

集合预报方法的全局研究*

范新岗

(南京大学大气科学系博士后流动站, 南京, 210093)

摘 要

应用一个最大简化的 6 变量非线性模式, 引入胞映射全局分析方法, 研究初值误差范围内不同扰动对预报结果的影响。在获得对其全局认识的基础上, 再通过理论与预报实践证明, 集合预报方法确能改进预报。理论研究表明, 可以求出最佳集合样本数, 而且由模式预报试验得到的最佳集合样本数与理论结果是一致的。另外, 观测精度、预报精度以及模式状态变量的个数对最佳集合样本数有着不同的影响。

关键词: 集合预报, 最佳集合样本数, 全局分析。

1 引 言

观测误差的存在, 使数值预报存在最大可预报期限, 预报准确率存在理论上限^[1]。为了减缓初值误差对预报结果的影响, 克服单一动力预报的随机误差, Leith C E 提出, 可以采用集合预报^[2]。集合预报最简单的方法就是 Monte Carlo 预报(MCF)方法, 即在初始观测场上叠加随机扰动, 产生 N 个初始状态的集合, 然后对集合的每个成员作时间积分, 用积分得到的预报集合计算所要求的估计值。因此, MCF 所需计算量是一个普通动力预报的 N 倍。设由第 i 个初始状态得到的预报为 \tilde{X}_i , 则集合平均预报为:

$$\hat{X} = \frac{1}{N} \sum_{i=1}^N \tilde{X}_i \quad (1)$$

Leith C E 证明^[2], MCF 方法优于常规的单一动力预报。

集合预报的理论和实践都表明^[3], 如果叠加的扰动能够正确地反映初始场的不确定性, 则集合预报能够比用原始初始场所做的控制预报提供更有用的信息。Hoffman N R 和 Kalnay E 于 1983 年提出了替代 MCF 的滞后平均预报(LAF)方法^[4], 这种方法是对间隔为 τ 的不同时刻的初始场做预报, 然后把由 $t = 0, -\tau, -2\tau, \dots, -(N-1)\tau$ 不同初始场所做的预报集合中同一时刻的预报值平均起来, 得到集合预报结果。LAF 方法继承了 MCF 方法的优点且充分利用了历史观测的信息, 因而, 在一定程度上优于 MCF 方法。

* 初稿时间: 1997 年 7 月 30 日; 修改稿时间: 1997 年 12 月 18 日。

资助课题: 国家自然科学基金 49735170。

近年来, ECMWF 和美国的 NMC 分别提出一种方法^[5,6]构造初始扰动场。ECMWF 利用 Lorenz E N 的线性共轭方程方法^[7], 直接计算切线性模式的奇异模, 把对应于前几个最大特征值的奇异矢量称为一组最优发展扰动矢量, 并利用它们作为集合扰动叠加在基本场上, 积分得到的结果将反映初始场的最大误差(不确定性)。从而使集合预报能提供更多更有用的信息。这一方法被称为最优发展扰动(FGP)法, 研究表明^[8,9], 这种方法所做的预报技巧高于控制预报。由 Toth Z 和 Kalnay E 提出的增长模繁育(BGM)法^[10], 可以从较简单的方法获得能够代表初始场分析中存在的误差, 从而得到可以从原初始场叠加或扣除的最佳扰动成分, NMC 已将此方法用于中期业务预报。特别是对于中期时效而言, 利用分辨率较低的模式做多成员的集合预报, 尤其是利用 BGM 方法产生小扰动所做的集合预报, 要较单纯提高分辨率耗费更少的计算机资源并获得更好的效果^[11]。

综观集合预报方法的发展过程, 集合预报的水平在不断提高, 这使人们不禁想问, 集合预报究竟能在多大程度上改进预报? 尤其在同一种预报方法中, 集合预报的样本量 N 对预报有较大影响, 比如最简单的 MCF 方法, 根据统计学原理, 当样本量趋于无穷时, 据样本平均做出的预报将是最优的, 但是对于有限的样本容量则不一定。实践中, 由于计算机资源的限制, 样本容量不可能趋于无穷, 那么, 样本容量 N 取多大才能在计算量最小的同时达到最佳预报效果? 要回答这个问题, 需要对非线性气候系统进行全局分析, 即考察在所有初始状态(亦即包含了所有可能的扰动)下系统的预报能力。这无法直接用数值模式对无穷多个初始状态进行数值试验得到, 幸运的是, 人们现在有了胞映射全局分析方法^[12], 它是对非线性系统进行全局分析的强有力工具。本文以 MCF 方法为例, 结合一个 6 变量最简模式, 应用胞映射全局分析方法对集合预报中存在的问题及其方法进行全局研究。

