

# 热带气旋频数的短期气候预测水平评估\*

雷小途

(上海台风研究所,上海 200030)

## 提 要

热带气旋频数的实况是一确定的整数,而预测仍具有一定的不确定性,其实质是一区间,亟待评估的问题是预测区间与一确定整数的接近程度。针对待评估对象的这种特征,定义了  $d$  指数、绝对误差 ( $E$ ) 和技巧水平 ( $S$ ) 3 个评估参数,分别从趋势预测、定量预测、相对于气候概率预测的技巧等方面客观地标度待评估方法的预测性能,并对“九五”攻关前 10 年 (1988~1997 年) 的业务预测性能进行了再评估:对 7 月 TCF 预测的  $d$  指数和  $E$  分别为 0.9 和 0.8 个,相对于气候概率(即无技巧)预测的技巧水平为 -0.06;8 月的  $d$ 、 $E$ 、 $S$  分别为 0.7、0.75 个和 0.13;9 月的  $d$ 、 $E$ 、 $S$  分别为 0.65、0.9 个和 0.01;台汛期的  $d$ 、 $E$ 、 $S$  分别为 0.3、1.7 个和 0.06;全年的  $d$ 、 $E$ 、 $S$  分别为 0.3、2.2 个和 0.07。

关键词:热带气旋 频数 短期气候预测 评估

## 引 言

上海台风研究所在“九五”期间承担了国家重点攻关课题“影响上海及华东地区热带气旋频数的短期气候预测研究”,其目标是建立性能优于“九五”攻关前长期业务预报水平的客观预测方法。显然,客观合理地评估热带气旋频数(以下简称 TCF)的预测水平具有重要意义。“我国短期气候预测系统的研究”项目统一了攻关前的年限为 5~10 年(即 1993~1997 年或 1988~1997 年),但并未定义 TCF 预测水平的评估参数<sup>[1]</sup>。作者<sup>[2]</sup>在 1998 年曾探讨过 TCF 预测水平的评估方法,指出:平均绝对误差、预测准确率、预测效率、TS 评分、技巧评分、一致系数、距平符号一致率及相关系数和解释方差等 9 个指标均可用于评估 TCF 的预测水平,并对上海中心气象台 1991~1997 年的业务预测水平进行了初步的评估。然而,这 9 个指标大多是评估预测方法在一段时间内的平均预测水平,即需要一定数量的预测和相应实况的样本,因此,并不完全适合于评估每年的 TCF 预测水平。

本文在文献[2]的基础上,分析了评估对象的特征,探讨了 TCF 预测水平的评估方法,并对攻关前 10 年的业务预测水平进行了再评估。

\* 本文受国家“九五”重点攻关课题(96-908-05-06-04)的资助。  
2000-02-22 收到,2000-07-10 收到修改稿。

## 1 评估对象的特征

由于以往没有开展影响华东地区 TCF 的短期气候预测工作,为比较起见,本文仅讨论评估影响上海地区 TCF 业务预测水平的方法。影响上海地区的 TCF 业务预测一直由上海中心气象台长期预报科制作,并于每年 4 月 12 日前后发布本年汛期(5~9 月)影响上海地区的 TCF,再在汛期(5~9 月)的每月 25 日前后制作并发布下个月影响上海地区 TCF 的更正预报。

按项目专家评估组对“九五”攻关前 5~10 年业务预测水平进行评估的要求,本文对业务预测水平的评估是对其攻关前 10 年(1988~1997 年)的预测进行的。表 1 列出了 1988~1997 年上海中心气象台长期预报科对汛期(5~9 月)影响上海的 TCF 预测结果。

