Monthly Extended Predicting Experiments with Nonlinear Regional Prediction. Part I: Prediction of Zonal Mean Flow*

CHEN Bomin[1,2] (陈伯民), JI Liren[2] (纪立人), YANG Peicai[2] (杨培才), and ZHANG Daomin[2](张道民)

1 *Shanghai Climate Center, Shanghai Meteorological Bureau, Shanghai* 200030

2 *LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing* 100029

(Received July 14, 2006)

ABSTRACT

Systematic errors have recently been founded to be distinct in the zonal mean component forecasts, which account for a large portion of the total monthly-mean forecast errors. To overcome the difficulty of numerical model, the monthly pentad-mean nonlinear dynamic regional prediction models of the zonal mean geopotential height at 200, 300, 500, and 700 hPa based on a large number of historical data (NCEP/NCAR reanalysis data) were constituted by employing the local approximation of the phase space reconstruction theory and nonlinear spatio-temporal series prediction method. The 12-month forecast experiments of 1996 indicated that the results of the nonlinear model are better than those of the persistent, climatic prediction, and T42L9 model either over the high- and mid-latitude areas of the Northern and Southern Hemispheres or the tropical area. The root-mean-square of the monthly-mean height of T42L9 model was considerably decreased with a change of 30.4%, 26.6%, 82.6%, and 39.4%, respectively, over the high- and mid-latitudes of the Northern Hemisphere, over the high- and mid-latitudes of the Southern Hemisphere, over the tropics and over the globe, and also the corresponding anomaly correlation coefficients over the four areas were respectively increased by 0.306-0.312, 0.304-0.429, 0.739-0.746, and 0.360-0.400 (averagely a relative change of 11.0% over the globe) by nonlinear correction after integration, implying that the forecasts given by nonlinear model include more useful information than those of T42L9 model.

Key words: dynamical extended prediction, zonal flow, persistent prediction, climatic prediction, nonlinear regional prediction, correction

1. Introduction

Although the zonal-mean flow (the zonal-mean component of height) is rather predictable theoretically, a large number of experiments indicated that there is a prominent systematical error in the above component forecasts, accounting for a large portion of the total forecast error (Chou, 1995; Lei and Zhang, 1991; Wang, 1996). That is, the zonal-mean flow should be a common source of the forecast error in different numerical models. A consistent low forecast skill of the monthly dynamical extended prediction has a relation with the prediction of the component by numerical models. Focusing on the difficulty, some studies proposed the solutions such as the modification of the forecast of the zonal-mean flow based on the systematic errors and their spatial and temporal pattern from a large number of forecasts, and on the climatic (or monthly-mean) value or the monthly tendency of the values, which made effects to some extent. While those methods are empirical, they clearly show the importance of the zonal-mean for forecasts.

Owing to the complex nonlinear feature of the atmospheric system, we will try the approach of nonlinear dynamical regional prediction (or nonlinear spatio-temporal series prediction) along with the phase-space reconstruction theory (Chen, 2000; Chen et al., 2002; Packard et al., 1980; Taken, 1981; Essex et al., 1987; Keppenne and Nicolis, 1989; Abarbanel, 1993; Yang, 1996; Yang et al., 2000) to predict the zonal-mean height that is dynamically-significant and leads to a considerable forecast error in current models, and furthermore substitutes the zonal-mean height by the numerical model with the corresponding result by the above nonlinear approach so as to eliminate the systematic forecast error of the numerical model related

*Supported by the National Natural Science Foundation of China under Grant No. 40175013, the National Key Project for Development of Science and Technology (96-908-02-01), and the Project of Chinese Academy of Sciences (ZKCX2-SW-210).

to the zonal mean. In contrast to numerical model, the nonlinear spatio-temporal series prediction describes dynamics of the system implicitly. Based on the phase-space reconstruction theory, the above approach is usually regarded as a dynamical one to some extent although the established predictive equation is still an approximation to the real (or original) dynamical system, by which the nonlinear characteristics of the real system are preserved and at the same time, the nonlinear feedback of physics in the system is thus included partly. Furthermore, in the theory, historical data can be used to the best, which has proven to be important and effective in the short-range climate prediction (Yang, 1996; Gong and Chou, 1999; Cao et al., 2000).

In this paper, the numerical prediction model used is the global spectral model based on ECMWF model and developed further by IAP (Institute of Atmospheric Physics, Chinese Academy of Sciences) (hereafter T42L9; Chen et al., 2004). The data are taken from the NCEP/NCAR reanalysis data of the pentad-mean zonal height from 1960 to 1996.

2. The nonlinear regional prediction of the pentad-mean height

2.1 *Phase space reconstruction theory and nonlinear regional (spatio-temporal series) prediction*

In the 1980s, Packard et al. (1980) and Takens (1981) proposed a concept representing the dynamics of the original dynamical system on the basis of one of its variables. When the spatial dimension m (The parameter m is generally called the embedding dimension and d the attractor dimension of the system. Here the relation $m > d$ is satisfied.) and the related time-delay parameter τ (τ should be divided exactly by observational interval of the series) are reasonably introduced for the temporal series $x(t_n)$(n=1, 2,$\cdots$, N. N is length of the series), and moreover the time-delay ordinate: $x(t), x(t+\tau), \ldots, x(t+(m-1)\tau)$ are set and the m-dimension embedding phase space is reconstructed, the attractor of the original dynamical system may be reestablished as follows:

$$\boldsymbol{X}_m(t_n) = \Big\{x(t_n), x(t_n+\tau), \ldots, x(t_n+(m-1)\tau)\Big\} \quad (n=1,2,\cdots,N-(m-1)\tau), \tag{1}$$

where $\boldsymbol{X}_m(t_n)$ is the nth state vector in embedding space, and $x(t_n)$, $x(t_n+\tau), \cdots, x(t_n+(m-1)\tau)$ are its ordinate (components).

