

基于 BGM 的不同繁殖长度对集合预报的影响^{*1}

于永锋¹⁾²⁾ 张立凤²⁾³⁾

¹⁾(海军南海舰队海洋水文气象中心,湛江 524001) ²⁾(解放军理工大学气象学院,南京 211101)

³⁾(中国科学院大气物理研究所大气科学和地球流体物理学数值模拟国家重点实验室,北京 100029)

摘 要

利用 T63L9 全球谱模式和 NCEP/NCAR 再分析资料,对 BGM 方法中增长模的繁殖长度对集合预报效果的影响进行研究。结果表明:与控制预报相比,不同繁殖长度的集合预报都能使预报效果得到一定程度的改进,特别是第 4 天预报以后,改进程度随预报时效而稳步提高。三组不同繁殖长度的集合预报对控制预报的改进存在差别,分析结果表明:繁殖长度为 2 d 的集合预报明显效果最差,而繁殖 3 d 和 4 d 的集合预报差别并不明显。对集合 Talagrand 分布以及离散度的初步分析表明,繁殖长度取为 3 d 似乎最为合理。

关键词: 增长模繁殖;繁殖长度;集合预报;初始扰动

引 言

为解决数值预报中的不确定性,集合预报方法应运而生,并于 20 世纪 90 年代在世界上多个国家投入业务使用。经典的集合预报方法是初值集合。对初值集合而言,如何构造初始扰动是至关重要的^[1]。集合预报的初始扰动技术研究始于 20 世纪 70 年代,扰动方法是估计分析误差的概率分布,如蒙特卡罗法(MFC)^[2]、时间滞后法(LAF)^[3]等。但是这些方法都不是产生初始扰动的最理想方法^[4-5]。进入 90 年代以后,基于“快速增长模”的思想^[4,6],提出增长模繁殖法(BGM)^[4]和奇异矢量法(SVs)^[5]。

用 BGM 法构造初始扰动与 SVs 方法可谓异曲同工。SVs 方法是通过切线性伴随方程直接求出奇异矢量,通过奇异矢量的叠加求出初始扰动,这需要大量的计算机资源;而增长模繁殖法是直接用数值模式通过繁殖循环估计出最快增长模。实际上用繁殖法得到的最快增长模近似等于主奇异矢量,但计算量却相对小得多^[7]。BGM 方法求取集合预报初始扰动实际上是模仿分析循环中误差(即扰动)的增长过程,即将一个任意的随机扰动叠加在分析场上,通过模式的循环积分使其得到“繁殖”。经过若干个

繁殖循环以后,初始扰动中的衰减分量大部分已经消耗殆尽,而主要留下(快速)增长分量,扰动的增长率也会逐渐稳定于某个值,或者说扰动达到了“饱和”。因此,可将扰动在某种意义上达到的饱和的状态作为对实际分析误差快速增长模的一个估计,并由此产生集合预报的初始扰动。

从 BGM 方法的基本思想可知,繁殖过程何时结束,即所谓的繁殖长度是一个很值得研究的问题。繁殖长度实际上就是在集合预报系统对初始扰动的繁殖总时间。确定繁殖长度的一般步骤是:① 定义增长模的饱和特征,即增长模繁殖需达到什么标准;② 考察繁殖循环中增长模的增长和饱和过程,大致确定饱和时间;③ 参考增长模饱和时间,选取不同繁殖长度进行大量的集合数值预报试验,根据预报效果决定繁殖长度。

关于增长模的增长和饱和问题已有一些研究^[4,8],结果表明繁殖过程中增长模在大小和形态两方面都有趋于饱和的过程,但不同物理量不同等压面上存在差别。总的来看,在取合理繁殖方案的情况下,增长模经过 3~4 d 左右的繁殖可达到饱和要求。本文将在以往研究结果的基础上设计试验,比较不同繁殖长度对集合预报效果的影响。

* 国家重点基础研究发展规划项目“我国南方致洪暴雨监测与预测的理论和研究方法研究”(2004CB418300)和中国科学院大气科学和地球流体物理学数值模拟国家重点实验室开放课题共同资助。

2005-09-23 收到,2006-08-15 收到再改稿。

1 资料选取及试验设计

本文所用的模式为全球中期数值预报谱模式 T63L9,模式分辨率为 $2.5^{\circ} \times 2.5^{\circ}$ 。有关模式的详细介绍参见文献[9]。试验初始资料来自 1998 年 1 月和 6 月的 NCEP/NCAR 逐 6 h 再分析场。以 1 月为例,分别从 5—20 日共 16 d 的 00:00(世界时,下同)起报,这样可获得 16 组集合预报试验个例,每组试验的预报时效均为 10 d; 6 月同样选取 5—20 日的 00:00 起报,得到 16 组集合预报试验个例,这样总共就是 32 组试验个例。预报场距平相关系数的检验资料采用 1958—1997 年 NCEP/NCAR 的月平均资料。

