

第四讲 中期数值预报的统计检验分析

刘还珠 张绍晴执笔*

(国家气象中心数值预报运控室, 北京 100081)

提 要

介绍了CBS标准化检验方法, 用该方法对 $T_{4.2}$ 的结果进行了检验, 并与国外有关模式的结果进行了比较。结果表明, $T_{4.2}$ 模式基本上达到了4天可用预报的水平。

关键词: 中期数值预报 统计检验 CBS标准化检验

引言

10年前, 世界气象组织(WMO)按照大气科学委员会(CAS)天气预报研究工作组建议, 要求各制作业务数值预报的气象中心参加“1979FGGE年日常业务分析预报资料的研究和对比”工作。这项工作的主要任务, 是对比各气象中心的数值预报结果, 经过检验比较而积累了大量的经验。据此, 基本系统委员会(CBS)1985年特别会议决定: 自1986年10月1日始, 数值预报业务中心执行统一的标准化检验方法。

1 CBS标准化检验方法简介

1.1 CBS标准化检验的有关规定

CBS确定的数值预报产品日常标准化统计检验包括两部分, 一部分是用各自的客观分析资料为实况, 另一部分是用探空观测为实况, 对预报产品分别进行检验。标准化检验的有关规定和内容由表1给出。

世界气象组织选定为标准检验的探空观测站共有365个, 其中北美97个, 欧洲85个, 亚洲153个, 澳大利亚和新西兰地区30个。在这365个探空标准站中, 我国有78个。用客观分析作检验时, 由于规定的资料为 2.5°

表 1 标准化检验的有关规定和内容

检验方案	用客观分析资料	用观测资料
要素	海平面气压 位势高度 气温、风	海平面气压 位势高度 气温、风
时效	24, 48, 72, 96, 120... 北半球(20N-90N)	24, 48, 72, 96, 120... 北美、欧洲、 亚洲、 澳大利亚新西兰
范围	赤道地区(20S-20N) 南半球(90S-20S)	
层次	海平面 500hPa 250hPa	850hPa 500hPa 250hPa
资料点	2.5×2.5 (0.0为原点)	365站
检验统计量	平均误差 均方根误差 相关系数 SI评分	平均误差 均方根误差 相关系数 风矢量均方根误差

$\times 2.5^\circ$ 等经纬网格, 为使区域检验评分更为合理, 计算时考虑纬度权重, 以达到等面积分布的检验效果。

CBS要求将标准化检验列入日常业务工作, 并把每月检验结果除上报世界气象组织外, 还参与分析预报中心(目前有6个)之间的实时交换, 以便各国之间的评估。北京气象中心由于种种原因, 目前尚未参加这项对比工作, 但已按照CBS标准化检验方法, 建立起数值预报日常业务检验系统, 并从

*参加该项工作的还有许媛、陆志善、张德新、杨云峰等。

1992年1月开始,每月向全国气象台提供检验结果。可望在不久的将来,能正式参加国际间的数值预报检验对比工作。

1.2 CBS 规定的检验指标的数学表达及天气学意义

倾向相关系数:

$$TEN \cdot COR = \frac{\Sigma(F - A_0 - M_{f_0})(A_v - A_0 - M_{v_0})}{[\Sigma(F - A_0 - M_{f_0})^2 \Sigma(A_v - A_0 - M_{v_0})^2]^{1/2}}$$

距平相关:

$$ANO \cdot COR = \frac{\Sigma(F - C - M_{fc})(A_v - C - M_{vc})}{[\Sigma(F - C - M_{fc})^2 \Sigma(A_v - C - M_{vc})^2]^{1/2}}$$

SI评分:

$$SI = 100 \times \frac{\Sigma(|F_x - A_{vx}|)(|F_y - A_{vy}|)}{\Sigma[\max(|F_x|, |A_{vx}|) + \max(|F_y|, |A_{vy}|)]}$$

