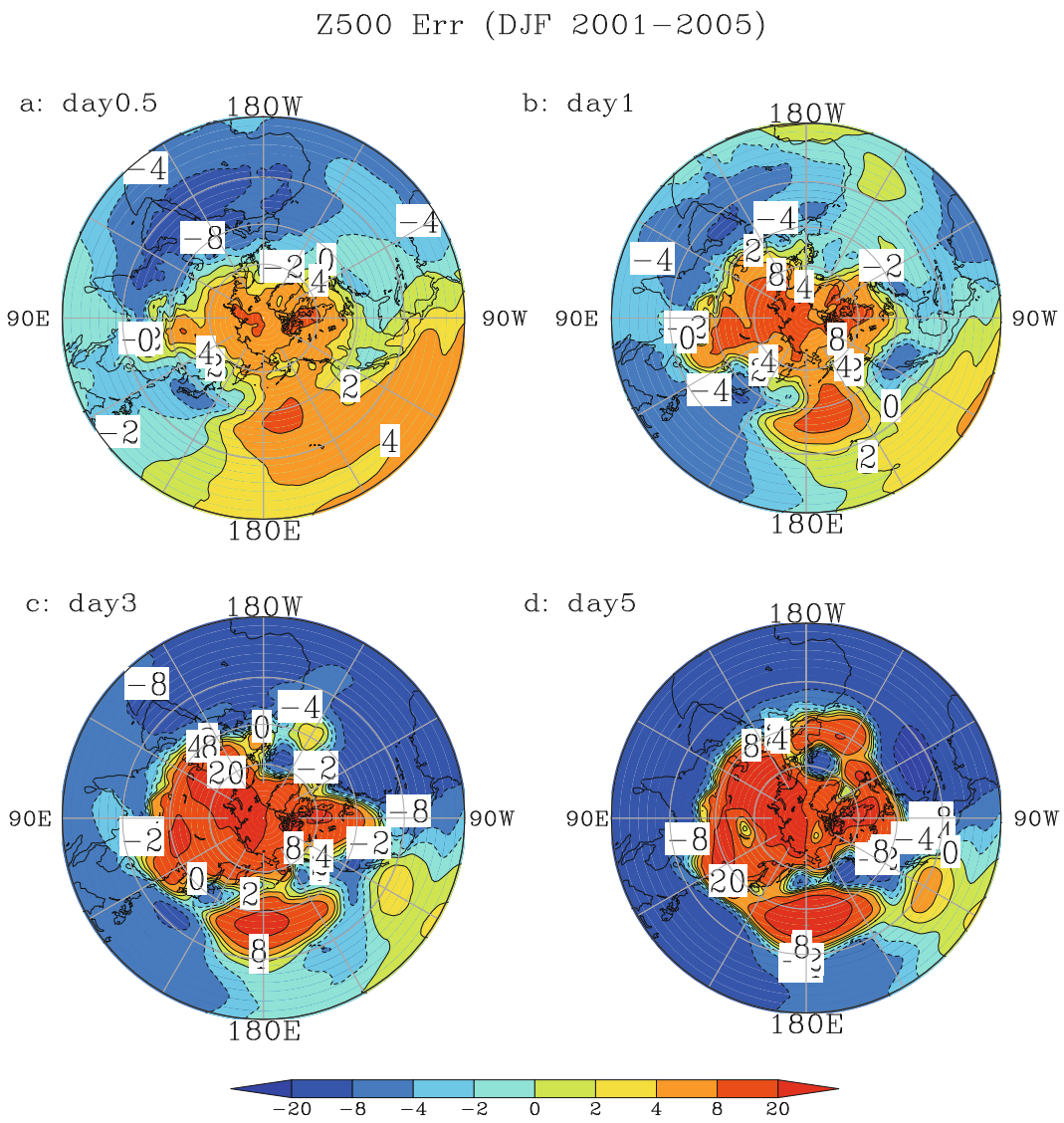

Fig. 1. Mean difference of 500 hPa geopotential height (gpm) between model forecasts of GRAPES-GFS and NCEP FNL for the winters (DJF) of 2001–2005: (a) hour 12; (b) day 1; (c) day 3; (d) day 5.

with negative bias. For summers, the results are similar but the patterns are different (not shown). For example, the divergent meridional wind bias center in the boundary layer moves between 5°N and 10°N, and a convergent meridional wind bias center is found at 500 hPa between 5°N and 10°N, which is different from the whole tropospheric layer convergent wind bias for winter. Note also that the zonal wind bias center in the boundary layer moves between 50°W and 100°W, and the upper tropospheric convergent zonal wind bias center is located above the equator.

Unfortunately, it is difficult to determine which one out of geostrophic wind bias or geopotential height bias in the extratropics is the cause of the other. However, the error source is not the point we focus on in this study; rather, we are interested in the patterns and evolution of the systematic errors. Note that the climatological systematic errors at 500 hPa and 850 hPa increase, with confirmed patterns, as forecast time increases (as shown in Fig. 1). It can be seen that the 500 hPa geopotential height errors in the tropics and both hemispheres evolve linearly, and the 200 hPa zonal wind errors also increase linearly (Fig. 3). For summers, the 500 hPa geopotential height errors in the Northern Hemisphere evolve linearly but the errors in the Southern Hemisphere and the tropics increase slowly after two days of integration (not shown). However, the 200 hPa zonal errors almost increase linearly. These results suggest that the systematic error can be linearly estimated by the mean MEs, for potential use in online correction.

3. Experimental approach and design

Suppose an unknown term $\boldsymbol{E}$ exists, which can be considered as the overall effect of different sources of model . As a result, the NWP model can be rewritten as an inverse problem

