



## 论文

## 常数型最优强迫在校正预报模式中的作用

封凡<sup>①②</sup>, 段晚锁<sup>①\*</sup>

① 大气科学和地球流体力学数值模拟国家重点实验室, 中国科学院大气物理研究所, 北京 100029;

② 中国科学院大学, 北京 100049

\* 联系人, E-mail: duanws@lasg.iap.ac.cn

收稿日期: 2012-03-23; 接受日期: 2012-11-07

国家基础研究发展计划(编号: 2012CB955202)、中国科学院知识创新工程重要方向项目(编号: KZCX2-YW-QN203)和国家自然科学基金(批准号: 41176013)资助

**摘要** 采用著名的 Lorenz 63 模式, 数值研究了常数型最优强迫在校正数值模式中的作用. 结果表明, 当数值模式仅考虑由于参数误差导致的随状态变量发展变化的模式误差时, 在数值模式倾向方程叠加常数型最优强迫能够很好地抵消该类模式误差对预报结果的影响; 当数值模式未考虑观测中依赖于时间的随机过程时, 常数型最优强迫也可以较好地抵消由随机过程导致的模式误差的影响. 实际情形中, 数值模式预报结果同时受到由随机过程和参数不确定性导致的模式误差及其相互作用的影响. 结果表明, 常数型最优强迫方法同样能够在很大程度上抵消该类混合型模式误差对预报结果的影响. 综上所述, 即使模式物理过程产生的模式误差是依赖于时间变化的, 在模式中叠加常数型最优强迫校正模式的方法也可以在很大程度上抵消模式误差对预报结果的影响. 常数型最优强迫方法可能是一个较好的校正模式和改进模式预报技巧的方法.

**关键词**可预报性  
预报误差  
模式误差  
最优强迫

预报结果不确定性(预报误差)的研究是数值天气预报和气候预测研究的核心内容. 预报结果不确定性通常是由初始误差和模式误差共同导致的. 为了分别研究初始误差和模式误差在产生预报结果不确定性中的作用, Lorenz<sup>[1]</sup>将可预报性问题分为第一类可预报性问题和第二类可预报性问题. 前者主要研究由初始误差引起的预报结果的不确定性, 后者则研究由模式误差产生的预报结果不确定性. 随着非线性理论<sup>[2,3]</sup>和气象资料、模式<sup>[4,5]</sup>的发展, 初始误差对可预报性的影响已经得到了广泛的关注<sup>[6-10]</sup>, 许多方法被先后提出, 如条件非线性最优扰动(CNOP)<sup>[11]</sup>、四维变分资料同化<sup>[12]</sup>和集合卡曼滤波<sup>[13]</sup>

方法等. 这些方法都是针对初始场的不确定性而提出的用于研究初始场不确定性的有效方法. 这些工作的一个共同点就是假定预报模式是准确的, 预报误差完全源自初始误差, 不存在模式误差. 然而, 在实际情形中, 模式方程和控制参数都不可能准确给出, 因而导致模式误差对模式的预报结果产生不可避免的负面影响. 因此, 对模式误差的研究和控制具有重要的实际意义.

大气、海洋、陆面及其耦合系统都是典型的非线性系统, 这使得任何数值模式的预报能力都存在一定的极限. 为使模式预报尽可能达到可预报的极限, 气象学家们想方设法努力减小预报模式的误差. 目

中文引用格式: 封凡, 段晚锁. 常数型最优强迫在校正预报模式中的作用. 中国科学: 地球科学, 2013, 43: 209-219

英文引用格式: Feng F, Duan W S. The role of constant optimal forcing in correcting forecast models. Science China: Earth Sciences, 2013, 56: 434-443, doi: 10.1007/s11430-012-4568-z

前关于减小模式误差所作的努力主要集中在提高模式的分辨率、改进物理过程的参数化方案、通过不同方法抵消数值模式的缺陷以及运用精度更高的计算机技术等方面. 文献[14]对国际上主要的抵消模式误差的方法做了评述, 并指出了这些方法的优缺点, 对我们进一步应用这些方法给予了指导.

D'Andrea 和 Vautard<sup>[15]</sup>建议, 假定某时段的观测值是已知的, 那么可以通过在模式倾向方程中添加一个合适的常数型外强迫用于校正模式, 使其模拟的结果与该观测最接近, 从而使得校正的模式有更好的预报结果. Roads<sup>[16]</sup>及 Vannitsem 和 Toth<sup>[17]</sup>也用类似的方法考虑模式误差的抵消. 事实上, 如果模式只存在不随时间变化的模式系统误差, 我们相信可以通过在数值模式叠加合适的常数型外强迫抵消模式误差的影响. 然而, 模式误差通常既包含不随时间变化的模式系统误差, 也包含随时间变化的其它类型的模式误差. 那么, 在此情形下, 尤其是模式存在随时间变化的模式误差时, 常数型外强迫能否有效抵消模式误差的影响, 提高模式的预报技巧? 本文将着重阐述这方面的问题.

## 1 常数型最优强迫

考虑非线性偏微分方程(组):

$$\begin{cases} \frac{\partial \mathbf{u}}{\partial t} = F(\mathbf{u}, t), \\ \mathbf{u}|_{t=0} = \mathbf{u}_0, \end{cases} \quad (1.1)$$

其中,  $\mathbf{u}(\mathbf{x}, t) = [u_1(\mathbf{x}, t), u_2(\mathbf{x}, t), \dots, u_n(\mathbf{x}, t)]$ ,  $\mathbf{x} \in \Omega \subset \mathbf{R}^n$ ,  $t < +\infty$  为状态向量. 方程(1.1)中的  $F$  是一个非线性算子,  $\mathbf{u}_0$  为向量  $\mathbf{u}$  的初始状态. 假定用方程(1.1)来预报大气、海洋等流体的运动, 但存在模式误差. 设  $M_\tau$  是方程(1.1)的传播算子, 即对于给定的初值  $\mathbf{u}_0$ , 方程(1.1)在预报时刻  $\tau$  的预报值为

$$\mathbf{u}(\mathbf{x}, \tau) = M_\tau(\mathbf{u}_0). \quad (1.2)$$

如果模式的初值是准确的, 那么由模式误差导致的预报误差可以定义为

$$E_\tau = \|M_\tau(\mathbf{u}_0^r) - \mathbf{u}_\tau^r\|, \quad (1.3)$$

其中  $\mathbf{u}_0^r$  是初始时刻状态的真值,  $\mathbf{u}_\tau^r$  是  $\tau$ 时刻状态的真值,  $\|\cdot\|$  为用于度量预报误差大小的范数. 大气、海洋运动的真值不可能确切的知道, 本文将把观测值

看作是真实值  $\mathbf{u}_\tau^r$  足够精确的近似<sup>[18]</sup>. 设 0 和  $\tau$ 时刻状态的观测分别是  $\mathbf{u}_0^{obs}$ ,  $\mathbf{u}_\tau^{obs}$ , 那么由模式误差导致的预报误差可以近似写成

$$E_\tau = \|M_\tau(\mathbf{u}_0^{obs}) - \mathbf{u}_\tau^{obs}\|. \quad (1.4)$$

在下文中所讨论的由模式误差导致的预报误差均由(1.4)式定义.