风矢量的检验指标用分量风的合成结果表征:

$$ME(\vec{V}) = [ME(u)^2 + ME(v)^2]^{1/2}$$

$$RMSE(\vec{V}) = [RMSE(u)^2 + RMSE(v)^2]^{1/2}$$

$$COR(\vec{V}) = \left\{ \frac{1}{2} [COR(u)^2 + COR(v)^2] \right\}^{1/2}$$

上列各式中, F 为预报值, A_v 为分析值, A_0 为预报初值, C 为气候平均值, N 为检验区域上的格点数。 $M_{f_0} = \frac{1}{N} \Sigma(F - A_0)$, M_{v_0}

$$= \frac{1}{N} \Sigma(A_v - A_0), M_{fc} = \frac{1}{N} \Sigma(F - C),$$

$$M_{vc} = \frac{1}{N} \Sigma(A_v - C), F_x = \frac{\partial F}{\partial x}, F_y =$$

$$\frac{\partial F}{\partial y}, A_{vx} = \frac{\partial A_v}{\partial x}, A_{vy} = \frac{\partial A_v}{\partial y}, \Sigma \text{ 为区域内所有格点求和。}$$

从计算式中可以看出,平均误差计算时,正负误差抵消,它反映的是统计区域内的某种系统性误差。均方根误差给出统计区域内误差幅度的平均情况,因而它是衡量预报误差最常用的一个统计参数。倾向相关系数反映预报场与实况场变化趋势的相似程度。从天气学意义上讲,它反映的是槽脊移

平均误差

$$M_E = \frac{1}{N} \Sigma(F - A_v)$$

均方根误差:

$$RMSE = [\frac{1}{N} \Sigma(F - A_v)^2]^{1/2}$$

动和强度变化的预报效果。随着预报时效的增加,倾向相关系数也随之减小,但有时会出现增加的情况,这是因为如果模式预报槽脊移动的速度误差较大,而预报时效较长时,会出现预报的槽脊与分析的非对应槽脊重合而造成虚假的倾向相关系数的上升,因而欧洲中期预报中心提出了距平相关系数的概念。距平相关系数也反映槽脊位置和强度的预报效果,但因它利用的是分析和预报的气候距平相关,避免了倾向相关系数随预报时效增加所出现的虚假增长的现象。SI评分主要反映高度场的梯度预报精度,但它确实抓住了高度场分布的核心内容,因而能够反映出模式预报的“技巧”。

随着预报时效的增加,预报精度下降,反映在统计检验指标上是均方根误差、平均误差增大,相关系数减小,SI评分增加,均方根误差和SI评分在预报时效较短时增加较快,而预报时效较长时,增长变慢。均方根误差随预报时效的继续增长,而趋于一个渐近值即大气的自然变化率。相关系数随预报时效增加减小的变率,表明了预报模式的好坏。一般把距平相关系数为0.60作为可用预报的界限,如在该值以上,则认为在日常预报中有参考价值,可利用的正确预报信息要多于错误信息。事实上,当距平相关系

数小于0.60时, 预报结果中仍有可利用的预报信息。

2 T₄₂中期数值预报模式的统计检验分析

2.1 距平相关逐月变化

图1给出了1990年1月到1991年12月T₄₂模式预报对客观分析的北半球距平相关逐月变化。图中表明, 距平相关系数在3个层次上的变化趋势比较一致, 只是愈近低层变化幅度愈大, 一年中6—9月相关系数较小, 11—5月较大。随着预报时效的增加, 相关系数的季节变化也加大。图1还表明了对流层中上层72小时以内的预报距平相关系数都超过了0.60的可用预报界值, 而96小时的预报在冬季达到0.60以上, 夏季(6—9月)略低于这一数值。850hPa的预报情况比对流层中上层的预报要差一些, 96和120小时预报都低于0.60。但总的来说T₄₂的96小时预报基本上达到了可用预报的水平。

