

预 报 “预 报 技 巧”

Eugenia Kalnay Amnon Dalcher

提要 我们已经指出, 预报数值天气预报的技巧水平(一个随时间和地点而变化的量)是可能的。这一点已经达到, 即使用由 5 种不同分析开始的预报集合, 以它们的集合元之间的离差(使用平均相关度量)作为预报因子。上述分析是先前研究卫星资料影响时导出的, 在北半球, 随着不同的观测系统而有适度的扰动。

如果使用北半球作为验证区域, 技巧水平的预报(效果)相当差。这是由于, 这样大的区域常常既包含预报得好的区域也包括差的, 因而难以鉴别其好坏。但是, 如果我们使用按区域验证, 则集合预报离差能非常好地预告各个预报的质量。

尽管本文所论时间仅涉及一个月, 但却包括了所考虑的四个区域中各区域预报水平有较大差异的个例。该方法能够在业务上使用更长时间的延迟预报集合进行试验, 以考查它对于不同的区域和天气状况的可应用性。

一、引 言

短期和长期数值天气预报的技巧水平都随时次和地区变化, 而且也与其它因素有关, 诸如季节、大气状态和边界异常等。因此, Tennekkes 等(1986)已经指出: “对预报技巧水平不作预报的预报是不完全的。”

已经对于数值天气预报进行了一些概率处理, 以便对预报技巧提供先验的估计。Epstein (1969) 研制了一个随机动力(stochastic-dynamic)预报方案, 它包括大气变量概率分布的预报方程。由于问题的规模, 除非对于最简单的模式; 这种方法是行不通的。Leith (1974) 提出使用象蒙特卡洛预报(MCF)那样的集合预报(ensemble forecasting)。他指出由随机扰动的初始条件进行的小量预报($N \sim 8$)的离差是平均预报的预报技巧的一种量度。Hoffman 和 Kalnay (1983) 用公式写成延迟的平均预报(LAF)方法, 它具有 MCF 的优点, 而在中期业务范围, 实际上能够无需费时就完成。它不仅使用最近的业务预报, 而且也使用前一两天开始的同一验证时间的预报。

在 LAF 预报中, 集合的单元被统计加

权。Hoffman 和 Kalnay (1983) 指出, 在低阶模式模拟中, LAF 预报比简单的动力预报或 MCF 预报稍精确。而且如所预料的那样, 预报的范围为集合平均失效时间*提供很好的评估。在 LAF 对 ECMWF 业务预报的最新应用中, Dalcher 等(1985)发现, 尽管 LAF 显示较之业务动力预报有明显的改进, 但是在预报个别预报技巧方面其成绩还是极小的。Dalcher 等把这种缺少成绩归因于使用全球验证, 这也许掩盖了按区域验证的技巧。

本文的目的是提出用集合预报来预报区域预报技巧研究所得到的结果。我们指出, 使用按区域验证会得到关于个别预报质量的很好预报。但是, 如果整个北半球用于验证的话, 则预报技巧的预报质量降低。

二、资料 and 结果

为了试验使用集合预报来预报区域预报技巧的可能性, 我们使用一组 5 种不同的分析和预报, 它们是以以前在资料影响研究中使用过(Halem 等, 1982; Kalnay 等, 1986)。用

* 由于集合误差也包含预报的离散, 所以这一结果不是结论性的。

“FGGE” “NOSAT” “NOTEMP” “NOWIND”和“NOCTW”表示的这些分析,是使用戈达德大气实验室(GLA)的四维分析/预报系统(Baher,1983;Kalnay 等,1983)导出的。在FGGE分析中,使用1979年1月5日至2月2日期间所有可以得到的常规的和卫星的资料。NOSAT分析,除了不使用卫星资料(TIROS-N温度反演和云迹风)之外,完成方式与上类似,在NOTEMP、NOWIND和NOCTW分析中分别删去某些资料集(温度、风和云迹风)。