对于一个给定的时刻  $T$ , 设该时刻的观测值为  $\mathbf{u}_T^{obs}$ , 初始时刻的观测为  $\mathbf{u}_0^{obs}$ , 则由方程(1.1)的模式误差导致的预报误差是  $E_T$ . 引言已经提到, 为了抵消由模式误差导致的预报误差对预报结果的影响, D'Andrea 和 Vautard<sup>[15]</sup>通过在原预报模式(1.1)上叠加一个不依赖时间变化的强迫项  $\mathbf{f}(\mathbf{x})$ (也可称为模式扰动)来校正模式, 即

$$\begin{cases} \frac{\partial \mathbf{u}}{\partial t} = F(\mathbf{u}, t) + \mathbf{f}(\mathbf{x}), \\ \mathbf{u}|_{t=0} = \mathbf{u}_0. \end{cases} \quad (1.5)$$

那么, 究竟叠加怎样的  $\mathbf{f}(\mathbf{x})$  可以使得模式有最好的预报效果? 事实上, 可以考虑将该类问题转化为一类非线性优化问题<sup>[15,17]</sup>, 即选择  $\mathbf{f}(\mathbf{x})$  使模式预报结果在预报时刻与已有的观测最接近. 假定  $M_T^f$  是方程(1.5)的从 0 时刻到  $T$  时刻的离散传播算子, 那么通过求解以下无约束优化问题

$$J(\mathbf{f}^*) = \min_f J(\mathbf{f}), \quad (1.6)$$

其中,

$$J(\mathbf{f}) = \|M_T^f(\mathbf{u}_0^{obs})(\mathbf{f}) - \mathbf{u}_T^{obs}\|, \quad (1.7)$$

所得到的  $\mathbf{f}^*$  就是使得模式在预报时刻的预报结果与已有观测最接近的模式扰动, 这里将其称为常数型最优强迫. 事实上, 我们也可以根据研究问题的不同, 定义其他类型的目标函数. 例如为了考察模式对所关心天气或气候事件的演变过程的预报能力, 我们可以考虑刻划模式在优化时间窗口内多个时刻的预报结果与已有观测最接近的目标函数. 总之, 计算常数型最优强迫的目标函数要依据物理问题而确定.

根据定义可以知道, 当采用叠加了常数型最优强迫  $\mathbf{f}^*$  的新模式方程(1.5)预报时, 在给定的优化时间窗口  $[0, T]$  的预报误差必然不会超过采用原模式时的预报误差. 然而, 我们通常要对未来时刻的未知结果进行预报, 即要对优化时间窗口以外的预报时刻进行预报, 在此情形下, 常数型最优强迫究竟可以在多

大程度上改进模式的预报效果? 另外, 在数值模式中的模式误差(记为  $R$ )一般是随时间变化的, 而我们用于抵消模式误差影响的初始模式扰动, 即常数型最优强迫是常数型的, 那么, 常数型最优强迫能否抵消或在多大程度上可以抵消随时间变化的模式误差? 在下文中, 我们将用 Lorenz 63 模型探讨上述问题.

由上述讨论可知, 当给定某一度量范数时, (1.6)和(1.7)式定义了一个无约束优化问题,  $\mathbf{f}^*$ 为目标函数极小值在空间相对应的点, 它可以通过 L-BFGS 算法<sup>[19]</sup>求解. 该算法的主要思想是梯度最速下降法, 其中的一个核心问题是求解目标函数关于外强迫的梯度. 国际上计算目标函数关于初始扰动的梯度常常使用伴随方法<sup>[20]</sup>, 那么我们是否也可以用伴随方法获得目标函数关于外强迫的梯度? 我们将在下节探讨该问题.

## 2 常数型最优强迫的求解

常数型最优强迫的求解是一个无约束优化问题. 在大规模的优化算法中目标函数关于初始扰动的梯度往往是由伴随方法得到的<sup>[20]</sup>, 国际上已有很多研究通过伴随模式求解目标函数关于初始扰动的梯度, 也建立了很多模式, 如二维准地转模式、Zebiak-Cane 模式<sup>[21]</sup>、MM5 模式、WRF 模式等的伴随模式. 常数型最优强迫也可以由优化算法通过梯度的计算而求得. 但需要注意的是, 常数型最优强迫的计算需要的是目标函数关于模式扰动  $\mathbf{f}$  的梯度, 与目标函数关于初始扰动的梯度有所不同. 事实上, 目标函数关于模式扰动  $\mathbf{f}$  的梯度可以转化成目标函数关于分量增加的初始扰动向量的梯度. 下面我们将探讨如何通过伴随方法求解梯度来计算常数型最优强迫.

为方便起见, 定义函数:

$$J_1(\mathbf{f}) = \frac{1}{2} \left\| M_T^f(\mathbf{u}_0^{obs})(\mathbf{f}) - \mathbf{u}_T^{obs} \right\|_0^2, \quad (2.1)$$

将其写成内积形式:

$$J_1(\mathbf{f}) = \frac{1}{2} \langle \mathbf{u}'(\tau), \mathbf{u}'(\tau) \rangle, \quad (2.2)$$

其中,  $\langle \bullet \rangle$ 表示内积,  $\mathbf{u}'(\tau)$ 表示由于模式扰动  $\mathbf{f}$  引起的模式预报与观测的误差. 那么, 通过(1.6)和(1.7)定义的常数型最优强迫可以通过计算  $J_1(\mathbf{f})$  的最小值求得.

$J_1(\mathbf{f})$ 的一阶变分为

$$\delta J_1 = \langle \mathbf{u}'(\tau), \delta \mathbf{u}'(\tau) \rangle = \left\langle \frac{\partial J_1}{\partial \mathbf{f}}, \delta \mathbf{f} \right\rangle, \quad (2.3)$$

另外,  $\delta \mathbf{f}$  满足切线性模式:

$$\begin{cases} \frac{\partial \delta \mathbf{u}'}{\partial t} = \frac{\partial F(\mathbf{u})}{\partial \mathbf{u}} \delta \mathbf{u}' + \delta \mathbf{f}(\mathbf{x}), \\ \frac{\partial \delta \mathbf{f}}{\partial t} = 0, \\ \delta \mathbf{u}'|_{t=0} = \delta \mathbf{u}'(0) = 0, \\ \delta \mathbf{f}|_{t=0} = \delta \mathbf{f}(\mathbf{x}). \end{cases} \quad (2.4)$$

引入两个 Lagrange 乘子  $\lambda_1$  和  $\lambda_2$ , 我们可以得到

$$\begin{aligned} \delta J_1 = \langle \mathbf{u}'(\tau), \delta \mathbf{u}'(\tau) \rangle & - \int_0^\tau \left\langle \lambda_1(t), \frac{\partial \delta \mathbf{u}'}{\partial t} - \frac{\partial F(\mathbf{u})}{\partial \mathbf{u}} \delta \mathbf{u}' - \delta \mathbf{f} \right\rangle dt \\ & - \int_0^\tau \langle \lambda_2(t), \frac{\partial \delta \mathbf{f}}{\partial t} \rangle dt. \end{aligned} \quad (2.5)$$

