

TRP 模型及其在航空气象中的应用

田俊杰 潘春生 赵颂华

(空军气象学院气象系, 南京 211101)

提 要

通过对有关原始资料,采用等价正态偏差统计处理方法,得到了一系列原始资料与其正态化值之间的正态化转换关系;又通过引进 TRP 概率模型,建立了南京地区关于低云量和能见度的 TRP 方程;并在此基础上讨论了 TRP 模型在航空气象中的应用。

关键词:等价正态偏差(END);TRP 模型;正态化转换;临界选择概率。

1 引 言

在航空气象要素预报方面,国内不少人对此作过大量而有成效的工作^{[1][2]}。随着航空事业的发展以及空军飞行任务的增多,预报的定量化、概率化成为航空气象要素预报的发展方向。但是,对于涉及到有关的临界气象要求的概率预报问题,需要我们给予更多的关注。譬如,对于机场而言,“在某一时刻,能见度大于1.5公里且低云量小于8成的概率是多少?”这样一类关系到机场开放条件的问题。而 Boehm(1976)等^[3]提出的 TRP(Trans-normalized Regression Probability)模型,恰恰能比较灵活地解决此类问题。本文基于 TRP 模型,以能见度和低云量作为航空要素,建立相应的 TRP 方程,并具体地讨论了在航空气象保障中的应用问题。

2 资料的来源及统计处理

本文所用资料选自有限区域细网格模式(LFM)的1986年24项产品^[4];此外,从《中国地面气象观测记录月报》(1979—1986年)中又选取了华东6站7个要素的地面观测资料。这6个站分别为:南京、合肥、蚌埠、芜湖、常州和高邮;7个要素分别为:地面气压(p)、气温(T)、水汽压(E)、总云量(N)、低云量(H)、能见度(V)及地面风(F),风向和风速分别记为 FD 和 FU)。为应用 TRP 模型还参考了有关飞机起飞的最低气象条件(略)。

1991年4月3日收到,1992年1月30日收到修改稿。

* 总参资助课题。

2.1 资料的 END 化处理

统计上,等价正态偏差^[5]是这样定义的:若 P 为某一随机变量 x 的累积概率,而变量 x_P 由下式确定:

$$P = \int_{-\infty}^{\ddot{x}_P} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \quad (1)$$

则称 x_P 为变量 x 关于概率 P 的等价正态偏差 (Equivalent Normal Deviation, 简称 END)。

为叙述方便,我们约定:某随机变量 x 所对应的 END 值记为 \ddot{x} , 下同。

对于原始资料求其相应的 END 值时分两步计算:首先根据多年要素资料来统计原始变量 x 样本的累积概率 P 值;再利用公式(2)由 P 求出对应的 END 值 \ddot{x}_P 。

Abramowitz^[6](1964)曾给出 \ddot{x}_P 的有理函数的近似表达式:

$$\ddot{x}_P = t - \frac{c_0 + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3} \quad (2)$$

其中, $c_0 = 2.515517$, $c_1 = 0.802853$, $c_2 = 0.01328$, $d_1 = 1.432788$,

$d_2 = 0.189269$, $d_3 = 0.001308$, $t = [\ln(1/P^2)]^{\frac{1}{2}}$ 。

反之,由 END 值 \ddot{x}_P 反求相应的累积概率 P 由公式(3)表达:

$$P(\ddot{x}_P) = \begin{cases} 1 - \frac{1}{2}(1 + C_1 \ddot{x}_P + C_2 \ddot{x}_P^2 + C_3 \ddot{x}_P^3 + C_4 \ddot{x}_P^4)^{-1} & \text{当 } \ddot{x}_P > 0 \\ \frac{1}{2}(1 - C_1 \ddot{x}_P + C_2 \ddot{x}_P^2 - C_3 \ddot{x}_P^3 + C_4 \ddot{x}_P^4)^{-1} & \text{当 } \ddot{x}_P < 0 \end{cases} \quad (3)$$

其中, $C_1 = 0.196854$, $C_2 = 0.115194$, $C_3 = 0.000344$, $C_4 = 0.019524$ 。

2.2 资料的正常化转换

对于任一随机变量 x , 无论其服从何种分布,在理论上总可以通过某个合适的变换 $y = y(x)$ 将变量 x 转换到 y 分布上^[7]。也就是说,这里所做的正常化转换就是将实际观测的原始要素的分布值转换为与之具有相同累积概率的正态分布的变量值。

对于实际观测的要素值,往往不是通过理论上求得变换 $y = y(x)$ 的,本文是通过原始样本序列及其相对应的 END 值 \ddot{x}_P 的序列经过多种拟合试验或函数逼近法来得到 x 和 \ddot{x}_P 之间的函数关系式,即正常化转换式的。

