

# 大气系统集合预报方法及应用研究的发展

吴增茂 张新玲

(青岛海洋大学, 青岛 266003)

**摘要** 文章介绍了大气系统集合预报方法研究的发展过程与现状, 集合预报初值集形成的基本方法, 并简单介绍了其应用研究发展的现状。

**关键词** 动力-统计预报 集合预报 增长模繁殖法 奇异矢量法

## 引言

“Ensemble Forecasting (Prediction)”在我国称之为“集合预报”, 于本世纪 70 年代由 Leith 等从大气运动的随机性角度首先提出。其理论基础是 Monte Carlo 统计试验法, 所以又常有人称之为 Monte Carlo 预报。目前有一股潮流就是将集合预报用到各种天气和气候预报领域中去, 以期尽可能地消除由观测、分析误差以及大气系统固有混沌引起的预报不确定性。

“混沌现象”在 Lorenz(1963) 的著名论文“决定论的非周期流”就已经提出, 时至 15 年之后才引起了人们的比较广泛的注意。至今非线性系统对初值的敏感性受到了物理、数学、工程、地学以及许多交叉性新生学科的重视。特别是在气候以及大气、海洋与生态系统动力学数值预测、预报研究中受到了极大的关注。

文献[1]中 Lorenz(1993) 指出, 新的“混沌”定义等价于: 对同一系统内变化的对初始条件的敏感依赖性。混沌系统是指敏感地依赖于初始条件的内在变化的系统。初值的敏感性问题, 即由于初值场的微小差别在一定长的积分时间之后, 将导致动力系统状态的显著差异, 这反映了现实动力学系统的内在复杂性。几乎没有例外, 现行世界各国的气候与大气预报模式的可预报性试验都指出: 初

始小的差别都将被扩大, 一直到它们不再小为止。因此得出结论: 大气本身是混沌的<sup>[1]</sup>。大气运动的混沌性大大地限制了我们的预报能力。

在天气与气候动力数值预报研究中, 长期以来是以确定论的思想为主导, 认为预报不准确是由于初始值场不够准确或者是由于对客观系统认知的局限性使物理模型没能精确描述客观系统以及数值求解误差所造成的。今天越来越多的人认识到, 确定论(deterministic)的数学物理方程已不足以确定地描述大气运动, 天气气候变化中还包含着超越物理因果关系的随机过程和混沌过程。一些新的科学思想和研究理论与方法, 诸如突变理论、非平衡态理论、灰色系统以及混沌和分数维数理论等被引入到天气、气候变化的动力学研究中<sup>[2]</sup>。站在随机论(Stochastic)角度, 初始状态与动力模型的解是“一多对应”的关系。动力数值模型预报出的每一时刻的天气状态都应该看成是统计学意义上的一个“实现”。对天气气候变化的这种随机性、混沌性的认识, 使得在近年来“集合预报”的方法与应用研究得到了快速发展。

不同时、空尺度与动力学特性的大气过程具有不同的可预报性。可预报性的分析方法有多种, 通常使用的有统计学方法、统计动力学方法以及数值试验方法。1964 年 Charney 等提出以温度或风场的小误差增大

1 倍的时间作为可预报性的标准量度<sup>[1]</sup>。后来有人从非线性理论出发,依据 Lyapunov 指数分析也认为系统的可预报极限可定为误差被放大 1 倍的时限。

当不断发展的高性能计算机系统用于气象预报中,预报模式的时空分辨率不断地提高,涉及到的物理过程的数学表述更加逼真。然而,统计分析得到的结果说明,世界上有代表性的模型的典型误差倍增时间为:60 年代中期 5 天,1970 年接近 3 天,80 年代初 2 天并且这个估计一直到 1990 年仍未变<sup>[1]</sup>。上述事实说明,误差倍增时间的长短(即可预报性),既是动力模型的属性,又同模型可分辨的天气过程尺度大小紧密相关,这是因为大气运动尺度越小,误差增长速率越大。

