

中期数值预报(T63 L16)

系统的统计性能评估

张绍晴 杨晓梅 黄卓

(国家气象中心, 北京 100081)

提 要

介绍了世界气象组织(WMO)基本系统委员会(CBS)标准化统计检验方法。用该方法对 T63 模式产品进行检验, 并与 T42 及日本模式的结果相比较。结果表明, T63 模式产品的可用预报时效可达 5—6 天, 其预报误差幅度比 T42 明显减小。

关键词: CBS 标准化检验 T63 模式预报 性能评估

引 言

1985 年 WMO 的基本系统委员会(CBS)特别会议决定, 各制作数值预报的业务中心执行统一的检验标准, 我国自 1990 年中期数值预报系统建立以来, 已建起了自己的标准化统计检验子系统, 刘还珠等^[1]曾对标准化检验的一些规范作过介绍, 其中表 1 为该规范的归纳, 这里再度引用。

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, 144, 168	24, 48, 72, 96, 120, 144, 168
范围	北半球(20N—90N) 赤道地区(20S—20N) 南半球(90S—20S)	北美、欧洲 亚洲、 澳大利亚新西兰
层次	海平面 500hPa 250hPa	850hPa 500hPa 250hPa
资料点	2.5×2.5(0, 0 为原点) 平均误差	365 站 平均误差
检验统计量	均方根误差 相关系数 SI 评分	均方根误差 相关系数 风矢量均方根误差

与 T42 L9 相比, T63 L16 除纬向截断波

数从 42 个波增加到 63 个波外, 模式中物理过程亦有所完善^[2], 同化方案也有很大改进^[3], 从图 1 可见, 目前 T63 L16 的 96 小时预报水平已与 80 年代的短期 B-模式预报系统的 48 小时预报水平相当。本文将对标准化统计检验的几个常用指标的数学物理意义作一些较为深入的探讨, 着重对 T63 L16 的性能作评估。

1 统计检验参数的数学物理意义

关于几个常用检验学指标的天气学意义, 已作过阐述^[1]。这里结合各参数的数学表达式对它们的物理图象作一个简要的描述。

1.1 平均误差

如其数学表达式 $M_F = \frac{1}{N} \sum (F - A_o)(F$ 表预报, A_o 表分析, N 为总格点数) 近述之意义, 即为区域内各点预报与实况差值之平均, 计算时, 正负误差会出现相互抵消的现象, 因而, 随着预报时效的增长, 平均误差并不是越来越大的^[4]。

1.2 均方根误差

均方根误差的数学表达为:

$$RMSE = \left[\frac{1}{N} \sum (F - A_o)^2 \right]^{\frac{1}{2}}$$

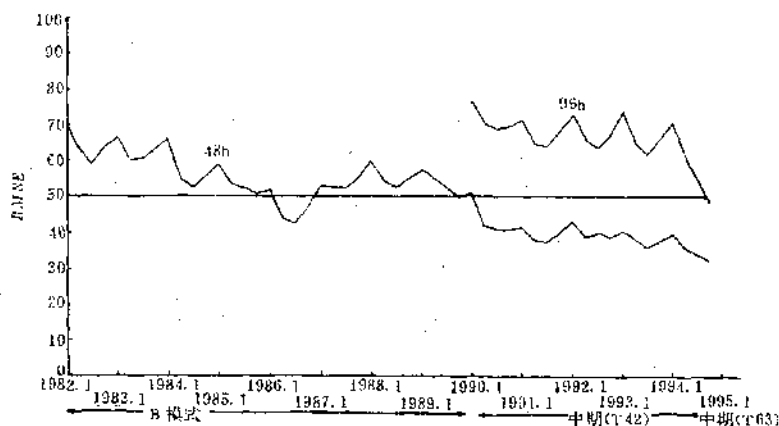


图1 国家气象中心数值预报误差演变

即检验区域内每点的预报与实况差值平方之平均求平方根,是预报场与实况场偏差程度的量度,一般而言,该统计量随着预报时效的增长而增长,但预报到某天左右时, $RMSE$ 呈缓变状态,之后增长较小,甚至有

时略有下降,这可能是模式预报的误差幅度已到达大气的自然变化值之缘故^[4]。

1.3 技巧评分

其数学表达式为:

$$SI = 100 \times \frac{\sum (|F_x - A_{tx}|)(|F_y - A_{ty}|)}{\sum [\max(|F_x|, |A_{tx}|) + \max(|F_y|, |A_{ty}|)]}$$