基于 BGM 方法生成集合预报的初始扰动时,繁殖循环的初始扰动、繁殖循环的周期和繁殖长度的确定都直接关系到集合预报的质量。本文的目的是分析繁殖长度对集合预报效果的影响,故试验设计中根据已有的一些研究成果将繁殖循环周期取为 6 h,繁殖过程中采用动态调整,即在繁殖过程中始终保持繁殖模与初始模大小相当(在均方根误差意义上)。繁殖循环的初始扰动即初始模参考文献[10-11]的作法,取随机分布,其随机数服从 $[-1, +1]$ 之间的均匀分布;文献[8]在对繁殖循环中增长模的饱和情况分析时指出,不同的物理量在不同的高度上,增长模的饱和时间和饱和增长率不同,而初始模的大小影响了增长模的饱和特征,故对不同的物理量和不同的高度应取不同的初始模,具体作法为:对位势高度场及

风场扰动,200 hPa 以上取为 24 h 预报误差的 100% (均方根误差意义上,下同),其余层次取为 24 h 预报误差的 50%;对相对湿度场,各层均取 24 h 预报误差的 25%。集合预报的成员取为 7 个,包含 3 对扰动预报以及 1 个控制预报,集合预报取 7 个成员的等权平均。

参考文献[8,10]关于增长模饱和的研究结果,分别取 2,3,4 d 作为繁殖长度。具体做法如下:以 5 日 00:00 起报的试验为例,若繁殖长度为 2 d,则繁殖循环就从 3 日 00:00 开始,同理,繁殖长度为 3 d 的繁殖循环从 2 日 00:00 开始,繁殖长度为 4 d 的繁殖循环就从 1 日 00:00 开始。

2 集合预报的有效率及加权评分分析

在一定数量的试验样本中,称集合预报优于控制预报的样本数为集合预报的有效率,为方便,这里给出的是具体数量而非百分比。表 1 给出了 32 组试验个例在给定的预报时效上,各集合预报优于相应控制预报(即 500 hPa 位势场均方根误差小于控制预报)的样本数。表中 F^0 代表控制预报, F^2, F^3, F^4 分别代表繁殖长度为 2,3,4 d 的集合预报。应注意的是进行比较的每组试验个例的每个预报时效(每天输出一次预报结果作为一个预报时效)都计为一个样本,这样表中预报第 3,6,9 d 的总样本数均为 32 个,5~8 d 就包含 4 个预报时效,其总样本数为 $32 \times 4 = 128$ 个。

总的来看,在各个预报时效上各组集合预报对

表 1 四组预报的 500 hPa 位势高度场均方根误差加权评分

Table 1 The root-mean-square error(RMSE) weight skill for 500 hPa geopotential height of four groups of forecasting experiments

范围	预报时效/d	优于控制预报的个例数			加权评分(考虑最好及次好)			
		F^2	F^3	F^4	F^0	F^2	F^3	F^4
全球	3	26	26	25	0.17	1.00	0.95	1.00
	6	28	30	32	0.00	1.00	1.33	1.55
	9	31	30	29	0.06	1.00	1.41	1.29
	5~8	115	119	123	0.03	1.00	1.60	1.67
北半球	3	22	26	21	0.67	1.00	2.13	1.54
	6	28	26	31	0.14	1.00	1.38	1.90
	9	26	27	28	0.33	1.00	1.70	1.70
	5~8	109	102	113	0.24	1.00	1.55	1.78
亚洲	3	22	24	27	0.48	1.00	1.74	1.52
	6	22	24	27	0.36	1.00	1.57	1.64
	9	28	27	31	0.14	1.00	1.28	2.00
	5~8	95	98	110	0.30	1.00	1.14	1.65

全球 500 hPa 形势场的预报效果均比控制预报有明显改进。其中预报第 6 天三组集合预报 F^2, F^3, F^4 好于控制预报的百分比依次为 88%, 94% 和 100%, 而所有 32 组试验个例在第 5~8 天的预报时效上的 128 个样本中, 三组集合预报优于控制预报的百分比分别达到 90%, 93% 和 96%。

从优于控制预报的样本数来看, 不同繁殖长度的各组集合预报在各个预报时效上也是各有优劣。然而对中期数值预报而言, 表中最具意义的第 6 天和第 5~8 天两项的比较数据都显示, 这三组集合预报按其改进控制预报的有效率由高到低依次是 F^4, F^3, F^2 。

为综合分析不同繁殖长度的三组集合预报对控制预报的改进程度, 还可以使用加权评分的方法。加权评分与上述集合预报的“有效率”分析的区别在于, “有效率”分析是将各组集合预报分别与控制预报作比较, 而加权评分则是对四组预报进行综合比较。