计算中拟合的方法分别采用了从一次到三次正交多项式拟合以及 Johnson^[8](1949)对数正态模式的拟合。

由于原始资料的处理量较大,仅仅对南京站 5 个要素分别在 02、08、14 及 20 时(北京时)的拟合结果进行了必要的对比计算。分析表明:除能见度和风向(度)用二次和一次正交多项式拟合效果较好外,其余诸要素以用三次正交多项式的拟合效果最佳。因此,文中其它大量资料的处理(如 LFM 数值产品等),均采用三次正交多项式作为正常化转换关系的逼近函数。

限于篇幅,这里只列出要考虑的预报要素能见度和低云量在4个时次的转换关系(等式左端为各时次的END值, x 表示原始值)。

(1) 南京站能见度(V)在四个时次的转换式:

$$\ddot{V}02 = -2.622 + 0.034x - 0.0001x^2$$

$$\ddot{V}08 = -1.993 + 0.032x - 0.0001x^2$$

$$\ddot{V}14 = -3.007 + 0.034x - 0.0001x^2$$

$$\ddot{V}20 = -3.173 + 0.038x - 0.0001x^2$$

(2) 南京站低云量(H)在4个时次的转换式:

$$\ddot{H}02 = 0.585 + 0.029x - 0.002x^2 + 0.0002x^3$$

$$\ddot{H}08 = 0.376 + 0.090x - 0.013x^2 + 0.001x^3$$

$$\ddot{H}14 = -0.110 + 0.162x - 0.014x^2 + 0.001x^3$$

$$\ddot{H}20 = 0.557 + 0.067x - 0.008x^2 + 0.001x^3$$

求取一系列正态化转换式的目的就在于:由TRP模型计算预报概率时,可由原始资料值通过正态化转换式方便地直接计算END值,而无需再通过大量的原始资料去统计了。由END理论和正态化转换试验表明:END技术具有一定的理论意义和实用价值。气象要素的原始资料一经END化处理,就能保证转换后的要素值服从正态分布^[3]。同时,END的统计方法是建立在以累积概率相等为转换原则的气候学意义基础之上的,故经过END处理过的要素值含有一定的气候因素。因此,在建立回归方程时为挑选预报因子提供了一定的气候学背景。

3 TRP模型

Boehm^[3](1976)的转换正态回归概率模型(下称TRP模型):

$$\ddot{P} = (\ddot{Y} - \ddot{M}) / \sqrt{1 - R^2} \quad (4)$$

其中, \ddot{P} 、 \ddot{Y} 及 \ddot{M} 分别为变量 P 、 Y 和 M 的END形式。

式(4)中: \ddot{Y} 为预报要素量临界气象要求(譬如能见度 $\leq 6\text{km}$)出现的气候概率。根据文献[3]的推导,经过END化之后的预报量 \ddot{Y} 和各预报因子 $\ddot{x}_i (i = \overline{1, m})$ 服从 $m+1$ 维的联合正态分布。因此,预报量 \ddot{Y} 的条件分布就是: $f(\ddot{Y} | \ddot{x}) = N(\ddot{M}, S^2)$,其中 $\ddot{M} = \sum_{i=1}^m A_i \ddot{x}_i$ (A_i 为偏回归系数, \ddot{x}_i 为第 i 个入选因子的END值),在统计意义上表示预报量 \ddot{Y} 的条件分布的均值, $S = \sqrt{1 - R^2}$ 表示方差, R 为复相关系数。最终,式(4)计算得到的 \ddot{P} 为预报要素量在多因子影响下大于(或小于)某一临界气象要求的条件概率。当然, P 是由 \ddot{P} 经过式(3)反求而得到的概率值。

引进TRP模型的意义就在于:在航空气象要素预报时,能够针对不同的临界气象要

求进行单要素或多要素的概率预报,其具体应用将在下一部分讨论。

由式(4)可知, \dot{P} (或 P) 值的大小取决于临界要求 \dot{Y} (或 Y) 与均值 (或回归值) \dot{M} 之差。若临界气象要求较低,为一容易发生的事件,相应地, $\dot{Y} (> 0)$ 相对较大 (或 $Y > 0.5$); 当 $\dot{M} (< 0)$ 较小时,将有 $\dot{Y} - \dot{M} > 0$ 。故由 TRP 模型得到的值: $\dot{P} > \dot{Y} - \dot{M} > 0$, 计算出的概率值 $P > 50\%$ 或更大,使符合临界气象要求的天气的发生是大概率事件。反之,若临界气象要求较高, \dot{Y} 较小 (< 0), 当 $\dot{M} (> 0)$ 较大时, $\dot{Y} - \dot{M} < 0$, 由此而计算得到的 $P < 50\%$ 甚至更小,未来天气的发生说明不能满足临界气象要求。当然, \dot{M} 值的大小取决于在预报时刻的回归值, \dot{M} 值越小,表明预报量 \dot{Y} 的条件分布愈接近标准正态分布。