由分部积分可得

$$\begin{aligned} \int_0^\tau \langle \lambda_1(t), \frac{\partial \delta \mathbf{u}'}{\partial t} \rangle dt & = \int_0^\tau \frac{\partial}{\partial t} \langle \lambda_1(t), \delta \mathbf{u}' \rangle dt \\ & - \int_0^\tau \langle \frac{\partial \lambda_1(t)}{\partial t}, \delta \mathbf{u}' \rangle dt \\ & = \langle \lambda_1(\tau), \delta \mathbf{u}'(\tau) \rangle - \langle \lambda_1(0), \delta \mathbf{u}'(0) \rangle \\ & - \int_0^\tau \langle \frac{\partial \lambda_1(t)}{\partial t}, \delta \mathbf{u}' \rangle dt \\ & = \langle \lambda_1(\tau), \delta \mathbf{u}'(\tau) \rangle - \int_0^\tau \langle \frac{\partial \lambda_1(t)}{\partial t}, \delta \mathbf{u}' \rangle dt. \end{aligned}$$

$$\begin{aligned} \int_0^\tau \langle \lambda_2(t), \frac{\partial \delta \mathbf{f}}{\partial t} \rangle dt & = \int_0^\tau \frac{\partial}{\partial t} \langle \lambda_2(t), \delta \mathbf{f} \rangle dt \\ & - \int_0^\tau \langle \frac{\partial \lambda_2(t)}{\partial t}, \delta \mathbf{f} \rangle dt \\ & = \langle \lambda_2(\tau), \delta \mathbf{f}(\tau) \rangle - \langle \lambda_2(0), \delta \mathbf{f}(0) \rangle \\ & - \int_0^\tau \langle \frac{\partial \lambda_2(t)}{\partial t}, \delta \mathbf{f} \rangle dt \\ & = \langle \lambda_2(\tau), \delta \mathbf{f} \rangle - \langle \lambda_2(0), \delta \mathbf{f} \rangle \\ & - \int_0^\tau \langle \frac{\partial \lambda_2(t)}{\partial t}, \delta \mathbf{f} \rangle dt. \end{aligned}$$

从而  $\delta J_1$  可以写成

$$\begin{aligned} \delta J_1 = \int_0^\tau \langle \frac{\partial \lambda_1(t)}{\partial t}, \delta \mathbf{u}' \rangle dt & + \langle \mathbf{u}'(\tau) - \lambda_1(\tau), \delta \mathbf{u}'(\tau) \rangle \\ & + \int_0^\tau \langle \lambda_1(t), \left[ \frac{\partial F(\mathbf{u})}{\partial \mathbf{u}} \right] \delta \mathbf{u}' \rangle dt + \int_0^\tau \langle \lambda_1(t), \delta \mathbf{f} \rangle dt \end{aligned}$$

$$\begin{aligned}
 & + \int_0^\tau \left\langle \frac{\partial \lambda_2(t)}{\partial t}, \delta \mathbf{f} \right\rangle dt + \langle 0 - \lambda_2(\tau), \delta \mathbf{f} \rangle + \langle \lambda_2(0), \delta \mathbf{f} \rangle \\
 = & \int_0^\tau \left\langle \frac{\partial \lambda_1(t)}{\partial t}, \delta \mathbf{u}' \right\rangle dt + \langle \mathbf{u}'(\tau) - \lambda_1(\tau), \delta \mathbf{u}'(\tau) \rangle \\
 & + \int_0^\tau \left\langle \left[ \frac{\partial F(\mathbf{u})}{\partial \mathbf{u}} \right]^* \lambda_1(t), \delta \mathbf{u}' \right\rangle dt \\
 & + \int_0^\tau \langle \lambda_1(t), \delta \mathbf{f} \rangle dt + \int_0^\tau \left\langle \frac{\partial \lambda_2(t)}{\partial t}, \delta \mathbf{f} \right\rangle dt \\
 & + \langle 0 - \lambda_2(\tau), \delta \mathbf{f} \rangle + \langle \lambda_2(0), \delta \mathbf{f} \rangle \\
 = & \int_0^\tau \left\langle \frac{\partial \lambda_1(t)}{\partial t} + \left[ \frac{\partial F(\mathbf{u})}{\partial \mathbf{u}} \right]^* \lambda_1(t), \delta \mathbf{u}' \right\rangle dt \\
 & + \langle \mathbf{u}'(\tau) - \lambda_1(\tau), \delta \mathbf{u}'(\tau) \rangle + \\
 & + \int_0^\tau \left\langle \frac{\partial \lambda_2(t)}{\partial t} + \lambda_1(t), \delta \mathbf{f} \right\rangle dt + \langle 0 \\
 & - \lambda_2(\tau), \delta \mathbf{f} \rangle + \langle \lambda_2(0), \delta \mathbf{f} \rangle, \tag{2.6}
 \end{aligned}$$

其中“[ ]<sup>\*</sup>”表示伴随算子, 而且在计算过程中已经考虑到  $\delta \mathbf{u}'(0) = 0$ . 通过对比(2.3)式和(2.6)式, 可知

$$\frac{\partial J_1}{\partial \mathbf{f}} = \lambda_2(0). \tag{2.7}$$

且(2.6)式中的  $\lambda_1(t)$  和  $\lambda_2(t)$  满足:

$$\begin{cases}
 \frac{\partial \lambda_1}{\partial t} + \left[ \frac{\partial F(\mathbf{u})}{\partial \mathbf{u}} \right]^* \lambda_1 = 0, \\
 \frac{\partial \lambda_2}{\partial t} + \lambda_1 = 0, \\
 \lambda_1|_{t=\tau} = \mathbf{u}'(\tau), \\
 \lambda_2|_{t=\tau} = 0.
 \end{cases} \tag{2.8}$$

(2.8)式是方程组(2.4)的伴随方程. 通过从预报时刻  $\tau$  到 0 时刻反向积分(2.8)式, 可以得到梯度  $\partial J_1 / \partial \mathbf{f}$ , 在此基础上, 我们可以通过该梯度用优化算法计算常数型最优强迫(此文中, 我们采用的是 L-BFGS 优化算法).

我们注意到, 伴随模式方程:

$$\begin{cases}
 \frac{\partial \lambda_1}{\partial t} + \left[ \frac{\partial F(\mathbf{u})}{\partial \mathbf{u}} \right]^* \lambda_1 = 0, \\
 \lambda_1|_{t=\tau} = \mathbf{u}(\tau).
 \end{cases} \tag{2.9}$$

包含于伴随模式(2.8)中, 是数值模式关于初始扰动的伴随模式. 前面已经提到, 国际上已经建立了很多模式的关于初始扰动的伴随模式. 从方程(2.8)可以看到, 数值模式关于模式扰动的伴随模式(2.8)可以建立在数值模式关于初始扰动的伴随模式(2.9)的基

础上. 事实上, 我们只要在伴随模式(2.9)之后添加一行关于  $\frac{\partial \lambda_2}{\partial t} + \lambda_1 = 0$  程序, 即可得到数值模式关于模式扰动的伴随模式.