2 最简模式与胞映射全局分析方法

2.1 最简地气耦合模式

最简地气耦合模式^[11]包含大气快变量和下垫面慢变量。其大气模式简化为一个三阶常微分方程组

$$\frac{d}{dt} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = \begin{bmatrix} 0 & a_{12} - \sigma_1 X_2 & a_{13} - \sigma_1 X_3 \\ \sigma_1 X_2 - a_{12} & 0 & a_{23} - \sigma_2 X_1 \\ \sigma_1 X_3 - a_{13} & \sigma_2 X_1 - a_{23} & 0 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} + \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad (2)$$

其中, $X = [X_1, X_2, X_3]^T$ 是无量纲大气变量, $a_{ij}(i, j)$ 和 $b_i(i, j = 1, 2, 3)$ 是无量纲下垫面温度 $T = [T_1, T_2, T_3]^T$ 的函数, $a_{ii}(i = 1, 2, 3)$, σ_1 和 σ_2 为常量。下垫面模式考虑为热传导方程的第三边值问题, 其解析解可写为

$$\begin{bmatrix} T_1 \\ T_2 \\ T_3 \end{bmatrix}_{(n)} = \mathbf{G}_1 \begin{bmatrix} T_1 \\ T_2 \\ T_3 \end{bmatrix}_{(n-1)} + \mathbf{G}_2 \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}_{(n)} + \begin{bmatrix} g_1(R_s) \\ g_2(R_s) \\ g_3(R_s) \end{bmatrix}_{(n)} \quad (3)$$

其中, \mathbf{G}_1 和 \mathbf{G}_2 是常矢量, $g_i (i = 1, 2, 3)$ 是太阳辐射 R_s 的函数, 下标 (n) 和 $(n-1)$ 为时间步数。式(2), (3) 构成了最简地气耦合模式, 在既定的参数混沌域中, 该系统存在混沌态。取定系统的控制变量, 对该模式进行足够长时间积分, 得到系统的六维状态空间为^[11]:

$$\begin{aligned} X_1 & [-3.1188, -0.9422], & X_2 & [-2.0747, 0.4987], \\ X_3 & [-1.2103, 1.9749] \end{aligned} \quad (4)$$

$$\begin{aligned} T_1 & [0.38567, 0.51203], & T_2 & [-0.01997, 0.08647], \\ T_3 & [-0.57188, -0.44552] \end{aligned} \quad (5)$$

2.2 广义胞映射方法的应用

实际中, 任何变量的观测误差都不可避免, 现设观测误差为 η , 则观测值 V 实际上代表了 $V \pm \eta/2$ 范围内无穷个实数。这样, 一个由 n 个变量 v_1, v_2, \dots, v_n 构成的 n 维系统的状态空间被观测误差划分成了 L 个尺寸为 η 的 n 维小立方体, 称为广义胞, 原来的状态空间变成了 n 维广义胞空间, 当用 L_1, L_2, \dots, L_n 分别表示各变量方向被 η 划分的区间数时, $L = L_1 L_2 \dots L_n$ 。当知道某时刻的观测值为 V 时, 还无法知道其处在广义胞中的具体位置, 这就出现了不确定性。图 1 给出一个简单的二维广义胞空间 S 的示意图, 该空间沿两个变量 v_1 和 v_2 方向分别被划分为 L_1 和 L_2 个区间, 称之为

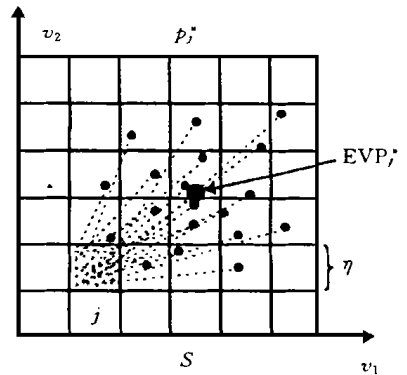


图 1 广义胞空间 S 的二维示意图

一维胞, 按从小到大的次序对一维胞排序, 序号用 $i_k (k = 1, 2)$ 表示。为了研究和叙述方便, 将所有二维广义胞也依次编号, 并用 j 表示, $j = i_1 + (i_2 - 1)L_1$ 。对于 n 维广义胞空间, 用 $i_k (k = 1, 2, \dots, n)$ 表示各方向的一维胞序号, 则