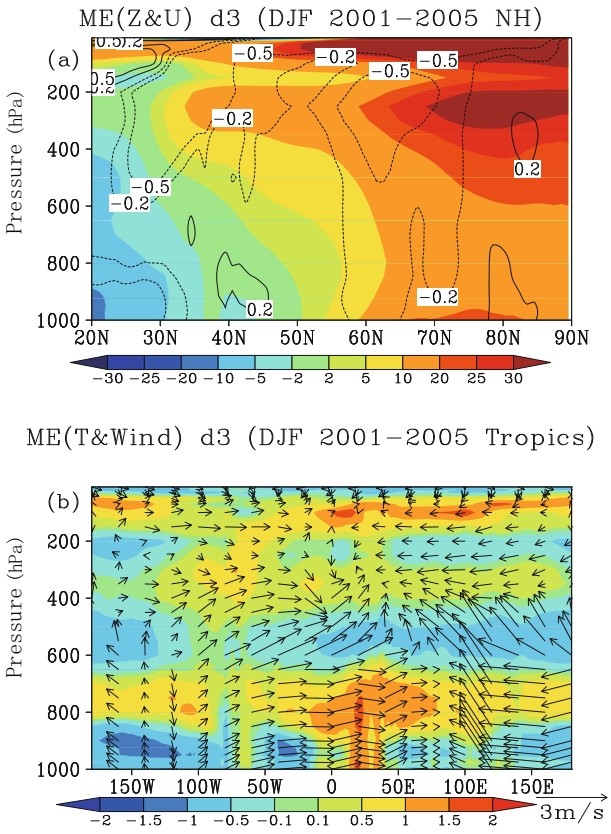

Fig. 2. (a) North Hemispheric error sections of Z (shaded; gpm) and u (contours; m s^{-1}), and (b) tropical region error sections of T (shaded; °C) and 2D wind vectors (arrows are displayed by v and $100\times w$). The errors are the mean differences between model forecasts of GRAPES-GFS and NCEP FNL in the period of winters (DJF) of 2001–2005, and the errors are averaged zonally for the Northern Hemisphere and meridionally for the tropic between 20°S and 20°N.

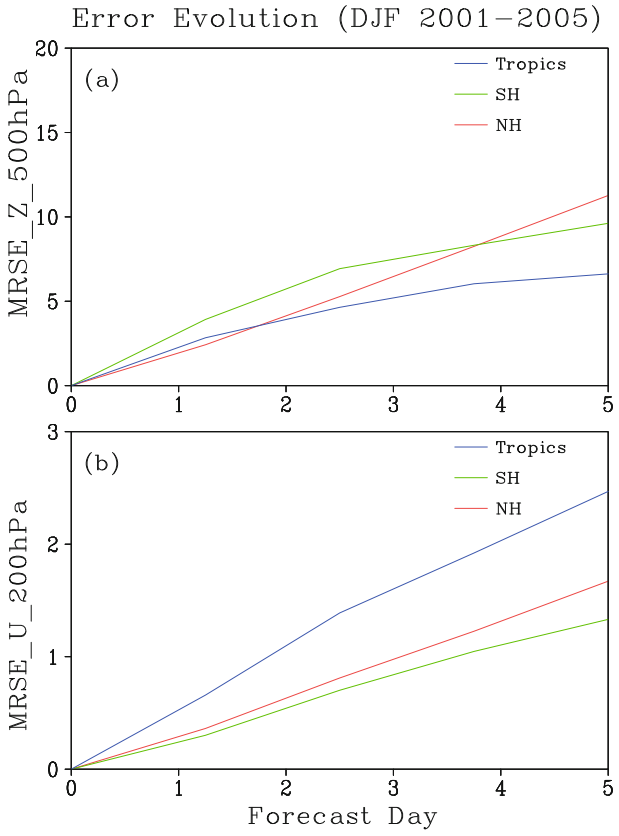

Fig. 3. RMSE evolution of (a) 500 hPa geopotential height (gpm) and (b) 200 hPa zonal wind (m s^{-1}) for the winters (DJF) of 2001–2005.

together with past observations described in Xue et al. (2015).

We further discretize the model error $\boldsymbol{E}$ into series of datasets and denote $\boldsymbol{E}_i$ $(i=-n,-(n-1),\cdots,-1)$ as the ME in the past interval between $i\delta$ and $(i+1)\delta$. For brevity, the following discussion focuses on one state variable of the model equation at the space-discretized point, as in Xue et al. (2015). In the past arbitrary interval $(-i\delta,-(i-1)\delta)$ [where $i=-n,-(n-1),\cdots,-1$], the ME $\overline{E}_i$ in the corresponding interval can be easily obtained iteratively by using the method described in Xue et al. (2015).

Suppose that, for a moment, the ME is dependent on the time and of the polynomial form of m order as $E(t)=a_0+a_1t+a_2t^2+\cdots+a_mt^m$. $a_0,a_2,\cdots,a_m$ are coefficients. Integrating the model Eq. (2a) in Xue et al. (2015) for each interval δ yields the following equations:

$$\overline{E}_i\delta=a_0\delta+a_1\frac{(i+1)^2-i^2}{2}\delta^2+\cdots+a_m\frac{(i+1)^{m+1}-i^{m+1}}{n+1}\delta^m, \tag{2}$$

where $i=-1,-2,\cdots,-n$, and m is a positive integer smaller than n. It is noted that there are n number of equations and $m+1$ number of unknown coefficients, which can be solved by the least-squares method. Minimizing the norm:

$$J(a_0,a_1,\cdots,a_m)=\sum_{i=-n}^{-1}\left[\overline{E}_i\delta-\left(a_0\delta+a_1\frac{(i+1)^2-i^2}{2}\delta^2+\cdots+a_m\frac{(i+1)^{m+1}-i^{m+1}}{n+1}\delta^m\right)\right]^2, \tag{3}$$

results in the following equation set:

$$\frac{\partial J(a_0,a_1,\cdots,a_m)}{\partial a_0}=0, \tag{4a}$$

$$\frac{\partial J(a_0,a_1,\cdots,a_m)}{\partial a_1}=0, \tag{4b}$$

$$\cdots\cdots$$

$$\frac{\partial J(a_0,a_1,\cdots,a_m)}{\partial a_m}=0. \tag{4bm}$$

Obviously, an explicit formula of ME can be solved from the equation set, which is actually a kind of curve-fitting result. One may argue that this formula works in the past model integration time, but whether it does in model forecast time is not confirmed. However, for some ME components of stable statistical character, e.g., climatological systematic errors, discussed in section 2, the statistical formula result from the

past will work in future. The climatological systematic error is the difference between model forecasts and observations, which maintain in a specific season and increase linearly in short-term prediction (Danforth et al., 2007). That means a constant online corrector is enough for the long-existent linear-increasing climatological systematic errors. It was discussed in Xue et al. (2015) that direct application of a constant model correction in the NWP model should impact the model at the very beginning of model integration, which makes the iteration divergent. Here, we are not concerned with the convergent problem, but the impact of the NWP model at the very beginning of model integration may also lead to some uncertainties. In reality, to deal with this problem technically, the correction is increased linearly from zero to its true magnitude in a few steps of model integration and then maintained at a constant in the following steps. Based on this discussion, the order of the polynomial should be zero ($m = 0$), and the solution of the equation set is easily solved as

$$E = a_0 = \frac{1}{n}\sum_{i=-n}^{-1}\overline{E}_i \,. \tag{5}$$

Actually, Eq. (5) shows that the systematic model error is none other than the time average of the MEs in past intervals. According to Eq. (1), the systematic errors are a kind of mean error, but the sample amount usually requires a long series of data (e.g., decades of annual data). However, to save computational resources (plus, the purpose of this study is not real operational use of the method), we use two-month samples to calculate the mean forecast error. Although this is too short to totally represent the systematic errors, the results show that the patterns and evolutions of the two-month mean errors are similar to the systematic error (5-year averaged errors). Therefore, the two-month mean errors are approximately represented by the systematic errors in the following sections.