表 1 1988~1997 年影响上海 TCF 的业务预测结果

年份	预 测 结 果		实 况	
	频数(个)	发生时段(月)	频数(个)	发生时段(月)
1988	2 左右	7 月下旬~8 月上半月;8 月下旬~9 月上半月	1	8 月 1 个
1989	2~3	7 月下旬~8 月初;8 月下旬;9 月上半月	6	7、8、9 月各 2 个
1990	1~2	8 月上半月;9 月	7	6、7、10 月各 1 个;8、9 月各 2 个
1991	2~3	7 月下旬后期~8 月初; 8 月中旬后期~下旬前期;9 月份	6	8 月 1 个;9 月 4 个;10 月 1 个
1992	2 左右	7 月下旬~8 月初;8 月下旬~9 月初	2	8 月 1 个;9 月 1 个
1993	2 左右	7 月下旬~8 月初;8 月底~9 月上半月	0	
1994	2~3	8 月下旬~9 月	6	6、7 月各 1 个;8 月 3 个;10 月 1 个
1995	1~2	7 月下旬;8 月下旬~9 月初	1	8 月 1 个
1996	2~3	7 月中;8 月下旬;9 月	2	8 月 1 个;9 月 1 个
1997	2 左右	7 月下旬~8 月初;8 月下旬~9 月上旬前期	3	8 月 2 个;9 月 1 个

由于该业务预测发布的汛期是指 5~9 月(即降水的汛期),而本课题中的汛期是指 7~9 月(台汛期),严格地讲两者是有差别的。从表 1 的预测发生时段可以发现,所有预测有影响的 TC 均发生在 7~9 月,所以,该业务预测实质上可以看成是对台汛期(7~9 月)影响上海的 TCF 的预测。实际应用中,还近似地将汛期 TCF 的预测用作全年 TCF 的预测。表 1 还同时列出了影响上海 TCF 的历史实况<sup>[3]</sup>。

从表 1 还可以看到,业务预测仍带有相当的不确定性和模糊性(或者说预测是具有相当的定性性质),即预测的结果形如  $F$  个左右或  $F_1 \sim F_2$  个,如 1988 年的预测为 2 个左右,1989 年的预测为 2~3 个……。其实质是一区间,对于后者, $F_1$ 、 $F_2$  分别为区间的下、上界;对于前者, $F_1$ 、 $F_2$  由下式计算<sup>[4]</sup>:

$$\begin{aligned}
 F_1 &= \begin{cases} F - 0.5 & (7、8、9 \text{ 月,台汛期}) \\ F - 1.0 & (\text{全年}) \end{cases} \\
 F_2 &= \begin{cases} F + 0.5 & (7、8、9 \text{ 月,台汛期}) \\ F + 1.0 & (\text{全年}) \end{cases}
 \end{aligned} \quad (1)$$

综上所述,我们要评估的是一预测区间与一整数(实况)的接近程度,两者越接近,即表明预测水平越高。文献[2]在预测区间内取两个整数值:其一与实况最接近(简称为预测 A,下同),其二与实况相差最大(简称为预测 B,下同),再分别考察 A、B 与实况的接近

程度即可评估出预测(区间)准确性的上、下界限。

## 2 评估参数

### 2.1 $d$ 指数

为评估预测方法对偏多、偏少、正常的趋势预测准确性,我们定义  $d$  指数(Determine Index):

$$d = \begin{cases} 1.0 & i = j \\ 0.0 & |i - j| = 1 \\ -1.0 & |i - j| = 2 \end{cases} \quad (2)$$

式中, $i$ 表示预测, $j$ 表示实况。且当预测为偏少时 $i$ 的取值为1,预测为正常时 $i=2$ ,预测为偏多时 $i=3$ ,同理,当实况为偏少、正常、偏多时, $j$ 分别取值为1、2、3。

当预测与实况一致(均为偏少、正常或偏多)时得1分, $d$ 指数的值为1;当预测与实况完全相反(预测偏少而实况偏多、预测偏多而实况偏少)时扣1分, $d=-1$ ;当预测与实况不一致但并不完全相反(预测和实况两者之一为正常,另一为偏多或偏少)时不得分, $d=0$ 。可见, $d$ 在(-1~1)之间取值,其值越大表明预测的定性趋势越准确。

