

# 南海台风路径预报试验<sup>1)</sup>

刘春霞

(广州热带海洋气象研究所, 广州 510080)

周家斌

(中国科学院大气物理研究所, 北京 100029)

## 提 要

利用预测误差平方和最小准则(PRESS)的逐步算法对台风路径进行拟合和预测试验, 试验结果表明 PRESS 方法得到的模型具有较强的预测能力。因而, PRESS 方法同逐步回归一样, 也是一种预报台风路径客观有效的方法。

**关键词:** 预测误差平方和 台风路径 预测

## 前 言

台风是影响广东工农业生产的一大自然灾害, 因而台风的预报具有非常重要的现实意义, 台风路径预报是台风预报中一个重要的方面。目前关于台风的预报方法有很多, 但多数是统计客观预报, 模型的建立也多采用逐步回归方法。本文采用 Allen(1971)提出的预测误差平方和最小准则及姚棣荣、俞善贤(1992)提出的此准则逐步算法对南海台风进行预报试验, 且与逐步回归的预报试验结果进行了比较分析。

## 1 资料与分析

### 1.1 资料及预处理

本文采用《台风年鉴》1980—1989年7—9月的南海台风(10—25°N, 125°E以西)作为个例, 把每个个例每日20时的台风作为样本, 一共选取196个, 用来建立预报模型。并选取1990—1992年的台风36个样本作预报试验。

选取89个预报因子, 其中12个台风本

身的因子(即起报时台风位置所在的经纬度, 起报6小时及12小时前台风位置所在的经纬度, 起报时台风中心气压及近中心最大风力, 及6小时和12小时经向、纬向移速), 67个由欧洲中心数值预报提供的物理量场, 即: 500hPa高度场24小时变高, 起报后24小时500hPa高度场形势(10—50°N, 90—150°E), 风场 $U$ 、 $V$ 及其24小时的变化场(0—25°N, 100—130°E)等的车贝雪夫正交展开系数。预报量为24小时的台风位置。

### 1.2 方法

本文采用预测误差平方和最小准则对196个样本作回归, 其基本原理如下:

对于 $P$ 个预测因子变量 $X$ 及预报量 $Y$ 的 $n$ 次观测值, 可建立如下模型:

$$Y = XB + E$$

式中,  $B$ 和 $E$ 分别为回归系数和残差矩阵。取样本中第 $i$ 个, 其模型记为:

$$y_i = x_i B + e_i$$

用最小二乘法可以找出回归系数的估计

1) 本文得到国家“八五”攻关项目资助。

值  $b(i)$ , 从而得到样本  $i$  的预测值  $x_i b(i)$  及预报误差  $f_i = y_i - x_i b(i)$ 。对每一个样本重复上述作法, 求预测误差  $f_1, f_2, \dots, f_n$ 。求出预测平方和最小值即为 PRESS。

由于此方法计算量很大, 为此采用姚祿荣等(1992年)提出的一种逐步算法。此算法具有逐步引入因子和剔除因子的功能, 计算简便易行。

同时, 也采用逐步回归建立模型, 便于比较两种方法的计算结果。

## 2 计算结果分析

为了较好地分析和研究 PRESS 方法进行了 3 类试验。

试验 I: 采用 PRESS 方法加赤池信息准则

试验 II: 在试验 I 的基础上人为控制因子数目的最大值

## 试验 III: 逐步回归方程模型

表 1 给出 3 种试验的回归方程复相关系数, 拟合最大、最小及平均误差。表中  $X, Y$  是台风位置所在的经( $^{\circ}E$ )纬( $^{\circ}N$ )度,  $R$  为复相关系数,  $U$  为回归方程的回归平方和,  $RS$  为残差平方和,  $E_{MIN}, E_{MAX}$  和  $E_{RM}$  分别为拟合的最小、最大及平均误差( $^{\circ}$ ),  $N$  为入选预报方程的因子数。

