

天气预报

经验选编

53

2

广东省气象局编

天气预报

一九七八年 第二期

经验选编

(总第二十期)

目 录

秋季台风趋势和时段长期预报.....	广东省气象台 (1)
运用台风进防环流分型与本站资料	
结合作台风预报的试验	开平县气象站 (7)
用地面天气形势分类作西太平洋台风登陆地段预报.....	海南行政区气象台 (10)
用相似过滤法作台风路径趋势预报.....	广东省热带气象研究所 (13)
气压双峰台风暴雨型	高州县气象站 (20)
太平洋台风路径的一种综合分型预报方法	佛山地区气象台 (22)
统计外推结合形势特征订正作	
西太平洋台风登陆纬度预报.....	海南区台风攻关会战组 (25)
一九七七年台风活动概述	广东省气象台 (28)
结合 500 mb 关键区讯号作低温阴雨过程的中期预报	阳山县气象站 (33)
应用起报讯号分型和多因子相关综合	
作低温阴雨中期预报	肇庆地区低温阴雨攻关组 (37)
用多因子综合相关法作春播期逐日降水预报	南雄县气象站 (40)

秋季台风趋势和时段的长期预报

广东省气象台预报科长期组*

做好秋季台风的预报服务工作是长期预报的重要内容之一。所谓秋季台风(简称“秋台”，下同)，是指“秋分”(九月下旬)以后登陆或严重影响我省的台风。

本文以“春分不暖，秋分不凉”，“春无三日晴，秋无三日落”等群众经验为线索，通过秋台前期 500 mb 环流特征的普查，挑选相关系数较优的预报因子，进而以分辨函数值组制散布图，制作“秋台”趋势预报。在趋势预报的基础上，采用韵律法，进一步作出“秋台”的时段预报。这套工具经过两年的实际使用，效果较好，如表一：

表一： 预 报 效 果 检 验 表

年份	趋 势		时 段	
	预 报	实 况	预 报	实 况
76	少(0—1次)	1次	共有三次特征，有两次消空，一次预报9月20—24日有台风	7619号台风于9月20日—25日先后在湛江、占县、万宁县三次登陆
77	少(0—1次)	1次	一次特征，预报9月20—30日有台风(因标准偏低，未达“秋台”，故预报9月未达“秋台”期)	7712号台风于9月25日在阳江县登陆

一 “秋台”的气候概况

为了统计方便，我们规定从九月二十日及以后，凡中心附近最大风力达6级或以上的热带低压或台风在我省登陆，或中心移到我省海岸线一个纬距之内者(称严重影响)都属于我们统计范围之列，统计结果见表二：

*此文由谢年奇同志执笔，中山大学滕成业老师参加一部分计算工作。

表二:

历年“秋台”次数表

年 份	1949	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960	1961	1962	1963	1964
秋台总次数	1	4	1	0	3	4	1	1	2	1	0	3	1	3	0	5
珠江口以东	1	0	0	0	0	1	0	0	0	0	0	0	1	1	0	1
珠江口以西	0	4	1	0	3	3	1	1	2	1	0	3	0	2	0	4

年 份	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	合计	平均			
秋台总次数	2	0	3	1	0	2	2	2	1	5	3	51	1.9			
珠江口以东	0	0	0	1	0	0	0	0	0	1	0	7	0.3			
珠江口以西	2	0	3	0	0	2	2	2	1	4	3	44	1.6			

从表二可以看出,自1949年至1975年的二十七年中,“秋台”共有五十一个,平均每年有两个,约占全年登陆和严重影响台风的四分之一;“秋台”最多的1964年和1974年各有五个,也有五年没有“秋台”。在时间上,“秋台”主要集中在九月下旬到十月中旬,占“秋台”总数的三分之二。登陆和严重影响的地段以珠江口以西最多。“秋台”的年际变化是每隔一至两年就会出一段多年(两个或以上),而多年一般只连续一至两年,但到了七十年代,多年延长到持续三年。

二 “秋台”的前期环流特征

实践证明,用环流形势特征作长期预报,是一种可行的方法。我们以“冬有奇寒,夏有奇涝”、“春分不暖,秋分不凉”的群众经验作为确定“前期”的线索,从分析典型年的环流特征入手,普查了“秋台”前一年的十二月至当年三月的500mb月平均高度图,从而制作了“秋台”多年和少年的逐月高度平均图和距平累积图,发现一月和三月的关系好,十二月和二月关系不好。

(一) 典型年的环流特征

当年一月份在乌拉尔山附近(50°E — 70°E)有一明显的深槽,“秋台”多(图一);而少年则在上述地区为一高压脊控制,低槽多在 40°E 以西的地区(图略)。

三月的环流特点,多年(图略)在欧洲到巴伦支海为一个明显的高压脊控制, 60°N 以北的 90° — 110°E 之间有一个浅槽;少年(图略)则在 90°E 以西往往有一个从高纬度到低纬的深槽控制。