2.2 1991年平均统计检验结果分析

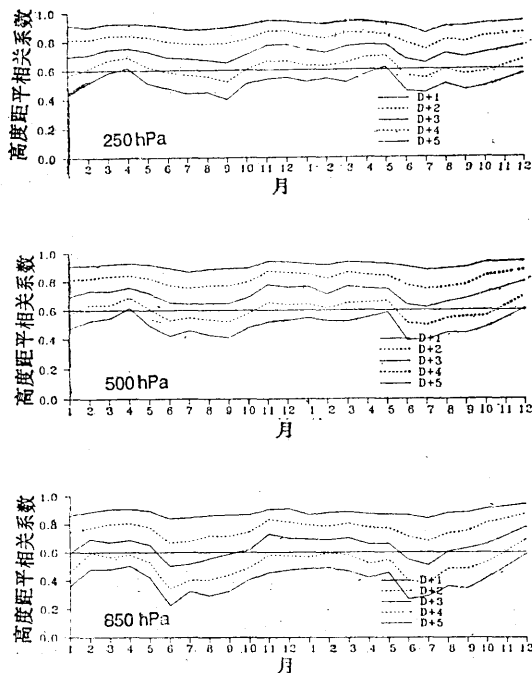


图1 1991年1月—1992年12月T₄₂北半球预报的距平相关系数

表2给出了T₄₂模式北半球1991年在250hPa、500hPa和850hPa3个层次上的预报对客观分析的年平均统计检验结果。可以看出, 高度场和风矢量预报的均方根误差随着预报层次的增高和预报时效的增加而增大, SI评分随预报时效的增加而增大, 低层数值较小, 中、高层比较接近。从这些统计量逐月变化情况(表略)还可以看到, 它们具有显著的季节变化特点, 冬半年预报误差大于夏半年, 以12月—3月的误差较大, 而6—9月较小, 相关系数的变化与均方根误差变化相同, 即冬半年相关大, 夏半年相关小。这是因为冬季要素的变化幅度大且系统较稳定而致使均方根误差大, 相关也大。一般而言, 冬夏季节间最大与最小之间的误差也是随着预报时效的增加而增大, 例如500hPa均方根误差全年月平均最大与最小之间的差额, 24至120小时分别是13, 20, 26, 30和32gpm。

表2 1991年北半球T₄₂预报对客观分析资料的检验结果

时效		24 ^h	48 ^h	72 ^h	96 ^h	120 ^h
H RMSE	250	36.7	57.9	77.3	95.5	111.1
	500	23.9	39.2	53.3	66.6	77.4
	850	19.7	31.0	40.8	49.6	56.4
H T.COR	250	0.826	0.817	0.779	0.722	0.673
	500	0.843	0.831	0.780	0.714	0.662
	850	0.807	0.799	0.740	0.662	0.607
H SI	250	31.3	40.2	47.5	53.5	58.4
	500	28.0	42.0	50.5	57.3	62.8
	850	40.9	52.9	61.3	67.8	73.0
V RMSE	250	7.6	10.4	12.8	14.8	16.5
	500	4.7	7.0	8.8	10.4	11.6
	850	3.9	5.6	6.8	7.7	8.4

为比较探空观测资料作检验的情况, 我们将1991年全年平均的500hPa高度场和250hPa风场两种方法的检验结果列成表3。由表可见, 用两种方法进行的统计检验结果与上述

表 3 1991年T₄₂北半球预报的均方根误差

时效	对客观分析		对观测资料	
	500hPa高度	250hPa风	500hPa高度	250hPa风
24 ^h	23.9	7.6	28.6	10.1
48 ^h	39.2	10.4	44.2	12.9
72 ^h	53.3	12.8	59.4	15.4
96 ^h	66.6	14.8	73.7	17.4
120 ^h	77.1	16.5	85.9	19.1

结论颇为一致，只是用探空观测资料作检验所得误差值大一些。

2.3 误差的分布

图2为500hPa和1000hPa1992年2月的72小时和96小时月平均预报误差图，图中表明，两个时效的预报误差都具有类似的特点。在500hPa上，中高纬带欧洲大部、西亚、太平洋到北美西部均为负偏差，其余为正偏差区，正误差中心分别位于北美北部(131gpm)和亚洲东部(105gpm)，负误差中心分别位于西欧(-191gpm)和北美西岸(-82gpm)。在低纬带，除大西洋中部的副热带地区和欧洲南部，北非西部为负偏差外，其余多为正误差区。在1000hPa