由于在北半球常规资料密度较大,在资料影响研究中看到,预报一般趋向于互相类

似,而在南半球,资料对预报的影响要大得多。因此可以把从5种分析导出的北半球预报看作是初始条件有适当扰动的预报集合的“元”。

在下面预报预报技巧的尝试中,我们把从NOSAT、NOTEMP、NOWIND、NOCTW(为了简化将分别以B、C、D、E表示)分析作的预报和从FGGE分析(用分析A表示)得到的预报进行比较。图1a、b、c、d给出1979年1月21日0000 UTC(世界协调时)开始的5天预报的500 hPa 和海平面气压距平的相关,验证区为北美和欧洲。下边的(虚)线表示相对于ECMWF分析作验证

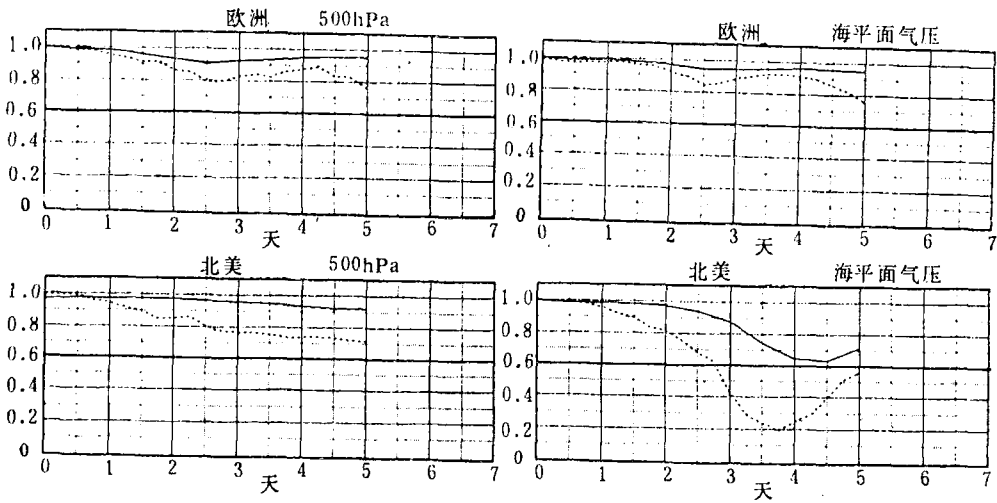


图 1 对应于1979年1月21日起的5天预报的距平相关。实线:平均的预报-预报相关;虚线:预报-分析相关

的,所要做的预报技巧的预报的距平相关。上边的(实)线表示,相对于从分析A得到的预报作为验证的,从B、C、D、E分析得出的预报的距平相关 ρ 的算术平均 R 。它可以在预报时间加以计算。

正如所预料的,从不同分析得出的预报相互之间一般比与从ECMWF分析得到的预报更为接近,这由图中实线比虚线有更高相关看出。不过两条曲线之间仍有些平行现象,这表明预报集合元之间的离差能够用于预报预报技巧的失效,例如,在欧洲距平相关 ρ 为0.8以上时,5天中大部分时间预报是好的,与预报的算术平均作比较,预报距平相关

的平均值仍然超过0.9。在北美500 hPa上预报依然是精确的,但是海平面气压预报失败,仅2.75天以后相关就小于0.6。预报对预报的验证反映出这种恶化,约3.2天相关就变得小于0.85。

下面,我们使用平均的预报/预报相关超过 $R=0.85$ 的时间作为预报技巧失效的预报因子。预报量是实测到的预报失效的时间,即分析/预报相关变为低于 $\rho=0.60$ 的时间。我们已经由1979年1月7日00 UTC—2月2日0000 UTC隔日开始的14个预报得出结果。验证是对四个区:北美、欧洲、北大西洋和北太平洋的海平面气压(SLP)和500

hPa 高度进行的。图 2 给出预报的预报失效时间和实测的预报失效时间的点聚图。由于预报时效仅 5 天，所以超过 5 天的仍然精确的预报就用 6 天失效时间去任意表示。图右上角的大圆点表示有 55 个个例，预报和实测的预报失效时间均在 5 天以后。仅 11 个例子是预报在 5 天以上仍有技巧，而不正确地预报成 5 天之前失效。46 个个例是初期就