## 1 集合预报方法的研究进展

1957 年 Malone 就曾说过,气象变量统计分布的透彻掌握似乎是天气预报的基础性问题。物理系统的每一个可能的初始状态都可以用抽象的相空间的一点表示,并且这些点的全部在受动力学规律支配的相域中运动<sup>[3]</sup>。确定性的动力学理论和模拟在许多方面是与统计学背景有关的。认识到了大气运动兼有确定性和随机性的双重特性,于是研究大气系统的统计-动力相结合的预报方法迅速发展起来。形成集合预报初值集的方法可归纳为以下主要 4 种。

(1) 随机动力预报与 Monte Carlo 预报法 Epstein(1969)提出了区别于惯常单算程动力预报的随机动力预报方法。从初值的不确定性出发,非均质统计流体动力学模型被直接用于均值预报及误差方差最优估计<sup>[4]</sup>,但制作这种预报工作复杂,计算量惊人。Leith(1974)为避免随机动力预报的困难且能有效地抑制初值场的不确定性和模式产生的随机误差,提出了以 Monte Carlo 统计试验法为理论基础的集合预报方法,被称之为“Monte Carlo Forecast(MCF)”方法。并且经试验结果分析得知,预报集合的样本个

数为 8 就可以得到很好的预报平均值估计<sup>[5]</sup>。

假定模式的状态向量以  $X$  表示,普通的单一算程的动力预报(ODF)的解记为:

$$X_1 = X(t; 0, x(0)) \quad (1)$$

其中  $x(0)$  表示最近的观测时刻  $t=0$  时的观测值场。

对于 MCF 的预报可形式地表示为:

$$\tilde{X}_i = X(t; 0, x(0) + z_i) \quad (2)$$

其中  $z_i$  是具有零均值且与  $x(0)$  无关的随机变量(通常为矢量场)。 $\tilde{X}_i$  为集合预报第  $i$  个成员。然后用线性回归得到预报均值  $\bar{X}$  最优估计,以及预报方差的最佳信息。试验证明,与 ODF 相比 MCF 的预报结果有很大的改进,这种改进是来自于 MCF 的滤波特性<sup>[5]</sup>。

(2) 滞后平均预报法 Hoffman 和 Kalnay(1983)提出可替代 MCF 的另一种形式的集合预报方法——滞后平均预报(LAF, Lagged average forecasting)<sup>[6]</sup>。该方法继承了 MCF 的优点,而与 MCF 不同的是, LAF 集合中的预报实现不是由同一时刻的初值个例实现的,而是对间隔为  $\tau$  的不同时刻的初值场作预报,然后把  $t=0, -\tau, -2\tau, \dots, -(N-1)\tau$  的  $N$  个由不同时刻初值场所作的同一时刻的预报结果进行统计分析,得出均值的最佳预报以及预报分散度等信息<sup>[7]</sup>。LAF 可一般地表示为:

$$\bar{X}_i = X(t; t_{0i}, x(t_{0i})) \quad (3)$$

LAF 方法在业务上是容易实现的,因为 LAF 集合中的每个组员都是在日常的业务预报循环中被产生出来的。

(3) 增长模繁殖法 Toth 和 Kalnay 的扰动繁殖法是基于他们 1993 年提出的大气增长矢量繁殖法(BGV—Breeding of the Growing Vectors)。该方法包括以下步骤:①将一个任意的小扰动加到大气的初始态上(或加入模式运行中的任一基本态上),其时间标为  $t_0$ ;②从受扰和未扰初态分别积分模式至一短时段  $t_1 - t_0$ (例如 1 天);③从一个积分结果中减去另一个得到差值场;④将差值

场按比例缩小,于是它有着同初始扰动相同的范数( $\text{norm}$ ),亦即  $\text{rms}$  振幅或转动动能,并以此作为初始扰动;⑤将该扰动加到对应  $t_1$  的分析场上,并且将②~⑤各步骤依时间向前重复进行<sup>[8]</sup>。