(X_s 表 X 对 s 的偏导)式中表明技巧评分计算时只用到变量的梯度值。当预报梯度与实况梯度完全相同时, SI 取最小值为零,梯度差越大, SI 越大。因而,随着预报时效的

增加,梯度误差越来越大, SI 亦随之增大,增大到一定程度趋于不变^[4]。

1.4 倾向相关与距平相关

它们的数学表达式为:

$$\text{倾向相关: } TC = \frac{\sum (F - A_0 - M_{f0})(A_t - A_0 - M_{t0})}{[\sum (F - A_0 - M_{f0})^2 \sum (A_t - A_0 - M_{t0})^2]^{\frac{1}{2}}}$$

$$\text{距平相关: } AC = \frac{\sum (F - C - M_{fc})(A_t - C - M_{tc})}{[\sum (F - C - M_{fc})^2 \sum (A_t - C - M_{tc})^2]^{\frac{1}{2}}}$$

式中, A_0 和 C 分别表示初始场和气候场, M 表示相应的全场平均。上述数学表达式表明,倾向相关和距平相关都是统计学上两个物理场的相关系数的算法,两者不同的是,它们各自减掉的基本场不一样,倾向是将预报和实况都与初值相减,而距平是与气候值相减。随着预报时效的增长,相关系数会随之下落,倾向相关下降得慢,距平相关下降得快。当下降到其最低点后,会出现回升现象,以后呈振荡形式前进^[4]。一般而言,距平

相关的最低点较倾向相关为低,其振幅亦比倾向为小。

相关系数的波动形式可能是大形势场(位势高度分布)的特有现象,我们知道,大气的基本运动形式是波动,造成两个波动场差距的原因不外乎是振幅和位相。在假定单波形式解的条件下,只考虑位相偏差时,如果预报相速为实况相速的一半,则倾向相关可简化为 $TC = \cos \frac{1}{2} \omega t$,说明在此条件下倾向相关将以比预报相速还小一倍的相速往前传

播,同样在一定假设条件下,由相速误差造成的距平相关表达式为 $AN = \frac{1}{2} [1 + \cos(\omega_a - \omega_f)]t$,说明距平相关亦是振荡前进的,但其振荡频率 $(\omega_a - \omega_f)$ 要比倾向相关来得大。这也许就是为什么距平要比倾向下降得快的原因。

2 T63 L16 系统准业务运行以来的统计性能评估

2.1 T63 系统与 T42 系统的比较分析

自 1994 年 7 月 1 日开始,第二代中期数

值预报系统(T63 L16)在银河-Ⅱ计算机上准业务运行,系统运行稳定。如前所述,该系统除采用较 T42 L9 更高分辨率的谱模式外,在物理过程的描写和资料分析技术方面均有较大的变化。因此,从初值到预报较 T42 L9 提供的产品均有较大的改进。图 2a、b 分别为 T63 和 T42 1995 年 1 月制作的 500hPa 96 小时月平均形势预报图,c、d 分别为它们的误差分布。与月平均分析场比较,T63 的预报无论从槽脊位置还是闭合系统的强度上都较