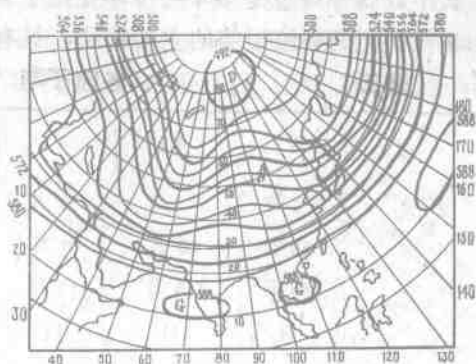
(二) 平均环流和距平累积图特征

选择“秋台”的五个多年和四个少年,分别制作了前一年十二月至当年三月的500mb逐

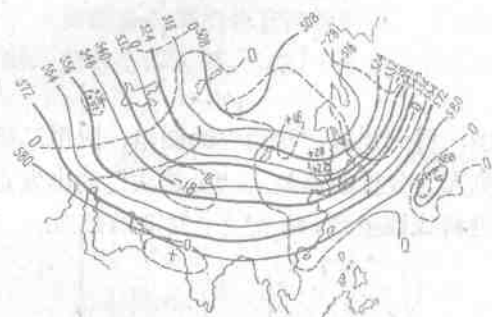
月平均环流和距平累积图。同样表明一月和三月的关系好，十二月和二月关系不好。现举多年的一月(图二)和少年的一月(图三)为例。

从图二、图三可以清楚地看出，“秋台”多年与少年的距平配置几乎全部相反。如：多年在贝加尔湖及我国东部为正区，少年则为负区；多年在鄂毕河以东及千岛群岛东侧是负区，而少年则是正区。

分析环流特征还可以看出：平均环流与典型环流相似，只是不及典型环流分明。秋台多年的年份在冬季以纬向环流为主；而少的年份在冬季则以经向环流为主。这种环流形势的差别在气温上的反映是，前者大多数年份较暖，而后者则比较冷。



图一 “多年”1月典型环流特征



图二 1660年、1964年、1967年、1974年、1975年五个多年的一月500mb图
实线：平均环流 虚线：高度距平累积



图三 1959年、1963年、1966年、1969年四个少年的一月500mb图
实线：平均环流 虚线：高度距平累积

三 “秋台”的趋势预报

(一) 预报因子的挑选

从前期环流特点和距平累积图的分析中，挑选出比较好的因子作图。在技术处理上，改变了过去用大多数年份拟合，以较少年份验证的办法，而采用前十一年拟合，后十年试报，挑选出拟合和试报都较好的因子。因子选出后，再进一步运用一元线性回归计算各单因子与“秋台”的相关系数，选出八个置信水平接近0.90或以上(表三)作为预报因子。

从八个预报因子所处的位置可以看出：三月的 50°N 、 160°E ， 35°N 、 125°E ，和一月的 40°N 、 120°E 都处于西风槽附近，前一点处于槽前，后两点处于槽后，并可以近以地代表北支西风急流的情况；前一点位置正是多年距平累积图上差异最大的中心；三月份的 25°N 、 145°E ， 20°N 、 110°E 和 25°N 、 65°E 的高度可以近似代表低纬度副热带高压的状

况，其中最后一点也是多年与少年在距平累积图上差异最大的中心。因此，可以认为上述因子能够代表前期环流形势的主要特征，具有一定的物理意义。

表三： 各因子与“秋台”次数相关系数表

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
	3月	3月	3月	1月	前一年12月	3月	3月	1月
因 子	50°N	25°N	20°N	55°N	45°N	35°N	25°N	40°N
	160°E	145°E	110°E	25°E	175°E	125°E	65°E	120°E
	H	H	H	H	H	H	H	H
与“秋台”次数 *相关系数Y	-0.30	-0.14	-0.30	-0.38	缺	-0.30	-0.34	0.60

$$* Y = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

(二) 图组分析和检验

将八个因子的每两个为一组，绘制成四张散布图，得到预报“秋台”的图组，应用二级分辨法确定“秋台”多年与少年的分界线。使之避免人为分析的误差。“分辨法”的出发点根据观测得到的特征量 $[x_i (i=1 \cdots P)]$ 预报因子] 进行线性组合，构成一个判别函数为：