$$j = i_1 + \sum_{k=1}^n \left[(i_k - 1) L_m^{n-k} \right] \quad (6)$$

设某时刻观测值处于第 j 广义胞(称此胞为观测胞), 但不知道具体是哪一个点。其中任一点都有可能是系统演变到此刻的真值, 如果假定每一点是真值的概率相等, 则可以由数值模式确定地预知观测胞中各点经时间 τ 演变后的像点, 显然, 所有这些像点在 L 个广义胞上的分布服从一个确定的概率分布, 记为 $p_j^* = (p_{1,j}^*, p_{2,j}^*, \dots, p_{L,j}^*)$ 。 p_j^* 可以近似地用样本法求得^[12], 即在 j 胞中均匀取 $\mathcal{L} \mathcal{B} > L$ 个点, 分别代入模式积分 τ 时间, 这样就得到了 \mathcal{L} 个经 τ 时间演变后的像点, 设 \mathcal{L} 个像点中有 \mathcal{L}_1 个落入第 1 广义胞, 有 \mathcal{L}_2 个落入第 2 广义胞, ..., 有 \mathcal{L}_L 个落入第 L 广义胞, 则

$$p_j^* \sim \frac{L_j}{L} \quad (i = 1, 2, \dots, L) \quad (7)$$

依据概率分布 p_j^* 可以做级别预报或量值预报^[13], 为了定量地研究预报效果, 以下将主要使用量值预报, 记依据 p_j^* 得到的量值预报为 $EVP_j^* = (v_{1,j}^*, v_{2,j}^*, \dots, v_{n,j}^*)$, 这是由 p_j^* 唯一确定的最佳量值预报^[13]。

在本文的研究中, 将式(4)中的区间(记为 X 空间)均匀分成 $M_1 = 10 \times 10 \times 10 = 10^3$ 个胞, 式(5)中的区间(记为 T 空间)均匀分成 $M_2 = 6 \times 6 \times 6 = 216$ 个胞, 构成广义胞空间。

3 观测误差内不同初值点的预报能力分析

如所知, 非线性系统对初值极为敏感, 由观测胞中不同点出发必将导致不同的结果。也就是说, 以观测胞中不同点作为初值, 其预报能力将是不同的。为了对此有一个全局的了解, 将观测胞再次均匀细分为 K 个小胞, 逐个考察它们的预报能力。这里仅以给定 (T_1, T_2, T_3) 时的 3 变量大气模式为例, 此时的状态空间为 X , 广义胞空间里有 M_1 个广义胞。图 2 给出任一广义胞 j (j 为 $1 \sim M_1$ 的整数) 作为观测胞时的二维放大示意图。将观测胞 j 沿 X_1, X_2 和 X_3 3 变量方向分别细分为 16 份, 用 m_1, m_2, m_3 表示序号, 则 $K = 16^3 = 4096$ 。对每一个小胞 i ($i = m_1 + (m_2 - 1) \times 16 + (m_3 - 1) \times 16^2$), 仍用样本法求取其内各点演变后像点的概率分布 $p_i = (p_{1,i}, p_{2,i}, \dots, p_{M_1,i})$, ($i = 1, 2, \dots, K$)。依 p_i 可得到以小胞 i 为初始点的量值预报 $EVP_i = (X_{1,i}, X_{2,i}, X_{3,i})$ 。下面给出两种不同参考态下不同小胞的预报能力。

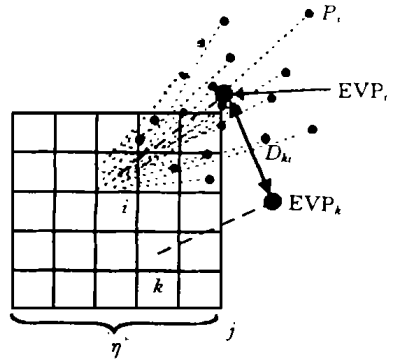


图 2 放大的观测胞 j

3.1 参考由观测胞 j 预报的概率分布 p_j^*

在由其中每个小胞 i 预报的概率分布 p_i 已知后, 假定 K 个小胞的每一个小胞 i 为“真值”的概率相等, 根据以上对观测胞再次划分的方法, 易得由观测胞 j 预报的概率分布 p_j^* :

$$p_j^* = \frac{1}{K} \left[\begin{matrix} K \\ i=1 \end{matrix} p_{1,i}, \begin{matrix} K \\ i=1 \end{matrix} p_{2,i}, \dots, \begin{matrix} K \\ i=1 \end{matrix} p_{M_1,i} \right] \quad (8)$$

每个 p_i ($i = 1, 2, \dots, K$) 与 p_j^* 的均方根差

$$R_i = \left[\frac{1}{M_1} \sum_{k=1}^{M_1} (p_{k,i} - p_{k,j}^*)^2 \right]^{\frac{1}{2}} \quad (9)$$