表 1 根据4个区域14个预报验证的 500hPa 和海平面气压的有效技巧($\rho \geq 0.60$)列联表

实测的失效	预 报 失 效	
	≥ 5 天	< 5 天
≥ 5 天	55	11
< 5 天	7	39

是正的，则方法是成功的，在表 1 的情况下， $p=0.59, p_1=0.89, p_2=0.22$ ，所以确有 $p_1 > p > p_2$ ，方法是成功的。使用两比例之间差异的标准检验，我们可以估计这种成功的统计显著性(Snedecor 和 Cochran, 1969 第 8 章)，检验是根据统计学的公式

$$Z = \frac{p_1 - p_2}{[p(1-p)(1/N_1 + 1/N_2)]^{1/2}} \quad (1)$$

在无预报技巧($p_1 = p_2 = p$)的零假设下， Z 是均值为零，标准差为 1 的正态分布。 $N_1 = 55 + 7, N_2 = 11 + 39$ 分别是预报为好的和差的预报的数目。当 Z 有连续分布时，由于表 1 矩阵的值是整数，则要从(1)式分子中减去 0.5 ($1/N_1 + 1/N_2$) 做连续修正 (如果 $p_1 > p_2$)。从表 1 中得出的 Z 值是 6.94，相当于单侧显著性概率 (one-sided significance probability) 为 10^{-11} 。即使因为我们没有使用独立样本，和 500 hPa 与 SLP 验证之间有某些相关而使自由度数减小，显然上述结果也仍然是高度显著的。

表 2 分别给出每一验证区的类似的列联表。对角线上的数值大表明预报因子在每一地区很好地判别出有技巧的预报 (无论是少

表 2 对各个指定区的验证列联表 (说明同表 1)