对于影响上海地区的热带气旋,其全年频数在2~4个之间为正常,台汛期2~3个为正常,7月和9月0~1个为正常,8月1~2个为正常<sup>[3]</sup>。

### 2.2 绝对误差( $E$ )

$d$ 指数能较好地评估TCF趋势预测的准确性,但不能反映预测与实况量值上的差异。如:1989、1994年对台汛期TCF预测的结果完全一样,均预测TCF正常( $i=2$ ),而实况分别为6个和4个,均为偏多( $j=3$ )。于是这两年的 $d$ 指数完全相等(为零),而无法反映出这两年实况的差别(因而预测性能的差别)。为此,我们采用绝对误差(Absolute Error)作为补充,其计算公式如下:

$$E = |x_f - x_o| \quad (3)$$

式中, $x_f$ 、 $x_o$ 分别为预测和实况, $E$ 即为预测的绝对误差。显然,绝对误差的值越小表明预测越好。

### 2.3 技巧水平( $S$ )

$d$ 指数和绝对误差( $E$ )已较好地评估了TCF趋势预测的准确性和预测与实况量值上的差异,但并未能体现预测方法的技巧。所谓技巧,是相对于某一特定预测方法而言的,通常将该特定的预测方法看作是无技巧的,该特定预测方法便简称为无技巧预测(如:评定短期TC路径预测方法性能时的CLIPER方法)。技巧水平(Skill Score)是描述预测方法有无技巧的定量指标,本文将其定义为:

$$S = \frac{1}{2} \left| \frac{d - d_0}{2} + \frac{E_0 - E}{T_b + E_0 + E} \right| \quad (4)$$

式中, $S$ 为技巧水平, $d$ 、 $d_0$ 分别为待评估预测和无技巧预测方法的 $d$ 指数, $E$ 、 $E_0$ 分别为待评估预测和无技巧预测方法的绝对误差, $T_b$ 为预测对象的气候平均值,本文取1949~1996年的平均<sup>[3]</sup>,即影响上海地区TCF的全年、台汛期7月、8月、9月的气候平均( $T_b$ )

分别为 2.90, 2.48, 0.67, 1.08 和 0.73 个。显然,  $S$  在  $(-1, 1)$  之间取值,  $S$  为正且越大表明相对于无技巧的预测越好, 反之亦然。

无技巧水平的预测是相对的, 可有很多种不同的取法, 本文的无技巧预测以 TCF 的气候分布概率为基础, 取为“最佳随机预测”——众多随机预测(区间长度为 1)中历史拟合最好的一种随机预测。对于影响上海的 TCF 预测, 其无技巧的全年和汛期预测为 1~2 个, 7 月、8 月和 9 月均为 0~1 个<sup>[5]</sup>。

### 3 攻关前的业务预测水平

利用以上定义的评估参数对“九五”攻关前 10 年(1988~1997 年)的业务预测性能进行评估, 具体的计算步骤如下:

① 分别从业务预测和无技巧预测的区间中取两个整数, 其一与实况最接近(简记为  $A, A_0$ ), 其二与实况相差最大(简记为  $B, B_0$ ); ② 按方程式(2)、(3)求出  $A, B$  和  $A_0, B_0$  的  $d$  指数、绝对误差( $E$ ); ③ 利用  $A, B_0$  的  $d$  和  $E$  值按方程式(4)求出  $S$ (为  $S$  的上限, 记为  $S_{上}$ ), 利用  $A_0, B$  的  $d$  和  $E$  值按方程式(4)求出  $S$ (为  $S$  的下限, 记为  $S_{下}$ ); ④ 求  $A, B$  的  $d, E$  及  $(S_{上}, S_{下})$  的中位数。

计算结果表明, 1988~1997 年业务预测的平均  $d$  指数值均为正, 即预测的总体趋势没有出现完全相反的情况。其中, 全年及台汛期预测的  $d$  指数均界于 0.2~0.4 之间, 7 月为 0.9, 8 月的  $d$  最大( $A$  预测等于 1, 即趋势预测完全准确)界于 0.4~1.0, 9 月界于 0.6~0.7 之间。