首先定义加权评分的方法, 在某个个例的某个预报时效上(如 1 月 5 日个例的第 6 天预报结果), 上述四组预报的预报效果总可以按某种评分(例如均方根误差)进行排序, 从而得到效果最好、次好的预报。下面加权评分的做法是: 效果最好的预报加 3 分, 次好的预报加 1 分, 其余不加分。这样, 对所有的个例在给定的预报时效上就会有总评分, 以此来反映各个预报的总体预报效果。

对均方根误差评分而言, 其值越小反映预报场越接近分析场, 说明预报效果越好。因此可根据均方根误差大小将四组预报即控制预报 F^0 和三组集合预报 F^2, F^3, F^4 , 按预报效果排序, 即预报均方根误差最小的为效果最好, 排在第 1 名, 以此类推。如对控制预报 F^0 , 假设第 i 组试验(32 个个例中的第 i 个)第 j 天的预报(预报 10 d 中的第 j 天), 其预报效果在四组预报中排在第 x 名, 则给该组预报加上 S_{ij}^x 分。其中:

$$S_{ij}^x = \begin{cases} 3, & x = 1 \\ 1, & x = 2 \\ 0, & x = 3, 4 \end{cases} \quad (1)$$

对 32 组试验个例, 若考虑预报第 5~8 天的预报效果, 则 F^0 的总评分可以表示如下:

$$S_1 = \sum_{i=1}^{32} \sum_{j=5}^8 S_{ij}^x \quad (2)$$

同理可得到其他三组预报的总加权评分。同时, 为便于比较, 本文对得到的加权评分数据进行了

标准化处理, 即以 F^2 的总得分 S_2 为基准分, 各预报得分与 S_2 相比得到最后的标准化评分。显然, 评分值越高代表预报效果越好。从表 1 给出的结果明显可见, 集合预报的加权评分远远超过控制预报, 这一点和上面关于集合预报的“有效率”的分析结论也是一致的。

对三组不同繁殖长度集合预报的加权评分进行比较的结论是: 总体来看, 在全球, 北半球($20^\circ \sim 90^\circ \text{N}$)和亚洲($0^\circ \sim 60^\circ \text{N}, 100^\circ \sim 160^\circ \text{E}$)的大部分预报时效上, F^4 和 F^3 要明显优于 F^2 , 而 F^4 则略优于 F^3 。例如对全球形势场第 5~8 天的预报, 控制预报和三组集合预报的标准化加权评分依次为 0.03, 1.00, 1.60 和 1.67, 非常典型的反映了四种情况的相对预报效果。

3 均方根误差和距平相关系数评分改进量及改进率分析

从上面的初步分析可看出, 三组集合预报均明显优于控制预报, 而繁殖 3 d 和 4 d 的集合预报又明显优于繁殖 2 d 的。为获得更加全面和直观的结论, 又计算了三组集合预报分别对控制预报均方根误差(RMSE)和距平相关系数(ACC)的改进量及相对改进率(32 组试验的平均值), 具体计算方法参见文献[12]。

3.1 均方根误差的改进情况

图 1a、图 2a、图 3a 分别给出了不同繁殖长度时全球、北半球和亚洲地区集合预报对控制预报的改进。从图上可见, 集合预报对控制预报的均方根误差改进有其“绝对”性, 即不论是全球、北半球还是亚洲, 从预报第 1 天到第 10 天三组集合预报均比控制预报有不同程度的改进, 且预报时效越长, 预报效果的改进越大。总体来看, 预报前 3 d 集合预报对控制预报的改进相当有限。从第 4 天开始, 集合预报的改进程度稳步提升。6 d 以后集合预报的全球形势场可比控制预报平均改进 5% 以上, 而北半球的改进更大, 6 d 以后可超过 10%。以第 6 天为例, 三组集合预报 500 hPa 全球位势高度场均方根误差的改进量分别为 3.8, 4.9 和 4.6 gpm。第 6 天集合预报比控制预报改进最为显著的是北半球形势场, 集合预报的均方根误差比控制预报的降低了 7 gpm。均方根误差可以表征预报场与分析场在数值上的近似程度, 因此集合预报对控制预报的均方根误差改