The presentation of the phase space reconstruction theory has greatly promoted the in-practice application of science of chaos and also makes it possible to constitute nonlinear predictive model with observational series. However, one of the problems faced all the time is that the size of observed data is almost not enough for prediction because the data are, in reality, of limited length and do not meet the theoretical requirement. As far as prediction concerned, foreseeing of the coming state of the dynamical system strongly depends on learning of its historically-occurred state, that is to say, the data with short-time history hardly provide a relatively complete description for the state set of the system (usually referred as the ergodicity problem). To alleviate the deficiency of the single-variable series, Essex et al. (1987), Keppenne and Nicolis (1989), Abarbanel (1993) and Yang et al. (2000) successively suggested a multi-variable series (or spatio-temparal series) analytic and predictive method, in which the observed series from various spatial positions were combined, and by which the ergodicity of the reconstructed system was improved in a large number of experiments. In this paper, analogous nonlinear prediction (i.e., regional prediction) will be made through incorporation of the zonal mean series on different positions (i.e., Gauss-latitudes of T42L9 model).

2.2 *Building of the nonlinear regional prediction model of the pentad-mean height*

The key of the regional prediction is the incorporation of observed series on different positions. Nevertheless, such combination is based on the premise that the series made up from different positions are all controlled by one dynamical system. So far, no commonly-accepted technique exists for the judgment. One usual empirical solution is based on whether the nonlinear dynamical invariables derived from different series are notably indiscriminate. In this paper, by using the 36-yr pentad zonal-mean height historical series on all Gauss-latitude of T42L9 model, calculating and comparing their minimum completely-embedding dimension, and analyzing the consistency and similarity of their wavelet transformation (Chen, 2000; Yang, 1995), we divide the nonlinear prediction region into three parts: 20°-70°N, 20°S-20°N, and 20°-70°S. For combination with the prediction of T42L9 model, the predicting length of the nonlinear predictive model over the above three regions is taken as a pentad.

Firstly, we take the $\Delta\bar{h}(j,t_n)$ as the spatio-temporal series of the zonal mean height departure at Guess-latitudes over the domains where $j=1,2,\ldots,J$ (J is equal to 64, the number of meridional points of T42L9), and $n=1,2,\ldots,N$ (N is the length of the series). Assuming that dynamics of the system has been reinstated in a embedding phase space with the dimension m (here the atmosphere is approximately regarded as a dynamical system), we can obtain its state trajectory in the following:

$$\begin{aligned}\boldsymbol{H}(j,t_n)=\big\{&\Delta\bar{h}(j,t_{n-(m-1)\tau}),\ldots,\Delta\bar{h}(j,t_{n-\tau}),\\&\Delta\bar{h}(j,t_n)\big\}\qquad (m-1)\tau+1\leqslant n\leqslant N.\end{aligned}\tag{2}$$

We further take $\boldsymbol{H}(j,t_N)=(\Delta\bar{h}(j,t_{N-(m-1)\tau}),\ldots,\Delta\bar{h}(j,t_{N-\tau}),\Delta\bar{h}(j,t_N)$ as the current state (or initial state) of the system, in which component m of $\boldsymbol{H}(j,t_N)$ is the initial value of the zonal mean height departure (the last pentad value of a month in the paper). Hence the prediction model to be built up is as follows:

$$\boldsymbol{H}(j,t_{N+k})=F_{k,j}(\boldsymbol{H}(j,t_N)),\tag{3}$$

where $k=1,2,\ldots,6$ is the step length of prediction, implying the longest leading time of 6 pentads for the monthly prediction, $\boldsymbol{H}(j,t_{N+k})=\Big\{\Delta\bar{h}(j,t_{N-(m-1)\tau+k}),\ldots,\Delta\bar{h}(j,t_{N-\tau+k}),\Delta\bar{h}(j,t_{N+k})\Big\}$, the coming state of $\boldsymbol{H}(j,t_N)$ for k pentads later, and $F_{k,j}$, the projection relation of the two states.

In practice, only $\Delta\bar{h}(j,t_{N+k})$, the value of $\Delta h(j,t_N)$ k pentads later should be predicted, thus Eq.(3) is replaced by

$$\Delta\bar{h}(j,t_{N+k})=F_{k,j}^{(m)}(\boldsymbol{H}(j,t_N)),\tag{4}$$

where $F_{k,j}^{(m)}$ is the m component of $F_{k,j}$. Hence, the problem is put in how to determine the projection $F_{k,j}^{(m)}$.

In virtue of the local approximation (Chen, 2000; Chen et al., 2002, 2003; Yang, 1996), the adjacent point set is chosen inside the neighborhood $\big\{\boldsymbol{H}_l(j,t_N),l=1,2,\ldots,L\big\}$ and L is the number or size of the adjacent points, then the adjacent point set (or members) inside the neighborhood near the current state is determined, and thus the information (or evolution) in all the adjacent points is used for the construction of the projection $F_{k,j}^{(m)}$, in which $\big\{\boldsymbol{H}_l(j,t_N),l=1,2,\ldots,L\big\}$ refers to all the state vectors $\boldsymbol{H}(r,t_n)$ (here r represents serial number of latitude, and r and n are respectively limited by condition $1\leqslant r\leqslant J$ and $(m-1)\tau<n<N$). In real operation, not the parameter ε but L is controlled.

$$\|\boldsymbol{H}(j,t_N)-\boldsymbol{H}(r,t_N)\|\leqslant\varepsilon.\tag{5}$$

It should be emphasized that, in the above approach, the adjacent points could be chosen from series at different latitudes, very different from the prediction based on the temporal series.

Firstly, the coming state of the current state can be easily given by means of the distribution of $\Big\{\boldsymbol{\Delta}\bar{h}_l(j,t_{N+k}),l=1,2,\ldots,L\Big\}$. The simplest zero-order projection relation is as follows

$$\Delta\bar{h}(j,t_{N+k})=\frac{1}{L}\sum_{l=1}^{L}\boldsymbol{\Delta}\bar{h}_l(j,t_{N+k})\tag{6}$$

or the similar relation to the average on the size of L' (the part of the concentrated-in-distribution number included in L),

$$\Delta\bar{h}(j,t_{N+k})=\frac{1}{L'}\sum_{l=1}^{L}\Delta\bar{h}_l(j,t_{N+k}).\tag{7}$$

In Eq.(7), the idea of cell-cell mapping is refered (Fan, 1996; Fan et al., 1999). Obviously both L and L' must be larger than the parameter m, and $L'<L$ is needed.