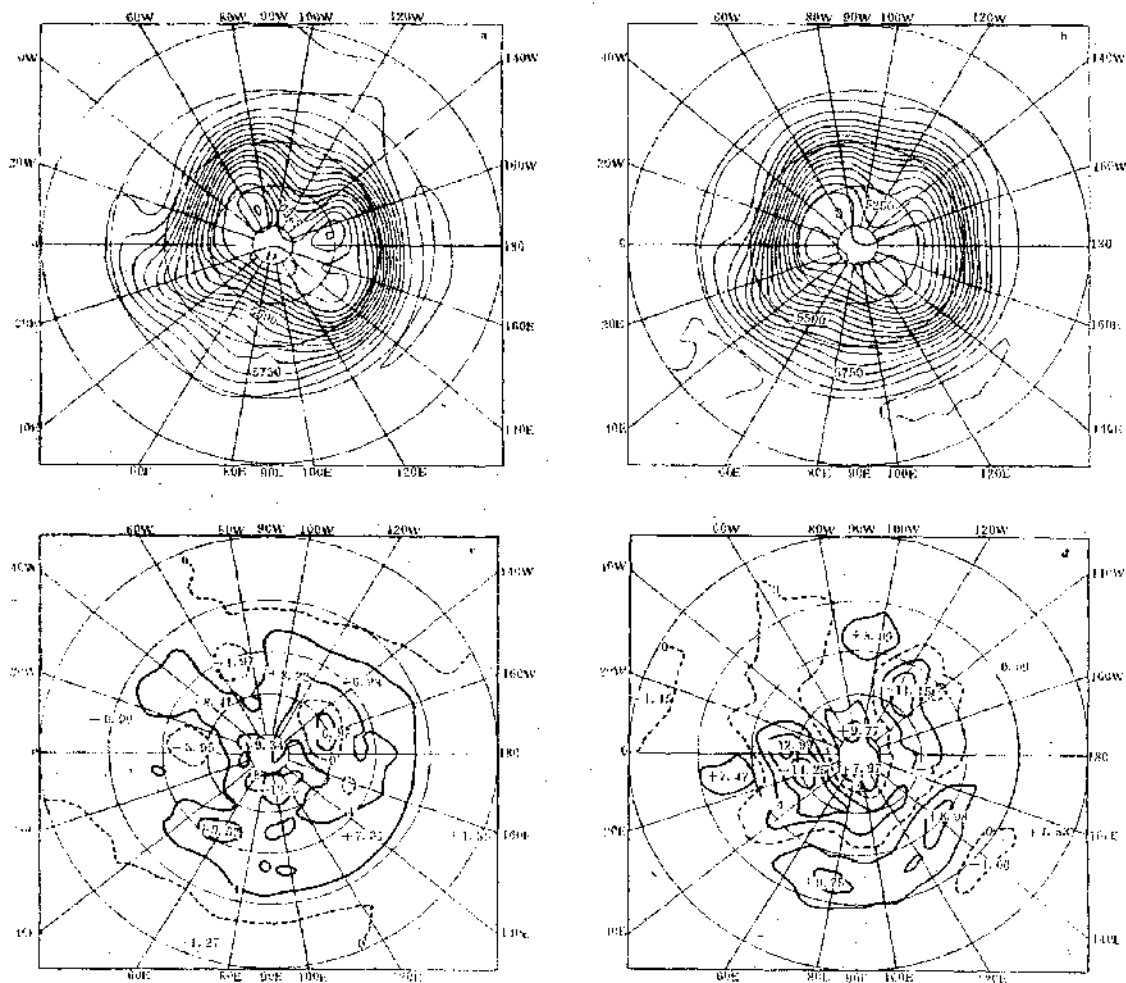


图2 T63 和 T42 1995 年 1 月 500hPa 96 小时预报月平均位势高度及其误差分布
a 为 T63 预报场, b 为 T42 预报场, c 为 T63 预报误差, d 为 T42 预报误差

T42 更加贴近于实况(分析),但仍然存在误差,其误差的正负中心强度较 T42 为弱,结合下面的统计评分,我们将会看到,T63 较 T42 在预报水平上有长足改进。

图 3a、b、c 分别给出 1994 年 7—12 月 T63 和 T42 系统预报 500hPa 高度的倾向相关、距平相关和均方根误差,图中所用的资料 T63 是第 4、5、6 天预报,而 T42 是第 3、4、5 天预报。整个半年来看,倾向相关(图 3a)、距平相关(图 3b)表明一个共同的特点:T63 第 4 天预报水平与 T42 的第 3 天预报水平相当,T63 第 5 天预报水平与 T42 的第 4 天预报水平相当,T63 第 6 天预报相关曲线与 T42 的第 5 天预报相关曲线有较好的重叠,而这种提前一天的预报相关曲线的重合,6

个月中,大约有 4 个月 T63 的预报水平略在 T42 之上。均方根误差(图 3c)的情况略差些(7、8、9 月 T42 略高,10、11、12 月 T63 略高)。图 4a、b、c 分别为 T63 与 T42 的 1994 年下半年平均的相关系数和均方根误差随预报时效的演变。另外,表 2 给出 1995 年 1 月 T63 和 T42 的位势高度月平均检验的对比情况。位势高度的均方根,除 250hPa 的 96 小时和 120 小时比 T42 高外,其它时效和层次,T63 预报均低于 T42 预报,尤其在低层(850hPa)。相同时效的相关系数 T63 比 T42 普遍要高得多,从距平相关上看,中高层的可用预报在 5 天左右,接近 6 天,低层接近 5 天,倾向相关 T63 的 7 天预报已超过 0.6。风场和温度场预报的改进也是很明显的(表略)。综合以上的图表及其分析,我们可以得到一个基本的概念:第二代中期数值预报系统(T63 L16)比第一代中期数值预报系统(T42 L9)的预报水平提高 1 天以上。