3.1 能见度和低云量的 TRP 模型

我们分别针对预报要素量能见度和低云量,建立了南京地区的 12 小时、24 小时和 36 小时的回归方程 \dot{M} 。这 6 个回归方程主要是通过普查华东 6 站 (包括南京) 的地面实况资料及其不同的组合因子的 END 值以及 LFM 数值预报产品的 END 化资料,经过逐步回归而得到的。实况资料的样本量为 7 年, LFM 数值产品的样本量为 1 年 (用于回归过程的只有 1 年), 在逐步回归过程中引进和剔除因子的标准分别为: $F_{91} = 2.5, F_m = 2.0$ 。限于篇幅,仅列出能见度 (V) 12 小时和低云量 (H) 24 小时回归方程:

$$\dot{V}_{12} = 0.6722 - 0.3984\dot{x}_1 + 0.2921\dot{x}_2 + 0.3253\dot{x}_3 - 0.1360\dot{x}_4 + 0.2146\dot{x}_5$$

$$- 0.0745\dot{x}_6 + 0.144\dot{x}_7 + 0.1724\dot{x}_8 + 0.2192\dot{x}_9$$

$$R_{12} = 0.4763, Q = 0.773, F = 14.97, D = 0.79$$

$$\dot{H}_{24} = 0.5779 + 0.2293\dot{x}_1 + 0.2469\dot{x}_2 - 0.1141\dot{x}_3$$

$$+ 0.0997\dot{x}_4 + 0.06224\dot{x}_5 + 0.0753\dot{x}_6$$

$$R_{24} = 0.46, Q = 0.79, F = 15.6, D = 0.81$$

上述回归方程中 R_{12} 和 R_{24} 分别表示 12 小时和 24 小时的回归方程的复相关系数; Q 为剩余平方和; F 为方差比, 即: $F = \frac{U/m}{Q/(n-m-1)}$; $D = n \cdot \frac{Q}{n-m-1}$; 另外, 预报因子 $\dot{x}_1, \dot{x}_2, \dots, \dot{x}_9$ 分别表示入选因子的 END 值 (其具体的物理意义限于篇幅省略)。

3.2 能见度和低云量的 TRP 模型的检验

我们选取了独立样本量为 1 年的原始资料进行分类 (级) 的试报, 并进行了 Brier^[9] 得分的计算, 如表 1 所示。结果表明: 针对不同的临界要求 (或类别), 其 Brier 得分 P_s 在三个预报时段内以 12 小时预报为最佳。能见度的预报能力要比同时效的低云量预报效果稍好, 且 36 小时的能见度预报及低云量的预报的 P_s 值较小。分析表明这主要归因于 36 小时预报方程入选的因子主要是数值预报产品缘故。此外, 试报中分级愈多, P_s 愈大, 即分类预报的区间愈细, 报错的概率愈大。归结起来, 所建立的能见度及低云量的 TRP 模型的预报能力受到限制的原因为 TRP 模式是建立在转换的正态概率分布基础之上的; 因此, 对

于预报要素低云量和能见度以及回归方程中的诸多入选因子要详细研究其概率分布特征;某些要素的概率密度分布在向正态化转换时可能不太适宜,需要从另外角度进行变换。我们曾对能见度、低云量和其它一些要素计算其概率密度函数及其 Pearson^[10] 类型,发现能见度均为正偏的第二类 B 分布(Pearson VI 型分布),而低云量则均为 Pearson I 型分布(β 分布)。低云量的这种“U”形分布对于 END 处理和正态化转换具有一定的影响。

表 1 TRP 模型的 Brier 得分 P_s

类别数 r	预 报 时 效					
	12 小时		24 小时		36 小时	
	低云量	能见度	低云量	能见度	低云量	能见度
2	0.79	0.79	0.94	1.10	1.08	0.70
3	0.81		0.95		1.12	
4		0.85		0.98		0.78
5	0.97		1.08		1.39	

注:(1)对低云量,其分类预报的临界值 H_c 取值如下: $r=2$ 时, $H_c=4$; $r=3$ 时, H_c 取 3 和 7; $r=5$ 时, H_c 取 3, 5, 8, 10。

