

预报模式识别的扰动方法

邱崇践 丑纪范

(兰州大学)

提 要

本文从数值天气预报模式并不完全准确描述实际大气过程的观点出发,提出利用观测的近期演变资料确定模式中的未知部份(模式识别),以改善预报结果。叙述了一种模式识别的扰动方法,用正压涡度方程模式作了模拟实验,证明方法是有效的。

一、引 言

现实世界各个方面的联系实在是无边无涯,无论怎样详尽复杂的数值天气预报模式都只能是实际大气的一种理想化,况且目前的模式离足够的精确还差得很远。例如大气中进行的一些重要的物理过程还不能准确地描述,处理摩擦、非绝热及各种次网格过程时引入的种种参数往往都不够可靠。一方面是方程的不精确性,而另一方面,大量的实况观测资料源源不断地提供着方程的解的信息,被提为微分方程初值问题的数值天气预报方法却未利用或未充分利用这些信息,因而,改善目前数值预报的一个可能途径应当是更好地利用观测资料提供的信息来提高预报模式的精确性,顾震潮最早提出了在预报模式中引入历史资料的重要性和可能性,随后丑纪范(1962)、郑庆林和杜行远等(1973)提出了在准地转模式中具体的实现方法^[1]。本文则将数值预报模式中利用近期实况演变资料的问题提为微分方程的反问题。这里涉及的是其中的一类——模式识别问题,即利用模式方程的已知解(或解的某种函数)去确定方程中的未知部份。这类问题广泛存在于自然科学和工程技术的许多领域。最近十多年来,在这方面已有大量的理论研究工作,并提出了一些数值处理方法^[2]。其中 Marchuk^[3]提出的识别具有已知结构算子的扰动方法是解决这类问题的有效工具。本文将此方法用于天气预报模式的识别问题,给出确定预报方程中未知参数的具体方法,并用一个正压涡度方程模式作了模拟实验。

二、预报模式及反问题

为简单起见,以下无量纲化的正压涡度方程模式^[4]为例叙述我们的方法:

$$\frac{\partial}{\partial t}(\zeta - \mu^2 \phi) = J(\zeta + h, \phi) - 2 \frac{\partial \phi}{\partial \lambda} + \varepsilon(\zeta_E - \zeta) \quad (1)$$

其中 ζ_E 为涡度强迫源, ε 为耗散系数, h 为无量纲地形高度,其余符号均为惯用。方程

中各个量的特征量同文献 [4], 这里不再列出.

将流函数按球谐函数展开

$$\psi = \sum_{r=|S|}^R \sum_{s=0}^S (F_r^s(t) \cos s\lambda + E_r^s(t) \sin s\lambda) P_r^s(\cos \theta) \quad (2)$$

这里 P_r^s 是缔合勒让德函数, F_r^s, E_r^s 是展开系数, 以下为书写方便, 以 G_k 表示所有的系数 F_r^s 和 E_r^s . ζ_E 和 k 也作同样的展开, 一并代入 (1) 式, 利用球函数的正交性得到如下形式的常微分方程组:

$$\frac{d}{dt} G_k = \sum_{i=1}^N \sum_{j=1}^N A_{k,ij} G_i G_j + \sum_{i=1}^N B_{k,i} G_i + C_k \quad k = 1, 2, \dots, N \quad (t > 0) \quad (3)$$

初值记为

$$G_k(0) = \varphi_k \quad (4)$$

这里, N 是 (2) 式中展开系数 E_r^s, F_r^s 的总数, A 是由方程中非线性项产生的相互作用系数, B 中包含有 k 的展开系数, C 中包含有 ζ_E 的展开系数, 为节省篇幅其具体表达式不再列出. 事实上, 在相当广泛的一类假定下, 数值预报方程经空间离散化后都可以写成上述形式的常微分方程组. 因此本文叙述的方法同样适用于多层的或差分形式的模式.