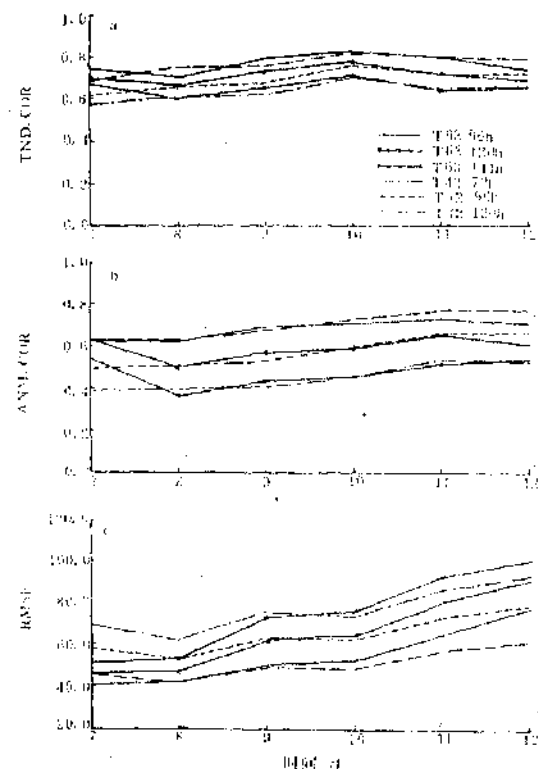


图 3 1994 年 7 月到 12 月 T63 和 T42 系统预报 500hPa 高度的倾向相关(a)、距平相关(b)和均方根误差(c)

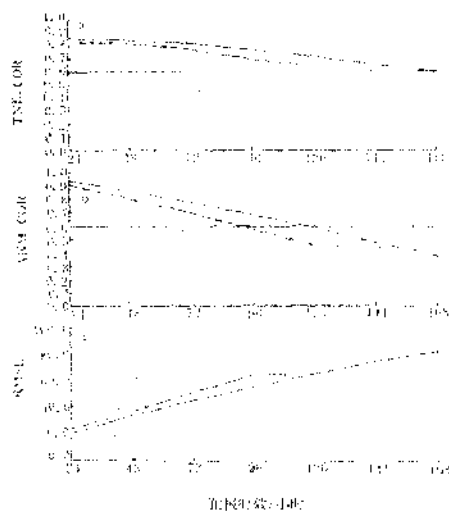


图 4 1994 年下半年平均 T63 与 T42 预报的 500hPa 位势高度检验结果的比较

2.2 T63 与日本的数值预报模式的比较

图 5 为 1994 年 11 月 T63、T42 和日本的数值预报水平比较,图中也给出 1992 年同月的 B 模式的情况,以作比较。由图 5 可看

出, T63 的 96 小时预报接近 B 模式的 48 小时预报, T63 的相关曲线介于 T42 与日本的相关曲线之间, 更靠近 T42, 而与日本的相距甚远, 尤其当预报时效增大时, 差距变得更大, 日本的第 7 天距平相关在 0.6 以上, 相当

于 T63 的第 5 天预报。由于国外的资料不全, 这里仅作了一月的情况比较, 还不足以作出结论, 但的确也说明, 尽管新一代的中期数值预报系统预报水平有较大的提高, 但与发达国家相比, 差距还是明显的。

表 2 1995 年 1 月北半球 T42 模式与 T63 模式高度场预报对分析标准检验结果对照表(月平均)

层	时效	RMSE		ANM. COR		SKL. SCR		TND. COR	
		T42	T63	T42	T63	T42	T63	T42	T63
250hPa	24	43.6	32.3	.963	.974	25.7	16.1	.852	.931
	48	72.0	63.1	.861	.924	34.7	26.8	.833	.908
	72	97.4	92.0	.773	.856	42.5	35.6	.781	.868
	96	117.2	119.1	.699	.766	48.1	43.2	.732	.817
	120	133.1	142.8	.614	.664	52.1	49.3	.720	.758
	144		163.5		.563		54.2		.716
	168		179.6		.477		57.8		.690
500hPa	24	30.9	22.6	.929	.968	26.3	19.2	.834	.917
	48	51.6	42.8	.856	.911	30.9	31.8	.821	.897
	72	69.9	61.4	.782	.836	45.6	41.7	.775	.854
	96	86.9	78.0	.690	.735	52.2	49.9	.731	.800
	120	98.0	93.9	.598	.625	56.7	56.1	.704	.748
	144		108.1		.518		61.5		.703
	168		118.6		.421		65.1		.675
850hPa	24	24.5	17.7	.900	.950	37.3	27.1	.816	.893
	48	40.9	31.6	.809	.880	49.0	40.8	.799	.873
	72	54.8	42.5	.703	.793	59.3	50.7	.719	.819
	96	67.7	52.7	.567	.696	66.4	58.7	.641	.772
	120	72.1	62.1	.484	.586	70.6	65.2	.619	.724
	144		70.6		.477		69.8		.671
	168		76.8		.383		72.5		.621