### 3 常数型最优强迫在校正预报模式中的作用

我们用 Lorenz 63 系统来分析常数型最优强迫在校正模式中的作用. Lorenz 63 系统方程为

$$\begin{cases}
 \frac{dx}{dt} = -\sigma x + \sigma y, \\
 \frac{dy}{dt} = rx - y - xz, \\
 \frac{dz}{dt} = -bz + xy.
 \end{cases} \tag{3.1}$$

$\sigma, r, b$  是模式参数, 取值为  $\sigma=10, r=28, b=8/3$ . Lorenz 63 模式是混沌系统, 常常被用来测试可预报性研究方法的有效性.

在这一节中, 我们主要考虑模式误差对预报结果的影响, 因此在以下的讨论中, 均假设初值是准确的. 上述已经提到, 模式误差通常是依赖时间变化的. 假定模式误差用  $R$  表示, 根据 D'Andrea 和 Vautard<sup>[15]</sup> 的研究结果,  $R$  可以剖分为

$$R = \bar{R} + R(\mathbf{u}) + R'. \tag{3.2}$$

其中,  $\bar{R}$  代表不依赖时间变化的模式误差, 即模式系统误差;  $R(\mathbf{u})$  代表与状态变量  $\mathbf{u}$  有关的模式误差,  $R'$  则代表不依赖于  $\mathbf{u}$  但随时间变化的部分, 它通常被看作是随机过程导致的不确定性. 我们有理由相信, 常数型最优强迫在抵消模式系统误差, 即不随时间变化的模式误差方面必然有好的作用. 在这篇文章中, 我们期望研究常数型最优强迫在抵消与状态变量有关的模式误差以及依赖时间变化随机过程导致的模式误差中的作用.

#### 3.1 模式只存在与状态变量有关的模式误差的情形

为了研究常数型最优强迫在改进模式预报技巧中的作用, 数值模式的预报结果必须与观测值进行比较. 在该节中, 为了阐明常数型最优强迫在抵消模式误差中的作用, 我们选取概念模型 Lorenz 63 模式, 所涉及的“观测”由 Lorenz 63 模式产生, 属于理想观

测. 为了体现模式预报和观测值之间的误差, 我们用模式(3.3)预报 Lorenz 63 模式(3.1)产生的“观测”.

$$\begin{cases} \frac{dx}{dt} = -10x + 10y, \\ \frac{dy}{dt} = 29x - y - xz, \\ \frac{dz}{dt} = -\frac{8}{3}z + xy, \end{cases} \quad (3.3)$$

预报模式(3.3)与 Lorenz 63 模式的差别在于 Lorenz 63 模式中参数  $r=28$ , 而预报模式中参数  $r=29$ , 所以预报模式(3.3)存在参数误差. 事实上, 我们注意到参数  $r$  在数值模式中 与状态变量  $x$  以乘积的形成存在, 所以参数误差虽然是常数, 但与状态变量相互作用后, 叠加在每一数值积分时间步的误差是依赖时间变化的. 因此, 该类误差属于(3.2)式中与状态变量有关的模式误差  $R(u)$ .

采用四阶 Rung-Kutta 方法离散方程组(3.1)和(3.3), 取初值  $x_0=12, y_0=2, z_0=9$ , 时间步长  $dt=0.01$ , 从  $t=0$  积分到  $t=5$ , 分别得到“观测”序列和预报模式(3.3)

关于该观测预报的结果(图 1). 我们注意到, “初始观测”和预报模式初值是相同的, 也就是说, 预报模式的初值是准确的, 在试验中仅考虑了模式误差的影响.

由图 1 可知, 在预报初期( $t \leq 0.5$ ), 由模式误差导致的预报误差比较小; 但大约从  $t=0.6$  开始, 预报误差明显变大, 而且随着时间的增加, 误差呈增长趋势; 当  $3 \leq t \leq 3.5$  时, “观测”和预报的时间序列呈现相反的位相, 此时预报模式(3.3)的预报结果是不可用的. 因此, 为了使模式的预报产品具有意义, 我们必须努力改进预报模式. 下面, 我们尝试用第一节提到的常数型最优强迫方法校正模式, 进而改进模式的预报技巧.

所谓常数型最优强迫方法, 即在预报模式(3.3)中叠加常数型最优强迫, 如方程(3.4)所示. 选取计算常数型最优强迫的优化时间区间为  $[0, 1]$ , 而预报的时间区间取为  $[1, 2]$ . 也就是说, 我们利用时间窗口  $[0, 1]$  的已知“观测”获得常数型最优强迫, 然后用常数型最优强迫校正的模式预报未来时间窗口  $[1, 2]$  内

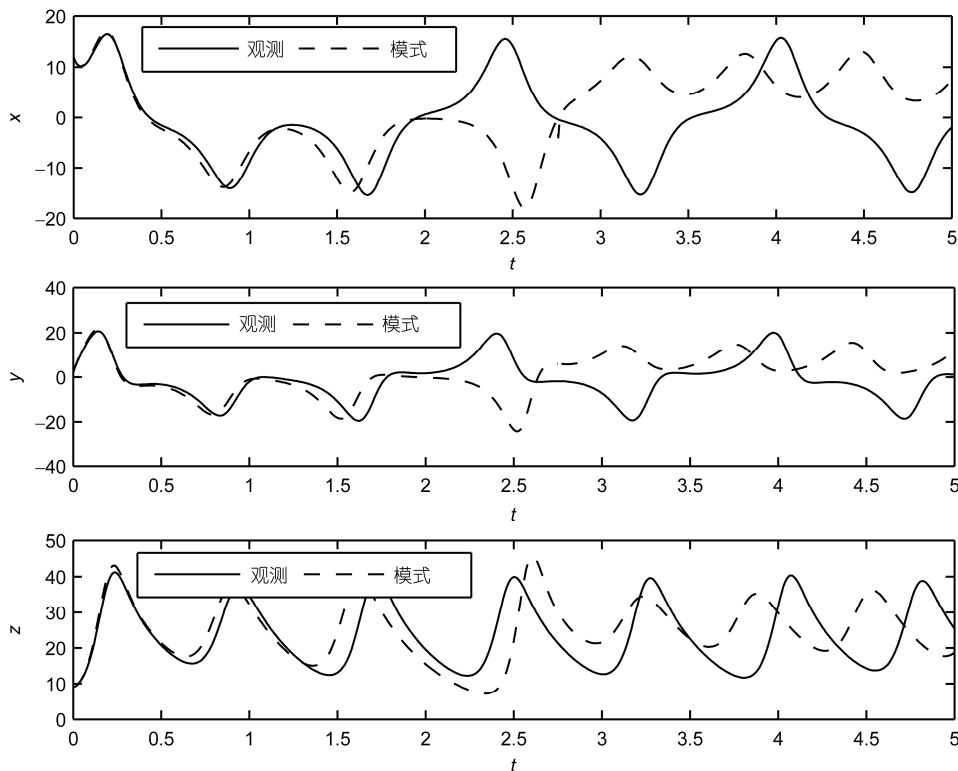


图 1 基于 Lorenz 63 模式产生的理想观测(实线)和预报模式(3.3)的预报结果(虚线)

$x, y, z$  分别是 Lorenz 63 模式的三个状态变量

变量(x, y, z)的状态.