$$y = \sum_{i=1}^P C_i x_i$$

P 代表 x_i 的个数，即因子个数； C_i 是常系数，根据历史资料采用数理统计原则确定。

为了使分界线 y_c 成为良好的预报指标，要求两类之间的差异尽可能大，而同类差异尽可能小。我们这里只讨论 $i=2$ 的简单情形，即

$$y = C_1 x_1 + C_2 x_2$$

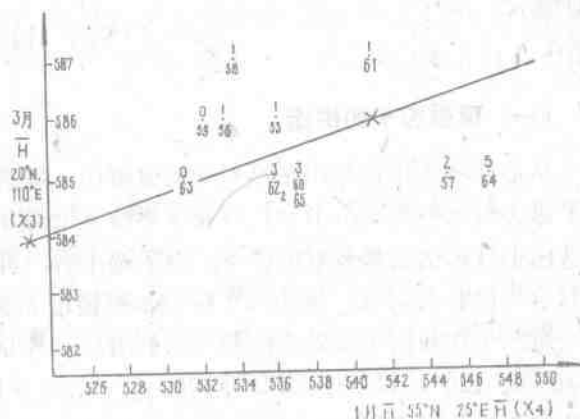
这样，每张散布图可以得到一个判别函数，现以图四为例。

判别函数系用 1955 年到 1965 年资料计算，其：

$$y_c = -340.4597$$

$$y = -0.6453 x_3 + 0.06966 x_4$$

y_c 、 C_1 、 C_2 都是已知，用 y_c 值代替 y 就可以得到分界线上任何一点的 x_3 和 x_4 ，做法是先设 a_1 点的 x_4 为 52 b，代入



图四 预报秋台的散布图组之一

上式, 得到 a_1 点的 $x_3 = 584.3445$, 将 x_3 值再代入上式又可得到 $x_4 = 522.8088$; 又设 a_2 点的 x_3 为 586, 分别得到 a_2 点的 $x_4 = 541.337$, $x_3 = 586.000$, 两点连成一条直线就是分界线。

其余三张图都照上法确定分界线 (图略), 要求关系都比较好后, 才直接应用各年的判别函数值组图, 其分辨函数为:

$$y_1 = -0.02072 x_1 - 0.10589 x_2$$

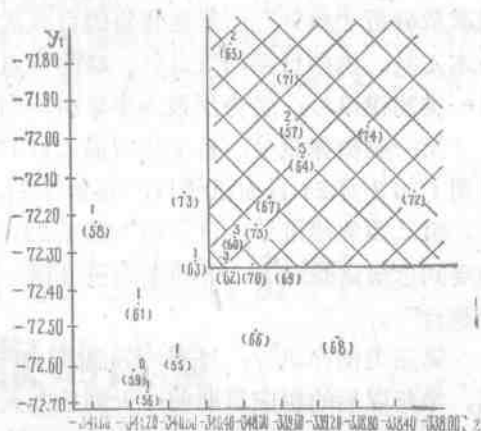
$$y_2 = -0.64534 x_3 + 0.06966 x_4$$

$$y_3 = -0.04194 x_5 - 0.0762 x_6$$

$$y_4 = -0.10038 x_7 + 0.03712 x_8$$

以四个 y 值组成两张散布图。预报时, 将 y 值点入图内, 当有两张图的点子均落入斜格区才预报“秋台”多, 否则报“秋台”少。图五是预报“秋台”的图例。

表四是全部拟合及试报情况, 可以从表四看出, 拟合全部正确, 试报十年也全部正确。



图五 “秋台” 趋势预报图

表四: “秋台” 拟合及试报情况表

年 份		1955	1956	1957	1958	1959	1960	1961	1962	1963	1964	1965
y_1 y_2 图		×	×	0	×	×	0	×	0	×	0	0
y_3 y_4 图		×	×	0	×	×	0	×	0	×	0	0
拟合 结 论		少	少	多	少	少	多	少	多	少	多	多
实 况	“秋台” 数	1	1	2	1	0	3	1	3	0	5	2
	多、少年	少	少	多	少	少	多	少	多	少	多	多
年 份		1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	
y_1 y_2 图		×	0	×	×	0	0	0	×	0	0	
y_3 y_4 图		×	0	×	×	0	0	0	×	0	0	
试 报 结 论		少	多	少	少	多	多	多	少	多	多	
实 况	“秋台” 数	0	3	1	0	2	2	2	1	5	3	
	多、少年	少	多	少	少	多	多	多	少	多	多	

四 “秋台”时段预报

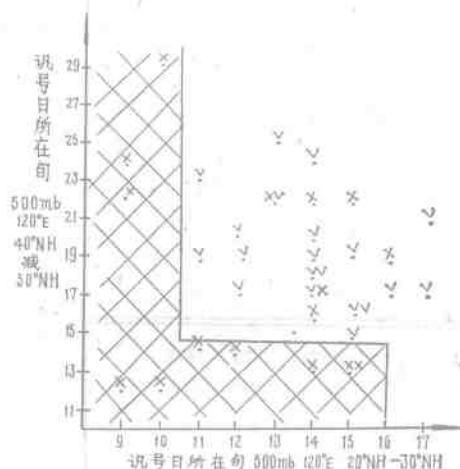
与趋势预报一样，仍采用春秋天气对应关系为思路（180天韵律），用“春分”后冷空气次数分析“秋分”后影响我省的台风次数的关系。以广州“春分”后逐日平均气温曲线为基本工具，依据日平均气温谷、峰的特点分为三类韵律讯号。当趋势预报为少年时，只使用第一类韵律讯号；趋势预报为多年时，则用第一、二、三类韵律讯号。具体规定是：