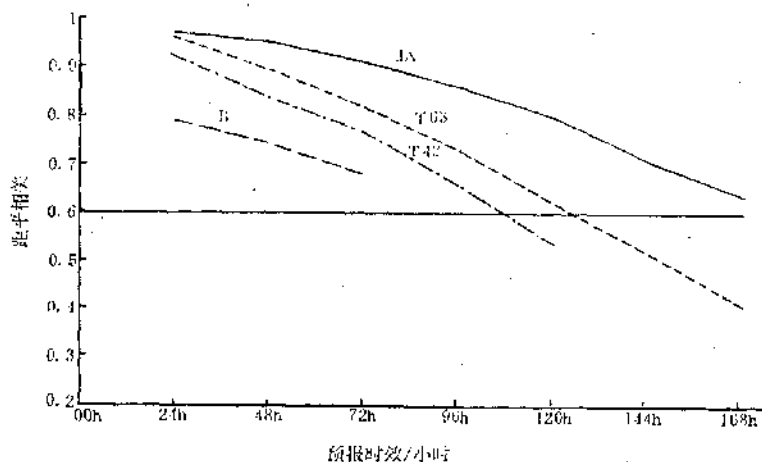


图 5 1994 年 11 月 T63、T42 和日本的数值预报 500hPa 位势高度距平相关比较, 圆线为 1992 年 11 月份的 B 模式相关系数

3 结论与讨论

通过以上的讨论,我们初步可得以下结论:

我国第二代中期数值预报系统(T63 L16)较第一代的中期数值预报系统,除较高的模式水平和垂直分辨率之外,模式物理过程描写更加合理,新系统的分析技术亦有较大改进,因而其预报水平有较大的提高,可用预报从4天半到5天提高到5天半到6天左右,提高幅度为一天或略超过一天。新系统已克服了老系统预报副高偏强的系统性偏差。总的来讲,新系统的误差幅度要比老系统小得多。当然,新系统也还存在着它本身固有的系统预报误差,但是对于这种误差的成因分析,统计学存在着局限性,我们将借助动力学的方法另文加以探讨。

近年来我国数值预报业务发展很快,从80年代初的短期B模式到目前的中期全球谱模式,可用预报从2—3天上升到6天左右,使我国步入了世界少数几个能制作和提

供中期数值天气预报产品的国家行列,大大地缩小了我国天气预报水平与发达国家之间的差距。但是就目前而言,我国数值预报水平与世界先进水平的差距还是很明显的,尤其在同化技术和中尺度数值模拟方面。

我们相信,随着我国计算机能力的提高,广大科研技术人员的不懈努力,建立在更好的初值基础上,中期数值预报将向更高分辨率和模式物理过程描写更加合理的发展方向发展,并将建立与之配套嵌套的中尺度数值预报业务系统,从而提高对灾害性天气的预报能力,充分发挥天气预报在国民经济建设中的作用。

参考文献

- 1 刘还珠,张绍晴.中期数值预报的统计检验分析.气象,1992(9).
- 2 杨学胜等. T63L16谱模式. T63产品推广应用讲义,国家气象中心技术报告,1995.
- 3 张德新等. 高分辨率全球客观分析方案及试验结果,第二次全国数值预报会议文集,1994.
- 4 郭肖容等. 统计检验一些问题的讨论. 待发表,1995.

Performance Assessment of the Medium-range NWP(T63 L16)

Zhang Shaoqing Yang Xiaomei Huang Zho

(National Meteorological Centre, Beijing 100081)

Abstract

The standard verification method of CBS of WMO is introduced. The products of T63 model are verified with the method, and are compared with the model products of T42 and that of the Japan model as well. The results show that the useful forecast effectiveness can reach 5—6 days. The prediction error range is much less than that of the T42 model.

Key Words: standare verification of CBS T42 model prediction performance assessment