由表 1 看出, 试验 I 入选回归方程的因子数太多, 因而在试验 II 中加入人工控制因子  $N$ , 使入选因子数适中。由此说明, 如果 PRESS 方法本身没法控制因子数, 可以通过加入一些其它的控制因子如本文所采用的赤地信息准则和人为控制等来控制入选因子数目。而逐步回归方法则通过  $F$  检验临界值来控制因子数。这里仅提供  $F=4.1$  时建立的回归方程, 便于与 PRESS 方法比较。

表 1 试验 I、II 回归方程的复相关系数、回归平方和及最大最小平均误差

	I		II		III	
	$\bar{X}$	$\bar{Y}$	$\bar{X}$	$\bar{Y}$	$\bar{X}$	$\bar{Y}$
$R$	0.9666	0.8895	0.9567	0.8839	0.9580	0.8839
$U$	5955.8177	1245.1322	5834.7989	1229.5707	5850.667	1229.5707
$RS$	418.9395	328.5150	540.0132	344.0765	524.1401	344.0765
$E_{RM}$	1.1515	0.9402	1.3100	0.9717	1.2775	0.9717
$E_{MIN}$	-0.0085	0.0066	-0.0038	0.0139	0.0152	0.0139
$E_{MAX}$	4.4686	-6.7817	4.6458	-6.8318	4.8829	-6.8318
$N$	17	10	8	7	8	7

试验 I、II 的回归方程复相关系数都在 0.9 左右, 拟合的平均误差比较小。试验 II 比试验 III 的最大、最小误差小。

这里仅给出试验 I、II 的预报方程:

试验 I:

$$X = -17.3261 + 1.0211X_{11} + 0.0136X_7 - 0.0370X_{39} - 0.0373X_{41} + 0.0466X_{43} + 0.0236X_{78} - 0.0329X_{82} + 0.0918X_{84}$$

$$Y = 3.5181 + 0.8795X_2 - 0.0315X_{50} + 0.0488X_{57} - 0.0413X_{59} + 0.0424X_{61} + 0.0375X_{72} - 0.0528X_{76}$$

试验 III:

$$X = -4.9699 + 1.0272x_{11} - 0.0232X_{16} - 0.0308X_{39} - 0.0431x_{41} + 0.0526x_{43} + 0.0386x_{78} - 0.0223x_{82} + 0.1004x_{84}$$

Y 与试验 II 相同。

上式中,  $X_1, X_2$  是起报时台风的经纬度;  $X_7$  为起报时台风中心气压;  $X_{16}$  为 500hPa 高度场的车氏系数  $A_{01}$ ;  $X_{30}$  为 24 小时 500hPa 预报场的车氏系数  $A_{22}$ ;  $X_{41}, X_{43}$  为 U 场的车氏系数  $A_{10}, A_{01}$ ;  $X_{50}, X_{57}$  为 U 场 24 小时变化场的车氏系数  $A_{10}, A_{22}$ ;  $X_{59}, X_{61}$  为 V 场的车氏系数  $A_{10}, A_{01}$ ;  $X_{72}, X_{75}$  为 V 场 24 小时变化的车氏系数  $A_{02}, A_{22}$ ;  $X_{79}$  为副高西伸脊点与台风所在经度的差值;  $X_{82}$  为副高面积指数;  $X_{84}$  为起报 24 小时前与台风同一经度副高脊

线的纬度与台风所在纬度的差值。

由入选回归因子可知, 物理量场因子在预报模型中占非常重要的地位。

表 2 给出了试验 I、试验 II 拟合及预报结果, 表中  $x_2, x_3$  表示试验 I、II 拟合及预报台风位置所在经度与实际值的绝对误差, Y 是两种方案拟合及预报台风位置所在纬度与实际位置的绝对误差。  $r_2, r_3$  是试验 I、II 预报台风位置与实际位置的距离平均误差,  $\theta_2, \theta_3$  表示试验 I、II 预报台风路径与实际路径的方向误差,  $\theta$  为方向误差。