If $F_{k,j}^{(m)}$ is defined as a polynomia, the first-order expression (the first-order approximation) of $F_{k,j}^{(m)}$ can be written as

$$\begin{aligned}&\Delta\bar{h}(j,t_{N+k})=\boldsymbol{H}(j,t_N)\boldsymbol{A}+a_0\\&\boldsymbol{A}=(a_1,a_2,\ldots,a_m)^{\mathrm{T}}.\end{aligned}\tag{8}$$

Its scalar presentation is

$$\Delta\bar{h}(j,t_{N+k}) = a_0 + \sum_{p=1}^{m} a_p \Delta\bar{h}(j,t_{N-(p-1)\tau}), \qquad (9)$$

where the coefficients a_0 and a_p will be determined by the following Eq.(10) that presents the evolution of the adjacent points of the current state.

$$\Delta\bar{h}_l(j,t_{N+k}) = a_0 + \sum_{p=1}^{m} a_p \Delta\bar{h}_l(j,t_{N-(p-1)\tau}), \qquad (10)$$

where $\Delta\bar{h}_l(j,t_{N+k})$ and $\Delta\bar{h}_l(j,t_{N-(p-1)\tau})$ ($p = 1,2,\ldots,m$) are both the components of the adjacent state points to the current state.

If taking the second-order relation (the second-order approximation) of $F_{k,j}^{(m)}$, we obtain

$$\begin{aligned}\Delta\bar{h}(j,t_{N+k}) = a_0 &+ \sum_{p=1}^{m} a_p \Delta\bar{h}(j,t_{N-(p-1)\tau}) \\ &+ \sum_{p=1}^{m} a_{pp} \Delta\bar{h}^2(j,t_{N-(p-1)\tau}) \\ &+ \sum_{p=1}^{m}\sum_{q=p+1}^{m-1} a_{pq} \Delta\bar{h}(j,t_{N-(p-1)\tau}) \\ &\cdot \Delta\bar{h}(j,t_{(N-(q-1)\tau)}), \end{aligned} \qquad (11)$$

where the coefficients a_0 and a_p can also be determined in similar manner to the first-order approximation.

In detail, the coefficients of the polynomial (the number of the coefficients is $\boldsymbol{C}_{m+\alpha}^{m}$, and α the order of the polynomial) are derived by the least square method. In operation, we used a square-root solution method (Xu, 1995).

The above predictive equations are usually called the direct approximation, in which $F_k^{(m)}$, the projection relation of the current state with the coming state k pentad later is directly constructed. Another one is the iteration approximation. Namely the corresponding prognostic value is taken as the new "current state" when one-step extrapolation (prediction) is done and the same extrapolation is done again, such extrapolation processes are repeated until the prediction is carried out (Essex et al., 1987), which corresponds to the (k-1)-times compound operation of the projection $F_l^{(m)}$ in the form

$$F_k^{(m)}(\boldsymbol{H}) = \underbrace{F_l^{(m)}(F_l^{(m)}(F_l^{(m)}\ldots))}_{k \text{ times}} \qquad (12)$$

Theoretical studies showed that the second-order polynomial is more accurate than the first-order polynomial and the iteration approximation is more effective than the direct approximation (Abarbanel, 1993; Casdagli, 1989). However, observational series were much more sophisticated than those derived from theoretical chaotic systems (Chen et al., 2002; Yang, 1996). As to the predictive model in the paper, how to choose the type of mapping relations and extrapolation (prediction) manners, and how to determine the parameters in predictive model are required to clarify the problems such as a large number of tests.

Twelve months of 1995 are used as the training cases by which the above nonlinear prediction models are adjusted and all the 36-yr data immediately before the 12 tuning cases (months) are used as the sample dataset for selecting the adjacent points of the current state. Finally, the optimal parameters in the nonlinear model at four levels and the formula of prediction (extrapolation) are determined as a result of a great deal of trainings (Table 1).

2.3 *A comparison of forecasts by nonlinear method with those by persistent and climatic prediction*

The 12-month predicting experiments of 1996, the last year of the above 37, are performed by using the nonlinear model. The following is the comparison of its forecasts with those of persistent forecast and climatic forecast at 500 hPa.

The 12-case-average forecasts of six pentads (Fig. 1) show that the zonal-mean height anomaly correlation coefficients (ACC) of the nonlinear prediction over three regions are mostly higher than those of persistent prediction and the ACC difference increases with the prediction time (steps). The ACCs of the two predictions tend to approach to each other when the former is lower than the latter and few exceptional results are presented.

We can see that the root-mean-square error

Table 1. The parameters of the 500, 700, 300, and 200-hPa pentad zonal-mean height departure nonlinear prediction model over the domain

Level	Domain	m	τ	Projection relation	Prediction manner	Adjacent point size (L, L')
700 hPa, 500 hPa	20°-70°N	6	2, 3	Zero-order	Direct	L=100, L'=50
	20°S-20°N	5	3	First-order	Direct	L=100
	20°-70°S	6	3	Zero-order	Direct	L=100, L'=50
300 hPa	20°-70°N	6	4	Zero-order	Direct	L=120, L'=60
	20°S-20°N	5	4	First-order	Direct	L=80
	20°-70°S	6	4	Zero-order	Direct	L=120, L'=60
200 hPa	20°-70°N	6	4	Zero-order	Direct	L=140, L'=70
	20°S-20°N	5	4	First-order	Direct	L=100
	20°-70°S	6	4	Zero-order	Direct	L=140, L'=70