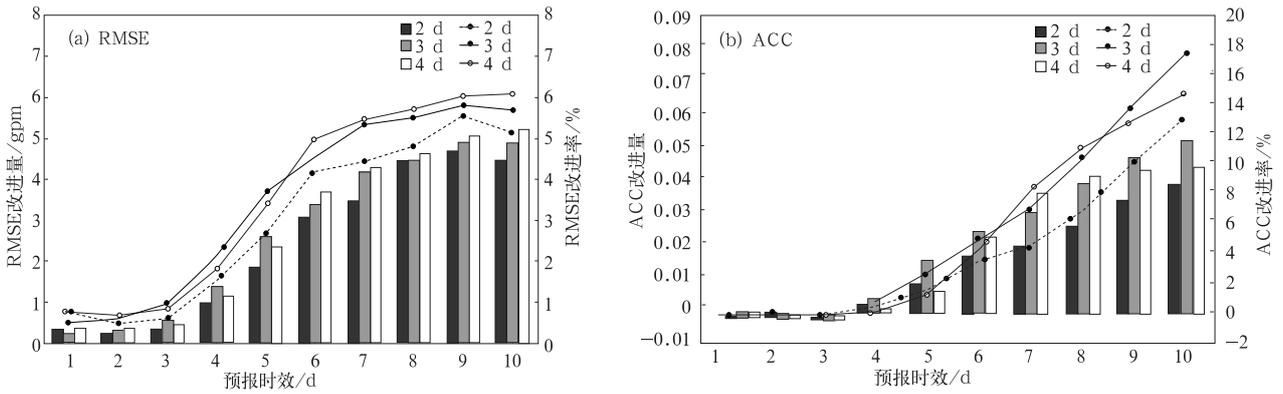


图 1 集合预报对控制预报全球 500 hPa 高度场的改进量(直方图)和改进率(曲线)

Fig. 1 Improvement quantity(column diagram) and quotiety (curve line) on the global 500 hPa geopotential height by the ensemble forecasts with respect to the control forecast

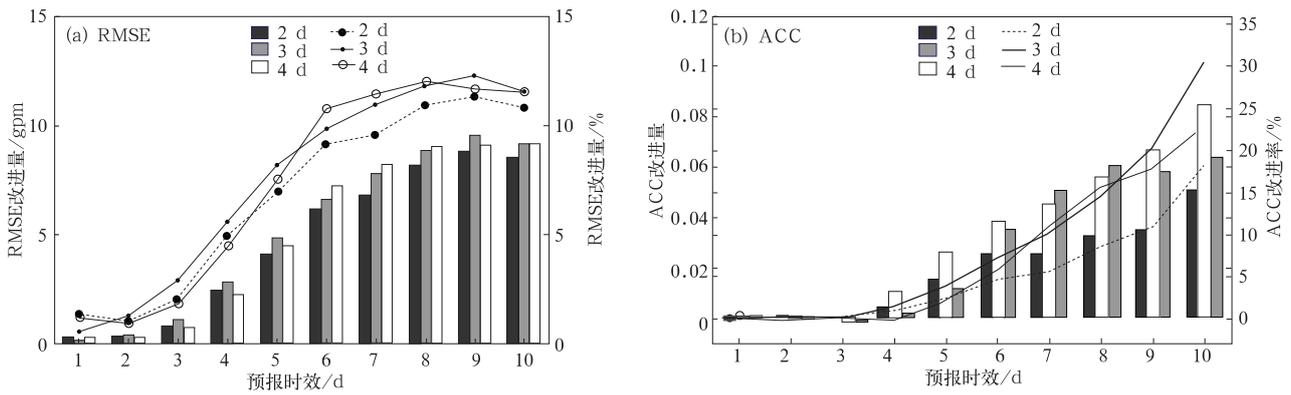


图 2 北半球 500 hPa 高度场的改进(其他同图 1)

Fig. 2 The improvement for the Northern hemisphere 500 hPa fields (others as in Fig. 1)

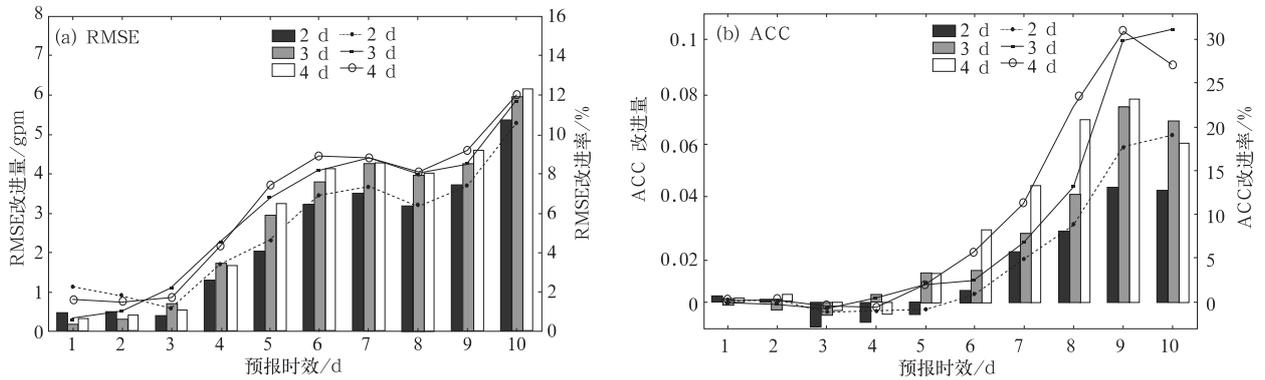


图 3 亚洲范围 500 hPa 高度场的改进(其他同图 1)