理论上讲,繁殖的扰动是同大气局地 Lyapunov 矢量(LLV)相关联的, Lyapunov 指数( $\lambda_i$ )已被广泛地用来表征简单动力学系统的行为:

$$\lambda_i = \lim_{t \rightarrow \infty} \frac{1}{t} \log_2 \left| \frac{p_i(t)}{p_i(0)} \right| \quad (4)$$

这里  $p$  为一个以正交矢量覆盖系统相空间的线性扰动。

Toth 和 Kalnay(1997)还讨论了繁殖法的其他技术要点,可归纳为:①集合预报是非线性积分,由于非线性特性阻碍了预报误差收敛于主要 LLV,因此使扰动集合要真实地表示初始分析场的不确定性,这是十分必要的。否则,使用的集合将不可能是最优的;②由于观测资料疏密不同而使资料分析误差大小不同,以及不同纬度带大气动力学特性的显著差异,因此按不同的地区给定不同的扰动振幅(如大洋上振幅大于陆区,赤道区小于中高纬区;③)既然起控制作用的分析场(control analysis)是大气真实状态的最佳估计,因此我们必须围绕该场来矫正集合扰动的中心。试验分析结果揭示出集合扰动表现出高于二阶矩的非线性影响<sup>[8]</sup>。

(4) 奇异矢量法 伴随方法近年来用于给出最优扰动矢量(Optional Vector, 现多见称之为奇异矢量 SV—Singular Vector)也日渐增多。SV 是作为有约束的最优化问题的解引入的,即在有限时段内最快增长的扰动。此类最优化问题通常是在线性范畴内予以论述的,并最终导向于本征值问题<sup>[9]</sup>。奇异矢量的定义可一般地表示为:  $x$  是模式有限维数的矢量,指标 0、1 分别标示  $t_0$ 、 $t_1$  时刻;  $L$  为积分切向线传播子(integral propagator),它依赖于包括两时间端点  $t_0$ 、 $t_1$  的全部被考虑时段,并给出由  $t_0$  时状态演变至  $t_1$  时刻的

全个相空间图像。方程(5)用一种普遍性形式表示出了切向线模式(TLM<sub>s</sub>)的数值(或解析)解。定义一个半正定标量为

$$\tilde{J}(x_1) = x_1^T P^T C P x_1 \quad (6)$$

这里  $C = A^T$ ,  $A$  为对称正定矩阵,  $P$  为转换矩阵,上标  $T$  表示矢量或矩阵的转置,矩阵  $C$  和  $P$  确定了状态矢量  $X_1$  各分量对目标函数  $\tilde{J}$  贡献的方式。

将式(5)代入式(6),即可得到以  $X_0$  表示的目标函数(以  $\hat{J}$  标明),并可用以下方式阐明引导出 SV 定义的极大值问题。

$$\hat{J}(x_0) = (Lx_0)^T P^T C P (Lx_0) \quad (7)$$

让式(7)达到最大值,并服从于

$$x_0^T P^T C P x_0 = 1 \quad (8)$$

$$x_0 = Qy_0 \quad (9)$$

这里  $Q$  为一矩阵,  $y_0$  为一矢量。鉴于有关系式(9),那么式(7)~(9)的极大值问题可以表示为

$$J(y_0) = (LQy_0)^T P^T C P (LQy_0) \quad (10)$$

使上式有极大值,并且服从于

$$y_0^T B^T B y_0 = 1 \quad (11)$$

这里矩阵  $B = APQ$  为满序矩阵,并因此是不可逆的。上述问题的解可以由本征值问题的解给出:

$$(B^T B)^{-1} (Q^T L^T P^T C P L Q) y = \lambda y,$$

$$\text{并有 } y^T B^T B y = 1 \quad (12)$$

由定义,本征矢量  $y$  被认为即为这里的 SV<sup>[9]</sup>。详细讨论可参见文献[9~11]。

根据其定义可知,SV 不仅依赖于数值预报模型,而且依赖于度量(Metric)的选取(常选用内积或相关范数( $\text{norm}$ )),以及最优化时间间隔( $t_1 - t_0$ )的选定<sup>[10]</sup>。当今天气预报研究中,最优化时段一般选用 24、36 或 48 h。Palmer 等(1998)及 Malteni 等(1996)都曾分析研究了使用不同度量(metric,如:总能量、总动能、涡度拟能、流函数方差)的 SV 的能谱分布等特征。他们的研究结果认为,当奇异矢量被用于产生集合预报初值时,以总能量内积作为度量似乎是最合适的。由于观测资料在资料同化过程中的作用是最重

要的,因此不难理解对不同类型的问题要求使用不同的度量<sup>[10,11]</sup>。

Gelaro 等(1998)考察了 ECMWF 的集合预报系统(EPS)用 SV 构造的扰动的最优程度。研究结果说明,虽说包含更多的奇异矢量看起来是有益的,但是基于 10~30 个奇异矢量的集合扰动应将给出中期预报不确定性的有据描述<sup>[12]</sup>。

在形成集合预报的初值集中,ECMWF 使用了伴随法扰动(即 SV),类似的方法最近也为日本气象厅所采用。在美国海军气象海洋中心(FNMO)及南非气象局的基于繁殖法的集合预报已经形成业务化。印度国家中期预报中心的繁殖法集合预报业务也正准备实施。在加拿大大气环境服务中心进行的集合预报研究试验,形成的预报集合中除了对初始大气条件及初始地表参数加扰动外,某些模式参数也加了扰动量<sup>[8]</sup>。最近对 NCEP 的全球集合预报检验表明,虽然单算程预报高分辨模式(T126)的准确率高于低分辨率模式(T62),但 T62 模式集合预报的总可用性仍高于 T126 模式的单算程预报<sup>[13]</sup>。由于 SV 方法给出的是动力学系统增长最快的扰动,同所考虑的天气过程紧密相关联,并且定域准确<sup>[9]</sup>。Ehrendorfer 和 Tribbia 的近期结果强调地指明,奇异矢量(SV)是最好的初值集。然而就实用而言,繁殖法不论其简便性还是节省计算耗费上都表现出十分明显的优越性<sup>[8]</sup>。

## 2 集合预报的应用发展

集合预报在其研究的前期多用于中长期预报、短期气候预报( $t \geq 1$  周或 2 周)<sup>[6,8]</sup>。在 80 年代后期我国气象工作者就使用一个简单的随机强迫模式进行了 500 hPa 月平均高度场的集合预报试验,结果表明它优于一般的单算程动力预报<sup>[15]</sup>。目前集合预报已被各种各样的预报及模拟研究所采用,而且预报(或初值)集形成方法趋向多样性。在 1996 年的美国气象学会年会上,中期集合预报

(MREP, Medium Range Ensemble Prediction)成了数值预报的热门内容之一,而在 1998 年会上短期集合预报(SREF)又成了数值预报的热门内容之一。样本生成的方法主要有:① 用不同的初值,包括增长模繁殖法(Breeding of growing modes)、奇异向量法(Singular Vector)以及 LAF 等;② 选择不同的物理过程参数化方案的组合;③ 用不同的分析预报模式(包括不同的分析同化、初值化、动力框架、物理过程等)。近期美国国家环境预报中心(NCEP)正在试运行的短期(0~3 d)集合预报(SREF)系统的预报集合有 25 个组元。包括全球模式繁殖的 5 个增长模、ETA 有限区域模式的 5 个增长模及 5 个不同分析场共计形成了 ETA 模式的 15 个初始值样本;另有来自全球模式和区域谱模式(RSM)的各 5 个增长模形成的 RSM 模式的 10 个初始值样本<sup>[13,14]</sup>。以下简单介绍集合预报方法在台风路径、暴发性气旋、中尺度对流降水以及污染物长距离输运预报研究中的应用发展。