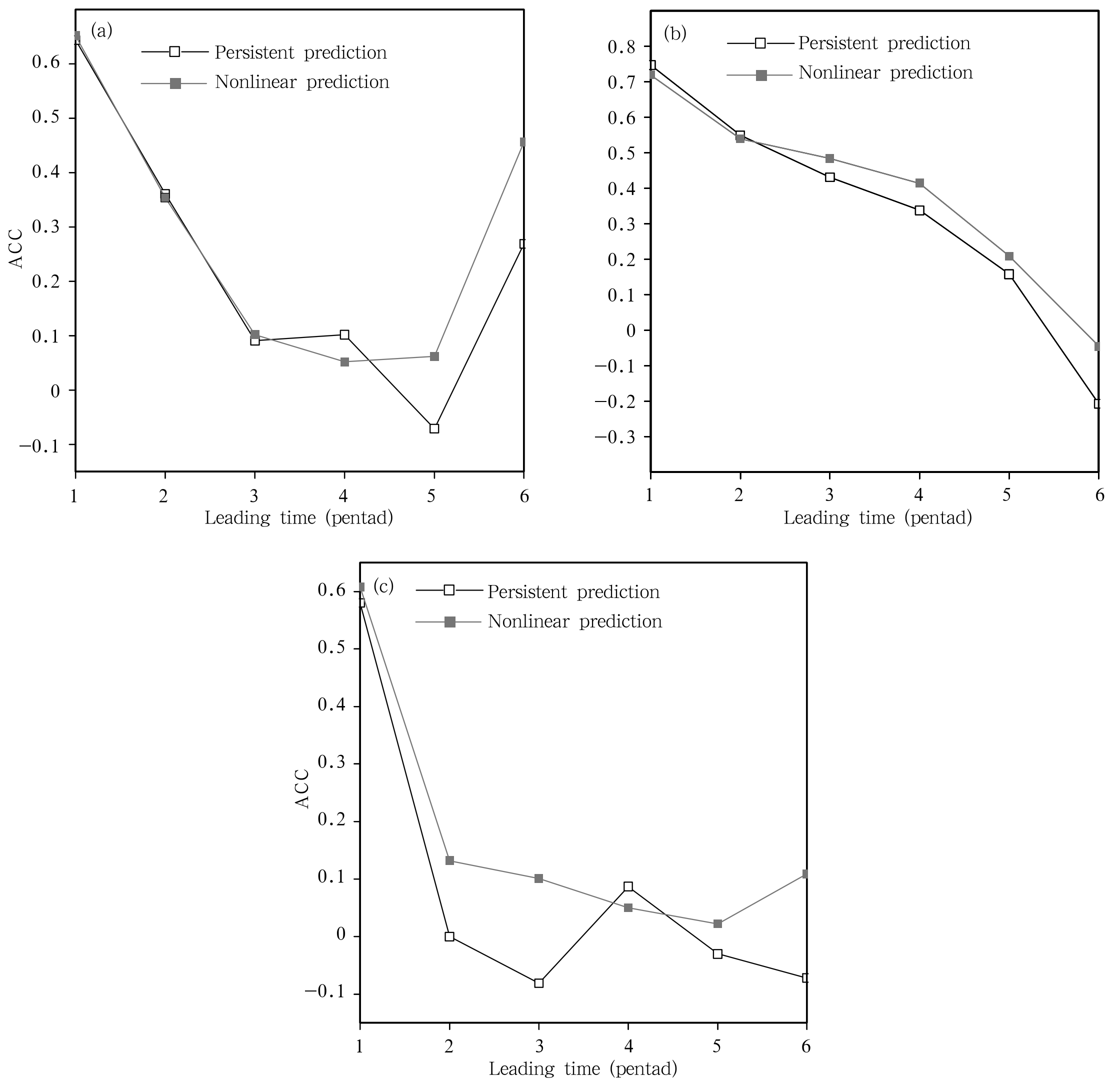

Fig.1. The anomaly-correlation coefficient (ACC) of the pentad zonal-mean height of the persistent and nonlinear forecasts over (a) the Northern Hemisphere (20°-70°N), (b) the Southern Hemisphere (20°-70°S), and (c) the tropics (20°S-20°N) (12-case average of 1996).

(RMSE) of the nonlinear prediction is considerably smaller than that of both the persistent prediction over the mid- and high-latitudes of the Northern and Southern Hemispheres, and the difference of RMSE gets clearer after the 2nd pentad. The RMSE of nonlinear prediction is also smaller than that of climatic prediction, although the difference is not more significant than that with persistent prediction. Over the tropics of both the hemispheres, the RMSEs are small for all the three approaches because the large-scale flow is relatively stable and let alone the differences of their skill.

Synthesizing all the forecast skill above and taking account that the climatic prediction cannot supply the pentad depature of the zonal-mean height, we can see that the nonlinear forecast is better than both of the persistent prediction and the climatic prediction. As far as the different regions are concerned, the above nonlinear method is relatively more effective over the Northern Hemisphere.