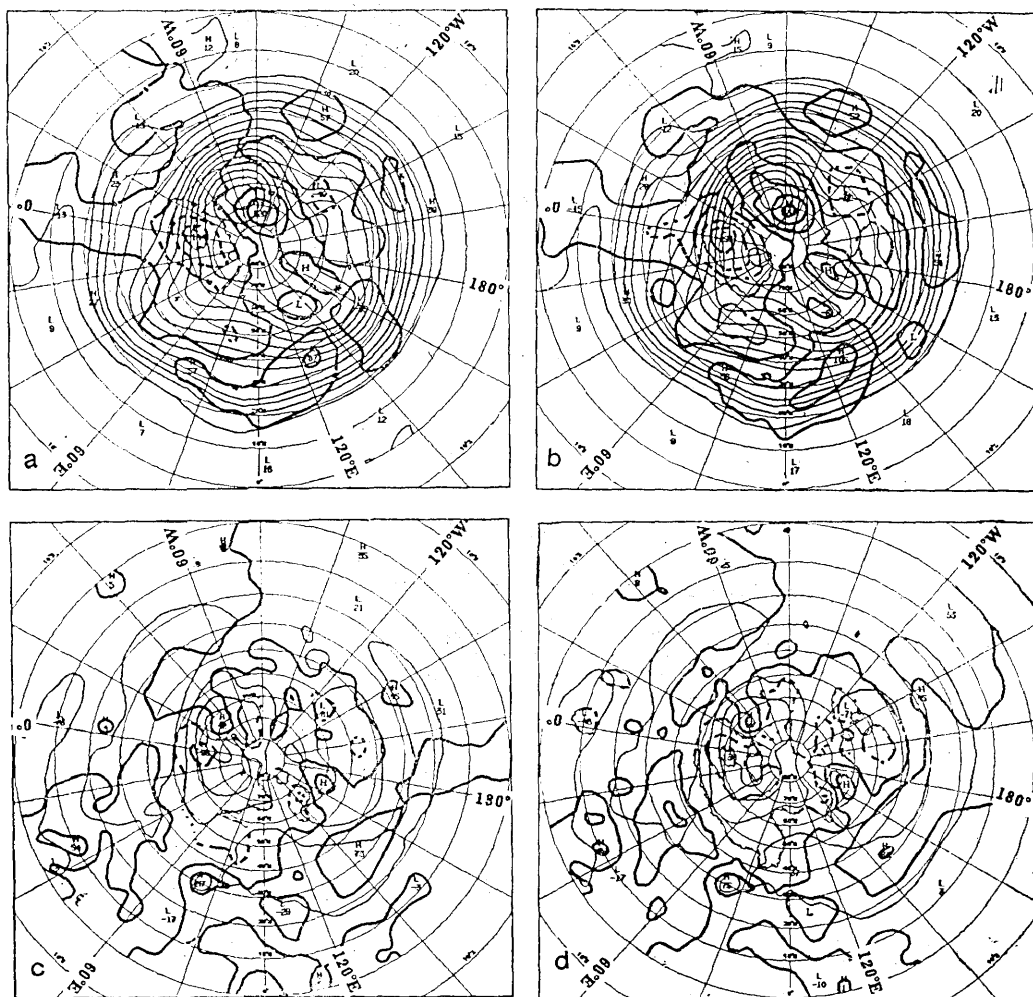


图2 1992年2月500hPa72小时(a)、96小时(b)、1000hPa72小时(c)、96小时(d)高度预报平均误差图

粗线为误差 细线为高度

上, 50°N以北除白令海峡和加拿大东部为正误差外, 其余地方均为负误差区, 最大误差中心分别位于西北欧(-136gpm)和北美西北部(-71gpm)。在50°N以南, 除美洲东部经大西洋到欧洲及南亚部分地区为负误差外, 其它地方都为正误差区, 最大中心位于中亚和太平洋西部。

2.4 T₄₂模式与有关模式预报检验结果的比较

为了进一步了解T₄₂的预报水平, 这里引入了加拿大(Canada), 英国(UK), 日本(JMA)及国家气象中心短期预报B模式1991年1月和7月的均方根误差和倾向相关