(2)对能见度,其分类预报的临界值 V_c 分别为: $r=2$ 时, $V_c=6\text{km}$; $r=4$ 时, V_c 分别取 1, 3, 6km。

4 TRP 模型在航空气象中的应用

由式(4)的 TRP 模型可知,预报变量是单个要素。在实际应用过程中,我们碰到的往往是多要素的概率预报。具体而言,譬如,设某机场开放所需要的低云量和能见度的最低临界气象条件为 (H_c, V_c) , 这就是一个关于 TRP 模型的二维问题。关于 TRP 模型的多维问题的讨论,我们将在以后的工作中进一步讨论。本文仅仅就低云量和能见度两个预报要素进行讨论。

在假定两个随机变量分别服从正态分布的条件下,可求得二维联合正态概率的分布函数及其级数展开^[6]。

因此,若假设低云量和能见度的临界气象条件为 (H_c, V_c) , 其联合概率函数记为: $P_{HV} = P_r\{(H:H \leq H_c) \cap (V:V > V_c)\}$, 则通过 TRP 预报值 \dot{P}_H 和 \dot{P}_V , 求得关于低云量和能见度的联合概率函数为:

$$P_{HV} = L(-\dot{P}_H, \dot{P}_V, -\rho_{\dot{H}\dot{V}})$$

其中 $\rho_{\dot{H}\dot{V}}$ 为 \dot{H} 和 \dot{V} 的相关系数。

至此,关于二维的 TRP 模型的计算问题已经解决。限于篇幅,不再列出计算实例。

在实际的航空气象保障中,实际发生的天气和 TRP 预报结果相比较共有四种情形:飞机起飞成功的概率(A),表示 TRP 预报可飞且实际天气也满足临界要求的概率;贻误时机的概率(B);保障失败的概率(C);余项(D),且 $A+B+C+D=1.0$; TRP 预报可飞的概率(X),天气满足要求的实际气候概率(Y)。

根据以上 A、B、C、D 的意义,我们选择 TRP 模型中的回归值 \dot{M} (或其对应的概率 M) 作为模拟的控制参量,称之为临界选择概率,记为 CP (或 END 值 $\dot{C}P$)。然后,给定一组 CP 值,对于不同的机型所要求的临界低云量和能见度的要求以及预报时效,分别计算 A、B、

C 和 D 值, 构成了 TRP 模型的模拟概率列联表。表 2 和表 3 分别给出了南京地区 TRP 模型模拟的概率列联表(起始时间 20 时; 预报时效为 12h; 预报时刻北京翌日 8 时; 相关系数 $R = 0.645$; 临界条件 $H_c = 8.00$; V_c 在表 2 为 7.00 相当于 0.1km; 在表 3 为 20.00 相当于 0.1km)。

表 2 喷气式轰炸机起飞的模拟概率表

临界选择概率 (CP)	模拟概率			
	A	B	C	D
0.010	0.732	0.000	0.268	0.000
0.050	0.731	0.001	0.267	0.001
0.100	0.732	0.000	0.262	0.006
0.150	0.734	-0.002*	0.251	0.017
0.200	0.737	-0.005	0.235	0.033
0.250	0.740	-0.008	0.215	0.053
0.300	0.742	-0.010	0.191	0.077
0.350	0.741	-0.009	0.164	0.104
0.400	0.737	-0.005	0.136	0.132
0.450	0.728	0.004	0.107	0.161
0.500	0.713	0.020	0.078	0.189
0.550	0.690	0.043	0.071	0.216
0.600	0.657	0.076	0.027	0.240
0.700	0.557	0.176	-0.007	0.275
0.800	0.404	0.329	-0.018	0.286
0.900	0.197	0.535	-0.004	0.272

表 3 直升机 1 号条件下起飞的模拟概率表

临界选择概率 (CP)	模拟概率			
	A	B	C	D
0.010	0.760	0.000	0.240	0.000
0.050	0.759	0.001	0.240	0.001
0.100	0.759	0.001	0.236	0.004
0.150	0.761	-0.001*	0.227	0.013
0.200	0.764	-0.004	0.214	0.026
0.250	0.767	-0.007	0.198	0.042
0.300	0.769	-0.009	0.177	0.063
0.350	0.769	-0.009	0.154	0.086
0.400	0.766	-0.006	0.129	0.111
0.450	0.758	0.002	0.103	0.137
0.500	0.745	0.015	0.077	0.163
0.550	0.724	0.036	0.053	0.187
0.600	0.694	0.066	0.030	0.210
0.700	0.599	0.161	-0.005	0.244
0.800	0.447	0.313	-0.017	0.257
0.900	0.231	0.529	-0.005	0.245