北 美	欧 洲	北大西洋	北太平洋
8 4	14 1	19 1	14 5
2 14	2 11	2 6	1 8

于 5 天还是大于 5 天)。另外，如果对整个北半球进行验证，有 $n_{11} = 11, n_{12} = n_{21} = 7, n_{22}$

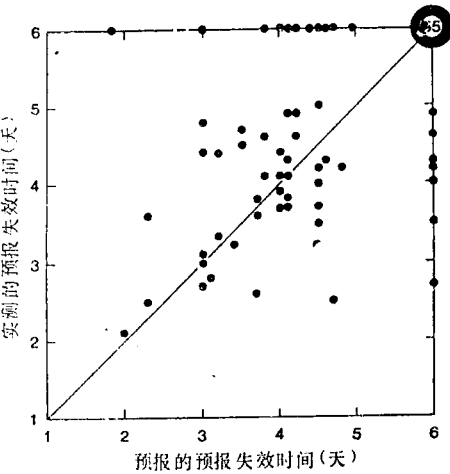


图 2 预报的预报失效时间和实测的 预报失效时间的点聚图。5 天以后失效的点在第 6 天

失效。39 个个例正确地预报出 预报技巧持续不超过 5 天，仅 7 个预报其技巧持续 5 天以上。这些资料在表 1 上列出。

表 1 假定，我们能够有技巧地做预报，不管个别预报在 5 天内是否仍然是好的 (也就是它对分析的相关是否保持在 0.6 以上)。为了判断预报质量的性能，需要计算

$$p = (n_{11} + n_{12}) / (n_{11} + n_{12} + n_{21} + n_{22})$$

$$p_1 = n_{11} / (n_{11} + n_{21}) \quad \text{和}$$

$$p_2 = n_{12} / (n_{12} + n_{22})$$

其中 n_{ij} 是表 1 中的元素， p 是所有预报中好的预报 (即 5 天中不失效的预报) 的比例； p_1 是预报报为好的中出现好的 预报的比例； p_2 是预报报为差的中好的 预报 的比例。如果 $p_1 > p > p_2$ ，则先验的技巧预报是成功的。直接计算表明不等式 $p_1 > p, p > p_2, p_1 > p_2$ ，每一个相当于另外两个不等式，且也相当于 $n_{11}n_{22} > n_{12}n_{21}$ 。因此，如果列联表的行列式

=3,判为失败。

在 $p=0.64, p_1=0.61, p_2=0.70$ 的条件下, $Z=-0.06$, 单侧概率为 0.52, 表明缺乏先验预报技巧, 甚至表明负的 Z 值与有技巧 $p_1 > p > p_2$ 条件相反的情况有关。表 3 综合了表 1、2 的条件概率 p_1, p, p_2, Z 统计量和显著性概率。

表 3 5 天以上仍然好的预报 ($\rho \geq 0.60$) 的部分概率, Z 统计量和显著性概率

区域	p_1	p	p_2	Z	显著性概率(Z)
4 个区	0.89	0.59	0.22	6.94	$2 \cdot 10^{-11}$
北美	0.80	0.43	0.22	2.96	$5 \cdot 10^{-3}$
欧洲	0.88	0.54	0.08	3.77	$8 \cdot 10^{-5}$
北大西洋	0.90	0.71	0.14	3.86	$6 \cdot 10^{-5}$
北太平洋	0.93	0.68	0.38	3.69	$1 \cdot 10^{-2}$
北半球	0.61	0.64	0.70	-0.06	0.52

我们看到,按区域进行的验证,在定性地预报预报技巧方面是很成功的,而对大的验证区域,该方法不能用,因为它受不同的天气系统影响并包括有好的和差的预报技巧的区域。也应指出,预报技巧对准确的 R 值(选作阈值)不敏感, R 改变在 0.05 内可以视为有类似的技巧。

指出下列情况是有意义的,当我们尝试使用象 rms 误差和 S_1 评分标准去量度预报技巧时,则不太成功。使用相关的优点可能在于相关是有界的。假若我们把每个预报视作多维矢量,则两预报间的相关,是相应矢量之间夹角的余弦。由于角度不太大, $\cos \alpha \approx 1 - \alpha^2/2$, 所以我们选作预报因子的平均相关 R 的确是刻划预报分散性的量度。

三、较短时间的预报

我们现在用类似的方法确定预报终止时是“很好”(“very good”)的预报的先验时间,把它规定作据 ECMWF 分析检验其区域的距平相关达到 $\rho=0.80$ 的时间。对 112 个预报调查的结果是,选用平均预报/预报相关达 $R=0.92$ 的时间作预报因子。

图 3 为预报终止时是很好的预报的时间

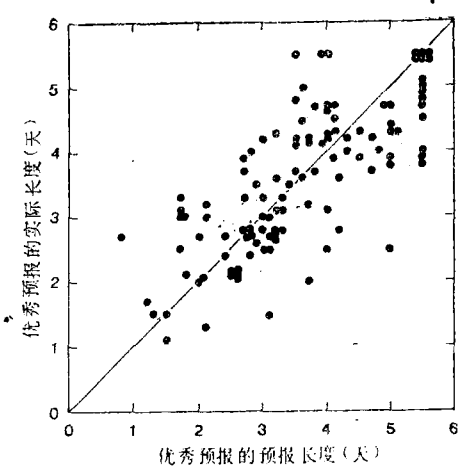


图 3 非常好的预报(相应距平相关大于 80%)的预报时间长短和实测长短的点聚图。把 5 天以上仍然非常好的预报随意点作 5.5 天

与实测时间的点聚图。时间超过 5 天的个例任意指定为 5.5 天。我们又一次看到先验的预报技巧。类似表 1,我们能够制作表(表 4),在其中我们考虑 3.5 天以上仍然很好的预报的和实测的预报的比例。