Another factor that should be considered is the increasing rate of systematic errors. We used 2-month MEs to estimate the mean errors and to approximately represent the systematic errors in Xue et al. (2015). The patterns and evolution of mean error corrections estimated from Eq. (5) have been compared offline to the mean forecast errors (Xue et al., 2015). The results show that the patterns of estimated error are well matched to the mean forecast errors for both the Northern Hemisphere and the tropics, but the increasing rates of estimated error are over-abrupt compared to mean forecast errors. Hence, a weight α is necessary to account for the overestimated error corrector, and the corrector becomes

$$E = \frac{\alpha}{n}\sum_{i=-n}^{-1}\overline{E}_i \,. \tag{6}$$

It is expected that a large amount of systematic errors will be offset by choosing the weight reasonably. However, there is another parameter that should be considered: the number of past MEs, n. Empirically, the climatological systematic error corrector should result statistically from the past data and model outputs of multiple years (Glahn and Lowry, 1972; Carter et al., 1989). Here, for the purpose of reducing the computational cost, we use the spectrum analysis method to obtain a proper amount of past data. Thus, averaged spectra of the MEs of horizontal velocity components u and v, perturbation from reference potential temperature θ' and Exner pressure perturbation Π' (blue) at 3rd model level and the 14th model level for GRAPES-GFS are shown in Fig. 4. Estimates are based on the MEs resulting from iteration for January–February 2010 with 6 h intervals. Apparently, the spectra of the four variables show three peaks in 60 days: the first is a one-day period, which might be associated with the diurnal cycle; the second is greater than a 3-day period, which may be related to the synoptic process; and the third one is greater than a 30-day period, which may represent the long-existent climatological systematic errors. Finally, we use 30 days' past data ($n = 120$) to determine the correction in the following forecast tests. In this part, for the purpose of simply testing the new method, only long-existent climatological systematic errors were considered in the following experiments.

To determine the weight factor α, a set of experiments with changing α from 0 to 1 were conducted for January 2010 firstly, based on the past MEs produced by iteration in Xue et al. (2015). Before applying the model correction through model integration, the ME for the forecast period should be calculated based on its previous 30-day MEs. Then, the ME is introduced into the model as a tendency term and forces the model at every time step. As different values of the weight act on the ME, the influence of the correction could be balanced. Secondly, according to the results of the first set of experiments, a linear decayed factor was used to weight the correction. Then, the linear weight correction was applied for July 2009 (JUL2009) and January 2010 (JAN2010). The experiments were all initialized by $1^\circ \times 1^\circ$ NCEP final (FNL) analyses at 0000 UTC and performed for 8-day forecasts. The model configuration and physical processes The resolution is set as 1°, and the model dynamics and physics are default.

4. Results

Although the online correction used here is deduced statistically from the past MEs, it is expected to work for the forecast time. To optimize the performance of the correction, a series of weights (0–1) were acted upon it, and the results are shown in Fig. 5. As the 500 hPa and 850 hPa levels are the most critical levels for large-scale and synoptic-scale processes, the anomaly correlation coefficient (ACC) and root mean square error (RMSE) of these levels are adopted here as the scores of the correction influence. The ACC is an index to measure how closely the forecasts match the analyses, while the RMSE is an index to weight the difference between the forecasts and analyses. It is assumed that the best model performance for the weight α is consistent regarding the two indexes. Fortunately, this is true for this study but, if it is not, we should balance the two indexes. As shown in Fig. 5, for the day 1 forecast, the best performance of the two levels was related to $\alpha = 0.8$; for the day 3 forecast, results

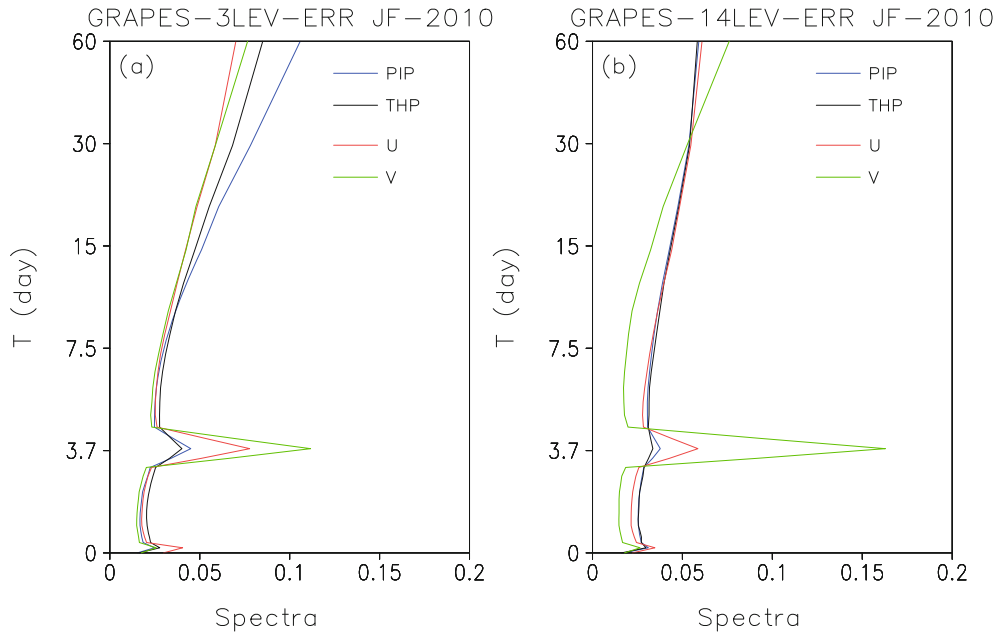

Fig. 4. Averaged spectra of the model errors of horizontal velocity components u (red) and v (green), perturbation from reference potential temperature θ' (black) and Exner pressure perturbation Π' (blue) at (a) 3rd model level and (b) the 14th model level for GRAPES-GFS. Estimations are based on the model errors resulting from iteration for January–February 2010 with 6 h intervals. The vertical axis is time.