$$\begin{cases} \frac{dx}{dt} = -10x + 10y + f_1, \\ \frac{dy}{dt} = 29x - y - xz + f_2, \\ \frac{dz}{dt} = -\frac{8}{3}z + xy + f_3. \end{cases} \quad (3.4)$$

取优化时刻为  $T=1.0$ , 用方程(3.4)和 $[0, 1]$ 时间区间的“观测”, 根据第一节中的(1.6)和(1.7)式, 用 L-BFGS 算法计算常数型最优强迫. 通过计算可知, 预报模式(3.3)的常数型最优强迫为  $\mathbf{f}^*=(f_1, f_2, f_3)=(2.8951, 3.0356, 2.9088)^T$ , 将其带入方程(3.4), 从而得到校正的预报模式, 然后用校正的预报模式预报时间区间 $[1, 2]$ 的状态(如图 2). 从图 2 可以看出, 无论是计算常数型最优强迫的时间窗口 $[0, 1]$ , 还是预报时间窗口 $[1, 2]$ , 校正模式的预报和“观测”都吻合得比较好, 而原预报模式(3.3)的预报结果与“观测”

却有较大偏差(图 3).

图 3 给出了预报模式和校正模式的预报误差的发展, 其中预报误差用 L2 范数度量, 即  $E_i = \sqrt{x'^2 + y'^2 + z'^2}$ , 其中  $x', y', z'$  分别代表各变量的预报误差,  $i=1, 2$  分别代表预报模式(3.3)和校正模式(3.4), 即  $E_1$  为原预报模式的预报误差,  $E_2$  为校正模式的预报误差. 可以看到, 在整个 $[0, 2]$ 时间段内, 校正模式的预报误差远远小于预报模式的预报误差. 该结果说明校正模式比原预报模式的预报技巧大大提高了. 也就是说, 通过在已知观测的时间窗口产生预报模式(3.3)的常数型最优强迫, 并将其叠加至预报模式的方法, 不仅可以使优化时间窗口内的模式预报和观测有很好的拟合, 抵消由参数误差产生的与状态变量有关模式误差的影响, 而且在预报时间窗口内, 该方法仍然可以很好地抵消该类模式误差的影响.

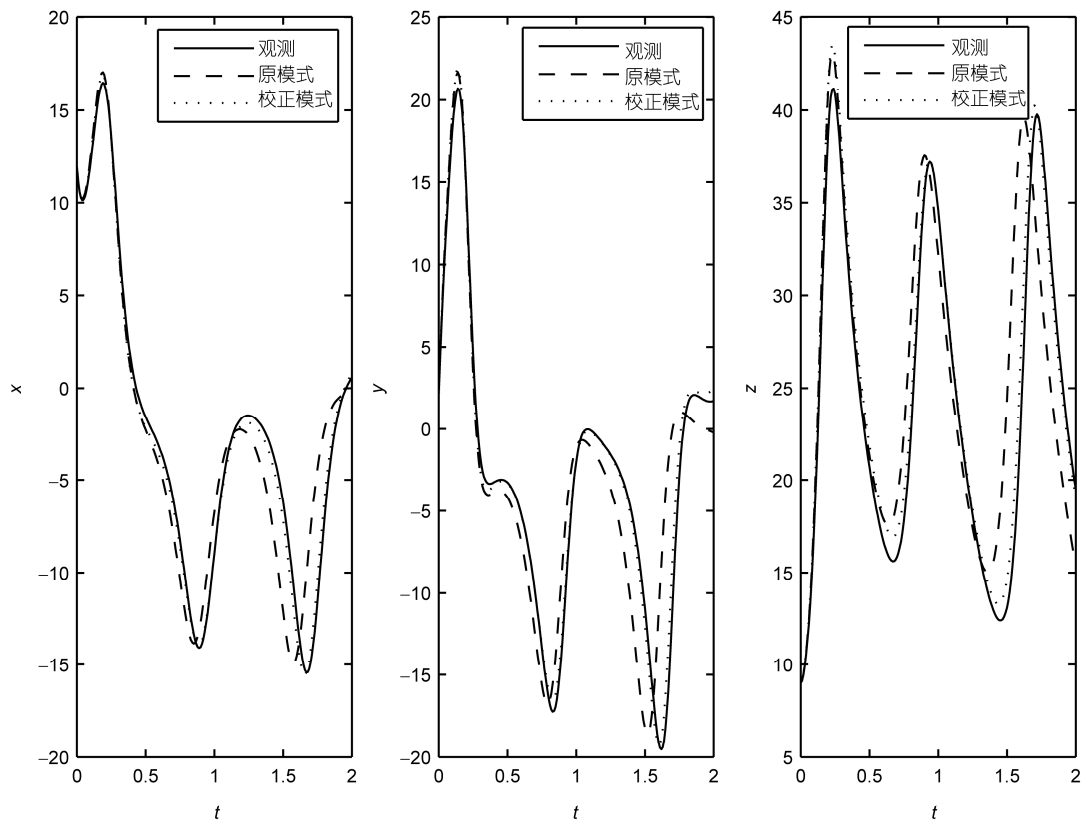


图 2 基于 Lorenz 63 模式的理想观测(实线)、模式(3.3)的预报(虚线)和校正后的预报(点线)

$x, y, z$  分别是 Lorenz 63 模式的三个状态变量

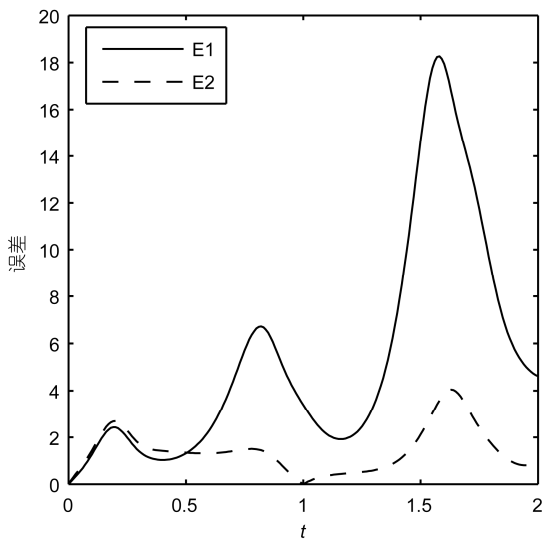


图3 预报模式(3.3)(实线)和校正后模式(虚线)的预报误差发展

### 3.2 模式只存在随机误差导致的模式误差

实际情形中, 观测往往包含随机过程的影响, 而预报模式却不能够确切描述随机过程, 因此, 预报模式通常存在由依赖时间变化的随机误差导致的模式误差. 本节将讨论常数型最优强迫在抵消随机模式误差中的效果. 为了体现模式预报和观测值之间的随机模式误差, 我们在 Lorenz 63 模式叠加随机项作为产生包含随机过程影响的“观测”的模式, 即用模式方程(3.5)产生“观测”, 而用 Lorenz 63 模式, 即用模式方程(3.1)预报该观测.

$$\begin{cases} \frac{dx}{dt} = -10x + 10y + R_1(t), \\ \frac{dy}{dt} = 28x - y - xz + R_2(t), \\ \frac{dz}{dt} = -\frac{8}{3}z + xy + R_3(t). \end{cases} \quad (3.5)$$

其中,  $R_1(t), R_2(t), R_3(t)$  是服从均匀分布的振幅为 [2, 4] 的随机项.