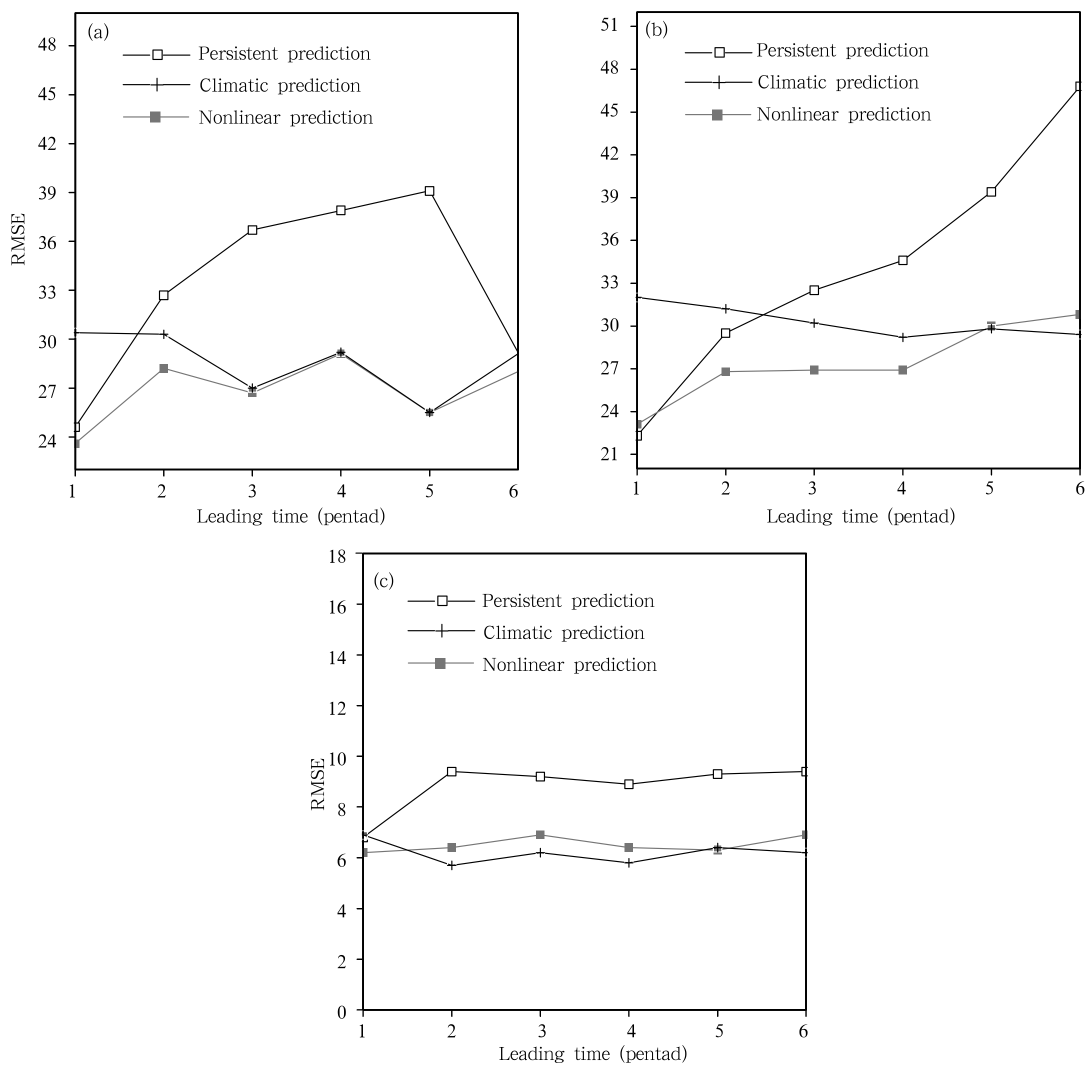

Fig.2. The root-mean-square error (RMSE, unit: gpm) of the pentad zonal-mean height of the persistent, climatic, and nonlinear forecasts over (a) the Northern Hemisphere (20°-70°N), (b) the Southern Hemisphere (20°-70°S), and (c) the tropics (20°S-20°N) (12-case average of 1996).

2.4 *A comparison of forecasts by nonlinear method with those by T42L9 model*

The following is a comparison of the forecast error distribution of the above 12-case-average zonal-mean height of 500 hPa by the nonlinear model and by T42L9 model (Fig.3). We can see that the zonal mean height of T42L9 model is underestimated over a majority of the domains except the north polar region, the reduction of the zonal height reaching a magnitude of -40 to -80 gpm, whereas that of the nonlinear model is very close to the reality, with the error totally less than ±10 gpm over the mid-latitude and tropics. The forecasts on the other levels are mostly analogous to that at 500 hPa (figure omitted).

3. Correction to height forecasted by T42L9 model according to nonlinear region prediction

The aim of predicting the zonal height by the nonlinear approach is to diminish the forecast errors of T42L9 model and then to improve its height forecast skill at each isobar surface (as shown in Fig.3). Hence we replace the 500-hPa zonal heights for pentads 2 to 6 produced with T42L9 model by the corresponding nonlinear results (hereafter post-correction). Concretely, every pentad nonlinear zonal-mean heights are firstly derived from the corresponding departure directly predicted by the nonlinear model and the climatic values, and then the zonal-mean heights of each day starting from the 2nd pentad to the 5th pentad, are substituted with the nonlinear values. We perform a weighed average of the two forecasts to the post-correction of the 5th day for smooth transition.

Consequently, the systematical forecasting error of T42L9 model is significantly reduced (Figs.1 and 2) and at the same time, its height forecast skill after correction is improved considerably. The following will discuss that of the 12-case skill scores of 1996.

3.1 *ACC and RMSE of the height forecasts averaged for day-1 to day-X in relation to the real heights for corresponding periods*

Firstly, we analyze the effects of the post-correction on the ACC of 500-hPa height by T42L9 model for the period of day 1 to day X (X=1, 2,$\cdots$, 30). Figure 4a depicts the 12-case-average differences of ACC with and without correction, which increases with the value of X, implying that the ACC skill after correction increases with the day number X for average. When X=30, those of corresponding monthly-mean (Table 2) can be obtained, i.e., with correction, the changes of ACC are respectively from 0.304 to 0.429 (a relative change of 41%) over the high- and mid-latitudes of the Southern Hemisphere, from 0.739 to 0.746 (a relative change of 1%) over the tropics,