系数。如图3所示。从图中可以看出, T₄₂比B模式的预报水平有了长足的提高, 其96小时预报比B模式48小时预报水平高; T₄₂的预报比加拿大、英国落后一天, 即T₄₂的96小时预报水平相当于这两国的120小时的预报水平, 而T₄₂比日本的预报就更差一些, 尤其是在冬季。表5是1992年1月T₄₂与欧洲中心(EC)的几个统计检验指标的北半球500hPa月平均情况对比, 表中说明, T₄₂在该月比欧洲中心落后两天左右, T₄₂5天预报的均方根误差是96.2, 欧洲中心的第7天预报均方根误差是126.8。

综上所述, T₄₂中期数值预报模式基本

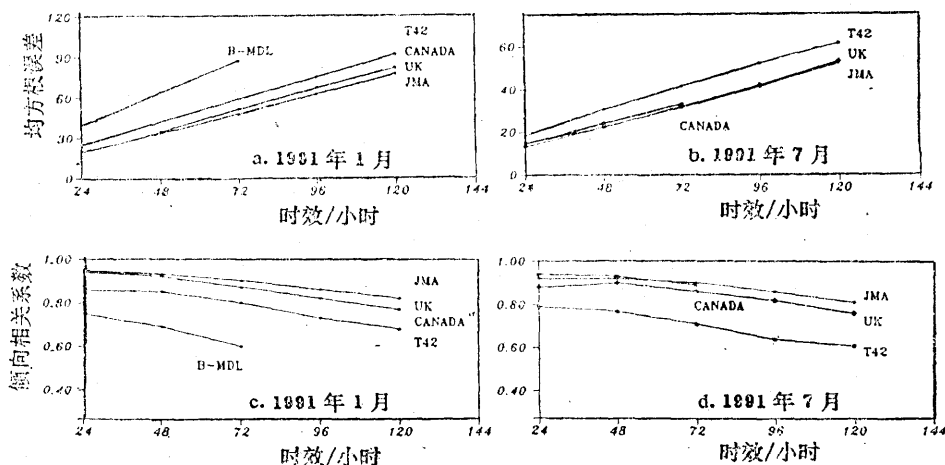


图3 T₄₂、B模式、加拿大(CANADA)、英国(UK)、日本(JMA)用客观分析资料检验的比较

表4 1992年1月欧洲中心(EC)与T₄₂500hPa月平均统计指标比较

时效	均方根误差		距平相关系数		技巧评分	
	EC	T ₄₂	EC	T ₄₂	EC	T ₄₂
24	15.9	27.2	0.99	0.95	17.7	25.1
48	27.0	46.2	0.97	0.87	26.3	35.0
72	40.3	63.7	0.94	0.77	33.7	43.0
96	55.2	80.4	0.89	0.69	40.3	47.8
120	72.3	96.2	0.81	0.54	46.8	52.3

上已达到了4天可用预报的水平, 比B模式的短期预报有了长足的提高, 比先进的日本、加拿大、英国的预报落后一天, 比欧洲中心预报落后两天。该模式对中高纬度天气

系统的预报较好, 但由于分辨率较低, 对一些较小的天气尺度系统预报较差; 对高原, 低纬带及极地附近的天气预报能力较差, 这可能是地形、辐射和地面过程等物理过程的缺陷。当分辨率更高, 各种物理过程处理得更完善的模式(T₆₃)投入业务以后, 可望这些问题会得到较好的解决。

3 国家气象中心模式评价检验的展望

目前, 国家气象中心已建立了比较完善的CBS标准化模式统计检验业务系统, 每月在《数值预报产品评价公报》上发布月平均

(下转第61页)

(上接第54页)

标准检验结果交流。但是,对模式的评价仅有统计检验是很不够的,目前正致力于从动力学的角度,对中期数值模式进行检验和评价,一些以大气环流统计量和西风加速为主要内容的动力检验子系统已建立起来,并

将在《公报》上逐渐充实。当中期数值预报的资料积累到一定程度,各种评价检验手段都发展得比较成熟的时候,模式的检验将为模式的改进提供意向明确的、客观的依据,并将为各级气象台站更好地应用数值预报产品提供更好更快的服务。