Fig. 3 The improvement for the Asian 500 hPa fields (others as in Fig. 1)

进的这种“绝对”性表明:集合预报至少在整体的数值大小方面使得预报场更接近实际分析场。

由于本文研究的是集合预报对中期数值天气预报的改进,因此重点关注预报第 4 天以后的改进情况。显而易见,三组集合预报中,繁殖长度为 2 d 的集合预报(F^2)效果最差。比较 F^4 和 F^3 两种情况,

分别统计这两组集合预报对控制预报的改进相对较好的预报日。例如对全球形势场的均方根误差改进,第 4~5 天的预报 F^3 明显好于 F^4 ,而第 6~10 天则 F^4 明显好于 F^3 (见图 1a)。表 2 给出的统计数据表明,就均方根误差而言,总的来说 F^4 要略优于 F^3 ,特别是 6 d 以后的预报。

表 2 F^4 和 F^3 对控制预报的改进相对较好的预报日
 Table 2 The forecasting days in which with 4-day breeding and with 3-day breeding have better improvements on the control forecast

范围	RMSE		ACC	
	F^3	F^4	F^3	F^4
全球	4~5	6~10	4~6,9~10	6~8
北半球	4~5,9	6~8,10	4~6,9~10	6~8
亚洲	4,7	4~10	4,10	5~9

注:表中有的预报日重复出现表示 F^3 和 F^4 在该预报日的改进程度不相上下。

3.2 距平相关系数的改进情况

同均方根误差的改进情况不同,距平相关系数通常是在预报第 4 天以后才有上述所谓的“绝对”改进,而对短期天气预报(1~3 d),试验结果显示集合预报并不比控制预报更好,这也说明中期集合预报的增长模繁殖方案并非通过简单移植就可用于短期集合预报,在一些关键技术上还需做进一步研究。

仍以繁殖长度为 3 d 的集合预报为例,从预报第 4 d 开始,集合预报的改进程度迅速提升。其中,北半球(图 2b)和亚洲(图 3b)形势场的改进最多,6 d 以后的改进率分别达到 8%~30% 和 5%~32%,全球(图 1b)形势场的平均改进程度约为 5%~18%。

显而易见,三组集合预报中, F^2 效果最差,表 2 给出了 F^4 和 F^3 的预报结果对距平相关系数相对改进情况的统计数据。就整个预报时效而言, F^4 与 F^3 的预报技巧不相上下。总体来看, F^3 似乎在 3~5 d 改进更多,而 F^4 在 6~8 d 更占优势。这一结论是否具有普遍性以及更深层次的原因,有待进一步研究。

4 Talagrand 分布检验

从平均统计意义上看,一个理想的集合预报系统中的每个成员的准确率应大致相同^[1],换言之,某个或某些预报成员不应该总是比其他一些成员准确,这可称为“成员等同性”(equal-likelihood)。这种“成员等同性”可以用所谓的 Talagrand 分布来度量^[13]。

设在某个预报时效,某个格点上 N 个集合成员对某气象变量的预报值按从小到大的顺序排列为 $x_1 \leq x_2 \leq \dots \leq x_N$,则观测值 x_a 必定落在某个区间

内。根据上述的“成员等同性”原理,对足够数目的样本,观测值落在 x_1 和 x_{i+1} 之间的概率 f_i 应趋于平均值 f ,即

$$f(x_a \leq x_1) = f(x_1 \leq x_a \leq x_2) = \dots = f(x_N \leq x_a) = f = \frac{1}{N+1} \quad (3)$$

取北半球 $20^\circ \sim 60^\circ N$ 为检验范围,共有 $G=2448$ 个格点。用于检验的实验个例为 $L=32$ 个,则实际有效样本为 $M=GL$ 。若观测值落在 $(N+1)$ 个区间中的频数分别为 S_1, S_2, \dots, S_{N+1} ,则可计算得频数 S_i 相对于期望值 $M/(N+1)$ 的均方差 D 以及观测值落在各区间的概率分布 f_i 和概率均方差 Q :

$$D = \frac{1}{L} \sum_{i=1}^L \left[\frac{1}{N+1} \sum_{i=1}^{N+1} (S_{il} - \frac{G}{N+1})^2 \right]^{1/2} \quad (4)$$