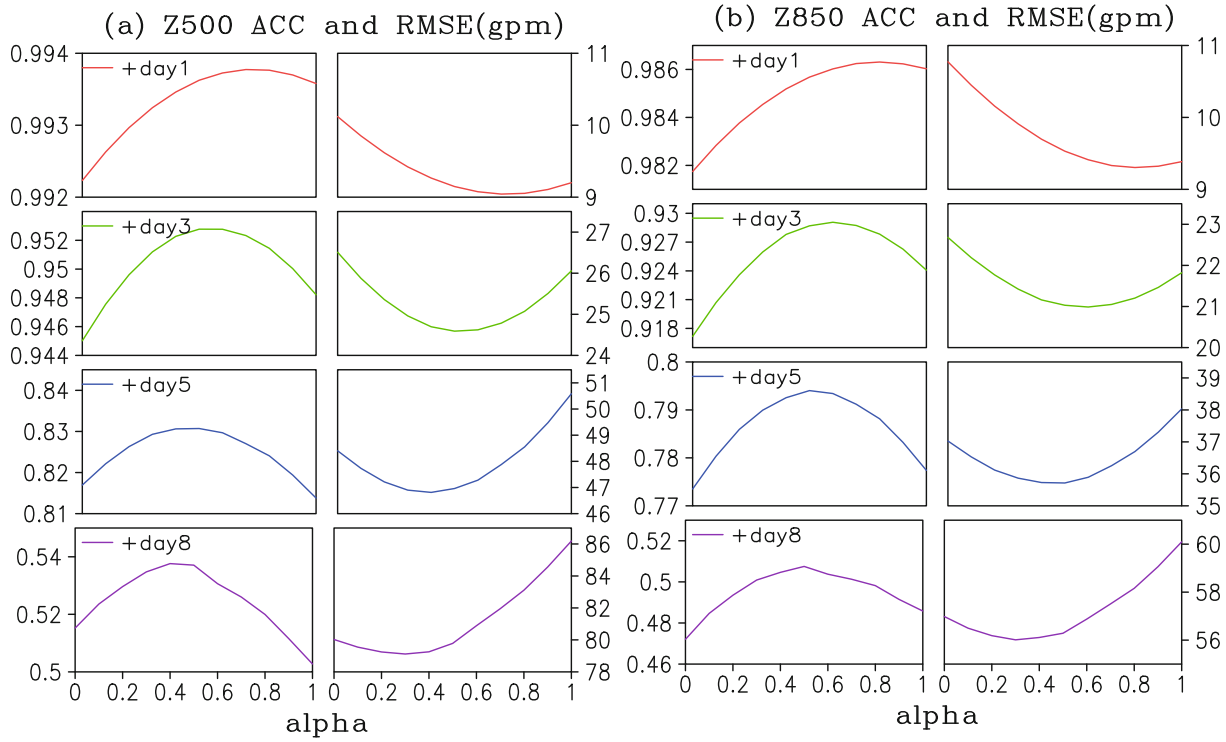

Fig. 5. ACCs and RMSEs of Z (gpm) at 500 hPa and 850 hPa for JAN2010. Panels in the top to bottom rows are hour 12, day 1, day 3 and day 5 forecasts, respectively.

showed high scores when $\alpha = 0.6$; for the day 5 forecasts, when the weight was about 0.4, the RMSE of both levels was lowest, but the ACC of the two levels was highest when the weight was about 0.5; and finally, for the day 8 forecast, the 500 hPa and 850 hPa ACC achieved higher scores when α was between 0.4 and 0.5 and the 500 hPa and 850 hPa RMSE achieved higher values when α was between 0.2 and 0.4. We can also see that, before the day 5 forecast, the results were

better when applying the correction, whatever the weight, but the performances were worse after the day 5 forecast when the weight was not reasonable. These results imply that the patterns and magnitudes of the correction are well matched with the model error before day 5 forecasts, but it is uncertain after day forecast. It was also noted that, as the forecast time increased, better model correction performances were related to decreasing weights, suggesting a decayed weight should be used. Therefore, a linear decayed weight (LDW) was applied to the correction from 1 at the very beginning of the forecast, to 0 after model integration. Before the forecast, the ME for the correction should first be calculated based on its previous 30-day MEs. Throughout the 8-day forecast, the ME forced the model at every time step with a decay factor that was determined by the weight α. That meant the influence of the correction was linearly decreased through the model integration and gradually became zero when the integration was complete. In the following test, the tests without correction ($\alpha = 0$) were considered as control tests, which were compared to the LDW correction tests.

Figures 6a and b present the mean $Z500$ (gpm) day 3 forecast for JAN2010 and the corresponding mean NCEP FNL. Together with the ACC of $Z500$, shown in Fig. 10a, it can be seen that the structure of the mean field forced by the LDW correction was more similar to the mean NCEP analyses. For clarity, the difference between Figs. 6b and a is shown as Fig. 6d. It is clear that the positive bias in the extratropics and negative bias in the tropics of $Z500$ was reduced by the LDW correction. That meant that the LDW correction enhanced the underestimated equator-to-pole geopotential gradient described in section 2. The result for JUL2009 was similar to JAN2010 (not shown).

To further verify the impacts of the LDW correction, Fig. 7 presents the mean $Z500$ and $Z850$ forecast errors for JAN2010. The pattern is almost similar to the systematic errors shown in section 2: overestimated Z in the extratropics and underestimated Z in the tropics. It can also be seen that the forecast errors increased from the 12th hour to day 5 forecasts, with a nearly maintained pattern. As expected, the bias was canceled to a very significant degree in most locations for the hour 12, day 1, 3 and 5 forecasts. For example, the positive bias over the extratropics and the Artic almost disappeared from the 12th h to day 5 forecasts. The negative $Z500$ bias from North Africa to the western Pacific was reduced by a large amount following the LDW correction in the forecasts at hour 12 but, because the phase of negative bias over North Africa and the Indian Ocean to South Asia changed to positive in the day 1 forecast, it was not corrected. For the day 3 and 5 forecasts, the negative bias over the Western Hemisphere and Eastern Hemisphere was reduced by the LDW correction, but still existed. The impact of the LDW correction on the mean error of $Z500$ and $Z850$ for JUL2009 was similar to the results of JAN2010 (not shown).