表 2 试验 I、II 拟合及预报各类误差分析

		拟 合		预 报	
		正常路径	特殊路径	正常路径	特殊路径
绝对值平均误差(km)	次数	125	71	13	23
	$x_2$	116	184	103	133
	$x_3$	126	141	82	149
	Y	105	111	149	166
距离平均误差	$r_2$	165	261	180	252
	$r_3$	171	218	182	257
方向平均误差	$\theta_2$	23	51	23.9	56.4
	$\theta_3$	24	46	24.4	52.2
最大方向误差	$\theta_2$	154	157	82	174
	$\theta_3$	151	176	84	153
最小方向误差	$\theta_2$	1	1	1	6
	$\theta_3$	1	1	1	6
$ \theta  \leq 45^\circ$ (次数)	$\theta_2$	70	34	8	11
	$\theta_3$	70	36	8	9
	$r_2 \leq 150\text{km}$		93		12
	$r_3 \leq 150\text{km}$		88		10

分析表 2 各项数据, 可以得到以下几点:

(1) 两种方法对正常路径的经纬度拟合和预报与实际位置的平均绝对误差及平均距离误差都较小。试验 I 的拟合及预报平均距离误差为 165km 和 180km, 试验 II 的拟合及预报平均距离误差都比 I 大, 分别为 171km 和 182km。对于正常路径来说, PRESS 平均距离预报误差小于逐步回归。

关于特殊路径的拟合及预报, 两种方案的结果不如正常路径好, 试验 I 的拟合及预报平均距离误差为 261km 及 252km, 试验 II 的拟合及预报平均距离误差为 218km 及 257km。

两种方案的平均距离预报误差 PRESS 为 226km, 逐步回归为 230km(表 3)。

表3 1990—1992年台风预报距离误差(km)

	平均绝对值误差(X,Y)		平均距离误差
试验 I	122	163	226
试验 II	126	163	230

进一步讨论误差情况, PRESS 及逐步回归的平均距离拟合误差在 150km 以下分别占总数的 48% 及 45%, 而平均距离预报误差则为 33% 及 28%。平均距离预报误差在 200km 以下各占 47%、50%。

(2) 两种方案对正常路径台风的拟合及预报的方向误差平均不超过 25° 及 60°, 最低方向误差甚至达到 1°。对于 PRESS 及逐步回归方法平均方向拟合误差低于 45° 的有 104 及 106 次, 分别占总样本数(196)的 53% 及 54%。低于 45° 的预报误差 PRESS 为 20 次, 占总样本(36)数的 56%, 而逐步回归为 17 次, 占总样本数(36)的 47%。

据上述对误差的分析, PRESS 方法建立的模型相对于逐步回归来说, 虽然拟合情况差不多, 但预报情况 PRESS 方法建立的模型效果较好。

### 3 结 论

3.1 在采用 PRESS 方法建立模型时, 若入

选因子太多时, 可采用一些其它控制条件如本文的赤池信息准则和人为控制。

3.2 由两种方案(PRESS 和逐步回归)误差情况来看, 两种方案的平均距离预报误差在 200km 以下占总样本数 50% 左右, 而且平均方向误差在 45° 以下占 50% 左右, 由此看来两种方案的预报结果比较好。据文中的分析, 利用 PRESS 方法建立的预测模型的预测能力比逐步回归强。

3.3 利用物理量场的典型场(正交展开系数)作为预报因子, 既可以引入物理量场的特征, 又可以方便计算, 不失为引入物理量场有效的方法。

3.4 据两种方案预报情况来看, 采用 PRESS 方法预报模型的平均距离预报误差为 226km, 平均方向预报误差为 45° 左右。

3.5 两种试验结果表明 PRESS 方法即预测误差平方和最小准则方法提供的预测模型在台风路径预报中是比较有效的。此方法不失为一种比较有效的客观预报台风路径的方法。

参考文献(略)

## The Prediction Experiments of Typhoon Tracks over South China Sea

Liu Chunxia

(Guangzhou Institute of Tropical and Oceanic Meteorology, 510080)

Zhou Jiabin

(Institute of Atmospheric, Academy Sinica)

### Abstract

The step method based on the least prediction error sums of square is used to predict and fit the typhoon tracks. The results show that the ability of prediction model is strong. The method is an objective one which could predice typhoon tracks as effectively as the stepwise regression equation.

**Key Words:** prediction error sums of square typhoon tracks predict