$$f_i = S_i / M \quad (5)$$

$$Q = \left[\frac{1}{N+1} \sum_{i=1}^{N+1} (f_i - f)^2 \right]^{1/2} \quad (6)$$

其中,平均概率为 $f=0.125$ 。图 4 是由概率分布 f_i 得到的 500 hPa 高度场北半球 $20^\circ \sim 60^\circ N$ 集合预报第 6 天的 Talagrand 分布检验图。可见不同繁殖长度的三组集合预报的 Talagrand 概率分布与理想情况下的概率期望值(图中以水平断线表示)都存在一定差距,相互间区别不明显,两端值特别是右端值均明显偏大。为量化表征集合预报 Talagrand 分布对平均概率的偏离程度,本文还计算了由式(4)、式(6)给出的 D 值和 Q 值。显然,这两个值应当越小越好。表 3 给出的计算结果表明,繁殖长度为 3 d 的集合预报(F^3)在 Talagrand 分布上最为合理,其次是 F^4 ,但优势并不明显。

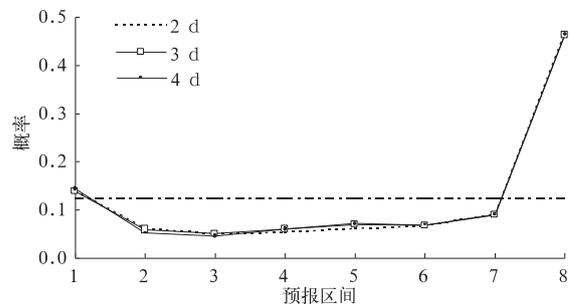


图 4 500 hPa 位势高度场的 Talagrand 分布图
(点划线表示期望概率 $f=0.125$)

Fig. 4 The Talagrand distribution diagram of 500 hPa geopotential height
(dash-dotted line represents the expected probability: $f=0.125$)

表 3 集合预报相对于期望值的频数均方差 D 和概率均方差 Q

Table 3 The frequency root-mean-square (RMS) D and probability RMS Q of ensemble forecasts with respect to the expected values

	F^2	F^3	F^4
D	327.7	322.1	325.5
Q	0.1326	0.1304	0.1315

5 离散度分析

集合预报离散度可定义为各预报成员相对集合平均值的标准差。一个好的集合预报系统,其成员间的离散度同集合平均预报的误差大小应大体上相当。这样当集合成员足够多时,大气的真实状态在大多数情况下就应该被包含在预报的集合当中了。但就现有的一些集合预报系统而言,它们的离散度一般都偏小^[1]。另外,集合预报的离散度应该反映真实大气的可预报性。如果集合成员间的发散程度(离散度)很大,则未来大气不可预报,反之则说明预报可信度高。遗憾的是,这种“离散度-可预报性”的相关关系在现有的大多数集合预报系统中表现并不是很明显^[13]。如何在不影响预报准确率的前提下适当增大集合成员间的离散度是目前很多集合预报系统需要解决的问题^[14]。

本文计算了北半球 500 hPa 位势高度场各集合成员相对于集合平均预报的平均离散度,同时为比较其与集合平均预报误差的相对大小,给出的是二者之间的比值:

$$r = S_{em} / E_{em} \quad (7)$$

其中, S_{em} , E_{em} 为集合成员的平均离散度和集合平均预报的均方根误差,分别由下式给出:

$$S_{em} = \frac{1}{L} \sum_{l=1}^L \left[\frac{1}{N} \sum_{i=1}^N (F_{il} - \bar{F}_l)^2 \right]^{1/2} \quad (8)$$

$$E_{em} = \frac{1}{L} \sum_{l=1}^L (\bar{F}_l - A_l) \quad (9)$$

式中, $L=32$ (试验个例数), $N=7$ (集合成员数), $\bar{F}_l = \frac{1}{N} \sum_{i=1}^N F_{il}$ 代表集合平均预报。计算结果由图 5 给出。平均而言,三组集合预报的离散度均小于其平均预报误差,和理想情况相比还存在一定差距。比较来看,繁殖长度为 3 d 的各集合成员间发散程度最为理想,同时又比控制预报有相对较大的改进。

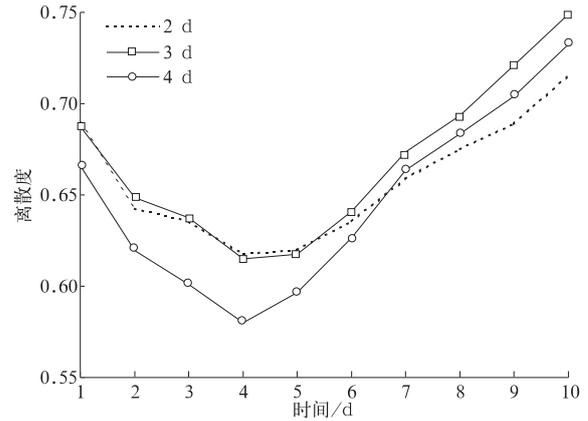


图 5 各集合成员相对于集合平均预报的离散度 (给出的是 500 hPa 位势高度场 32 组试验个例的平均值,计算范围取北半球)