则反映了小胞 i 的预报与观测胞 j 的预报间的误差。由于所研究的观测胞 j 中的小胞空间仍是三维胞空间, 为直观起见, 逐个考察每个平行于坐标轴的二维剖面上 R_i 的分布情况。图 3 给出当观测胞为广义胞 $j = 556$ (X_1, X_2, X_3 方向的一维胞序号分别为 $i_1 = 6, i_2 = 6, i_3 = 6$) 时, R_i 在 $m_1 = 4, 12, m_2 = 4, 12, m_3 = 4, 12$ 处的 6 个剖面图, 由图 3 可见, 即使在

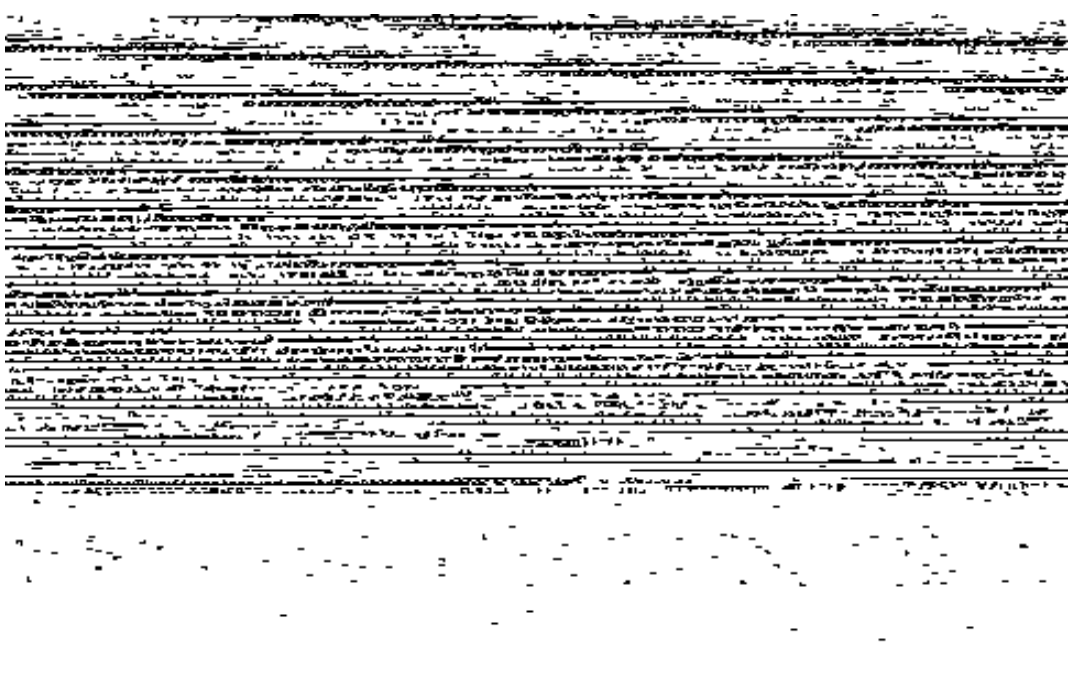


图3 由观测胞 j 得到的概率分布与其内各点得到的概率分布间的均方差 R_i 的剖面分布

同一观测胞中,当初值取不同的点时,所得预报与该观测胞的预报间的差距也是很不相同的,其分布呈现一种整体有一定规律而局部混乱的状态,这是系统的演变和非线性特征所决定的。其它各剖面图结果以及观测胞 j 为其它广义胞时所得结论与此类似。

3.2 参考由观测胞 j 中第 k 小胞预报的值 EVP_k

如图2所示,设真正的初值在观测胞 j 中的第 k 小胞,以第 k 小胞得到的预报为 EVP_k 。当初值取其它的小胞,设取 $i(i = 1, 2, \dots, K)$ 时,得到预报值 EVP_i, EVP_i 与 EVP_k 之间的距离

$$D_{ki} = EVP_i - EVP_k$$

$$\left[(X_{1,i} - X_{1,k})^2 + (X_{2,i} - X_{2,k})^2 + (X_{3,i} - X_{3,k})^2 \right]^{1/2} \quad (10)$$

反映了由于初值选取误差所导致的预报误差。图4给出当观测胞为广义胞 $j = 556, k = 1366(m_1 = m_2 = m_3 = 6)$ 和 $k = 2731(m_1 = m_2 = m_3 = 11)$ 时, D_{ki} 在通过这两点且平行于坐标轴的6个剖面上的分布图。图4结果表明,以离第 k 胞越近的小胞为初值所作的预报误差越小,显然,小的初值选取误差导致的预报误差也小。同时可看出,误差分布在空间上有一定的走向及规律性,这是系统演变至该真值点附近时所具有的特征。由其它各点得到的截面图结果与此类似(图略)。这一点具有一定的启发意义,即利用系统演变的规律及历史演变的相似性有助于找到离真值点较近的预报初值点,从而改善预报,这是有待研究的课题。