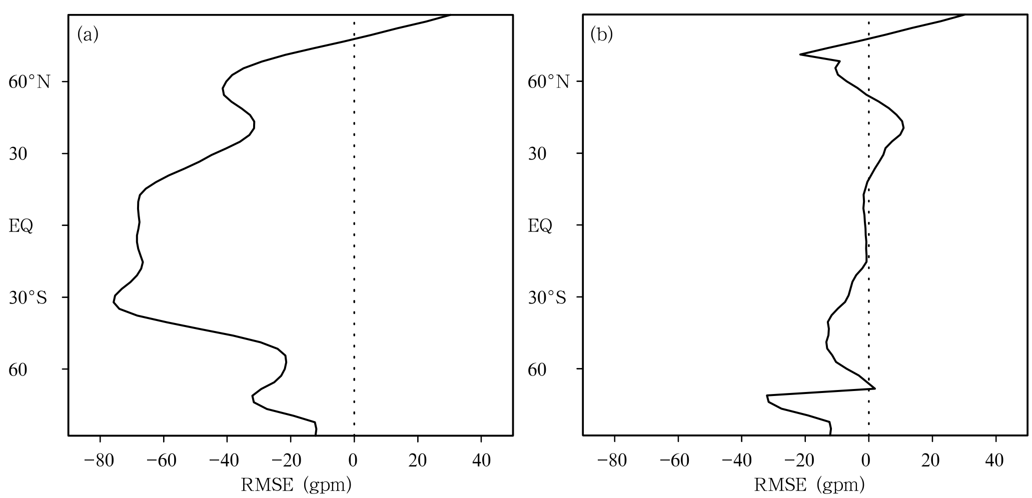

Fig.3. The RMSE (gpm) of the monthly mean zonal height field for 12 cases of 1996 predicted by (a) T42L9 model and (b) the nonlinear model. The initial and actual fields are taken from NCAR/NCEP reanalysis data.

from 0.306 to 0.312 (a relative change of 2.0%) over the high- and mid-latitudes of the Northern Hemisphere, and from 0.36 to 0.400 (a relative change of 11%) over the globe. In sum, as far as the ACC skill was concerned the correction is considerably more effective over the Southern Hemisphere than over the North Hemisphere. Nevertheless, an impartial appraisal of influence of the correction still depends on the change of RMSE skill of height.

Table 2. ACC of days 1-30 averaged 500-hPa height predicted by T42L9 model with and without nonlinear correction (12-case results of 1996)

Month	The Northern Hemisphere (20°-90°N)		The Southern Hemisphere (20°-90°S)		The tropics (20°S-20°N)		The globe (90°S-90°N)	
	without correction	with correction	without correction	with correction	without correction	with correction	without correction	with correction
Jan.	0.212	0.170	0.301	0.555	0.803	0.817	0.273	0.35
Feb.	-0.015	0.126	0.109	0.264	0.761	0.770	0.147	0.244
Mar.	0.264	0.158	0.060	0.426	0.816	0.814	0.241	0.299
Apr.	0.683	0.637	0.340	0.507	0.810	0.812	0.538	0.594
May	0.263	0.266	0.546	0.559	0.731	0.741	0.425	0.469
Jun.	0.242	0.231	0.307	0.336	0.727	0.730	0.397	0.381
Jul.	0.437	0.509	0.118	0.088	0.589	0.611	0.331	0.302
Aug.	0.210	0.389	0.540	0.390	0.634	0.627	0.468	0.435
Sept.	0.487	0.412	0.620	0.548	0.564	0.588	0.535	0.461
Oct.	0.350	0.403	0.170	0.519	0.790	0.786	0.289	0.482
Nov.	0.227	0.199	0.419	0.518	0.826	0.824	0.392	0.418
Dec.	0.310	0.248	0.115	0.441	0.813	0.829	0.280	0.360
Mean	0.306	0.312	0.304	0.429	0.739	0.746	0.360	0.400

The relative change in RMSEs of 500-hPa height for day 1 to day X with and without correction is demonstrated in Fig.4b. It is clear that the RMSEs of height forecasted over the three regions with correction is much smaller than those without correction, and the more the day number X, the more significant the reduction of the RMSEs. The monthly-mean RMSEs over the high- and mid-latitudes of the Northern Hemisphere, the Southern Hemisphere and over the tropics, are respectively reduced from 71.2, 73.0, and 68.5 gpm (without correction) to 52.3, 51.0, and 11.9 gpm (with correction), with corresponding relative changes of 26.5%, 30.3%, and 82.6%, and those over the globe from 71.7 gpm to 43.4 gpm (a relative change of 39.4%, see Fig.3). That is to say, over all the regions related, the RMSEs of 500 hPa height by T42L9 model are remarkably decreased by means of the above nonlinear correction.

As for the different months (cases), over the high- and mid-latitudes of the Northern Hemisphere the ACCs before correction are, roughly by half, larger than those after correction, on average, the latter slightly larger than the former, over the high- and mid-latitudes of the Southern Hemisphere the nine-case ACCs after correction larger than those before correction with an exception to case 7 (July), case 8 (August), and case 9 (September) where the correction made effect more considerably during summer, and over the tropics the ACCs of most cases are also increased through correction although their relative change is small due to the fact that the skill by T42L9 model is large more or less.

The RMSEs of all cases over three regions with correction are smaller than those without correction, and further their difference is notable (Table 3), with a maximum relative change over the tropics. Concerning the different seasons, over the high- and mid-latitudes of the two hemispheres the reduction of RMSE during summer is more considerable than that during winter, and the reverse is over the tropics.

Table 3. A relative decrease ($|\Delta$RMSE/ RMSE(T42)%$|$) of RMSE for days 1-30 averaged 500-hPa height predicted by T42L9 model with nonlinear correction (12-case results of 1996)

Month	The Northern Hemisphere (20°-90°N)	The Southern Hemisphere (20°-90°S)	The tropics (20°S-20°N)	The globe (90°S-90°N)
Jan.	29.4	28.6	80.1	38.5
Feb.	30.4	32.2	84.8	40.3
Mar.	24.5	40.5	86.5	40.2
Apr.	21.0	46.4	84.3	49.4
May	17.6	33.8	79.8	39.7
Jun.	21.7	23.0	80.4	32.7
Jul.	30.6	22.0	75.7	32.7
Aug.	33.9	26.1	76.7	36.9
Sept.	29.8	26.2	83.0	42.7
Oct.	39.8	40.8	86.8	53.0
Nov.	24.8	21.5	87.2	35.8
Dec.	16.7	26.3	84.9	33.7
Mean	26.5	30.3	82.6	39.4