Fig. 5 Spread of the ensemble forecasts relative to respective ensemble mean

(shown are average values of 32 examples for geopotential height fields at 500 hPa and the Northern hemisphere is selected to be the analysis area)

6 小结

本文利用 T63L9 全球谱模式和 NCEP/NCAR 再分析资料对增长模繁殖法中增长模的繁殖长度对集合预报效果的影响进行了分析,结果表明:不同繁殖长度的集合预报都能够使控制预报的效果得到一定程度的改进,特别是第 4 天以后的预报,改进程度随预报时效稳步提高。但是集合预报对预报场的均方根误差和距平相关系数的改进有所不同,从平均情况来看,对前者的改进贯穿于整个预报时效,而对后者的改进集中在第 4 天以后。三组不同繁殖长度的集合预报对控制预报的改进还存在差别,繁殖长度为 2 d 的集合预报 (F^2) 效果最差, F^4 与 F^3 在整个预报时效上的预报效果不相上下,但对 6 d 以后的预报 F^4 略优于 F^3 。本文还对集合 Talagrand 分布以及离散度情况进行了初步分析。结果表明,繁殖长度取为 3 d 似乎最为合理。

需要说明的是,本文的试验虽然得到了一些定性的结论,但要使集合预报有更好的预报效果,还可在增加集合成员数、改变初始扰动大小、采用区域加扰方式、引入物理过程扰动、进行多模式超级集合等方面作工作。这也是今后的工作方向。

致谢:感谢南海舰队海洋水文气象中心的刘玉国和谢璐分

担了论文的一部分试验个例的计算。

参考文献

- [1] 杜钧. 集合预报的现状和前景. *应用气象学报*, 2002, 13(1): 16-28.
- [2] Epstein E S. Stochastic dynamic prediction. *Tellus*, 1969, 21: 739-759.
- [3] Hoffman R N, Kalnay E. Lagged average forecasting, an alternative to Monte Carlo forecasting. *Tellus*, 1983, 35A: 100-118.
- [4] Toth Z, Kalnay E. Ensemble forecasting at NMC: The generation of perturbations. *Bull Amer Meteor Soc*, 1993, 74: 2317-2330.
- [5] Molteni F, Buizza R, Palmer T N, et al. The ECMWF ensemble prediction system; methodology and validation. *Q J R Meteor Soc*, 1996, 122: 73-119.
- [6] Palmer T N, Molteni F, Mureau R, et al. Ensemble prediction, ECMWF research department. *Tech Memo*, 1992, 188: 1-45.
- [7] Houtekamer P L, Derome J. Methods for ensemble prediction. *Mon Wea Rev*, 1995, 123: 2181-2196.
- [8] 于永锋, 张立凤. 基于增长模繁殖法的集合预报初始扰动饱和分析. *大气科学*, 2005, 29(6): 955-964.
- [9] 付顺旗, 张立凤, 张铭. 谱模式 T63L9 正规初值化方案及试验. *大气科学*, 2001, 25(5): 661-675.
- [10] 于永锋. 用增长模繁殖法生成集合预报初始扰动的关键技术研究. 南京: 解放军理工大学气象学院, 2005.
- [11] Toth Z, Kalnay E. Ensemble forecasting at NCEP and the breeding method. *Mon Wea Rev*, 1997, 125(6): 3297-3319.
- [12] 柳崇健, 刘英, 康红文. 一种新的物理耗散技术及其在中尺度数值天气预报模式中的应用. *中国科学(D辑)*, 2002, 32(3): 255-264.
- [13] Talagrand O, Vautard R. Evaluation of probabilistic prediction systems. Proc of the ECMWF Workshop on Predictability, Reading, United Kingdom, 1997: 1-25.
- [14] Hersbach H, Mureau R, Opsteegh J D. A short-range to early-medium-range ensemble prediction system for the european area. *Mon Wea Rev*, 2000, 128: 3501-3519.