Figure 8 shows the zonally averaged latitude–height cross sections of Z and u errors of the control experiments (left column) and the LDW correction (right column)

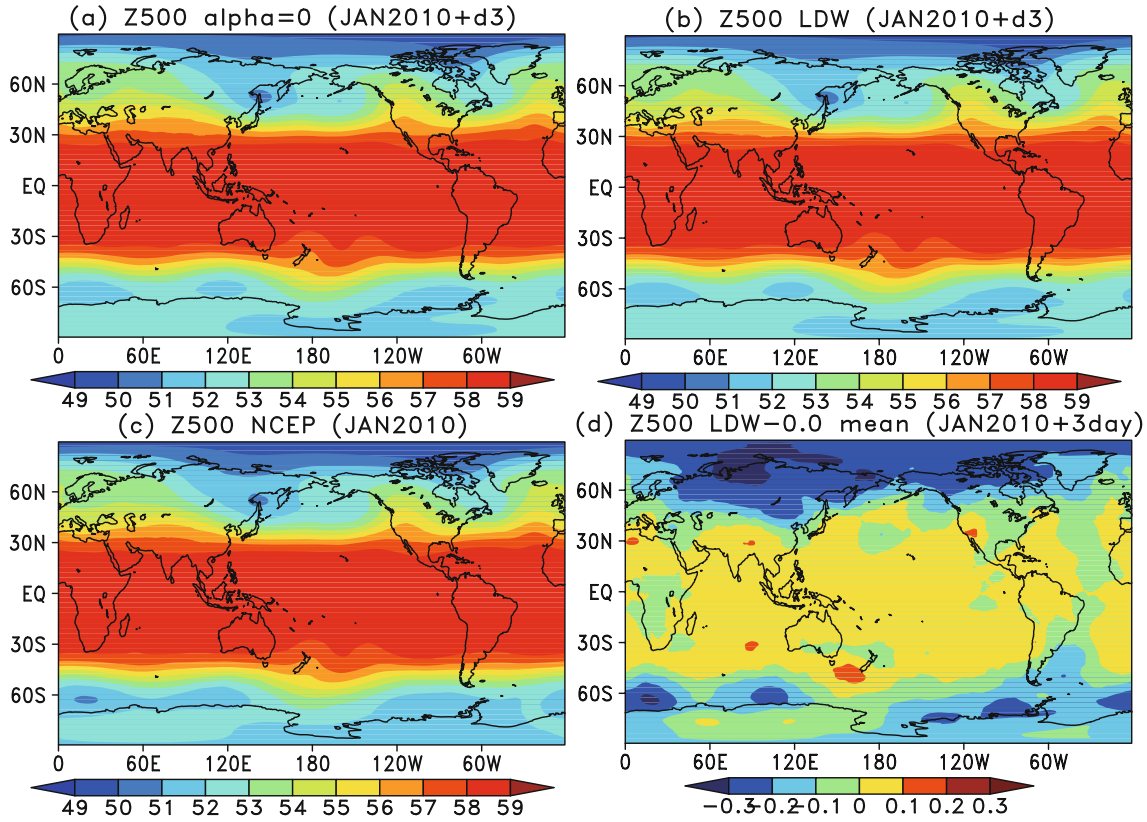

Fig. 6. Mean Z (gpm) at 500 hPa for day 3 forecast of (a) control experiments and (b) LDW correction for JAN2010. Panel (c) shows mean Z (gpm) at 500 hPa for the NCEP analyses and (d) is the difference between (b) and (a).