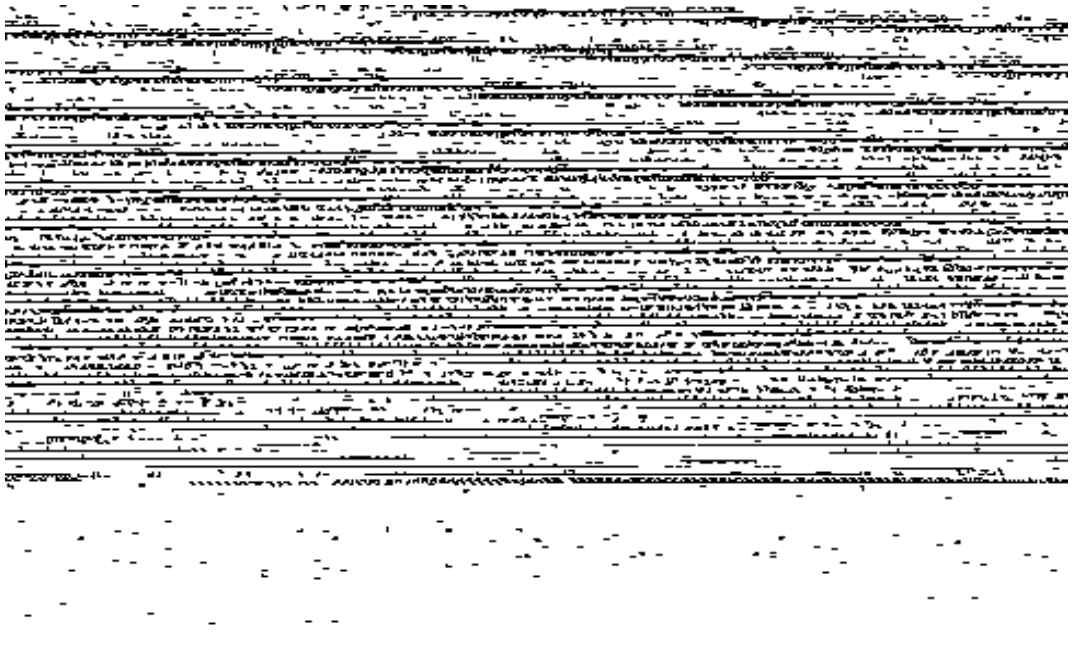


图 4 由观测胞 j 中第 k 点得到的预报值与其它各点得到的预报值间的距离 D_{ki} 的剖面分布

4 集合预报

上节结果表明,即使在观测误差内,初值取法不同时预报效果仍有很大差异。为避免个别预报的随机误差,诞生了 MCF 方法。下面对 MCF 方法的预报效果,尤其是集合样本数进行全局研究。

4.1 集合预报的最佳样本数

从观测胞 j 的 $K = 4096$ 种初值取法中,随机取出 $N (N = 1, 2, \dots, K)$ 种,将各自的预报点 $EVP_i (i = 1, 2, \dots, N)$ 求平均,就得到本次集合样本数为 N 的集合预报点 ENP 。为了考察 N 值取法的全局特性,现将 K 种初值中取 N 种的所有可能组合得到的集合预报点进行平均,得

$$\overline{ENP}_N = \frac{1}{C_K^N} \sum_{i=1}^{C_K^N} ENP_i \quad (11)$$

以 \overline{ENP}_N 代表集合样本数为 N 时的集合预报全局结果。由于观测精度的限制,对预报结果也只能要求与观测有同样的精度。根据全局研究结果,由已知观测值所在的胞所能得到的最佳量值预报是 EVP_j^* (参见 2.2 节),而集合预报的每一个样本的初值是从观测胞中随机选取的单个的点,正是由于这种初值选取的误差导致了集合预报 \overline{ENP}_N 与最佳量值预报 EVP_j^* 之间的差距。因此,只要 \overline{ENP}_N 落入 EVP_j^* 所在的胞,就可以认为集合预报已经克服了初值选取误差且达到了相对最佳预报效果,为此,定义

$$A^N = \frac{2}{\eta^*} \overline{ENP}_N - EVP_j^* \quad N = 1, 2, \dots, K \quad (12)$$

来衡量预报的相对准确性, 此处用广义胞的对角线长度 η^* 来表征胞的大小, 简单定义当 $A_N < 1$ 时, 就得到了最佳集合预报。在所有 $A_N < 1$ 的集合预报中的最小 N 值即为最佳集合样本数。图 5a 给出 A_N 的演变曲线, 可以看出, 随着集合样本数的增加, 预报误差逐渐减小, 且当 $N = 15$ 时, 集合预报点 $\overline{ENP_N}$ 落入 EVP_j^* 所在的胞, 最佳集合样本数即为 $N = 15$ 。