第一类韵律讯号：日平均气温 $<17^{\circ}\text{C}$ 的谷日为讯号日。这类讯号出现后，半年后的相应日期（如3月24日为讯号日，则9月24日为相应日期，下同）的 ± 8 天内将有“秋台”。

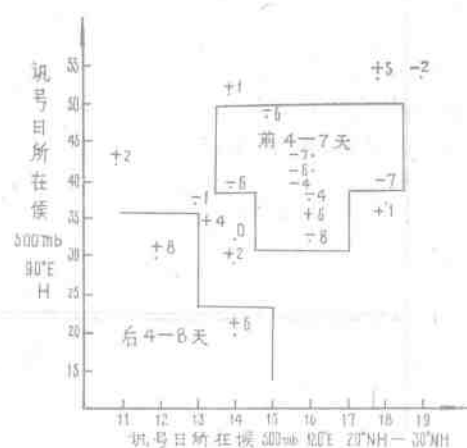
第二类韵律讯号：日平均气温在 $23.4\text{--}26.0^{\circ}\text{C}$ 范围内的完整峰日（是指峰的两侧下降区均要到达规定线以下，下同）为讯号日。这类讯号出现后，半年后的相应日期 ± 7 天内有“秋台”。

第三类韵律讯号：日平均气温 $27.1\text{--}28.4^{\circ}\text{C}$ 范围内的完整峰日为讯号日。这类讯号出现后，半年以后的相应日期前一天到后九天内有“秋台”。

上述三类讯号概括了1955—1975年43次“秋台”中的41次，概括率为95%。但由于预报时段的范围较宽（有11—16天），且有空报的问题，需要先进行消空，尔后才作时段预报。具体用讯号日所在旬、候看500mb图上关键区高度为预报因子做散布图。现以第一类韵律讯号为例说明：如果两张散布图都通过非斜方格区时则不能消空，否则作消空论。当不能消空时再利用散布图作时段预报。图六和图七是消空和时段预报的图例。



图六 消空预报图组之一



图七 “秋台”出现时段预报图

五 小结

（一）目前制作长期预报时，各项具体预报对象之间往往缺乏联系，因而结论难下。例如，趋势工具预报少年，时段工具可以报出属于多年，这时主要靠主观综合分析判断，而

有时由于分析错误使预报归于失败。如果能将趋势和时段的预报工具互相补充,就可以尽量避免上述问题。

(二) 趋势和时段预报工具真正能够起到互相补充的作用,选择被挂钩的基础工具特别重要。最好能选择有一定物理意义,而且经过实际使用证明比较稳定的工具。时段预报用的韵律法要重视对于韵律讯号因子和预报对象的共同特点的分析。

(三) 时段预报工具还比较粗糙,尚不够客观,特别当预报范围太窄时,需要适当外延,或进行具体分析。例如1977年,工具预报9月17—20日有“秋台”,这时就得“扩展”在9月中旬末至下旬前期。

运用台风进防环流分型与本站资料 结合作台风预报的试验

开平县气象站

一 基本思路

七三年以来,我站对台风预报曾进行过多次改革,虽取得了一定效果,但鉴于过去工具都囿于本站资料上,有其局限性,远不能满足防台抗台服务的要求。近两年我站与省地台协作,在省台台风环流分型的基础上,利用单站资料并结合群众测台经验,做台风的具体预报。分别制作出5—11月西太平洋台风和南海台风的登陆地段、风雨影响程度等的预报工具。

二 技术规定

根据我县的地理位置,将台风登陆地段分为惠东以东、宝安到阳江、电白以西等东、中、西三段。风雨影响程度分为有影响和无影响二类,并把有影响类按影响程度分为一般影响(6、7级风或出现大雨)与严重影响(8级风以上或暴雨以上)两种。

南海台风以生成日、西太平洋台风以进入省台第Ⅲ防线(自 25°N 、 122°E ,经 20°N 、 120°E , 17°N 、 115°E ,到 15°N 、 109°E)的当天作为起报零天。

三 省台台风环流型

西太平洋台风进入Ⅲ防时,根据影响台风移动的主要系统和500 mb环流形势,分为四个型:

① **强副高型** 太平洋副高 588 线西伸到 105°E 以西。

② **冷空气型** 在 $100-126^{\circ}\text{E}$ 范围内, 地面冷高压脊的 1020 mb 等压线南伸到 40°N 以南。

③ **东亚大槽型** $115^{\circ}-126^{\circ}\text{E}$ 有较深的槽, 槽底南伸到 30°N 或以南, 500 mb 青岛高度 < 588 位势什米, 且台风进Ⅲ防时的经度 $> 117^{\circ}\text{E}$ 。

④ **其它型** 凡不符合上述三型者, 都划归其它型。

根据南海台风形成时的环流特征分三个型:

① **双台风型** 南海台风形成时, 在西太平洋上已有台风存在, 且两台风间距 ≤ 20 纬距。

② **华南东风型** 在 $21-25^{\circ}\text{N}$ 、 $105-118.5^{\circ}\text{E}$ 范围内, 500 mb 上空盛吹 SSE—NNE 风。

③ **华南其它风型** ② 中所指范围内, 500 mb 上空是其它风向者。

四 环流分型下单站预报工具的制作

我们把单站预报工具建立在环流分型的基础上, 分析不同的气象要素值及其变化, 来区分台风的不同移向和影响程度。在处理方法上, 南海台风是用各种指标来决定其登陆地段和风雨影响程度的, 而西太平洋台风则采用点聚图法。

现以西太平洋台风强副高型为例, 具体说明其做法。如根据预报经验, 选取台风进防纬度与台风进防前 12 小时的移向角组合, 根据“热极生风、闷极生雨”, 选取台风进防日及其前一天最高温度之和与进防日 14 时 $e-T$ 24 小时变量组合, 如此共选了十个因子组成了五张点聚图(图略)来预报台风登陆地段。预报时以多数表决法作出结论, 如 7703 号台风五张图中有三张是报西路的, 实况是在琼海县登陆。

同样地建立了五张预报风雨影响程度点聚图组, 并以多数出现为其结论意见。

五 试报和使用结果

本工具概括了自 60—76 年 5—11 月 168 个台风中的 90% 以上。试报的 8 个台风中, 路径报对 6 个, 风雨影响程度全部准确。七七年投入业务使用, 9 个台风路径报对的有 8 个, 风、雨影响程度的预报准确率分别达 78% (附表)。

附表:

试 报 和 使 用 情 况

试 报	台 风 编 号	预 报			实 况			评 定		
		路 径	风	雨	路 径	风	雨	路 径	风	雨
	6025	西	无	无	荣 市			+	+	+
	6309	西	无	无	文 昌			+	+	+
	6517	西	无	无	电 白			+	+	+
	6903	东	有	有	惠 来	6 级	71.5	+	+	+
	7106	西	无	无	万 宁			+	+	+
	7202	西	无	无	电白—阳江			+	+	+
	7607	西	无	无	转东北出南海			—	+	+
	7616	中	有	有	海 丰	7 级	84.1	—	+	+
七 七 年 使 用	7701	东 路	无	无	福 建			+	+	+
	7702	西	无	无	吴 川		43.6	+	+	—
	7703	西	无	有	琼 海		63.9	+	+	+
	7704	东	无	无	福 建			+	+	+
	7707	东	无	无	转东北出南海			+	+	+
	7709	西	无	无	越 南			+	+	+
	77 0	西	有	无	越 南			+	—	+
	7712	西	无	无	电 白	7 级	46.5 40.9	+	—	—
	7716	西	无	无	转东北减弱			—	+	+

用地面天气形势分类 作西太平洋台风登陆地段预报

海南行政区气象台

一 气候过滤

我们将 1957—1976 年 4—11 月过 120°E 的共 110 个西太平洋台风, 以过 120°E 的纬度(台风路径与经线的交点)和日期进行气候过滤, 滤去不影响本区或在珠江口以东登陆的台风共 30 例, 又把剩下的 80 个台风按季节分为过渡季节(4—5 月、9—11 月)和盛夏季节(6—8 月)两大类。前者有 50 例, 后者共 30 例。

二 过渡季节的地面天气形势分类

根据台风移近 120°E 的 08 时东亚 ($15^{\circ}-51^{\circ}\text{N}$ 、 $95^{\circ}-145^{\circ}\text{E}$) 地面天气形势, 依其主要特征分为两类(第 I 类、第 II 类)。

1. 第 I 类: 东西向脊(图 1)

其主要特征是变性高压脊从日本海一带伸向华南或西南地区。在 $25^{\circ}-45^{\circ}\text{N}$ 、 $100^{\circ}-140^{\circ}\text{E}$ 范围内, 120°E 以西的高压中心值低于或等于以东的高压中心值。 $25^{\circ}-35^{\circ}\text{N}$ 、 $105^{\circ}-120^{\circ}\text{E}$ 范围内无冷锋。

当 588 线西脊点(包括华南高压, 下同)在 116°E 或以西时, 若台风中心过 120°E 的纬度 $\geq 15.2^{\circ}\text{N}$, 则台风登陆海康至本岛, 或从 17°N 以北的本岛南部海面经过(5/5); 若台风中心过 120°E 的纬度 $< 15.2^{\circ}\text{N}$, 则台风从 16.4°N 以南过 110°E (4/4)。

当 588 线西脊点在 117°E 或以东时, 则台风登陆阳江以东或在南海东部转向(6/7, 这 6 次在 $105^{\circ}-120^{\circ}\text{E}$ 范围内, 1020 mb 等压线南伸至