1988~1997 年, 对全年 TCF 预测的平均绝对误差为 1.7~2.7 个, 台汛期为 1.2~2.2 个, 其余均为 1 个左右(0.5~1.3 个), 其中, 8 月预测  $A$  的绝对误差最小, 仅为 0.2 个。而且 8 月的预测技巧也较高, 其上限定 0.51, 但下限仅 -0.25, 其可能的变动区间范围达 0.76, 表明其稳定性较差。技巧水平居其次的是对全年及台汛期的预测, 最高可达 0.29; 而 7 月的预测技巧水平最低, 其可能的取值区间为 -0.28, 0.16。

为避免预测性能稳定性的差异给实际应用带来不便, 作为简化, 表 2 列出了预测性能

表 2 1988~1997 年影响上海地区 TCF 业务预测的性能指标

年份	7月			8月			9月			台汛			全年			
	$d$	$E$	$S$	$d$	$E$	$S$	$d$	$E$	$S$	$d$	$E$	$S$	$d$	$E$	$S$	
1988	1.0	0.5	0.0	0.5	1.0 - 0.12	1.0	0.5	0.0	0.0	1.5 - 0.24	0.0	1.5 - 0.24	0.0	1.5 - 0.24	0.0	
1989	0.0	1.5	0.0	1.0	0.5	0.29	0.0	1.0	0.07	0.0	3.5	0.18	0.0	3.5	0.18	
1990	1.0	1.0 - 0.15	1.0	1.0	1.0	0.19	0.0	1.0	0.07	-0.5	3.5	0.0	-0.5	5.5	0.0	
1991	1.0	0.5	0.0	1.0	0.5	0.13	0.0	3.0	0.03	0.0	2.5	0.04	0.0	3.5	0.13	
1992	1.0	0.5	0.0	1.0	0.5	0.13	1.0	0.5	0.0	1.0	0.5	0.13	1.0	0.5	0.26	
1993	1.0	0.5	0.0	0.5	1.0 - 0.04	1.0	0.5	0.0	0.0	2.5 - 0.26	0.0	2.5 - 0.20	0.0	2.5 - 0.20	0.0	
1994	1.0	1.0 - 0.15	0.0	1.5	0.27	0.5	1.5 - 0.18	0.0	1.5	0.21	0.0	3.5	0.17	0.0	3.5	0.17
1995	1.0	1.0 - 0.15	0.5	0.5	0.0	1.0	0.5	0.0	0.5	0.5	0.0	0.5	0.5	0.0	0.5	
1996	1.0	1.0 - 0.15	1.0	0.0	0.25	1.0	0.0	0.15	1.0	0.5	0.33	1.0	0.5	0.13	1.0	
1997	1.0	0.5	0.0	0.5	1.0	0.23	1.0	0.5	0.0	1.0	0.5	0.24	1.0	0.5	0.24	
平均	0.9	0.8 - 0.06	0.7	0.75	0.13	0.65	0.9	0.9	0.01	0.3	1.7	0.06	0.3	2.2	0.07	

指标取值区间的中位数。

由表 2 可见,全年及台汛期预测的  $d$  指数均为 0.3(相当于趋势预测的准确率为 30%),而汛期各月(7、8、9 月)的  $d$  指数高达 0.65~0.9(相当于趋势预测的准确率为 65%~90%),表明月趋势的业务预测较为准确,全年及台汛期趋势的预测则较差。从表 2 还可见,各月预测的绝对误差均接近 1 个,全年及台汛期预测则为 2 个左右。此外,表 2 还表明 8 月的预测技巧较高,技巧水平达 0.13,其次是全年及台汛期预测均接近 0.1 左右(0.06、0.07),9 月、7 月的预测则较差,其中 7 月为负技巧,9 月也接近零技巧。