为叙述方便, 将 (3) 式中的所有参数 A, B 和 C 记为向量 $\vec{a}$, 用向量 $\vec{G}$ 表示因变量. 预报问题是在给定的 $\vec{a}$ 下求问题 (3)、(4) 的解 $\vec{G}(t)$. 现在, 我们还知道 $\vec{G}(t)$ 的 M 次观测值 (不妨设观测时间是等间隔的)

$$\vec{G}(t_m) = \vec{\beta}_m \quad (t_m = m\Delta t, m = 1, 2, \dots, M) \quad (5)$$

反问题就是在允许的参数范围 Q 内找到 $\vec{a}$, 使 $\vec{G}(t)$ 除满足 (3)、(4) 外还满足条件 (5). 一般说来, 同时满足 (3)、(4)、(5) 的函数 $\vec{G}(t)$ 不是唯一存在的, 即反问题的解不是唯一存在的, 这时必须研究反问题的辅助问题.

对于 Q 中的一个 $\vec{a}$, 令 $\vec{G}(t, \vec{a})$ 表示相应的 (3) 和 (4) 的唯一解. 定义误差函数

$$E_{\beta}(\vec{a}) = \sum_{m=1}^M \eta_m^2 \|\vec{G}(t_m, \vec{a}) - \vec{\beta}_m\| \quad (6)$$

η_m^2 是适当的权重系数. 如果在 Q 中有 $\vec{a}_*$ 使

$$E_{\beta}(\vec{a}_*) = \inf_{\vec{a} \in Q} E_{\beta}(\vec{a}) \quad (7)$$

则称 $\vec{a}_*$ 为反问题的辅助问题的解. 显然, 若 $E_{\beta}(\vec{a}_*) = 0$, $\vec{a}_*$ 就是反问题的严格解. 可以证明, 问题 (3)、(4)、(7) 至少有一个解.

三、解反问题的扰动方法

Marchuk^[3] 根据共轭函数理论最早提出识别具有已知结构的线性算子的扰动方法. 对非线性算子可利用小扰动原理将其线性化, 再结合迭代过程作同样的处理.

设方程 (3) 中的所有参数 $\vec{A}, \vec{B}, \vec{C}$ 均待定, 其第 r 次估计值已知, 记为 $\vec{A}^{(r)}, \vec{B}^{(r)}, \vec{C}^{(r)}$, 相应的状态 $\vec{G}^{(r)}$ 满足

$$\frac{d}{dt} G_k^{(r)} = \sum_{i=1}^N \sum_{j=1}^N A_{k,i,j}^{(r)} G_i^{(r)} G_j^{(r)} + \sum_{i=1}^N B_{k,i}^{(r)} G_i^{(r)} + C_k^{(r)} \quad (t > 0) \quad (8)$$

$$G_k^{(r)}(0) = \varphi_k \quad k = 1, 2, \dots, N \quad (9)$$

令第 $r+1$ 次参数估计值为 $\vec{A}^{(r+1)} = \vec{A}^{(r)} + \delta \vec{A}$, $\vec{B}^{(r+1)} = \vec{B}^{(r)} + \delta \vec{B}$, $\vec{C}^{(r+1)} = \vec{C}^{(r)} + \delta \vec{C}$, 相应的解是 $\vec{G}^{(r+1)} = \vec{G}^{(r)} + \delta \vec{G}$. 代入 (3) 式并与 (8) 式相减. 若 $\|\delta \vec{A}\| < \|\vec{A}^{(r)}\|$, $\|\delta \vec{B}\| < \|\vec{B}^{(r)}\|$, $\|\delta \vec{C}\| < \|\vec{C}^{(r)}\|$, 可认为 $\|\delta \vec{G}\| < \|\vec{G}^{(r)}\|$, 略去扰动量的二阶以上小量, 得到 $\delta \vec{G}$ 满足的线性方程:

$$\begin{aligned} \frac{d}{dt} \delta G_k - \sum_{i=1}^N \sum_{j=1}^N (A_{k,i,j}^{(r)} + A_{k,i,j}^{(r)} G_i^{(r)} G_j^{(r)}) \delta G_j - \sum_{i=1}^N B_{k,i}^{(r)} \delta G_i \\ = \sum_{i=1}^N \sum_{j=1}^N G_i^{(r)} G_j^{(r)} \delta A_{k,i,j} + \sum_{i=1}^N G_i^{(r)} \delta B_{k,i} + \delta C_k \quad (t > 0) \end{aligned} \quad (10)$$

$$\delta G_k(0) = 0 \quad k = 1, 2, \dots, N \quad (11)$$

将上述方程简记为向量方程:

$$\frac{d}{dt} \delta \vec{G} + \vec{A} \delta \vec{G} = \vec{F} \quad (12)$$

$$\delta \vec{G}(0) = 0 \quad (13)$$

这里矩阵 $\vec{A}$ 的元素即 (10) 式左端中的系数, 向量 $\vec{F}$ 的分量即 (10) 式的右端项. 将其共轭问题提为

$$\begin{aligned} -\frac{d}{dt} \delta \tilde{G}_{i,m} + \vec{A}^* \delta \tilde{G}_{i,m} = \tilde{F}_{i,m} \\ (0 \leq t < t_m, i = 1, 2, \dots, N, m = 1, 2, \dots, M) \end{aligned} \quad (14)$$

$$\delta \tilde{G}_{i,m}(t_m) = 0 \quad (15)$$

这里上标“*”表示矩阵或向量的转置. 向量 $\tilde{F}_{i,m}$ 的形式是:

$$\begin{aligned} \tilde{F}_{i,m} = (0, 0, \dots, 0, \delta_m, 0, \dots, 0) \\ 1, 2, \dots, i-1, i, i+1, \dots, N \end{aligned}$$

$$\delta_m = \begin{cases} 1 & t_m - \delta t \leq t < t_m \\ 0 & 0 \leq t < t_m - \delta t \end{cases}$$

δt 是一很小的正数, t_m 即有 $\vec{G}^{(r)}$ 观测资料的时刻. 依次取不同的 i 和 t_m , 方程组 (14) 的总数是 $N \times M$ 个. 将 (12) 式与 $\delta \tilde{G}_{i,m}$ 取内积, (14) 式与 $\delta \vec{G}$ 取内积, 相减得到:

$$\delta t \cdot \delta \tilde{G}_{i,m} = (\delta \tilde{G}_{i,m}^*, \vec{F}) \quad (16)$$

内积定义为

$$(\vec{A}, \vec{B}) = \int_0^{t_m} \vec{A}^* \cdot \vec{B} dt \quad (17)$$

(16) 式中的 $\delta \tilde{G}_{i,m}$ 是在区间 $t_m - \delta t \leq t \leq t_m$ 内 $\delta \vec{G}$ 的第 i 个分量的平均值. 由于 δt 很小, 可视为 t_m 时刻的瞬时值, 即 $\delta \tilde{G}_{i,m} = \delta G_i(t_m)$. 向量 $\vec{F}$ 中包含有待求的参数扰动量 $\delta \vec{A}$, $\delta \vec{B}$ 和 $\delta \vec{C}$, (16) 式是对参数扰动量的线性方程组. 记 t_m 时刻观测值 $\vec{p}_m$ 与预报值 $\vec{G}^{(r)}(t_m)$ 之差为 $\delta \vec{G}^{ob}(t_m)$, 即 $\delta \vec{G}^{ob}(t_m) = \vec{p}_m - \vec{G}^{(r)}(t_m)$, 在辅助条件

$$\sum_{m=1}^M \eta_m^2 \|\delta \vec{G}^{ob}(t_m) - \delta \vec{G}(t_m)\| = \min \quad (18)$$

下可求得(16)式的解。得到参数扰动量后,叠加在它的第 r 次估计值上得到新的估计值。重复上述过程直到收敛为止。

需要指出,在(18)式的条件下解方程组(16),所得的法方程组往往是病态的,必须作专门处理才能得到满意的结果。这方面已有不少成熟的工作^[5]。

从上面叙述的过程可看到,此方法的计算量很大,每次迭代过程都需解一系列共轭问题。若待定的参数多,最后求解的代数方程组有很高的维数,给计算带来困难。不过在下面的例子里会看到,实际运用时大大简化的可能性是存在的。