从表 2 得出的统计值是 $Z=7.73$, 它有单侧显著性概率 1.2×10^{-15} , 对于指定区域的类似表(表 5)也表明有好的技巧。在这种情况下,甚至对北半球热带以外的整个区域进行验证,也表明有某些技巧($n_{11}=11, n_{12}=3, n_{21}=5, n_{22}=9$)。

表 4 依据 4 个区域 500hPa 和 SLP 验证的 14 个预报资料作的列联表, 它表明 3.5 天以上非常好的预报 ($\rho \geq 0.80$) 对预报其仍然非常好的预报的比例

实测到非常好的技巧	预报有非常好的技巧	
	≥ 3.5 天	< 3.5 天
≥ 3.5 天	54	8
< 3.5 天	6	44

表 6 总结了对列联表 4、5 的条件概率值、 Z 统计量和它的显著性概率。

当 R 值摆动到 0.02 左右,可以得到类似的技巧。在这种较短的期间, R 接近于它的最大值,因此很少有变动的余地。

表 5 同表4但为给定个别区域验证的列联表

北 美		欧 洲		北大西洋		北太平洋	
7	2	13	2	18	2	16	2
1	18	4	4	1	7	0	10

表 6 同表3,但是是对于 3.5 天以上
仍然非常好的预报的部分

区域	p_1	p	p_2	Z	显著性概率(Z)
4 个区	0.90	0.55	0.15	7.73	$5 \cdot 10^{-15}$
北美	0.88	0.32	0.10	3.52	$2 \cdot 10^{-4}$
欧洲	0.76	0.54	0.18	2.63	$4 \cdot 10^{-3}$
北大西洋	0.95	0.71	0.22	3.52	$2 \cdot 10^{-4}$
北太平洋	1.00	0.64	0.17	4.16	$2 \cdot 10^{-5}$
北半球	0.69	0.50	0.25	1.91	$3 \cdot 10^{-3}$

四、预报的和实测的相关之间的关系

根据至今介绍的结果,我们使用下列方法构成预报/分析和预报/预报相关之间的对应关系;对于平均的预报/预报相关 R 的每一个值,确定跨过中值时间 T ,即有半数预报至少一次达到 R 值的时间。然后确定中值预报/分析相关 ρ ,使得在时间 T ,半数预报仍没有达到 ρ 。

图 4 给出导出的关系 $\rho(R)$,由于预报时间仅有 5 天,所以仅有对 $\rho \geq 0.66$ 和 $R \geq 0.86$ 的资料。但是,在图 4 中两相关之间显现为几乎线性关系,我们采用下列的校准公式:

$$\rho = 2.205R - 1.215 \quad (2)$$

在 $R = 0.86$ 以下,我们也用它外推。

现在我们给出欧洲 14 个预报的每一个预报的实测的距平相关 ρ 和从(2)式导出的预报的相关 $\rho(R)$ 。图 5 是 500hPa 高度的情况。如表 3 和表 6,对欧洲验证的结果是 4 个区域的代表,它比北美的验证结果稍微好一些,比北大西洋的结果稍差一些。应该指出两层(如 hPa 和海平面)的技巧预报精度是类似的,而且样本包含有高的及低的可预报性的情况。在其它三个区域,这些观测结果