通过对比方程(3.1)和(3.5)可知, Lorenz 63 模式因为不能刻划方程(3.5)中的随机过程而存在模式误差. 不难看出, 这里的模式误差是依赖时间变化的随机误差, 属于(3.2)中的随机误差  $R'$ . 那么, 常数型最优强迫能否抵消该类模式误差对预报结果的影响呢?

与 3.1 节类似, 采用四阶 Rung-Kutta 方法离散方

程组(3.1)和(3.5), 取初值  $x_0=12, y_0=12, z_0=9$ , 时间步长  $dt=0.01$ , 从  $t=0$  积分到  $t=5$ , 分别得到“观测”序列和 Lorenz 63 模式关于该“观测”预报的结果, 由于随机误差的影响, 预报模式 Lorenz 63 模式的预报结果与观测有较大差别(图略). 下面, 我们使用常数型最优强迫校正模式, 考察该方法是否可以用于改进模式的预报技巧.

计算常数型最优强迫的时间区间仍然取为 [0, 1], 预报的时间区间取为 [1, 2], 即假定 [0, 1] 时间区间的“观测”为已知的. 用已知的“观测”计算 Lorenz 63 模式的常数型最优强迫为  $\mathbf{f}^* = (f_1, f_2, f_3) = (2.8951, 3.0356, 2.9088)^T$ , 将其叠加到 Lorenz 63 模式, 从而得到校正的 Lorenz 63 模式, 然后用校正的 Lorenz 63 模式预报时间区间 [1, 2] 变量的状态(如图 4). 从图 4 可以看出, 无论是优化时间窗口 [0, 1], 还是预报时间窗口 [1, 2], 校正模式的预报和“观测”都几乎重合, 而原预报模式, 即 Lorenz 模式的预报结果与“观测”有较大偏差.

图 5 给出了 Lorenz 63 预报模式和校正模式的预报误差的发展. 结果表明, 在整个 [0, 2] 时间段内, 校正模式的预报误差基本维持在 0 附近, 而预报模式 Lorenz 63 模式的预报误差却显著偏大. 该结果说明在 Lorenz 预报模式叠加常数型最优强迫的方法提高了模式的预报技巧, 在所研究的时间区间内, 常数型最优强迫几乎完全抵消了依赖时间变化的随机模式误差对预报结果的影响.

### 3.3 模式同时存在随机误差和参数误差导致的模式误差

在 3.1 和 3.2 节中, 我们的结果表明, 当预报模式仅存在由参数误差导致的依赖于状态变量的模式误差或者只存在随机误差时, 常数型最优强迫方法对于模式的校正以及模式预报技巧的改进是相当有效的. 然而, 在实际情形中, 这两种模式误差常常同时存在, 而且两者的相互作用也在很大程度上影响着预报结果的不确定性. 我们自然会问, 当预报模式同时存在随机误差和依赖状态变量的模式误差时, 常数型最优强迫能够在多大程度上抵消模式误差的影响呢?

参照 3.2 节中的做法, 仍用模式(3.5)产生“观测”, 用下面的模式(3.6)预报该观测.

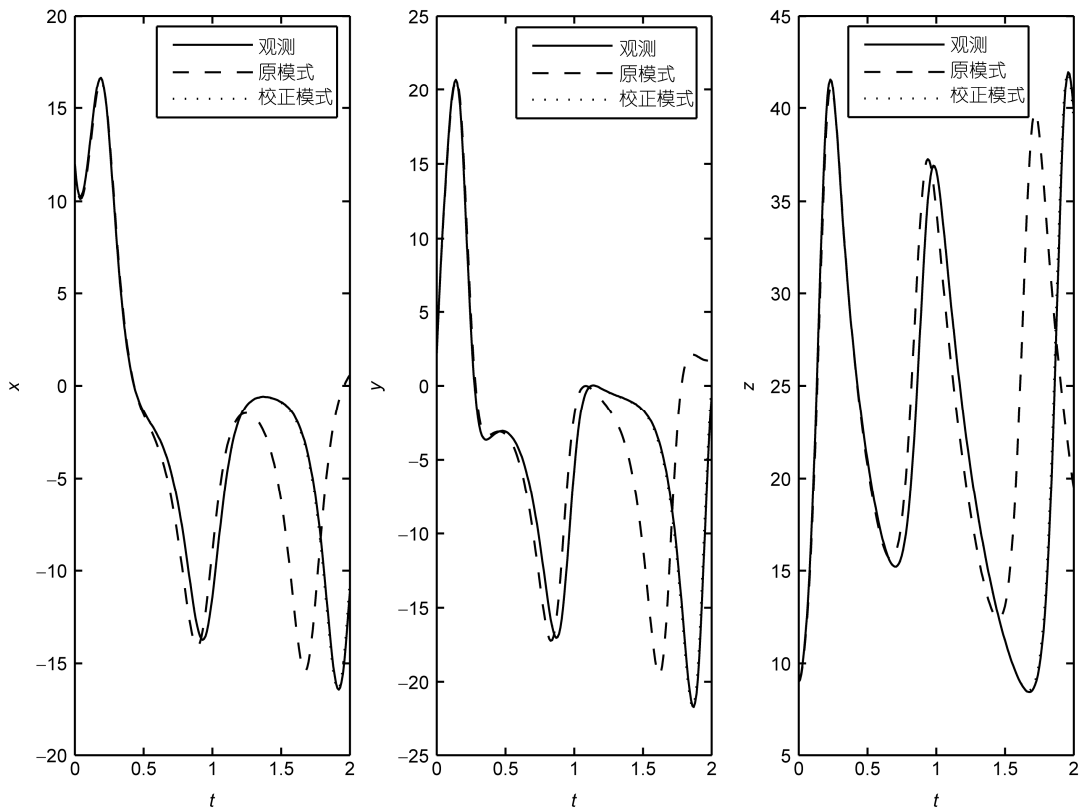


图 4 基于模式(3.5)产生的理想观测(实线)、模式(3.1)的预报(虚线)和校正后的预报(点线)  
 x, y, z 分别是 Lorenz 模式的三个状态变量

$$\begin{cases} \frac{dx}{dt} = -10x + 10y, \\ \frac{dy}{dt} = 29x - y - xz, \\ \frac{dz}{dt} = -\frac{8}{3}z + xy. \end{cases} \quad (3.6)$$

通过与模式(3.5)比较, 我们发现预报模式(3.6)不仅没有考虑由随机误差导致的模式误差, 而且数值模式中的参数也存在误差, 即参数  $r=29$ , 而不是模式(3.5)中的  $r=28$ , 从而导致模式(3.6)也存在依赖状态变量的模式误差.