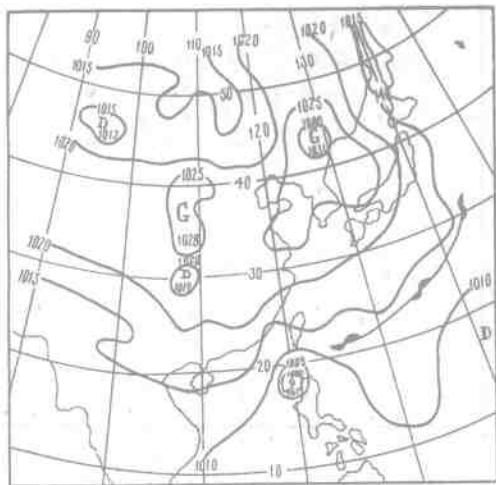


图 1: 第 I 类北渡季节东西向脊

28°N 或以北；而另一次南伸至 26°N，却在文昌县登陆）。

2. 第Ⅱ类：南北（或西北—东南）向脊

我国大陆为南北向的冷高压脊控制，或黄河到南岭之间有冷锋存在。此类又以日本海一带（35—51°N、124—145°E）有反气旋中心分为两副类。

① 第Ⅱ₁类日本海有气旋（图2）

其指标条件如表一。

表一

指标条件	符合条件之一：a, 105—120°E范围内，1020mb等压线南伸至31°N及以南；b, 长江至华南沿海(20—27°N、105—120°E)有锋面；c, 592线东、西脊点分别伸过140°E、120°E。	
	河套西侧(30—45°N、95—110°E)有≥1040mb的冷高中心。	其它
结论	台风从≤11.6°N过110°E或在越南东南部近海消失(3/3)	台风登陆徐闻至本岛或从17°N以北的本岛近海过110°E(14/15、另一次急来)

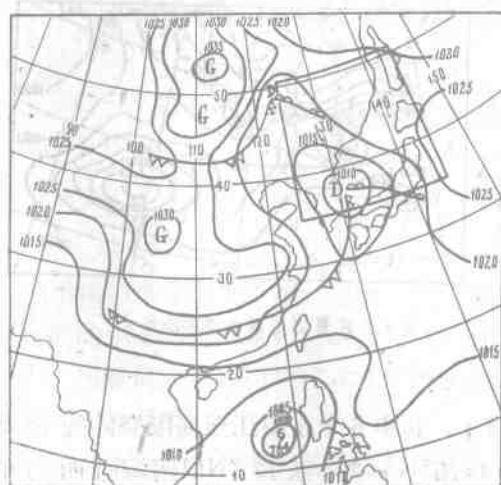


图2：第Ⅱ₁过渡季节南北向脊日本海有气旋

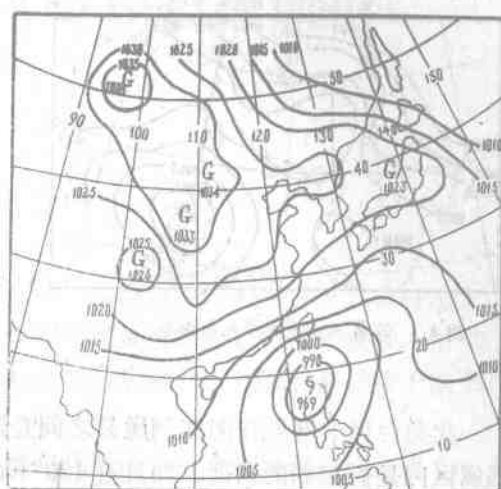


图3：第Ⅱ₂类过渡季节南北向脊日本海无气旋

② 第Ⅱ₂类 日本海无气旋（图3）

其指标条件以 105—120°E 范围内 1020 mb 等压线南伸程度分三种情况（表二）。

表二

1020 线南伸程度			
$\leq 24^{\circ}\text{N}$			
台风从 $10-15^{\circ}\text{N}$ 过 120°E	台风从 $15.7-18.9^{\circ}\text{N}$ 过 120°E	$25-32^{\circ}\text{N}$	32°N 以北
台风在 17°N 以 南西行 (2/2)	台风登陆雷州半岛 (2/3)	台风在南海消失或从本岛 东部海面北上 (5/5)	台风登陆海丰或 在海上转向 (2/2)

三 盛夏季节的地面天气形势分类

将江南至华南 ($20-30^{\circ}\text{N}$ 、 $105-120^{\circ}\text{E}$) 范围内地面天气形势分为三类 (第Ⅱ、Ⅳ、Ⅴ类)