4.2 预报精度不同时的最佳集合样本数 N

实际预报中, 有些情况下只需较粗的预报精度, 这意味着减少预报空间中胞的个数, 相应地胞的体积增大, 实验中将式 (4) 空间划分为 $M_p = 8^3$ 个胞检验预报结果, 而初始时刻胞的大小仍取 $M_1 = 10^3$ 不变, 即观测精度不变。对观测胞 $j = 556$ 的实验结果见图 5b, 可以看出, 集合预报结果与样本数间的关系与图 5a 一致, 而此时的最佳集合样本数 $N = 13$, 说明当降低预报精度要求时, 也就降低了预报难度, 使最佳集合样本数减少。

4.3 观测精度不同时的最佳集合样本数 N

另外一些情况下, 变量的观测精度有限, 设初始时刻观测胞空间为 $M_1 = 8^3$, 既然观测精度较低, 没有理由要求较高的预报精度, 因而预报空间胞的数目也取为 $M_p = 8^3$ 。对观测胞 $j = 293$ 的实验结果见图 5c。可见, 最佳集合样本数 $N = 34$, 说明观测精度对最佳集合样本数的影响更为显著, 即使这样, 集合预报的改进效果仍是明显的。

4.4 状态变量个数不同时的最佳集合样本数 N

上述实验都是在大气模式中进行的, 只有 3 个变量。在变量个数增多时, N 是如何变化的呢? 现在采用耦合模式, 6 个变量, 式 (4), (5) 中的空间分别划分为 $M_1 = 10^3, M_2 = 6^3$

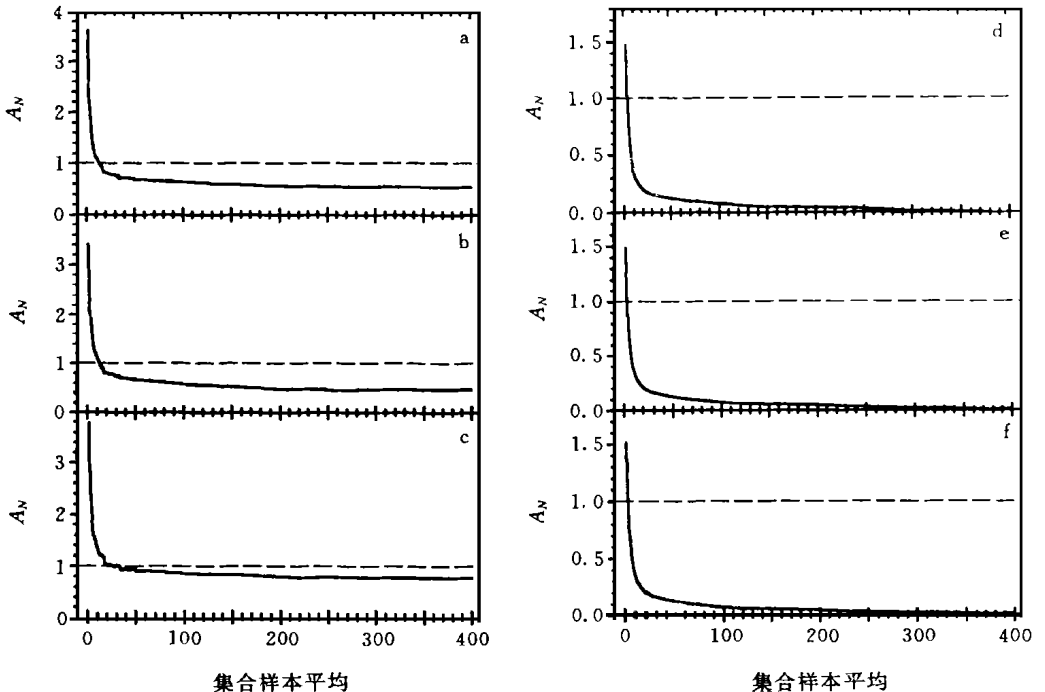


图 5 A_N 随集合样本数变化的演变曲线

(其中: j 为试验的观测胞; M_1 为观测空间胞数; M_p 为预报空间胞数; N 为最佳集合样本数)

个胞。在固定 T_1 时, 有 5 个变量, 固定 T_2, T_3 时有 4 个变量, 实验结果见图 5d-f, 在 4 变量、5 变量和 6 变量情况下, 最佳集合样本数分别为 $N = 17, N = 17$ 和 $N = 18$ 。可见, 变量个数增加对 N 的影响不大, 仅有微小的增长。