由于初值资料的不确定性,在今后的飓(台)风路径业务预报中将追从集合预报已是不可避免的。Zhang 和 Krishnamurti (1997)<sup>[16]</sup>利用 T63 全球谱模式,基于在最初积分中模式扰动线性增长的前提,为了得到飓风预报集合扰动的全体,不仅对飓风中心初始位置,而且对其结构及环境都需叠加上扰动。初始位置加扰是将其中心原始位置顺次沿北、南、东、西各移 50 km。飓风环境场及其结构的加扰可通过对基于规则分析的预报与基于随机扰动分析预报的差值场实施 EOF 分析来产生。仅仅是温度场和风场叠加上量级正比于相应观测误差的扰动。该方法对每个飓风产生 15 个集合预报成员。研究结果表明,对于大多数个例,集合预报显著地减少了路径预报误差。还值得注意的是,集合预报结果的分散度(Spread)对于评价预报可靠性是有用的<sup>[16]</sup>。Mullen 和 Baumhefner (1994)对 10 个快速发展的海洋气旋分别用

11 个成员的 Monte Carlo 方法进行了模拟。集合预报结果证明, 初始状态的不确定性是导致预报结果变异的主要原因, 并且指出了对初始扰动位置与结构的微小差异的高度敏感性<sup>[17]</sup>。

Straume 等(1998)使用集合预报法分析研究了气象输入资料的不确定性对核泄漏污染物长距离输运预测的影响。集合预报成员是由天气预报给出的初值场叠加扰动形成的。扰动是用奇异矢量法计算得到, 它们表现出了前期预报积分中由气流不稳定引起的各种可能的预报发展。当使用了集合预报的气象场输入, 污染物喷射物流的预报显示出了很大的分散度。这说明气象输入资料质量对于扩散模型的成功预报是至关重要的<sup>[18]</sup>。短期集合预报是基于中尺度过程与降水过程的落区、出现时间以及强度预报存在着极大的不确定性而发展起来的。NCEP 使用区域谱模式的集合预报, 改进了定量降水预报(QPF)。英国气象局用 NWP 模式输出推算对流降水发生概率, 作日常业务降水概率预报。日本气象厅自 1998 年春开始依据集合预报产品为预报员提供强中尺度现象(强降水、疾风)的信息<sup>[8]</sup>。我国学者近年来在集合预报理论及实践研究方面的发展也很快。范新岗(1999)通过集合预报的全局理论及预报实践研究认为, 可以求出最佳集合样本数, 集合预报总能用有限的样本数达到最佳预报结果<sup>[19]</sup>。龚建东等(1999)提出了一种利用四维变分同化方法吸纳 Monte Carlo 法具有明确统计意义的特点和滞后平均预报法包含多时刻初值信息的长处, 来形成一组与动力模式相协调的集合预报初值的方法, 改善了集合预报效果<sup>[20]</sup>。

### 3 结语与展望

集合预报是基于大气运动的随机性以及大气预报系统对初值敏感的混沌特性而发展起来的一种动力-统计相结合的预报方法。它

不同于 MOS 预报。由于集合预报能显著地提高预报的可用性, 延伸预报的实效, 为预报员提供远比确定论预报丰富得多的信息, 因此集合预报已经在短期气候、台风路径、爆发性气旋、中尺度风暴、对流性降水定量预报以及大气污染等的预报与模拟分析中日渐广泛地被采用。其非线性动力学理论基础及集合扰动集的生成方法研究近年来也发展迅速。我国在这一领域的方法研究与应用的发展都明显滞后, 其原因主要是计算机资源的限制, 也有思想认识上的差距。集合预报的方法研究应该在我国得以迅速推广和发展。