The Effect of Different Breeding Length upon Ensemble Forecasting Based on BGM

Yu Yongfeng¹⁾²⁾ Zhang Lifeng²⁾³⁾

¹⁾ (Hydrometeorological Center of South China Sea Navy, Zhanjiang 524001)

²⁾ (Institute of Meteorology, PLA University of Science and Technology, Nanjing 211101)

³⁾ (State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics, Institute of Atmospheric Physics, CAS, Beijing 100029)

Abstract

Ensemble of initial conditions is the most important method in ensemble prediction, and how to generate the initial perturbations is crucial. Among all the methods to create perturbations for ensemble forecast, the "breeding of growing modes" (BGM) method has gained more and more favor for its good performance and almost "free" cost of computation. The breeding method simulates how fast-growing errors are "bred" in the analysis cycle. When a breeding cycle starts, a random initial perturbation field is added upon the control analysis. After some time's breeding, the most non-growing or decaying perturbations are filtered and the remainders can be mainly the fast-growing modes (perturbations). Accordingly, the perturbation's growth rate reaches certain value and its shape is also getting to the phase of slow change. In some sense, the perturbation in the breeding cycle reaches "saturation". This saturated status of perturbation is regarded as the estimate of the fast-growing modes in the realistic analysis error, thus is applied as the initial perturbation.

From the idea of BGM method, one of the crucial questions may be when the breeding cycle ends, which is called "the breeding length" problem. In fact, the breeding length is the total time of breeding of growing modes (initial perturbations) in the ensemble prediction system. The general steps to determine the breeding length are: defining the saturated character of growing modes; investigating the increasing and saturating process of the growing modes, and defining an approximate saturated time; according to the

saturated time, doing a great lot of experiments with different breeding lengths, and determining the final length by the forecast skill.

The model is the global media-range forecast spectral model T63L9 and the NCEP/NCAR 6-hour re-analysis data in 1998 are used. The initial modes in the beginning of a breeding cycle are random fields with unified distribution in $[-1, 1]$. The value of 6-hour is chosen as the length of the integration in each cycle. Besides, "rescaling" method to make the magnitude of growing modes comparable with the initial analysis error in RMSE sense is used.

Some researches show that, if an appropriate breeding scheme is adopted, the breeding of the growing modes can present a clear saturation character in both the norm (magnitude) and the form (shape) after 3—4 days of breeding. Based on these results, numerical experiments are done with breeding length of 2, 3 and 4 days. Conclusions from these experiments show that the ensemble prediction with different breeding length can all reach a certain extent improvement upon the control forecast, especially the betterments increase steadily after 4 days of forecast lead time. Results also show the differences among the three breeding lengths in the improvement upon the control forecast. It shows a better improvement in ensemble mean with a breeding length of 3 days than 2 days, but the 4-day breeding seems pretty much the same as the 3-day breeding. A pilot study on the spread and the Talagrand distribution of the ensemble members with different breeding lengths is also made. It seems that the most appropriate way is to choose 3-day as the breeding length in the system.

Key words: breeding growing modes (BGM); breeding length; ensemble forecasting; initial perturbation

国家 973“我国南方致洪暴雨监测与预测的理论和方法研究”项目 七个课题通过中期评估

2006 年 12 月 11—12 日,由中国气象科学研究院承担的国家 973“我国南方致洪暴雨监测与预测的理论和方法研究”项目七个课题中期评估会议在北京举行,经过专家组的严格评审,顺利通过了中期评估,获得专家的好评。

陈运泰院士、李泽椿院士、周秀骥院士、吴国雄院士、伍荣生院士、丑纪范院士、黄荣辉院士、郑国光研究员、张人禾研究员、郭亚曦研究员、倪允琪教授 11 位验收专家组成员出席中期评估会议。

会议由项目首席科学家张人禾研究员主持,由国家科技部基础司批准的验收专家组分别对七个课题进行评估。张人禾研究员担任验收专家组组长,周秀骥院士担任副组长。会议期间,七个课题分别对组织落实情况、研究工作的主要进展、课题的组织管理和队伍建设、人才培养的措施和成效及课题研究中存在的主要问题等方面进行汇报和演示。专家组对一些科学问题进行了提问和质疑。专家组对照课题的任务书,认真评议了每个课题的完成情况,最后给出了权威而又客观公正的验收意见。

项目自启动以来,得到课题承担单位的积极支持和配合。经过两年的努力,已经在南方致洪暴雨的中尺度结构、机理以及监测与预报理论和方法研究方面取得重要进展,研究成果显示出良好的业务应用前景,圆满完成了头两年任务书所规定的目标与任务,为后三年的研究奠定了良好的基础,在验收会上,七个课题均获得了验收组的一致肯定。

最后,张人禾研究员感谢国家科技部、中国气象局对项目研究工作的大力支持,并对全体参加评估会的评估专家和管理专家的亲临指导表示诚挚的谢意。他说,项目组一定不辜负国家的期望,虚心接受评估专家的宝贵意见和建议,并将按照国家科技部的要求和评估专家组的评估意见做好后三年项目计划任务的调整方案,特别是项目的研究成果要在气象业务上有所应用,在解决我国暴雨预报中所面临的关键科学问题,提高我国气象业务水平上发挥积极的作用。

(中国气象科学研究院国家 973 南方致洪暴雨项目管理办公室 王德英)