## 4 小 结

TCF 的实况是一确定的整数,而预测仍具有一定的不确定性,其实质是一区间,即待评估的问题是:预测区间与一确定整数的接近程度。据此,本文定义了 3 个评估参数: $d$  指数、绝对误差( $E$ )和技巧水平( $S$ ),分别从趋势预测、定量预测、相对于气候概率预测的技巧等方面客观地尺度待评估方法的预测性能。并对“九五”攻关前 10 年(1988~1997 年)的业务预测水平进行了再评估,结果表明:

(1) 全年, $d$  指数为 0.3(界于 0.2~0.4),绝对误差为 2.2 个(界于 1.7~2.7 个),相对于气候概率预测(每年均报 1~2 个)的技巧水平为 0.07(界于 -0.15~0.29);

(2) 汛期, $d$  指数为 0.3(界于 0.2~0.4),绝对误差为 1.7 个(界于 1.2~2.2 个),相对于气候概率预测(每年均报 1~2 个)的技巧水平为 0.06(界于 -0.16~0.29);

(3) 7 月, $d$  指数为 0.9,绝对误差为 0.8 个(界于 0.5~1.1 个),相对于气候概率预测(每年均报 0~1 个)的技巧水平为 -0.06(界于 -0.28~0.16);

(4) 8 月, $d$  指数为 0.7(界于 0.4~1.0),绝对误差为 0.75 个(界于 0.2~1.3 个),相对于气候概率预测(每年均报 0~1 个)的技巧水平为 0.13(界于 -0.25~0.51);

(5) 9 月, $d$  指数为 0.65(界于 0.6~0.7),绝对误差为 0.9 个(界于 0.5~1.3 个),相对于气候概率预测(每年均报 0~1 个)的技巧水平为 0.01(界于 -0.21~0.24)。

本文所定义的 3 个评估参数不仅能用于评估具有一定历史样本的 TCF 预测方法的性能,而且适用于仅有短样本(如一个样本)的 TCF 预测方法的性能评估。

## 参 考 文 献

- 1 王绍武,朱锦红.短期气候预测的评估问题.应用气象学报,2000,11(增刊):1~9.
- 2 雷小途.热带气旋频数预测的研究进展和业务预测水平.大气科学研究与应用,1998,(14):196~202.
- 3 冯泾贤,杨自植,邓之瀛.影响上海市及长江三角洲地区热带气旋气候规律的研究.大气科学研究与应用,1998,(14):36~41.
- 4 雷小途.热带气旋短期气候趋势的预测量研究.大气科学研究与应用,2000,(18):51~59.
- 5 雷小途.影响华东的热带气旋短期气候趋势的可预报性分析.气候学研究-气候与环境,1998,212~217.

## ACCURACY EVALUATION OF SHORT-CLIMATIC FORECAST ON TROPICAL CYCLONE FREQUENCY

Lei Xiaotu

(*Shanghai Typhoon Institute, Shanghai 200030*)

### Abstract

The tropical cyclone frequency (TCF) is actually a certain integer, but there exists uncertainty in the subjective and objective short-climatic forecasts, so the contiguous degree must be evaluated. In order to evaluate the accuracy of short-climatic forecasting on TCF, firstly three parameters, determine index ( $d$ ), absolute error ( $E$ ) and skill score ( $S$ ), are defined comparatively with climatic probability forecast (CPF) which was regarded as naught forecast. Then, the accuracy was evaluated of subjective forecast from 1988 to 1997. The results indicate:  $d$ ,  $E$  and  $S$  of typhoon season are 0.3, 1.7 and 0.06;  $d$ ,  $E$  and  $S$  of July are 0.9, 0.8 and 0.06;  $d$ ,  $E$  and  $S$  of August are 0.7, 0.75 and 0.13;  $d$ ,  $E$  and  $S$  of September are 0.65, 0.9 and 0.01, respectively.

**Key words:** Tropical cyclone    Frequency    Short-climatic forecast    Accuracy evaluation.