海洋环境要素(如海冰、海浪、海温等)预报, 海洋生态系统动态演变的预报(预测), 以及赤潮预报等不仅在系统本身的随机性和对初值场敏感性上同大气预报有很大的相似性, 而且还有着对动力模型参数的高度敏感性, 这些动力学系统都表现出了很强的非线性特性。由此可以推知, 集合预报的思想和方法必将会在海洋环境要素预报、海洋生态系统动态演变预测预报、赤潮预报等研究中起到重要推进作用并得以发展。

### 参考文献

- 1 Lorenz E N. 混沌的本质. 刘式达等译. 北京: 气象出版社, 1997
- 2 史国宁. 从美国气象科技发展探讨气象科技现代化中的方法论问题. 美国气象科技发展动向的研究与分析. 北京: 中国气象科学研究院情报所, 1999
- 3 Gleeson T A. Statistical-dynamical prediction. J. Appl. Meteorol., 1970, 9: 333- 344
- 4 Epstein E S. Stochastic dynamic prediction. Tellus, 1969, 21: 739- 759
- 5 Leith C E. Theoretical skill of Monte Carlo forecasts. Mon. Wea. Rev., 1974, 102: 409- 418
- 6 Hoffman R N, Kalnay E. Lagged average forecasting, an alternative to Monte Carlo forecasting. Tellus, 1983, 35A: 100- 118
- 7 曹鸿兴, 魏凤英. 滞后平均预报的原理及其应用. 气象科技, 1989, (1): 32- 36
- 8 Toth Z, Kalnay E. Ensemble forecasting at NCEP and

- breeding method. *Mon. Wea. Rev.*, 1997, 125(12): 3297 - 3319
- 9 Ehrendorfer M, Errico R M. Mesoscale predictability and the spectrum of optimal perturbations. *J. Atmos. Sci.*, 1995, 52(20): 3475- 3500
- 10 Buizza R. Impact of horizontal diffusion on T 21, T 42, and T 63 singular vectors. *J. Atmos. Sci.*, 1998, 55(6): 1069- 1083
- 11 Palmer T N, Gelaro R, Barkmeijer J. Singular vectors, metric, and adaptive observations. *J. Atmos. Sci.*, 1998, 55(4): 633- 653
- 12 Gelaro R, Buizza R, Palmer T N. Sensitivity of analysis of forecast errors and the construction of optimal perturbations using singular vectors. *J. Atmos. Sci.*, 1998, 55(6): 1012- 1037
- 13 曹鸿兴. 第 12 次 AMS 数值天气预报学术会议简介. *气象科技*, 1999, (1): 1- 4
- 14 陈德辉. 美国数值预报研究动态与业务进展. *美国气象科技发展动向的研究与分析*. 北京: 中国气象科学研究院情报所, 1999
- 15 统计-动力预报研究组(史久恩等). 冬季月平均环流的集合预报研究. 长期天气预报和日地关系研究. 北京: 海洋出版社, 1992
- 16 Zhang Z, Krishnamurti T N. Ensemble forecasting of Hurricane tracks. *Bull. AMS*, 1997, 78(12): 2785- 2796
- 17 Mullen S, Baumhefner D P. Monte Carlo simulations of explosive cyclogenesis. *Mon. Wea. Rev.*, 1994, 122(7): 1548- 1567
- 18 Straume A G, Dri Koffi E N, Nodop K. Dispersion modeling using ensemble forecasts compared to ETEX measurements. *J. Appl. Meteor.*, 1998, 37(11): 1444 - 1456
- 19 范新岗. 集合预报方法的全局研究. *气象学报*, 1999, 57(1): 74- 83
- 20 龚建东, 李维京, 丑纪范. 集合预报最优初值形成的四维变分同化方法. *科学通报*, 1999, 44(10): 1113- 1115