四、模拟实验

现用数值模拟方法检验本方法的效能。

1. “观测值”的计算

适当给出方程(1)中的各个参数值作为准确值,对一定的初值计算出预报值,将其视为“实况观测”而构成模式识别及检验识别效果的依据。具体取 $\mu^2 = 25$, $\varepsilon = 0.05$ 。作球函数展开时截取 $S = P = 8$,并认为南北半球对称,总的展开系数 $N = 45$ 。地形高度采用截取同样波数的北半球实际值。根据北半球一月份500 hPa位势高度气候值利用地转关系得到流函数,在平衡条件

$$J(\zeta + h, \phi) - 2 \frac{\partial \psi}{\partial \lambda} + \varepsilon(\zeta_E - \zeta) = 0 \quad (19)$$

下得到 ζ_E 作为强迫源。在上述气候值上随机地迭加25%的扰动(指扰动的平均振幅为原波动振幅的25%)作为初始值。以 $\Delta t = 2$ 小时(无量纲数0.525)为时间步长计算出60小时的预报。将第12小时的预报作为模式识别时已知的“实况”,后48小时的预报在检验模式识别效果时用作对照。亦即在我们的试验中只取了一个时次的观测值用于识别模式($M = 1$)。

2. 共轭问题的解法

不难证明共轭方程(14)等价于

$$\begin{aligned} & -\frac{\partial}{\partial t}(\xi_{i,m} - \mu^2 \tilde{\phi}_{i,m}) + J(\zeta^{(r)} + h^{(r)}, \tilde{\phi}_{i,m}) + J(\xi_{i,m}, \phi_{i,m}^{(r)}) \\ & - 2 \frac{\partial \tilde{\phi}_{i,m}}{\partial \lambda} + \varepsilon^{(r)} \xi_{i,m} = \begin{cases} Y_i(\lambda, \cos \theta) & (t_m - \delta t \leq t < t_m) \\ 0 & (0 \leq t < t_m - \delta t) \end{cases} \end{aligned} \quad (20)$$

Y_i 是球函数的第 i 个分量, $\xi_{i,m}$ 的展开系数即是(14)式中的 $\delta \tilde{G}_{i,m}$ 。这样共轭问题和预报问题一样可用通常的谱格点变换方法求解。具体求解时取小量 $\delta t = \Delta t \times 10^{-3}$ 。这时所得的 $\tilde{\phi}$ 与 ϕ 有相同的量级,解共轭问题时可以取与预报时相同的时间步长 Δt 。

计算表明,在我们取 $t_1 = 12$ 小时时,平流项的贡献较小,共轭函数 $\tilde{\phi}_i$ 的各个系数

$\delta\tilde{G}_{i,k}$ 中, 只有纬圈波数 S_k 等于 S_i 的两个绝对值最大, 其余几乎要小一个量级。例如 $i=6$ 时, $k=6$ 和 $k=7$ 的两项对应相同的纬圈波数 $S_6=2$ 。 $t=0$ 时, $\delta\tilde{G}_{6,6}=0.514 \times 10^{-3}$, $\delta\tilde{G}_{6,7}=-0.166 \times 10^{-3}$, 而其余系数中最大的 $\delta G_{6,8}=-0.030 \times 10^{-3}$ 。因而可以近似认为 $\delta\tilde{G}_{i,k} (k=1, 2, \dots, N)$ 中, 除两项最大的以外其余均为零。这样可以大大减少求解代数方程组 (16) 时的困难。