* 计算中 B 或 C 栏负概率的出现主要是采用 Abramowitz(1964) 公式截断误差所致, 不过量级很小, 可忽略不计; 表 3 情况同此。

表 2、表 3 的模拟结果表明: 在地域、时间和临界气象要求一定的情况下, CP 愈大, 飞机起飞的成功率 A 愈小, 保障失败的概率 C 愈大; 反之, CP 愈小, A 则愈大; 即模拟表中 A 、 B 、 C 、 D 值对临界选择概率 CP 的变化较为敏感。在具体的航空气象保障的实际工作当中, 航调指挥人员首先可根据需要, 选择表中某一合适的横行中的成功概率 A , 然后, 沿着 A 所在的横行, 找到对应的临界选择概率 CP ; 此时再调用计算机系统中已建立的回归方程的预报值 \bar{M} 及其对应的概率值 $P(\bar{M})$, 若 $P(\bar{M}) \leq CP$, 则判定所要起飞的飞机在指挥员所选择的危险意义上能够安全起飞或降落; 反之, 若 $P(\bar{M}) > CP$, 则飞机不可飞。以表 2 为例, 飞行指挥员傍晚 20 时计划翌日 08 时进行喷气轰炸机起降任务, 模拟结果表明, 指挥员若选 $A = 72.8\%$, 则查得 $CP = 0.45$; 假若第二天 08 时南京机场的能见度 $> 2\text{km}$ 且低云量 < 8 成的回归预报概率 $P(\bar{M}) < 0.45$, 则飞行任务可以完成。在实际应用中, 能否出现“指挥人员为求保险而选表中最大的成功率 A ”这种情况呢? 事实上, 这还牵涉到回归预报概率值 $P(\bar{M})$ 大小的问题, 即指挥员所选择的 A 只有在 $P(\bar{M}) \leq CP$ 的情况下方认为飞行任务能够完成; 否则, 作这种所谓 A 最大的选择是毫无意义的。实际上, 当时当地以及在预报时刻实际所发生的天气制约着指挥员对成功率 A 的选择。

5 结 语

本文利用 TRP 概率模型, 建立了南京地区关于低云量和能见度的 TRP 方程; 并在此

基础上进行了模拟试验,讨论了TRP模型在航空气象中的应用。对于实际的航空气象保障,在天气多变的情况下,调航或机场指挥员针对不同的临界气象要求如何确定航班计划,无疑为客观定量决策提供了一个新思路。

参 考 文 献

- 1 王炳仁,刘健文.应用GEM模式作北京地区航空气象要素预报的探讨.气象学报,1987,45(2):202—209.
- 2 王炳仁.低云、雾和能见度预报和现状的研究.航空气象科技,1984,(5).
- 3 Boehm,A. R. . Transnormalization Regression Probability,AWS-TR-75-259,1976,(AWS).
- 4 张绍晴.北京气象中心有限区域模式软盘资料.技术报告,第8704号,北京气象中心.
- 5 Kendall Maurice G. and Buckland,R. William. A Dictionary of Statistical Terms, Oliver and Boyd. London: 1957. 493.
- 6 Abramowitz M. , et al. . Handbook of Mathematical Function, National Bureau of Standards, Applied Math. Series 55, Washington D. C. :1964.
- 7 么枕生.气候统计学基础.北京:科学出版社,1984.
- 8 Johnson, N. L. . Systems of frequency curves generated by methods of transiation. *Biometrika*, 1949,(36):149—176.
- 9 Brier, G. W. . Verification of forecasts expressed in term of probability. *Monthly Weather Review*, 1950,78 (1): 1—3.
- 10 .Sir Maurice Kendall and Alan Stuart. The Advanced Theory of Statistics, London and High Wycombe;Charles Griffin and Company Limited, 1977.

THE TRANSNORMALIZED REGRESSION PROBABILITY (TRP) MODEL AND ITS APPLICATION IN AERONAUTICAL METEOROLOGICAL SERVICE

Tian Junjie Pan Chunsheng Zhao Songhua

(Air Force Institute of Meteorology, Nanjing 211101)

Abstract

In the paper, an Equivalent Normal Deviate (END) statistical method is introduced in order to process relative data and establish corresponding transnormalized functions by using the climatological data at Nanjing and surrounding five locations and the data of LFM numerical products of NMC in 1986. TRP equations of low cloud cover and visibility at Nanjing are also developed using TRP model. Finally, the applications of TRP model in aeronautical meteorological service are discussed.

Key words: Equivalent Normal Deviate (END); Transnormalized Regression Probability (TRP) model; Critical meteorological requirements; Critical selected probabilities.