

# 台风暴雨的逐步 Logit 动力统计释用预报试验<sup>†</sup>

吕纯濂<sup>1)</sup> 陈舜华<sup>2)</sup> 范淦清<sup>3)</sup>

(南京气象学院 1) 基础科学系, 2) 大气物理学系, 南京 210044, 3) 江苏省气象局, 南京 210008)

**摘要** 在对 Logit 判别系数  $\beta$  作似然比检验的基础上, 用数值预报产品作为外生变量(预报因子), 用极大似然(ML)估计, 对 Logit 判别分析的外生变量进行逐步筛选, 建立了南京台风季节(6~9月)逐日的逐步 Logit 动力统计释用天气预报模型。

**关键词** 数值预报产品, 台风暴雨, 动力统计释用

**分类号** P456

## 1 背景

根据文献 [1~4], 设  $X = (x_0 = 1, x_1, \dots, x_p)$  为  $(p+1)$  维外生变量(预报因子)向量;  $Y$  为一类别型内生变量(预报对象), 可为 1, 或为 2, ..., 或为  $m$ ;  $y_i, x_i (i=1, \dots, n)$  分别为  $Y, X$  的  $n$  个独立观测的一个样本。

Logit 判别分析(LDA)由后验条件概率(PCP)

$$\pi_s = P(Y|x_i) = \frac{\exp(u_s(x_i, \beta))}{\sum_{r=1}^m \exp(u_r(x_i, \beta))}$$

Logit 判别函数(LDF)

$$u_s(x_i, \beta) = \sum_{j=0}^p \beta_j x_{ij}, \quad s = 1, \dots, q; \quad q = m - 1; \quad (u_m = 0)$$

Logit 判别系数(LDC)

$$\beta = (\beta^1, \dots, \beta^q); \quad \beta^s = (\beta^{s0}, \dots, \beta^{sp}), \quad s = 0, \dots, q$$

和条件似然函数(CLF)

$$l(\beta|X) = \prod_{i=1}^n \prod_{s=1}^m \pi_{is}^{y_{is}}, \quad y_{is} = \begin{cases} 1, & y^i = s \\ 0, & y^i \neq s \end{cases}$$

描述。LDC  $\beta$  用 Newton-Scoring 迭代法算得。

## 2 逐步 Logit 判别分析(SLDA)

根据文献 [5~9], LDC 可重写为

$$\beta = (\beta_0, \dots, \beta_p), \quad \beta_j = (\beta_{1j}, \dots, \beta_{qj}), \quad j = 0, \dots, p$$

第 0 步 当仅考虑变量  $x_0$  时, ML 估计  $\beta_0$  由使 CLF 达极大而得到。

<sup>†</sup> 八五科技攻关(85-906-05-02-06)资助项目

收稿日期: 1995-11-02; 改回日期: 1996-01-06

第1步 当  $x_0, x_j$  在 LDF 中时,  $(\hat{\beta}_0, \hat{\beta}_j), j = 1, \dots, p$  由使 CLF 达极大得到,  $l(\hat{\beta}_0, \hat{\beta}_j)$  表示 CLF 值

$$E_{i_1} = -2 \ln \frac{l(\hat{\beta}_0)}{l(\hat{\beta}_0, \hat{\beta}_{i_1})} = \max_j (-2 \ln \frac{l(\hat{\beta}_0)}{l(\hat{\beta}_0, \hat{\beta}_j)})$$

若  $E_{i_1} > K = \chi^2_{(1-\alpha)}(m-1)$ , 则  $x_{i_1}$  在第1步引进, 否则停止。

第2步 对  $j = 1, \dots, p, j \neq i_1$ , 计算  $(\hat{\beta}_0, \hat{\beta}_i, \hat{\beta}_j)$ ,  $l(\hat{\beta}_0, \hat{\beta}_i, \hat{\beta}_j)$

$$E_{i_1, i_2} = -2 \ln \frac{l(\hat{\beta}_0, \hat{\beta}_{i_2})}{l(\hat{\beta}_0, \hat{\beta}_{i_1}, \hat{\beta}_{i_2})} = \max_{j \neq i_1} (-2 \ln \frac{l(\hat{\beta}_0, \hat{\beta}_j)}{l(\hat{\beta}_0, \hat{\beta}_{i_1}, \hat{\beta}_j)})$$

若  $E_{i_1, i_2} > K$ , 则  $x_{i_2}$  引进, 否则停止。

$$R_{i_1} = -2 \ln \frac{l(\hat{\beta}_0, \hat{\beta}_{i_2})}{l(\hat{\beta}_0, \hat{\beta}_{i_1}, \hat{\beta}_{i_2})}$$

若  $R_{i_1} < K_2(K)$ , 则  $x_{i_1}$  在下一步剔除

.....