对于上述情形, 我们仍用常数型最优强迫方法对模式进行校正. 与 3.1 和 3.2 节类似, 取计算常数型最优强迫的优化时间区间为  $[0, 1]$ , 预报时间区间取为  $[1, 2]$ . 根据  $[0, 1]$  时间区间的“观测”以及第 1 节中的(1.6)和(1.7)式, 计算预报模式(3.6)的常数型最优强迫. 结果表明, 预报模式(3.6)的常数型最优强迫为  $\mathbf{f}^* = (f_1, f_2, f_3) = (-1.6297, 7.0150, 5.2176)^T$ , 将其叠

加到预报模式(3.6)上, 即得校正的预报模式, 然后用该校正模式预报时间区间  $[1, 2]$  变量  $(x, y, z)$  的状态.

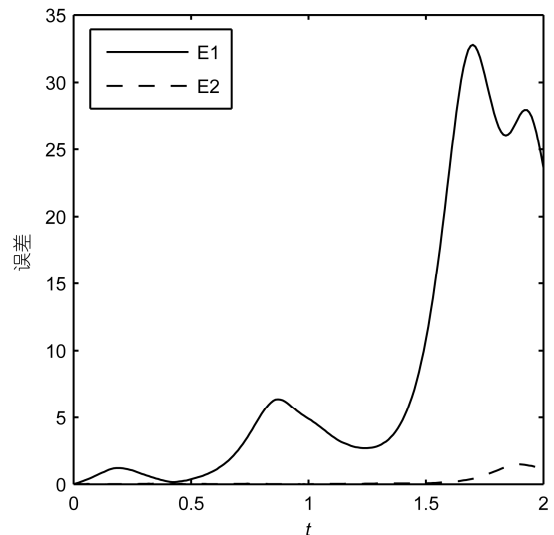


图 5 预报模式(3.1)(实线)和校正后模式(虚线)的预报误差发展



图 6 给出了由预报模式(3.6)和校正模式得到的关于状态变量( $x, y, z$ )的预报结果. 容易看出, 叠加常数型最优强迫后的模式, 即校正的预报模式在 $[0, 2]$ 时间区间内与“观测”的吻合程度远远好于原预报模式, 尤其在优化时刻  $t=1.0$ , 校正模式的预报误差几乎为零(图 7), 即校正模式的预报结果与“观测”几乎重合. 图 7 给出了当预报模式同时存在参数误差和随机误差时, 原预报模式和校正模式的预报误差的发展. 可以看出, 在整个 $[0, 2]$ 时间区间内, 校正模式的预报误差相对于预报模式 Lorenz 63 模式的预报误差有了明显减小. 这些结果说明, 对于模式中同时存在依赖状态变量的模式误差和随机误差的情形, 常数型最优强迫也是一种较为有效的校正模式和改进模式预报技巧的方法.

应该注意到, 当预报模式同时存在参数误差和随机误差时, 常数型最优强迫对模式误差的校正明显不如 3.1 和 3.2 节中只存在参数误差或随机误差情

形下的校正效果. 另外, 在数值试验中, 我们也发现, 当模式只存在随机误差时, 在不同时间区间优化所得的常数型最优强迫区别不大, 这暗示着常数型最优强迫可能在较长的预报时段对具有随机误差的模式校正有效, 即可能使模式在较长的预报时段内有较高的预报技巧; 而对于模式同时存在参数误差和随机误差时, 不同时间区间计算所得的常数型最优强迫有较大区别. 也就是说, 此情形下, 对于给定的常数型最优强迫, 其可能只在相对较短的预报时段内对模式的校正有效. 这实际上体现了不同种类模式误差相互作用的复杂性. 同时也说明, 尽管常数型最优强迫能够在一定程度上抵消模式误差的影响, 但是仅由一个常数型最优强迫项难于同时校正复杂的多个模式误差源. 无论如何, 上述结果给了我们很大的鼓舞, 常数型最优强迫方法可以抵消一些特定类型的模式误差, 可以在一定程度上校正模式, 进而提高模式的预报技巧, 是一个值得尝试的方法.

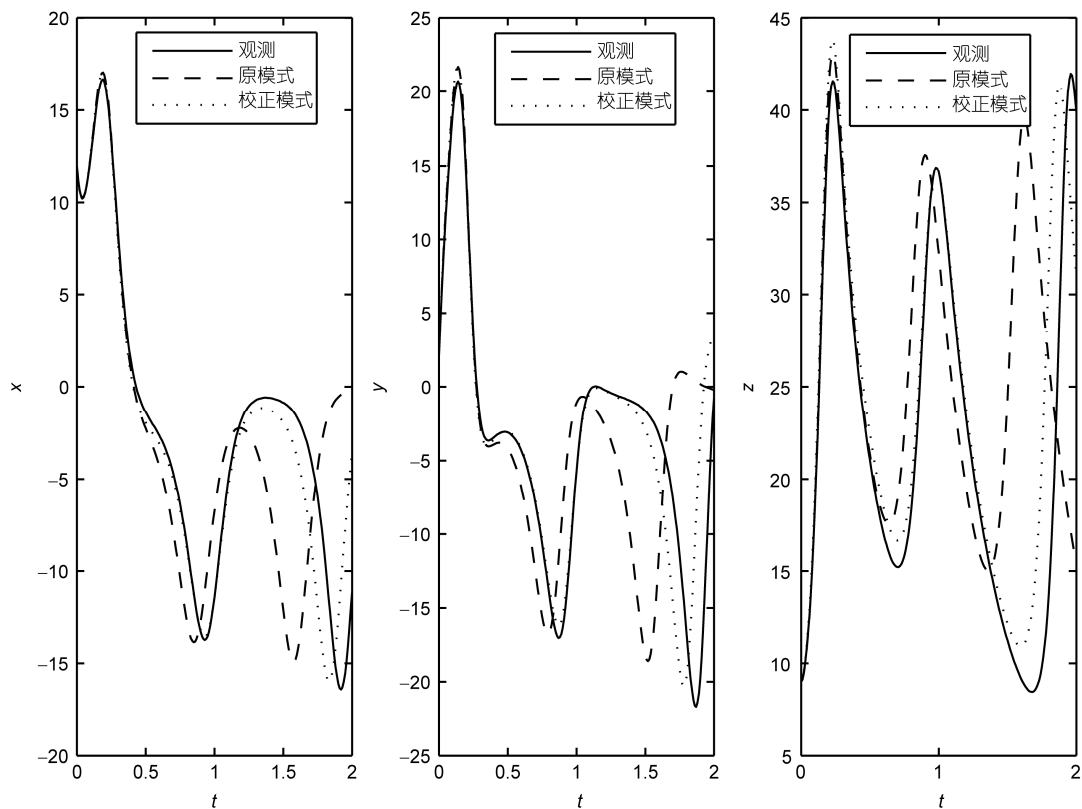


图 6 基于模式(3.5)产生的理想观测(实线)、模式(3.6)的预报(虚线)和校正后的预报(点线)