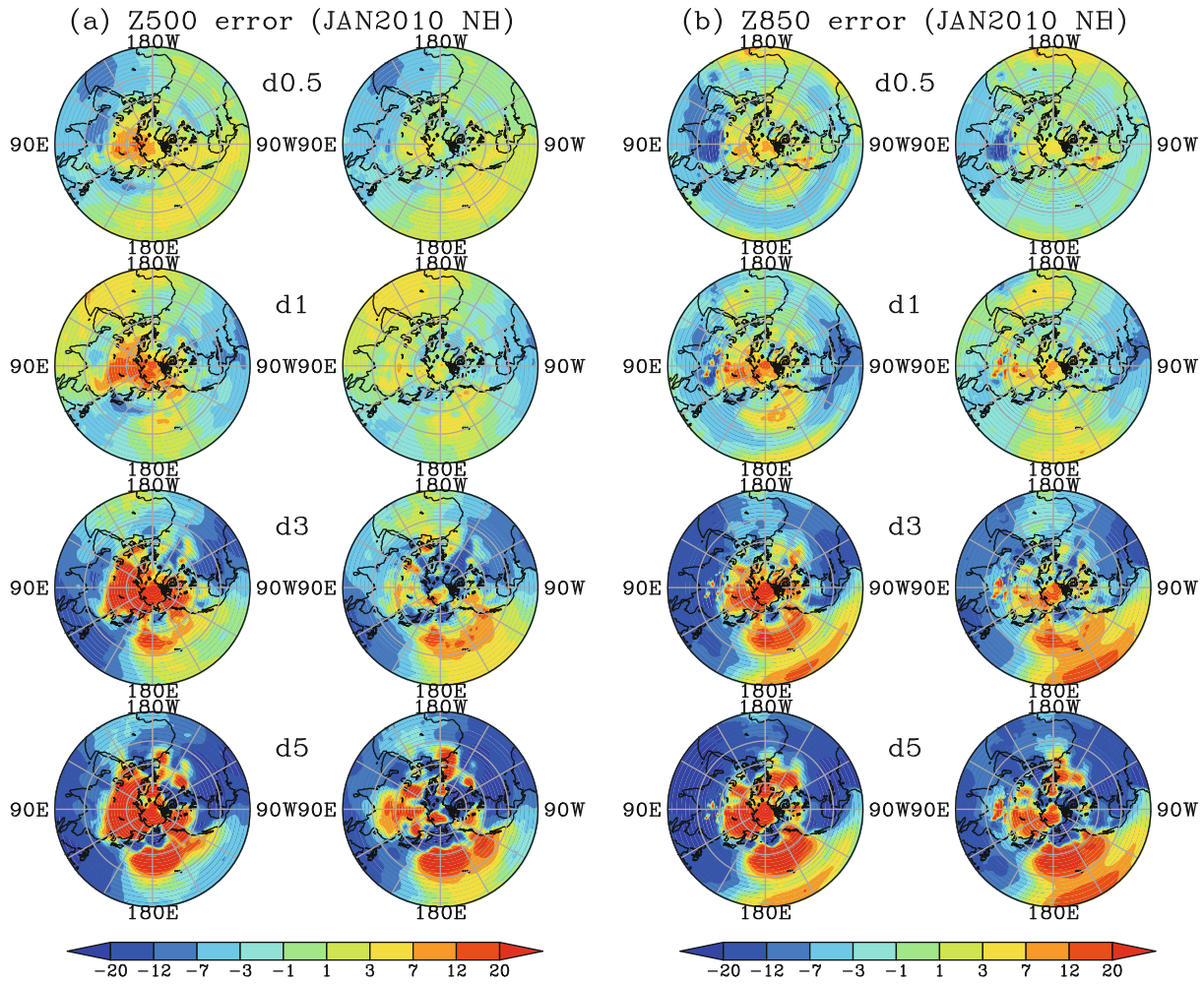

Fig. 7. Mean forecast Z errors (gpm) of the control experiments (left column) and LDW correction (right column) at 500 hPa and 850 hPa for JAN2010. Panels from the top to bottom row are hour 12, day 1, day 3 and day 5 forecasts, respectively.

for JAN2010. The underestimated westerly wind, associated with the geopotential relationship between Z and wind, was observed in section 2. As shown in Fig. 8 (left column), the overestimated mean Z over the Arctic and extratropics, and the underestimated mean Z over the low-latitude area, were clearly seen from the 12th h to day 5 forecasts, accompanied by the underestimated westerly wind over most parts of the Northern Hemisphere. Influenced by the LDW correction, the systematic geopotential bias almost disappeared, except in the overcorrected upper troposphere in the day 3 and 5 forecasts (shown in the right column of Fig. 8). Sequentially, the negative bias westerly wind over the Northern Hemisphere was nearly canceled completely in the 12th h and day 1 forecasts, and the underestimated westerly winds were also enhanced by the correction in the day 3 and 5 forecasts. It was also noticed that overcorrected westerly wind bias was found at a few locations; for example, the troposphere over 60°N, in the day 3 and 5 forecasts.

Figure 9 shows the meridionally averaged (from 20°S to 20°N) longitude–height cross sections of temperature (T) and 2D wind vector errors of the control experiments (left column) and the LDW correction (right column) for JAN2010. In the tropics, large-scale average wind is governed by atmospheric convergence or divergence, driven by the nonuniform heating. Hence, the wind bias is of course associated with the temperature bias, i.e., convergent wind bias accompanied by positive temperature bias and divergent wind bias with negative temperature bias. As shown in Fig. 9 (top row), there were two bias centers in the forecast at hour 12: the positive bias was located above the surface from 100°W to 150°W, accompanied by a convergence of the 2D wind bias; while the negative bias was located in the surface layer between 0° and 50°E, where a divergent 2D wind bias was found. After the correction, the temperature and 2D wind biases were all reduced. For the day 1 to 5 forecasts (Fig. 9, rows 2–4), two positive temperature biases were found: one located in the surface layer and lower part of the troposphere between 50°W and 100°E, and the other extended from the upper troposphere to the stratosphere. Sequentially, two 2D wind convergence bias centers were found with respect to the temperature biases. It should be noted that the positive temperature biases were greatly reduced by the LDW correction. Impacted by the correction, the 2D wind bias nearly disappeared in the day 1 forecast, and the positive temperature

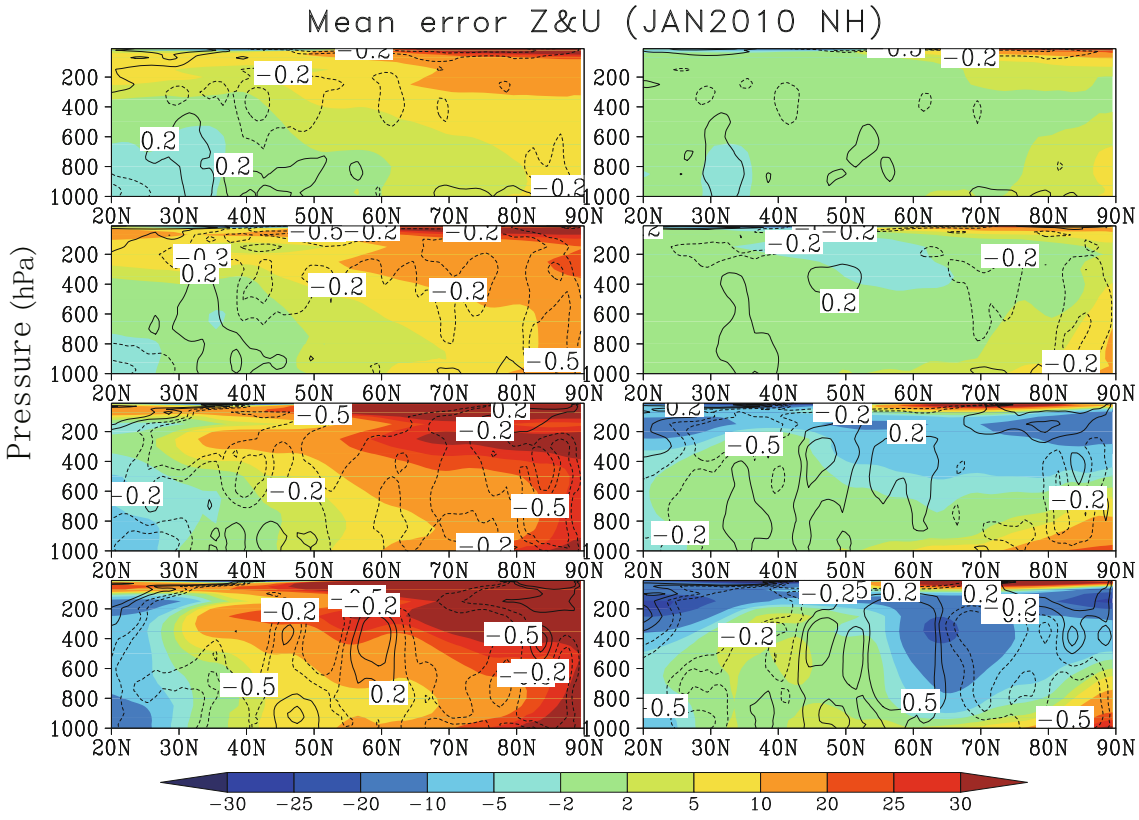

Fig. 8. Zonally averaged latitude–height cross sections of Z (gpm) and zonal wind u (m s^{-1}) errors of the control experiments (left column) and LDW correction (right column) for JAN2010. Panels from the top to the bottom row are hour 12, day 1, day 3 and day 5 forecasts, respectively. The vertical axis is pressure.

biases and 2D wind biases also decreased in the day 3 and 5 forecasts. However, the negative temperature bias band in the upper troposphere from 100°E to 100°W increased, along with the 2D wind divergence bias. The pattern of the North Hemisphere Z and wind errors for July 2009 was different from that for January 2010, but the influence of the correction on the errors was similar (not shown). The pattern of the tropical temperature and wind errors for July 2009 was similar to January 2010, and the impact of the correction was similar too (not shown).