3. 强迫源识别试验

现假定强迫源 ζ_E 未知, 其余参数的准确值均已知。对 ζ_E 的识别相当于对 (3) 式中参数 C 的识别。这时方程 (10) 右端仅保留 δC_k 一项, 参数识别的基本方程 (16) 成为:

$$\delta t \cdot \delta G_k(t_1) = \sum_{i=1}^N \left(\delta C_i \cdot \int_0^{t_1} \delta\tilde{G}_{i,k} dt \right) \quad k=1, 2, \dots, N \quad (21)$$

这里 $t_1=12$ 小时, $\delta\tilde{G}_{i,k}$ 即 (14) 式中取 $m=1$, $\tilde{e}$ 的第 i 个分量为 δ_i 所得的共轭函数的第 k 个分量。直接取

$$\delta G_k(t_1) = \delta G_k^{ob}(t_1) \equiv G_k^{ob}(t_1) - G_k^{(r)}(t_1) \quad (22)$$

方程组 (21) 有 N 个方程 N 个未知数, 根据前面叙述的共轭函数的性质, 方程组系数矩阵的对角线元素有最大绝对值, 解具有很好的稳定性。甚至可以在每一方程中只保留两个绝对值最大的系数, 而其余均取为零。这样可分别解一系列二元一次方程组而得到解。本实验即采取这种近似方法。

(22) 式中的 $G_k^{ob}(t_1)$ 是 $t=t_1$ 时的观测值, 模拟试验中即是根据强迫源的“准确值”算得的 12 小时预报。将流函数按地转关系转换为 500 hPa 等压面高度 (有量纲量) 网格点值, 结果如图 1。

现取 ζ_E 的初始估计值为 0, 作出 $t=t_1$ 时的预报 $G_k^{(0)}(t_1)$, 从而得出误差 $\delta G_k(t_1)$, 结果也在图 1 中绘出。根据这一误差提供的信息, 按前面叙述的方法对强迫源作出修正。以修正前强迫源的误差为 1, 表 1 给出了各次修正后强迫源的均方误差和最大误差。可以看到, 识别的效果是相当好的。

再以 $t=t_1$ 时的“观测值”为初值, 分别采用修正前后的强迫源作 48 小时预报与“实况”比较, 以检验修正后对预报的改

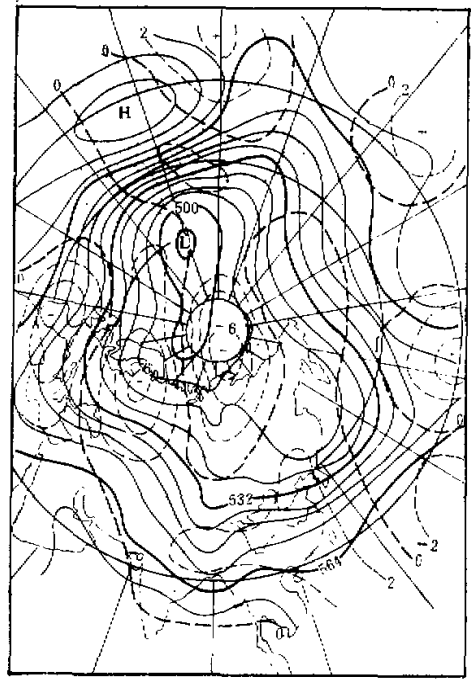


图 1 $t=12$ 小时的 500 hPa 高度“实况”(实线, 间距 80 米)和无强迫源的预报误差(虚线, 间距 2 米)