$x, y, z$  分别是 Lorenz 6.3 模式的三个状态变量

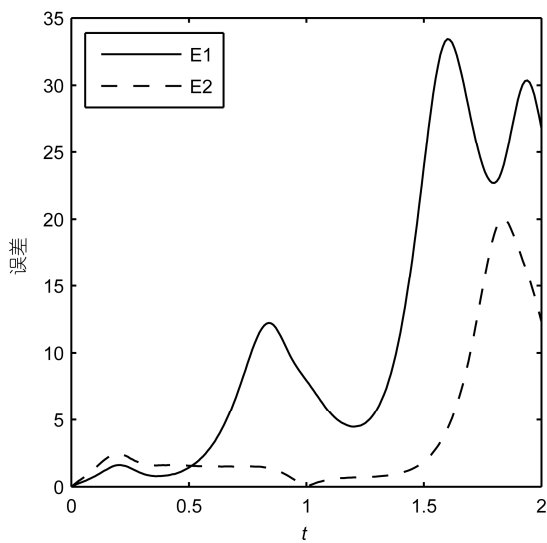


图7 预报模式(3.6)(实线)和校正后模式(虚线)的预报误差发展

#### 4 讨论和总结

为了研究常数型最优强迫方法在校正预报模式中的作用, 本文首先探讨了用伴随方法计算目标函数关于外强迫的梯度的一般公式, 并应用该公式求解了数值模式的常数型最优强迫; 然后以 Lorenz 63 模式为例, 数值研究了常数型最优强迫在抵消不同种类模式误差影响中的作用. 首先, 文章考察了当数值模式仅考虑由于参数误差导致的随状态变量发展变化的模式误差的情形. 结果表明, 常数型最优强迫能够很好地抵消该类模式误差对预报结果的影响, 数值模式在预报时段的预报技巧大大提高; 其次, 研究了当数值模式未考虑观测中依赖于时间的随机过程的情形. 此种情形下, 常数型最优强迫不仅能够很好地抵消由随机过程导致的模式误差对预报结果的影响, 而且可以在较长的时间段内使模式的预报技巧显著提高; 最后, 文章考虑了更为实际的情形, 即数值模式预报结果同时受到由随机过程和参数不确定性导致的模式误差及其相互作用的影响. 在此情形下, 常数型最优强迫方法也明显地抵消了该类混

合型模式误差对预报结果的影响. 综上所述, 尽管模式物理过程产生的模式误差依赖于时间变化, 在模式中叠加常数型最优强迫校正模式的方法依然能够显著地抵消模式误差对预报结果的影响. 所以, 常数型最优强迫是一个较为有效地校正模式和改进模式预报技巧的方法.

关于由参数误差导致的依赖状态变量的模式误差的研究, 在文中我们只假定了一个模式参数存在误差, 而其他两个参数都是准确的. 事实上, 对于其他参数也存在误差, 甚至三个参数同时存在误差, 以及参数选取不同值的情形, 我们也做了类似的数值试验, 并得到了类似的结果, 即叠加常数型最优强迫的方法对于减小该类模式误差, 进而校正模式、改进模式预报技巧的效果是显著的, 尽管改进的程度在一定程度上依赖于参数的取值. 为不赘述, 本文仅针对只有一个参数存在误差的情况来描述试验结果. 另外, 需要指出的是, 针对数值模式中的参数, 我们也可以通过直接优化参数, 确定使数值模式模拟结果和观测最接近的参数, 进而改进模式及其预报能力<sup>[22]</sup>. 不过, 该方法的计算比常数型最优强迫的计算更加复杂, 在复杂预报模式中的应用更加困难.

在用优化算法计算常数型最优强迫时, 我们的方法仅仅是令在优化时刻的预报与“观测”最接近, 因此这里的目标函数只度量这一时刻的由模式误差引起的预报误差的大小. 在实际应用中, 我们可以通过令优化时间窗口内多个时刻的预报和“观测”最接近的思想来构造更为合理的目标函数.

还要说明的一点是, 为了简明地阐释常数型最优强迫在改进模式预报技巧中的作用, 在数值试验中, 本文使用的是简单的概念模型——Lorenz 63 模式, 并利用此模式来产生理想观测值. 本文的结果鼓舞我们进一步采用更加实际的预报模式, 如三层斜压准地转模式、Zebiak-Cane 模式和 MM5 模式, 甚至 GCM 模式等来考察常数型最优强迫在校正模式及改进模式预报技巧中的作用. 期望常数型最优强迫方法能够成为在实际业务预测中可选择使用的改进模式预报技巧的有效方法.

#### 参考文献

- 1 Lorenz E N. Climate predictability. In: *The Physical Basics of Climate Modeling*. WMO GARP Publ Ser, 1975, 16: 132–136
- 2 Lorenz E N. Deterministic nonperiodic flow. *J Atmos Sci*, 1963, 20: 130–141

- 3 Ott E. Chaos in Dynamical System. Cambridge: Cambridge University Press, 1993. 397
- 4 Schubert S D, Suarez M. Dynamical predictability in a simple general circulation model: Average error growth. *J Atmos Sci*, 1989, 46: 353–370
- 5 Simmons A J, Mureau R, Petroligis T. Error growth and estimates of predictability from the ECMWF forecasting system. *Q J R Meteorol Soc*, 1995, 121: 1739–1771
- 6 Leith C E. Numerical simulation of the Earth's atmosphere. In: Alder B, ed. *Methods in Computational Physics*. New York: Academic Press, 1965. 1–28
- 7 Charney J G, Fleagle V E, Lally V E, et al. The feasibility of a global observation and analysis experiment. *Bull Amer Meteorol Soc*, 1996, 47: 200–220
- 8 Lorenz E N. A study of the predictability of a 28 variable atmospheric model. *Tellus*, 1965, 17: 321–333
- 9 Lorenz E N. Atmospheric predictability experiments with a large numerical model. *Tellus*, 1982, 34: 505–513
- 10 Smagorinsky J. Problems and promises of deterministic extended range forecasting. *Bull Amer Meteorol Soc*, 1969, 50: 286–312
- 11 Mu M, Duan W S, Wang B. Conditional nonlinear optimal perturbation and its applications. *Nonlinear Process Geophys*, 2003, 10: 493–501
- 12 Navon I M, Zou X, Derber J, et al. Variational data assimilation with an adiabatic version of NMC spectral model. *Mon Weather Rev*, 1992, 120: 1433–1446
- 13 Evensen G. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to prediction error statistic. *J Geophys Res*, 1994, 99: 10143–10162
- 14 任宏利, 丑纪范. 数值模式的预报策略和方法研究进展. *地球科学进展*, 2007, 22: 376–385
- 15 D'Andrea F, Vautard R. Reducing systematic errors by empirically correcting model errors. *Tellus*, 2000, 52A: 21–41
- 16 Roads J O. Predictability in the extended range. *J Atmos Sci*, 1987, 44: 1228–1251
- 17 Vannitsem S, Toth Z. Short-term dynamics of model errors. *J Atmos Sci*, 2002, 59: 2594–2604
- 18 Qiu C J, Chou J F. A new approach to improve the numerical weather prediction. *Sci China Ser B*, 1987, 17: 903–908
- 19 Liu D C, Nocedal J. On the limited memory BFGS method for large scale optimization. *Math Program*, 1989, 45: 503–528
- 20 LeDimet F, Talagrand O. Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus Ser A-Dyn Meteorol Oceanol*, 1986, 38: 97–110
- 21 Zebiak S E, Cane M A. A model El Nino-Southern oscillation. *Mon Weather Rev*, 1987, 115: 2262–2278
- 22 Qiu C J, Chou J F. The method of optimizing parameters in numerical prediction model. *Sci China Ser B*, 1990, 20: 218–224