According to the results on isobaric surfaces and in the cross sections, it can be concluded that the systematic errors can be sharply reduced using the LDW correction. However, not everywhere benefits from the LDW correction; for example, the negative forecasting temperature biases in the tropics. This may be related to the other components of the errors. The forecast errors are composed of multiple scales of components; for example, the climatological systematic errors and diurnal errors. In some situations, the phases of the two components are the same, and the systematic error correction cannot cancel the error completely. In other situations, the phases are opposite between the climatological systematic errors and diurnal errors, and the systematic error correction causes the sign of error to be the opposite. Therefore, more attention should be paid to the elimination of other error components in future research. Besides, to properly eliminate the climatological systematic errors may require long-term statistics, so another point to be noted is the sample size of the past data is relatively small in this study.

Finally, mean forecast scores of the model performance are given: the mean bias, RMSE and ACC. The 8-day forecast RMSE and bias of u, v, T and Z at 500 hPa and 850 hPa for JAN2010 have been examined (not shown). The RMSE of Z and T at 500 hPa was reduced before the day 5 forecast by the LDW correction, and were similar to the control experiments after the day 5 forecast. It is clear that the RMSE and bias of Z and T at 850 hPa were decreased for all 8-days forecasts. It should be noted that the impacts of the correction on the wind RMSE and bias were negligible and the RMSE of u and v after the day 6 forecast increased slightly. Regardless, it is confirmed that the wind prediction and its correction are more difficult to achieve than for temperature and Z. Figure 10 shows the mean ACC between the forecast results of GRAPES and the analyses of NCEP FNL. The ACC was calculated based on the Z fields at the 500 hPa and 850 hPa levels for JAN2010. It is shown that the confidence intervals were small, so the mean ACC was credible. From Fig. 10 we can see that GRAPES-GFS obviously possesses a finer capability for prediction as a result of the correction. The ACC between GRAPES-GFS hindcasts and NCEP analyses increased in all the 8-days forecasts. When taking an ACC of 0.6 as a threshold value for the effectiveness of the forecasts, GRAPES-GFS extended the capability of the effective forecast time by more than 7 days.

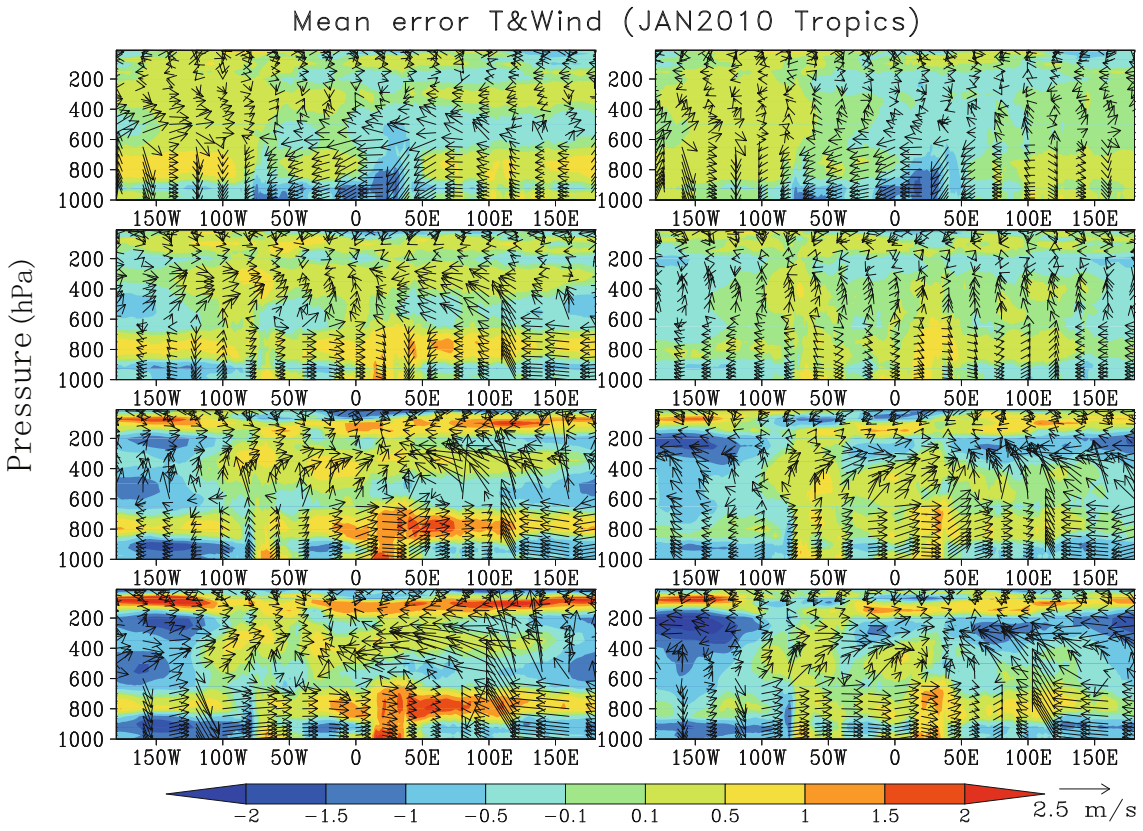

Fig. 9. Meridionally averaged longitude–height cross sections of T (K) and 2D wind vector errors of control experiments (left column) and LDW correction (right column) for JAN2010, where vertical speed has been amplified by 100. Panels from the top to the bottom row are hour 12, day 1, day 3 and day 5 forecasts, respectively. The vertical axis is pressure.