4.5 模式预报试验

以上从理论上研究了集合预报的效果及最佳集合样本数的求法, 本节用 3 变量大气

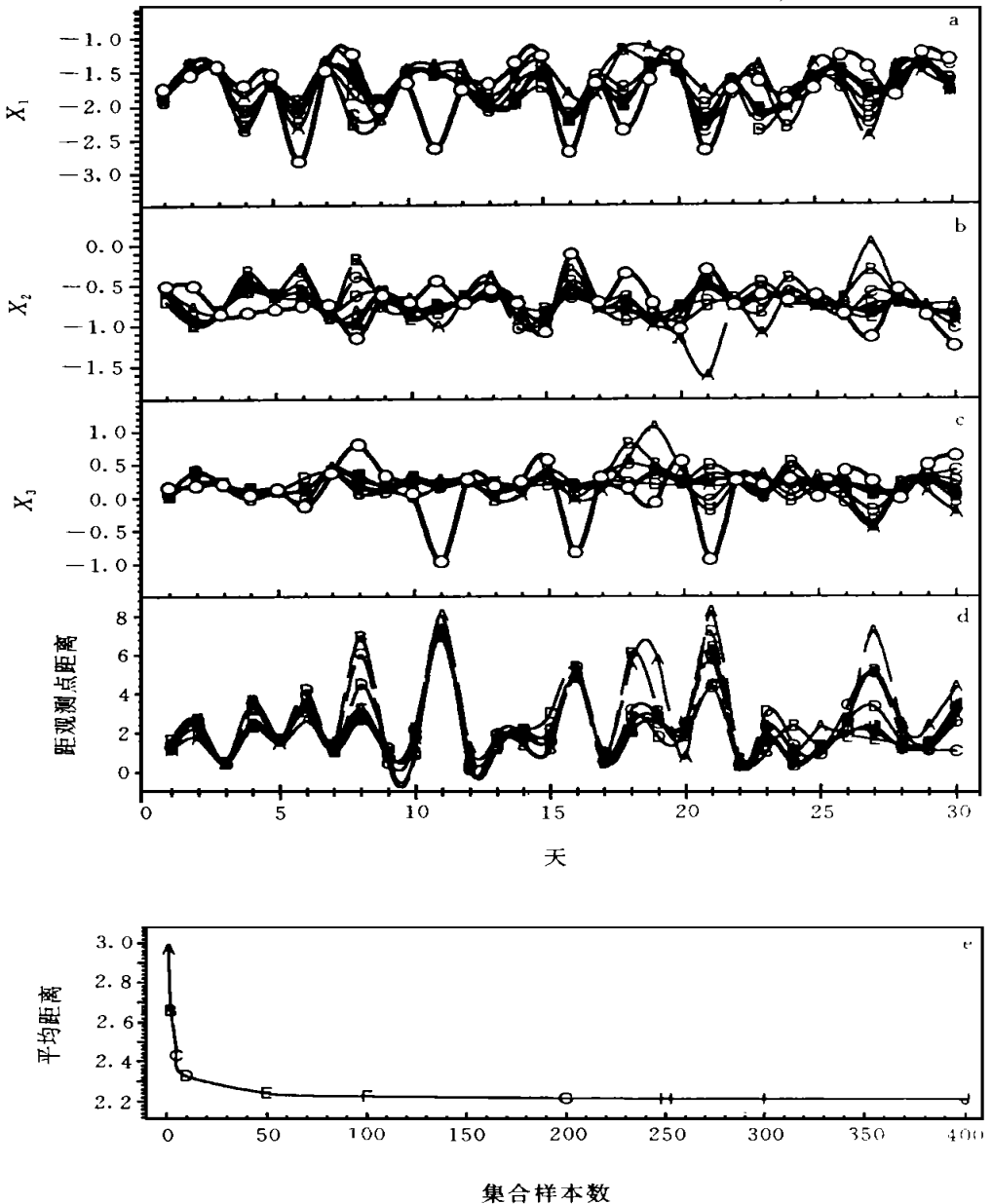


图 6 逐日集合预报的数值试验结果

(a-c. 观测与 10 组预报的 X_1, X_2, X_3 序列; d. 10 组预报值与观测值间的距离; e. 平均预报误差随集合样本数的变化)