1. 第Ⅱ类: 江南至华南有冷锋存在, 或有自东向西伸展的高压脊 (脊内无闭合小高压中心) 控制该区大部 (图4)。

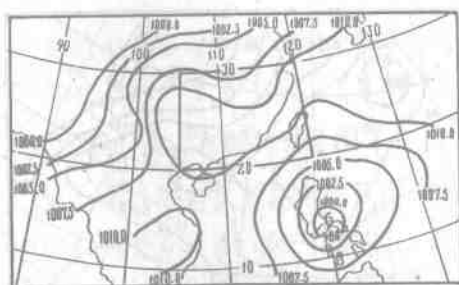


图4.1 第Ⅱ类盛夏季节东西向脊

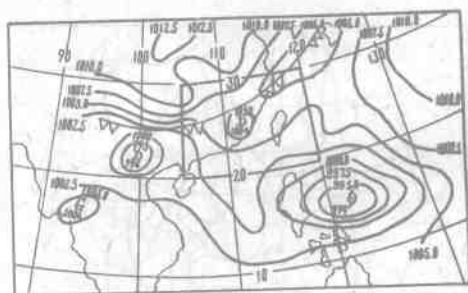


图4.2 第Ⅱ类盛夏季节有冷锋

此类台风大部份在阳江到崖县之间登陆 (12/14)。其中5个在阳江至海康登陆的台风都属区内是脊控制的情况, 而且过 120°E 的纬度 $\geq 14.5^{\circ}\text{N}$; 2个在 17°N 以南海面西行的台风, 其进 120°E 的纬度都较低, $\leq 13.0^{\circ}\text{N}$; 其它7例则都在徐闻至海南岛登陆。

2. 第Ⅳ类: 江南至华南有单独的高压存在, 或低槽控制华南大部份地区 (图5)。

此类台风大多数登陆珠江口以东到闽南地区, 或在南海东北部转向消失 (10/11)。

3. 第Ⅴ类 (其它类): 这类台风共5例, 包括鞍形场、内有闭合小高的高压脊 (呈东西向) 和静止锋 (江南至华南有非东面向脊控制) 等几类形势。本类台风登陆珠海至台山 5/5。

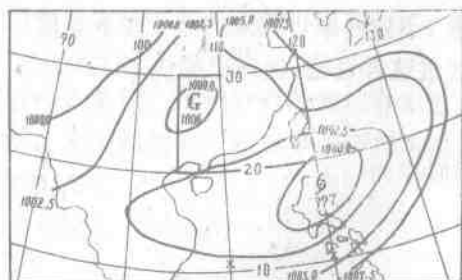


图 5.1 第Ⅳ 盛夏季节单独高压

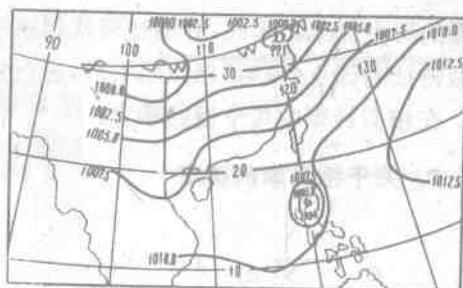


图 5.2 第Ⅳ 盛夏季节低槽

四 试 报

用 1955—1956 年十个进 120°E 的西太平洋台风试报，其中用气候过渡掉 3 个，余下的 7 个台风中，过渡季节的 5 个台风报对了 4 个，盛夏的 2 个台风试报均准确。

（编者：我们对来文作了较大修改，原文中的各类个例表均略）

用相似过滤法作台风路径趋势预报

广东省热带气象研究所 史董

一 前 言

用相似法作天气预报，已为广大气象工作者所熟知。虽然方法多种多样，各具特点。但其研究的中心问题，是如何确定相似条件和判据，希望得到最佳的相似样本和效果。在作预报时，则是从大量的历史样本中寻找出相似样本，并以它们后期的天气实况来推断预报日后天气演变的趋势。

本方法是在学习有关相似方法的基础上，在探索台风路径预报中反复改进后完成的。因为我们对相似的运用主要是在分型（分类）基础上进行的，虽然通过了一些定量的处理，但其结论仍然是一种定性的趋势预报。

二 思路与方法

找相似的整个过程，是围绕新样本的主要特征而进行的，为从大量的历史样本中寻找出与之较相似的样本，我们通过分型处理，首先剔除大量明显不相似的样本，取出与新样本同

类型的历史样本。因此,分型过程就相当于一次相似初选,尔后再通过选定的若干个表征它们相似特征的因子(或叫“相似因子”),进行定量的计算,经过这样过滤,选择出最佳相似样本。

在相似计算中几个量的确定:

1. 关于相似率的确定

(1) 我们假定了相似率为

$$D = \sum_{i=1}^N a_i \frac{|x_{i1} - x_i|}{x_{iM} - x_{im}}$$

式中 N 为相似因子个数。 $|x_{i1} - x_i|$ 为新样本与历史样本之间某相似因子的数值差,称之为“相似距”,取绝对值。 $x_{iM} - x_{im}$ 为全部历史样本中某相似因子数值的变化幅度,即极端值之差,称为“级差”。