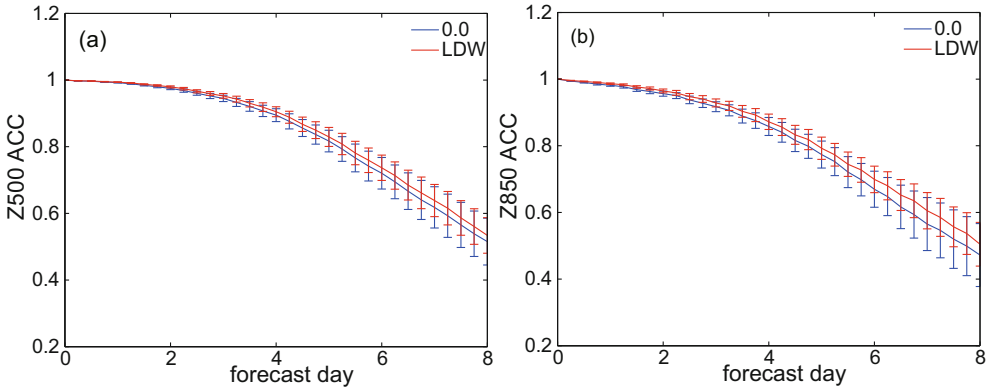

Fig. 10. Mean ACC and its confidence intervals of Z (gpm) at 500 hPa and 850 hPa for JAN2010.

5. Discussion and conclusions

As a key strategy in weather forecasting, NWP still requires an elimination of its systematic errors due to model deficiencies. To improve the skill of model performance, besides traditional approaches, an alternative method is to consider the NWP as an inverse problem in order to utilize numerous past analyses, which was first raised by Chou (1974). The main idea is solving an inverse problem to obtain an unknown term in the prediction equations by using the past analyses, which is presumed to represent the imperfection of the NWP model. However, it is hard to deduce the ME term inversely because the formula of the ME is uncertain. Fortunately, the error evolves linearly in a short interval, i.e., 6 hours, so the ME term is assumed to be constant. The ME terms in past intervals can then be obtained iteratively (see Xue et al., 2015). Furthermore, based on the patterns and evolution of the GRAPES-GFS forecast error, an online model correction formula is deduced by using the datasets of the MEs. To test the correction, the climatological systematic errors of GRAPES are considered as the object to be corrected. Two-month (January 2010 and July 2009) experiments were

carried out to determine the weight and check the impact of the correction. The conclusions can be summarized as follows:

(1) The short-term systematic errors of GRAPES-GFS have been measured and the results indicate that positive Z bias in the extratropics and negative Z bias in the subtropics are dominant, with the negative westerly wind bias in the extratropics. For the tropics, the divergent wind bias is dominantly forced by the heterogeneous heating bias. Even more importantly, the growth rate of systematic errors is almost linear in 5-day forecasts.

(2) Based on the results regarding different weights (from 0 to 1), the larger the weight, the better the model performance is for the early days of the forecasts (for example the first 3 days of the forecasts); the smaller the weight is, the better the model performance is for the later days of the forecasts (e.g., day 5 to 8). This suggests an LDW should be used as integration.

(3) According to the results of the corrected experiments at 500 and 800 hPa, the systematically underestimated equator-to-pole geopotential gradient over the Northern Hemisphere is sharply enhanced due to LDW correction. Sequentially, the negative westerly wind bias is reduced as well.

(4) The temperature bias over the tropics is also reduced by the LDW correction, as well as the convergence and divergence wind biases associated with an unrealistic heating source and sink.

(5) The LDW correction leads to better mean forecast scores: the RMSE and bias are reduced for short-term forecasts (first 5 days of forecasts), and the ACC is increased for all the 8-day forecasts.

Although the overall forecasts of GRAPES-GFS largely benefit from the correction, the LDW correction cannot improve the model performance everywhere. The forecast errors consist of multiple scales of components, e.g., climatological systematic errors, diurnal errors and state-dependent errors. As shown in Xue et al. (2015), the diurnal error oscillation is obvious before iteration. Hence, the systematic error correction alone is not enough, and more attention should be paid to the elimination of other error components in future research. Besides, to properly eliminate the climatological systematic errors may need years of statistics, so another point to note is the sample size of the past data is relatively small in this study.

Acknowledgements. The data for this paper are available at CISL Research Data Archive Atmospheric Re-analysis Data. Data set: NCEP FNL. Dataset name: NCEP FNL (Final) Operational Global Analysis data. We would like to acknowledge that this work is funded by the National Natural Science Foundation Science Fund for Youth (Grant No. 41405095), the Key Projects in the National Science and Technology Pillar Program during the Twelfth Five-year Plan Period (Grant No. 2012BAC22B02), and the National Natural Science Foundation Science Fund for Creative Research Groups (Grant No. 41221064). We especially thank YU Haipeng for his help about data collection.

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)

Berner, J., T. Jung, and T. N. Palmer, 2012: Systematic model error: the impact of increased horizontal resolution versus improved stochastic and deterministic parameterizations. *J. Climate*, **25**, 4946–4961.

Carter, G. M., J. P. Dallavalle, and H. R. Glahn, 1989: Statistical forecasts based on the national meteorological center's numerical weather prediction system. *Wea. Forecasting*, **4**, 401–412.

Chou, G. 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, 100 pp. (in Chinese)

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

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

Jung, T., 2005: Systematic errors of the atmospheric circulation in the ECMWF forecasting system. *Quart. J. Roy. Meteor. Soc.*, **131**, 1045–1073.

Jung, T., and A. M. Tompkins, 2003: Systematic errors in the ECMWF forecasting system. Technical Memorandum, No. 442, ECMWF, Shinfield Park, Reading RG 29AX, U. K., 72 pp.

Murphy, A. H., 1988: Skill scores based on the mean square error and their relationships to the correlation coefficient. *Mon. Wea. Rev.*, **116**, 2417–2424.

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

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

Xue, H. L., X. S. Shen, and J. F. Chou, 2015: An online model correction method based on an inverse problem: Part I—Model error estimation by iteration. *Adv. Atmos. Sci.*, **32**(10), 1329–1340, doi: 10.1007/s00376-015-4261-1.

Zheng, F., J. Zhu, R.-H. Zhang, and G.-Q. Zhou, 2006: Ensemble hindcasts of SST anomalies in the tropical Pacific using an intermediate coupled model. *Geophys. Res. Lett.*, **33**, L19604, doi: 10.1029/2006GL026994.

Zheng, F., J. Zhu, H. Wang, and R.-H. Zhang, 2009: Ensemble hindcasts of ENSO events over the past 120 years using a large number of ensembles. *Adv. Atmos. Sci.*, **26**(2), 359–372, doi: 10.1007/s00376-009-0359-7.