第t步 直至所有变量都引进或  $E_{i_1, i_2, \dots, i_t} < K$

### 3 逐步 Logit 动力统计释用逐日天气预报的试验

内生变量  $Y$ : 南京 1989 年 6 月 1 日~9 月 30 的天气情况,  $Y=1$  无雨;  $Y=2$  小到中雨( $0 < R \leq 25 \text{ mm}$ );  $Y=3$  大到暴雨( $R > 25 \text{ mm}$ )。从 6 月 1 日~9 月 30 日, 影响预报地区(南京)位于  $(110^\circ \text{E}, 38^\circ \text{N}), (110^\circ \text{E}, 29^\circ \text{N}), (125^\circ \text{E}, 34^\circ \text{N}), (122^\circ \text{E}, 26^\circ \text{N})$  矩形区域上四等分行和五等分列的网格点上的预报因子为

W1 700 hPa $\theta_e$ 36 h 预报	W10 700 hPa 全风速 36 h 预报
W2 500~850 hPa $\theta_e$ 36 h 预报	W12 700 hPa 水汽通量 36 h 预报
W3 700 hPa $T - T_d$ 36 h 预报	W13 850 hPa 水汽通量 36 h 预报
W4 850 hPa $T$ 36 h 预报	W14 500 hPa 涡度 48 h 预报
W5 500~1000 hPa 厚度 36 h 预报	W15 700 hPa 垂直运动 48 h 预报
W6 500 hPa (36h~12h) 变高预报	W16 700 hPa 水汽通量 36 h 预报
W7 500 hPa 涡度 36 h 预报	W17 500 hPa 高度场 48 h 预报
W9 700 hPa 垂直运动 36 h 预报	W18 700 hPa 高度场 36 h 预报
	W19 海平面气压场 48 h 预报

#### 3.1 24 h 预报试验

从 360( $18 \times 20$ ) 个数值预报产品中, 用 SLDA 筛选 6 个因子:  $W1(18)$  为在第 18 格点上的  $W1$ (说明均同, 以下类推);  $W1(20)$ ;  $W3(9)$ ;  $W7(18)$ ;  $W9(2)$ ;  $W9(14)$ 。对数似然函数  $\ln L = -72.343$ , 拟合率为 71.074 38%, 用于预报的 Logit 判别函数为

$$\begin{aligned} u1(x) = & -26.654 + 1.404 \times 10^{-2} W1(18) + 7.100 \times 10^{-2} W1(20) + \\ & 3.092 \times 10^{-2} W3(9) - 1.548 \times 10^{-1} W7(18) + 1.943 \times 10^{-2} W9(2) + \\ & 1.721 \times 10^{-1} W9(14) \end{aligned}$$

$$\begin{aligned} u2(x) = & -97.423 + 1.021 \times 10^{-1} W1(18) + 1.879 \times 10^{-1} W1(20) + \\ & 3.736 \times 10^{-2} W3(9) - 7.281 \times 10^{-2} W7(18) + 3.210 \times 10^{-2} W9(2) + \\ & 1.018 \times 10^{-1} W9(14) \end{aligned}$$



9月	8日	9日	10日	11日	12日	13日	14日	15日	16日	17日	18日
实况	1	2	2	2	1	2	2	2	2	1	1
预报	1	2	2	2	1	2	2	1	1	2	1

### 3.3 72 h 预报试验

用 SLDA 筛选  $W1(10)$ 、 $W11(19)$ 、 $W14(18)$ 、 $W16(15)$  和  $W18(8)$  5 个因子, 对数似然函数  $\ln L = -87.589$ , 拟合率为 64.706%, 用于预报的 Logit 判别函数为

$$u_1(x) = 40.759 - 1.050 \times 10^{-1} W1(10) - 4.908 \times 10^{-2} W11(19) - 2.232 \times 10^{-1} W14(18) + 4.170 \times 10^{-3} W16(15) - 1.070 \times 10^{-2} W18(8)$$

$$u_2(x) = 35.344 - 8.765 \times 10^{-2} W1(10) - 3.400 \times 10^{-2} W11(19) - 1.851 \times 10^{-1} W14(18) + 4.187 \times 10^{-3} W16(15) - 2.706 \times 10^{-2} W18(8)$$

则对上述同时段的回报和预报(准确率为 53.92%)情况分别为(( )内为 9 月份的日期)