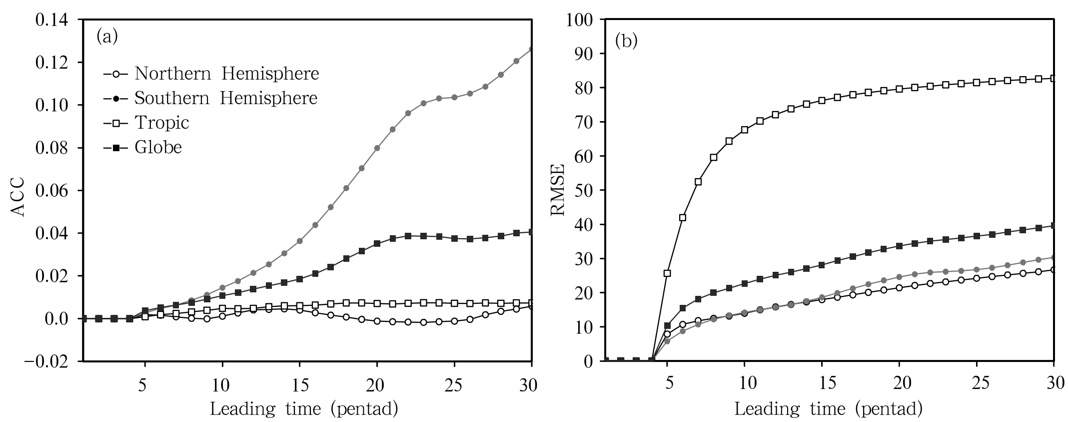

Fig.4. The change of ACC (a) and the relative change of RMSE (b) in 500 hPa height fields predicted by T42L9 model with correction based on the 12-case results of 1996.

3.2 *ACC and RMSE of the height forecasts averaged for day-1 to day-X in relation to the real heights for 30 days*

Most of numerical models are able to present a relatively reasonable prediction at the beginning of integration, and taking average is equivalent to the filtration for prediction, thus the average from day 1 to day X will approach, by degrees, the real 30-day-average at the beginning with the daily-forecasts gradually included in the average (or the value of X increases from 1). However the average tends to drift off the real 30-day-average due to the fact that myriads of errors are more and more introduced with X increasing further. Namely the ACC of the forecast average for day 1 to day X related to the real 30-day-average rises firstly and then falls down little by little with the number of lead day (X). In such distribution of ACC, a parameter $X_{\max}$ exists, which makes the average of the forecasts for day 1 to day X correspond to the maximum value of ACC at $X = X_{\max}$.

From 108-case numerical forecasts, Tracton (1989) argued that $X_{\max}$=8. In addition, a six-case-average numerical result from Zhang et al. (1996) showed that $X_{\max}$=13 over the Northern Hemisphere, and $X_{\max}$=8 over the Southern Hemisphere. In the paper, $X_{\max}$ of each case forecast was firstly noted, and then $X_{\max}$ of 12-case average was reckoned that

$X_{\max}$=12.8 over the Northern Hemisphere, and $X_{\max}$=9.1 over the Southern Hemisphere (Table 4).

Contrary to ACC, in the distribution of RMSE related to the real 30-day-average, there exists a parameter $X_{\min}$, which the average of the forecasts for day 1 to day X match the minimum value of RMSE provided that X=$X_{\min}$. We obtained that $X_{\min}$=18 over the Northern Hemisphere, and $X_{\min}$=11 over the Southern Hemisphere. Obviously, the larger the values of $X_{\max}$ and $X_{\min}$ are, the better the performance of model will be.

Table 4. $X_{\max}$ of 500-hPa height predicted by T42L9 model over the Northern Hemisphere (20°-90°N), the Southern Hemisphere (20°-90°S), and the globe (90°S-90°N) with and without correction (12-case results of 1996)

Month	The Northern Hemisphere (20°-70°N)		The Southern Hemisphere (20°-70°S)		The globe (70°S-70°N)	
	without correction	with correction	without correction	with correction	without correction	with correction
Jan.	10	11	5	8	8	10
Feb.	1	24	3	4	1	1
Mar.	3	3	5	5	3	3
Apr.	9	8	21	22	20	21
May	6	6	4	4	5	5
Jun.	12	12	4	4	4	4
Jul.	10	10	6	6	7	7
Aug.	11	13	11	10	10	10
Sept.	28	27	10	21	7	7
Oct.	15	30	17	23	16	23
Nov.	19	10	14	14	14	14
Dec.	30	30	9	13	17	26
Mean	12.8	15.3	9.1	11.2	9.3	10.8

Based on a case-by-case counting of $X_{\max}$, we found that $X_{\max}$ values of most cases are increased to some different extent.

The 12-case-average results further showed that the values of $X_{\max}$ with the correction are increased by factors 2.5, 2.1, and 1.5 respectively over the three regions (Table 4). $X_{\min}$ with correction is considerably changed, being from 18 to 25 over the Northern Hemisphere and from 11 to 30 over the Sorthern Hemisphere with no exception.

The promotion of $X_{\max}$ and $X_{\min}$ after correction along with the alteration of the ACC and RMSE skill implies that the forecasts presented by nonlinear model include more useful information than those of the numerical model.

4. Concluding remarks

The paper applies the nonlinear spatio-temporal series prediction method associated with the phase-space reconstruction theory to the zonal-mean flow (height) to overcome the universal forecast error of the component in numerical prediction model. The monthly pentad-mean nonlinear dynamic prediction models of the zonal mean geopotential height at 200, 300, 500, and 700 hPa over the Northern and Southern Hemispheres as well as the tropics are firstly constructed on the basis of 36-yr historical data (NCEP/NCAR reanalysis data), and the parameters of the nonlinear models are then determined through a large number of experiments.