表 1 修正前后强迫源均方误差 (E_M) 和极大误差 (E_A) 相对值

修正次数	0	1	2	3
E_M	1.0000	.1049	.0326	.0061
E_A	1.0000	.6605	.0149	.0043

进。图 2 是修正前的 48 小时预报及误差图,图 3 是一次修正后的预报及误差图。“实况”与图 3 所示的预报无明显区别,故略。

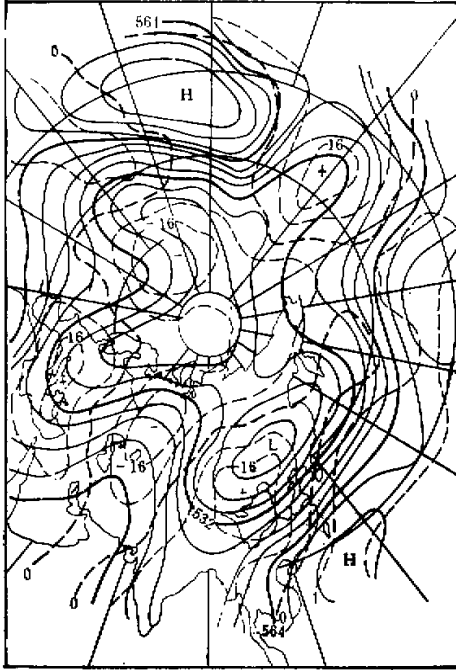


图 2 $t = 48$ 小时无强迫源预报(实线)及误差(虚线,间距 8 米)

可以看到,若认为 $\zeta_B = 0$, 积分 12 小时的误差并不显著(见图 1),大部份地区高度误差不到 20 米,最大在极地附近约为 70 米。48 小时的预报误差就大大增加了,5 个大的误差中心绝对值都在 300 米以上,至使预报的高度场与“实况”相比有了很显著的差异。对强迫源修正一次后结果就大为改观,48 小时的预报误差比修正前 12 小时的预报误差还小,绝大部分地区低于 20 米。

表 2 给出了 48 小时内强迫源修正前后的预报均方误差统计。可以看出,修正前后的预报误差随时间都是近乎线性增长的,但修正后误差减少的相对量几乎不随时间变化。经过一次修正,预报误差降低约 84%,第二次修正又降低约 77%。两次订正后的预报误差已不足原来的 4%。

表 2 强迫源修正前后预报均方误差比较(单位:米)

时 间	12 小时	24 小时	36 小时	48 小时
修正前	27.50	52.50	75.08	95.49
第一次修正	4.49	8.39	12.09	15.21
第二次修正	0.98	1.95	2.73	3.51

4. 地形误差修正试验

设地形梯度存在误差。此时相当于(3)式中一次项系数 $B_{k,i}$ 有误差。由(16)式得到识别 $\delta B_{k,i}$ 的基本方程

$$\delta t \cdot \delta G_j(t_1) = \sum_{k=1}^N \sum_{i=1}^N \left(\delta B_{k,i} \int_0^{t_1} \delta \tilde{G}_{j,k} \cdot G_i^{(r)} dt \right) \quad j = 1, 2, \dots, N \quad (23)$$

这里有 N^2 个未知数 N 个方程, 有无穷多组解。为使解唯一还应增加约束条件, 一个合理的约束是取

$$\|\delta \tilde{B}_*\| = \min_{\delta \tilde{B} \in S} \|\delta \tilde{B}\| \quad (24)$$

这里 S 是方程 (23) 的解集。取 $\delta \tilde{B}_*$ 作为待求的解。不过在我们的试验中作了简化, 在每个方程的系数 $\int_0^t \delta \tilde{G}_{i,k} \cdot G_i^* dt$ 中只保留两个最大的, 即只修正对结果影响最大的两组参数, 这样大大简化了求解过程, 避免了可能出现的解的不稳定现象。

模拟实验是在地形梯度的准确值上叠加 16% 的随机扰动作为有误差的地形。由于这里不是直接修正地形值, 不便对修正前后的参数直接比较, 但可通过比较修正前后的预报误差来说明修正的效果, 结果见表 3。尽管我们作了很大的简化, 经过一次修正后 48 小时的预报误差仍降低了一半以上。不过与前一个实验不同的是修正带来的好处随预报时间的延长而下降, 继续迭代也不会收到明显效益, 如何既有效又简便地处理方程组 (23) 看来还须进一步研究。