对单个相似因子,判定两个样本是否相似的直接方法,可用相似距 $|x_{i1} - x_i|$ 表示之,视其数值愈小则愈相似,反之则愈不相似;用多个相似因子来判定两个样本相似的程度,也可由多个因子的相似距总和 $\sum_{i=1}^N |x_{i1} - x_i|$ 来表示。然而,在多个相似因子的情况下,单考虑

相似距是不够的,这是因为相似程度不但取决于因子的相似距大小,而且还与因子的级差有关。举个例子来说,对于不同的正态分布的因子,虽然它们各取的相似距数值相等,如图1所示,但对于级差大的 x_1 因子, x_{11} 就较接近众数状态,而对于级差较小的 x_2 因子, x_{21} 则已趋于极端值。因此,直接用“相似距”表示两个样本的相似程度是不尽合理的,不同因子间难予比较,为此采用

比值 $\frac{|x_{i1} - x_i|}{x_{iM} - x_{im}}$, 以消除各相似因子“级差”差异。显然,

$$\frac{|x_{i1} - x_i|}{x_{iM} - x_{im}} \text{ 愈大,表示两样本愈不相似; } \frac{|x_{i1} - x_i|}{x_{iM} - x_{im}} \text{ 趋向于 0, 表示两样本愈相似。}$$

a_i 为相似订正参数。

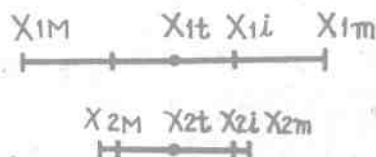
(2) 相似订正参数(简称相似参数)

$$a_i = \frac{2|x_{i1} - x_i|_M}{x_{iM} - x_{im}}$$

式中 $|x_{i1} - x_i|_M$ 为同类(专指分型)样本中,某相似因子的最大相似距。

从历史样本中寻找相似时,可选上的相似样本的多寡,还与同类样本的分布状况有关,

故引入相似参数 $\frac{2|x_{i1} - x_i|_M}{x_{iM} - x_{im}}$, 表示同类样本的分散与集中状况,如图2所示。



图一注: X_{1M}, X_{2M} 分别为相似因子 X_1, X_2 的历史最高值, X_{1m}, X_{2m} 分别为相似因子 X_1, X_2 的历史最低值, X_{1t}, X_{2t} 分别为相似因子 X_1, X_2 的新样本数值, X_{1i}, X_{2i} 分别为相似因子 X_1, X_2 某历史样本数值。

相似参数 $\frac{2|x_{it}-x_{im}|}{x_{im}-x_{im}}$ 愈大, 表明部分同类样本偏

离新样本愈远, 可能取到的相似样本愈少; 反之, 相似参数愈小, 表明同类样本愈集中于新样本附近, 可能选取的相似样本愈多。

相似率 D 原式为

$$D = \sum_{i=1}^N \frac{2|x_{it}-x_{im}|}{x_{im}-x_{im}} \cdot \frac{|x_{it}-x_{im}|}{x_{im}-x_{im}}$$

为实际操作方便, 可将相似率 D 改写为

$$D = \sum_{i=1}^N \frac{2|x_{it}-x_{im}|}{(x_{im}-x_{im})^2} \cdot |x_{it}-x_{im}| = \sum_{i=1}^N A_i |x_{it}-x_{im}|$$

$$A_i = \frac{2|x_{it}-x_{im}|}{(x_{im}-x_{im})^2}$$

当确定了历史样本后, 便有了 $x_{im}-x_{im}$ 。从而 $\frac{1}{(x_{im}-x_{im})^2}$ 可事先计算好, 当求得同类样本中的 $|x_{it}-x_{im}|$ 后, A_i 也就可确定了。

2. 关于相似水准的确定

同时要求多个因子相似, 比只考虑单个因子相似来得不易。如由多个因子求得的相似率, 和由单个因子求得的相似率假设都是 0.5, 表面看来数值一样, 但表示两个样本的相似程度实际上前者就优于后者。

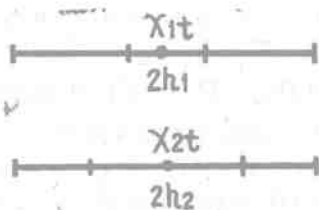
在选择相似样本时, 哪些样本可取, 哪些样本不取? 必须有个准则, 这个准则不妨叫它为“相似水准”, 它与相似因子个数和相似率的大小有关, 即相似水准 $D_a = f(N, D_m)$ 。在我们的试验中假设

$$(D_m + D_a) \cdot \frac{1}{2} = \frac{N^2}{20} \quad \text{移项得}$$

$$D_a = \frac{N^2}{10} - D_m$$

式中 D_m 是同类样本中相似率最小值, N 为相似因子的个数。

凡相似率 $D \leq$ 相似水准 D_a 的样本, 即为相似样本, 而被录取; 否则为不相似, 被剔除。当 $D_m > D_a$ 时, 则认为找不到相似样本。



图二注: h_1 h_2 分别表示与 X_{Nt} X_{2t} 同类样本中的最大相似距。