The 12-month pentadly forecasting experiments of 1996 have demonstrated that the results of the nonlinear model are better than those of persistent, climatic prediction and T42L9 model, implying that section division of nonlinear prediction and the values of its parameters are overall reasonable. That is to say, the nonlinear model can be used to hindcast the monthly pentad-mean zonal-mean height departure.

Not only the systematical forecast error in the zonal mean height by the numerical model was diminished, but also its forecast skill of 500-hPa height was altered by nonlinear correction after integration, in this way, ACCs over the high- and mid-latitudes of the Southern Hemisphere, over the tropics, over the high- and mid-latitudes of the Northern Hemisphere and over the globe were increased respectively from 0.304 to 0.429, from 0.739 to 0.746, from 0.306 to 0.312, and from 0.360 to 0.400, in turn with a relative change of 41%, 1.0%, 0.2%, and 11%, and the corresponding RMSEs reduced with a relative change of 26.5%, 82.6%, 30.3%, and 39.4%. It implies that the forecasts given by nonlinear model include indeed more useful information than those of T42L9 model.

REFERENCES

Abarbanel, H. D. I., 1993: The analysis of observed chaotic data in physical system. *Rev. Modern Phys.*, **65**(4), 1331-1392.

Baumhefner, D. B., 1996: Numerical extended-range prediction: Forecast skill using a low-resolution climate model. *Mon. Wea. Rev.*, **124**, 1965-1980.

Bian Jianchun, 1999: Nonlinearly statistical analysis and prediction of some atmospheric processes. Ph.D dissertation of Peking University, 6-7. (in Chinese)

Cao Hongxing, Feng Guolin, and Wei Fengying, 2000: Modeling of a regional climate prediction. *Acta Meteorologica Sinica*, **58**(2), 159-167. (in Chinese)

Casdagli, M., 1989: Nonlinear prediction of chaotic time series. *Phys. D.*, **35**, 335-356.

Chen Bomin, 2000: An approach to improving the skill of dynamic extended prediction. Ph.D dissertation of Institute of Atmospheric Science, Chinese Academy of Sciences. (in Chinese)

Chen Bomin, Ji Liren, Yang Peicai, et al., 2002: Systematic errors of zonal-mean flow in dynamical monthly prediction and its improvement. *Advances in Atmospheric Sciences*, **20**(1), 17-27.

Chen Bomin, Ji Liren, Yang Peicai, et al., 2003: An approach to improving the skill of dynamic-extended (monthly) prediction. *Chinese Science Bulletin*, **48**(7), 696-703.

Chen Bomin, Ji Liren, Yang Peicai, et al., 2006: Monthly extended predicting experiments with non-linear regional prediction. Part II: Improvement of wave component prediction. *Acta Meteorologica Sinica*, **20**(3), 295-305

Chou Jifan, 1995: *Long-term Numerical Weather Prediction (revision)*. China Meteorological Press, Beijing, 353 pp. (in Chinese)

Essex, C., T. Lookman, and M. A. H. Nerenberg, 1987: The climate attractor over short timescales. *Nature*, **326**, 64-66.

Fan Xingang, 1996: Global study on climate predictability and prediction approach. Ph.D dissertation of Lanzhou University. (in Chinese)

Fan Xingang, Zhang Hongliang, and Chou Jifan, 1999: Global study on climate predictability. *Acta Meteorologica Sinica*, **57**(2), 190-197. (in Chinese)

Gong Jiandong, 1999: The study on conjoining statistical with dynamical extended-range forecast. Ph.D dissertation of Lanzhou University. (in Chinese)

Gong Jiandong and Chou Jifan, 1999: The theories and methods of utilizing historical data in numerical weather forecast. *Plateau Meteorology*, **18**(3), 392-399. (in Chinese)

Keppenne, C. L., and C. Nicolis, 1989: Global properties and local structure of the weather attractor over western Europe. *J. Atmos. Sci.*, **46**, 2356-2370.

Lei Zhaochong and Zhang Jijia, 1991: *Spectral method in numerical model*. Chinese Meteorological Press, Beijing, 1-17. (in Chinese)

Li Weijing, 2000: Progress of the monthly dynamical extended-range prediction. *Study on Monthly Dynamic Extended-range Prediction*. Li Weijing and Ji Liren, Eds., China Meteorological Press, Beijing, 1-9. (in Chinese)

Packard, N. H., J. P. Cretchfield, J. D. Farmer, and R. S. Shaw, 1980: Geometry from a time series. *Phys. Rev. Lett.*, **45**, 712-715.

Saha, S., 1992: Response of the NMC MRF model to systematic error correction with integration. *Mon. Wea. Rev.*, **120**(2), 345-360.

Taken, F., 1981: Detecting strange attractors in turbulence. *Dynamical Systems and Tuebulence*. L. S. Young, Eds., Springer-Verlag, 366-381.

Wang Shaowu, 1996: Progress and status of short-range climate prediction. *Study on Climate Prediction*. China Meteorological Press, Beijing, 1-17.

Xu Shiliang, 1995: *Common algorithm for computer (revision)*. Qinghua University Press, Beijing, 26-28. (in Chinese)

Yang Peicai, 1996: Field-temporal series, spatio-temporal chaos and regional prediction −A new stage of chaotic temporal series analysis. *Precedence and Prospect of Model Atmospheric Sciences*. China Meteorological Press, Beijing, 189-193. (in Chinese)

Yang, P. C., X. J. Zhou, and J. C. Bian, 2000: A nonlinear regional prediction experiment on a short-range climatic process of the atmospheric ozone. *J. Geophys. Res.*, **105**(D10), 12253-12258.

Zhang Daomin, Li Jinlong, Ji Liren, Huang Boyin, Wu Wanli, Chen Jiabin, and Song Zhengshan, 1995: A global spectral model and test of its performance. *Adv. Atmos. Sci.*, **12**(1), 67-78.

Zhang Daomin, Ji Liren, and Li Jinlong, 1997: Seasonal variability and systematical error in dynamical extended-range forecast experiments with the IAP global spectral model. *Chinese Journal of Atmospheric Science*, **21**(4), 141-150. (in Chinese)