5. 次网格过程误差修正试验

实际上预报方程中有的未知项连函数形式都不知道, 对它的识别也许更令人感兴趣。方程作离散化处理引起的误差就是这种情况。可以将这种未知项近似展为某一正交函数系的级数, 问题即可转化为参数识别的问题。我们作的第三个试验就是模拟次网格过程误差的修正。以截断波数 $R = S = 11$ 时作出的预报为“实况”, 而将 $R = S = 8$ 的模式看为有次网格过程误差的模式。二者的初值及其余参数均相等。现取零级近似, 将次网格过程误差看为定常的, 即认为 (3) 式中的系数 C_k 有误差而加以修正。这相当于对一段时间内的平均误差作出订正。修正

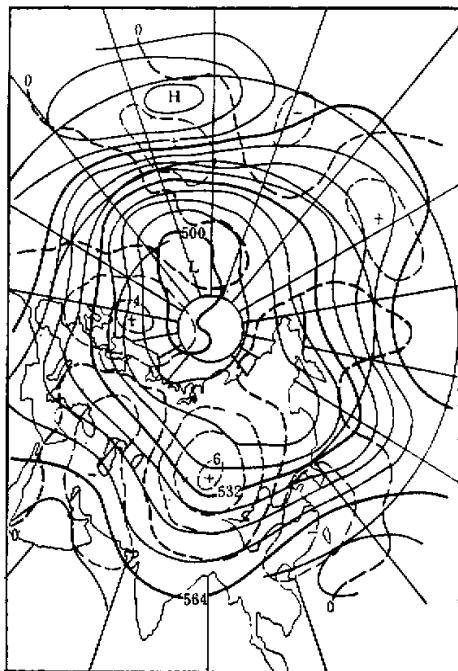


图 3 $t = 48$ 小时强迫源修正一次后的预报及误差(间距 2 米)

表 3 地形误差修正前后预报均方误差比较(单位: 米)

时 间	12 小时	24 小时	36 小时	48 小时
修正前	1.46	2.79	4.02	4.93
第一次修正	0.57	0.87	1.60	2.35
第二次修正	0.33	0.92	1.64	2.46

前后的预报误差见表 4。修正一次后 48 小时预报误差降低 2/3 以上。但这里也存在与第二个试验相同的问题,即修正的好处随时间下降,欲取得更好的效果可能应增加未知项的阶数。

表 4 次网格过程误差修正前后预报均方误差比较(单位: 米)

时 间	12 小时	24 小时	36 小时	48 小时
修正前	6.36	12.05	16.96	21.02
第一次修正	1.37	2.91	4.62	6.36
第二次修正	0.82	2.13	3.69	5.30

五、结 语

我们所作的试验表明,模式识别的扰动方法用于预报模式的参数修正是相当有效的。它也很容易推广到对边值和初值的修正。我们认为,如果每次预报时都能用近期观测资料校准模式,将会使预报准确率提高。对此我们将作进一步的试验。此外,本方法也可用来解决预报模式参数的调试问题。当前存在的两个主要问题是:第一,计算工作量大。每次迭代过程都需解一系列共轭问题,模式中状态变量维数很高时,其工作量是很大的。第二,最后得到的计算模型常常是不适定的,为了保证计算稳定性要注意采取慎重的处理方法。在我们的试验中采用了简化的方法,缩短了计算时间,也避开了解高维代数方程组可能遇到的问题。这在用于常数项参数修正时有相当好的结果,但对一次以上项参数的修正还不尽如意。更有效的处理方法尚待进一步研究。

参 考 文 献

- [1] 曾庆存, 1979, 我国大气动力学和数值天气预报研究工作的进展, 大气科学, 第 3 卷, p. 256—269.
- [2] 刘家琦, 1983, 微分方程的反问题及其数值方法, 科学探索, 第 3 卷 p. 105—118.
- [3] Marchuk, G.I., 1975, Methods of Numerical Mathematics, Springer-Verlag, New York Inc.
- [4] Källén, E., 1981, The Nonlinear Effects of Orographic and Momentum Forcing in a Low-order, Barotropic Model, *J. Atmos. Sci.*, Vol. 38, P. 2150—2164.
- [5] A. H. 吉洪诺夫等, 1979, 不适定问题的解法(中译本), 地质出版社.