1989 年 8 月	3	4	5	6	7	8	9	(13)	(14)	(15)	(16)	(17)
实况	1	3	3	2	1	1	1	1	2	2	3	1
回报	1	3	3	1	1	1	1	1	2	3	3	1
1990 年 8 月	1	2	3	4	5	6	7	8	9	10	11	12
实况	1	1	2	2	2	1	2	2	1	2	1	2
预报	1	2	2	2	2	2	2	1	2	1	2	1

### 3.4 对比

用完全相同的数值预报产品数据和南京地区台风季节的天气资料, 在 SAS 统计分析系统的支持下, 调用 SAS 的 Fisher 逐步判别过程作了同样的 24、48 和 72 h 的变量筛选、回报和预报工作, 并与逐步 Logit 判别试验作了比较(见表 1)

表 1 两种方法结果对比

Table 1 Comparison of forecastings in two methods

	24 h				48 h				72 h			
	Logit(6)		SAS(8)		Logit(5)		SAS(10)		Logit(5)		SAS(10)	
	回报	预报	回报	预报	回报	预报	回报	预报	回报	预报	回报	预报
拟合(准确)率	0.71	0.60	0.83	0.63	0.74	0.58	0.91	0.65	0.65	0.54	0.82	0.49
对降水过程预报效果	好	好	好	不好	好	好	一般	一般	好	好	好	不好

说明: 拟合(准确)率只统计每天的实况与回(预)报完全一致的拟合情况; 预报效果指某一降水过程的回(预)效果, 即能否报出该过程, 两者不一定一致; ( )为所选因子数

## 4 讨论

根据本文分析结果, 不论是回报还是预报, 逐步 Logit 模型虽有时拟合率不一定很高, 但在一定程度上能反映出台风的降水过程, 比通常的以 Wilks-A 准则的逐步判别要好, 作为台风降水动力统计释用预报的新方法和新技术的尝试, 效果还是令人满意的。

致谢 本工作得到上海气象局朱永<sub>丰</sub>教授的支持帮助, 在此谨表谢意。

## 参 考 文 献

- 1 Anderson J A. Separate sample logistic discrimination. *Biometrika*, 1972, 59: 19 ~ 35
- 2 Fahrmeir L. *Multivariate statistische verfahren*. de Gruyter. Berlin New York, 1984
- 3 吕纯濂, 陈舜华. 二次 Logistic 判别分析及其气象应用. *气象学报*, 1986, (3): 299 ~ 306
- 4 吕纯濂, 姚鸿滨, 陈舜华. 用 Quasi-Newton 迭代法解二次 Logistic 判别系数超越方程组. *数值计算与计算机应用*, 1987, 8(1): 19 ~ 26
- 5 吕纯濂, 陈舜华. 逐步 Logistic 判别分析. *南京气象学院学报*, 1991, 14(3): 316 ~ 324
- 6 L C L(吕纯濂), Mathes H A. A quadratic logit discrimination algorithm in matrix form. *Computational statistics quarterly*, Physica-verlag, Heidelberg, 1991, (4): 269 ~ 280
- 7 吕纯濂, 陈舜华. 论 Logistic 判别分析中极大似然估计的存在性. *南京气象学院学报*, 1992, 15(3): 373 ~ 385
- 8 L C L(吕纯濂). The logit regression model with errors in variables and its calculation. *Chinese Journal of numerical mathematics and applications*, Allerton Inc. USA, 1994, 16(4): 29 ~ 38
- 9 吕纯濂, 陈舜华. 带有变量中误差的 Probit 模型. *系统科学与数学*, 1993, 13(4): 338 ~ 344

## FORECASTING EXPERIMENT ON STEPWISE LOGIT DYNAMIC-STATISTIC INTERPRETATION WITH NWP PRODUCTS FOR TYPHOON RAINSTORM

L Chunlian<sup>1)</sup> Chen Shunhua<sup>2)</sup> Fan Ganqing<sup>3)</sup>

(1) Department of Basic Courses, 2) Department of Atmospheric Physics, NIM, Nanjing 210044

3) Jiangsu Meteorological Bureau, Nanjing 210008)

**Abstract** A stepwise selection is made of exogenous variables (predictors) for Stepwise Logit Discrimination Analysis (SLDA) in terms of ML estimations and likelihood ratio tests for Logit coefficients  $\beta$ , whereby is constructed a forecasting model for stepwise Logit dynamic-statistical interpretation on a daily basis for the June ~ September typhoon season at Nanjing.

**Keywords** NWP products, typhoon rainstorm, dynamic-statistic interpretation