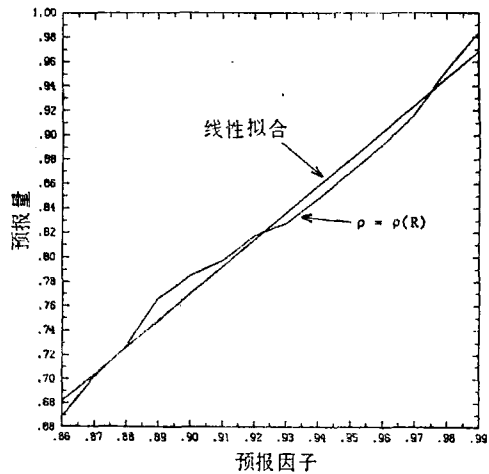


图 4 平均预报/预报相关 R 与使用中值法导出的预报/分析相关 ρ 的关系及它的线性拟合

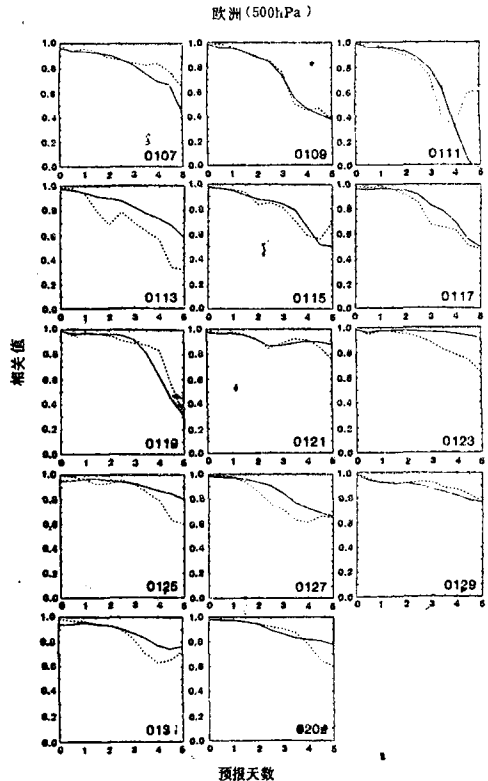


图 5 在欧洲验证的 14 个 5 天预报的观测的(虚线)和预报的(实线)距平相关的比较。横轴是预报时间,纵轴为相关。初始条件的月、日用每张图中的 4 个数字表示

也是这样。尽管有太乐观和太悲观的技巧预报的情况,但综合起来,图 5 表明,我们的方法提供了一个逐日技巧的评估,这在业务上是有用的。

五、讨 论

由于试验仅用一个月时段,我们的结果显然是初步的,且可能并非在所有的环境条件下成立。尽管如此,这些结果也很令人鼓舞,并且的确指出,即使集合小至仅 3、4 个,小的预报集合其集合元的离差也能是预报技巧的好的先验预报因子(在我们的集合中,由 NOCTW 分析得到的预报几乎总是非常接近 FGGE 预报,因此影响无差别)。

使用相同的判据 $R(\rho)$,且结果对于所选的精确值不敏感这一事实指出,该方法在超出这种资料集时可以应用。

应该深入讨论预报技巧变动的物理学原因。能够引起区域可预报性下降的两类过程是不稳定性(诸如气旋新生),和天气形势改变(象阻塞高压的发展或衰减)。在这两种情况下,可以预料,预报突然转变所增加的困难

会伴随一预报集合中集合元间离差增大。当然,这是使用集合预报进行技巧预报得出的基本看法。研究大气不稳定性(Tennekes等,1986)和天气形势特征典型存在时间(Mo 和 Ghil,1985)的时、空变化,能更加理解所涉及的过程。

所介绍的预报预报技巧的程序已被用于有中等扰动分析的集合,它类似于应用蒙特卡洛预报(Leith,1974),虽然分析间的差别反映的是资料的不精确性而不是随机误差。Hoffman 和 Kalnay(1983)使用一低阶模式的结果表明,使用延迟平均预报集合同样有效,尽管它可能需要进行更频繁的中期预报而不是一天一次,该方法应在业务范围内进行试验,使用更长的时期以试验它对不同区域、不同季节和不同天气形势的可应用范围。

胡圣昌节译自 «Monthly Weather Review» Vol.115, No. 2 文及校