模式的预报实践对理论结果进行验证,将理论与实践结合起来。

首先将模式(2)积分3 a,所得 X_1, X_2, X_3 的演变作为历史观测资料。用此资料做逐日预报,以前一天的观测作为初值,当天的观测用来检验,这样可以做总共3 a的逐日预报试验。将式(4)空间划分为 $M_1 = 10^3$ 个胞作为观测空间,每次预报,将前一天的观测值所在广义胞划分为 $K = 4096$ 个小胞,然后随机地取其中的 N 个小胞的中心值作为初值积分模式一天,再将得到的 N 个预报值点求平均,即得到集合样本数为 N 时的集合预报结果。图6给出 N 分别取1, 2, 5, 10, 50, 100, 200, 250, 300, 400时的集合预报结果,其中图6a-c为 X_1, X_2, X_3 的观测序列及上述10组(用A-J标示)集合预报序列,可以看出集合预报的效果是明显的。集合样本数越多,效果越好,且预报趋势正确。图6d给出10组集合预报值分别与观测值点的距离,可见,集合样本数越大,预报误差越小,但在观测序列有较大异常处,所有的预报误差都很大,这是MCF方法本身的缺陷所在。图6e给出3 a共1080次逐日预报的平均误差随集合样本数的变化曲线,对照图5a,两图的演变趋势非常相似,理论结果与模式的实际预报试验得到的结果是一致的。

5 结论与讨论

在获得了观测误差对预报能力影响的全局认识的基础上,通过理论与预报实践表明,集合预报方法确能消除数值预报中的初值选取误差而改进预报。研究表明,理论上可以求出最佳集合样本数,而且由模式预报试验得到的最佳集合样本数与理论结果也是一致的。这一结论表明,除了本文提出的全局分析方法外,还可用实际数值模式的预报试验反推最佳集合样本数。

虽然本文只是以最简单的MCF方法为例进行研究,但其结果具有普遍意义。据研究,LAF,FGP,BGM方法都能不同程度地改进集合预报^[4,9,10],也就是说,用这3种方法做集合预报,均只需不大于MCF方法最佳集合样本数的样本量就能达到同样的预报水平。因此说,MCF方法得到的最佳集合样本数是集合预报样本数的上限。由此可以认为,集合预报总能用有限的样本数达到最佳预报效果。而且改进集合方法,可以减少最佳集合样本数,达到在降低计算量的基础上改进预报的效果。

参考文献

- 1 郭秉荣,江剑民,范新岗,张红亮,丑纪范著.气候系统的非线性特征及其预测理论.北京:气象出版社,1996年.254pp
- 2 Leith C E. Theoretical skill of Monte Carlo forecasts. *Mon Wea Rev*, 1974, 102: 409- 418
- 3 丑纪范,郗吉东著.长期数值天气预报(修订版).北京:气象出版社,1995年.300~304
- 4 Hoffman R N and Kalnay E. Lagged average forecasting, an alternative to Monte Carlo forecasting. *Tellus*, 1983, 35: 100- 118
- 5 Moteni F. Predictability and finite-time instability of the northern winter circulation. *Quart J Roy Meteor Soc*, 1993, 119: 269- 298
- 6 Kalnay E and Toth Z. Estimation of the growing modes from short range forecast errors. *Research Highlights of the NMC Development Division: 1990- 1991*, 160- 165 [A available from the National Meteorological Center, Washington D. C. 20233], 1991
- 7 Houtekamer P L. The construction of Optimal perturbations, *Mon Wea Rev*, 1995, 123: 2888- 2898

- 8 Lorenz E N. A study of the predictability of a 28-variable atmospheric model. *Tellus*, 1965, 17: 321- 333
- 9 Moteni F, Buizza R, Palmer T N and Petroliaigis T. The ECM WF ensemble prediction system: methodology and validation. *Quart J Roy Meteor Soc*, 1996, 122: 73- 119
- 10 Toth Z and Kalnay E. Ensemble forecasting at NMC: the generation of perturbations. *Bull Amer Meteor Soc*, 1993, 74: 2317- 2330
- 11 李小泉. 美国国家气象中心中期预报时段内的集合预报. *气象科技*, 1994, (2):7- 10
- 12 Hsu C S. A theory of cell-to-cell mapping dynamical systems. *AM SE Journal of Applied Mechanics*, 1980, 47: 931- 939
- 13 范新岗, 丑纪范. 基于概率分布的级别预报和量值预报方法. *南京大学学报(自然科学)*, 1997, 33(专辑): 277 ~ 279

A GLOBAL STUDY ON ENSEMBLE PREDICTION METHOD

Fan Xingang

(Department of Atmospheric Sciences, Nanjing University, Nanjing, 210093)

Abstract

By introducing cell-to-cell mapping method of global analysis and using a six-variable nonlinear model, the influence of different perturbation within observational error on prediction is studied first in this paper. Based upon such a global understanding, it is proved that ensemble prediction method does improve the prediction by both theoretical research and forecasting practice. The optimal number of ensemble samples can be obtained theoretically, and it is consistent with the result of model forecasting experiment. Moreover, the observational precision, forecasting precision, and variable number of a model have different influences on the optimal number of ensemble samples.

Key words: Ensemble prediction, The optimal number of ensemble samples, Global analysis.