

提高小概率事件预报成功率的一条途径^①

彭治班¹ 吴宝俊¹ 江剑民² 王淑静³

¹(中国气象科学研究院,北京 100081)
²(中国气象局培训中心,北京 100081)
³(国家气象中心,北京 100081)

提 要

在给出小概率事件定义与评估小概率事件预报成功率参数的基础上,根据参数变换后的表达式推断:变小概率事件为条件概率下的大概率事件,是提高小概率事件预报成功率的一条途径。

关键词: 小概率事件 临界成功指数 CSI 强降雪

引 言

在气象服务业,用户关心的暴雨、强降雪、雷雨大风、龙卷等许多灾害性天气现象,都属于发生次数较少的小概率事件。提高小概率事件的气象服务质量,虽然与许多环节有关,但提高预报成功率却是其非常关键的一环。如何提高?我们认为变小概率事件为条件概率下的大概率事件,是提高小概率事件预报成功率的一条途径。下面以强降雪为例,说明我们对该问题思考的结果。

1 小概率事件

1.1 小概率事件的界定

小概率事件一词,在我国天气预报人员中虽颇为流行,但却很难找到具体界定标准。在文献[1]中,按事物发生可能性分为5类:(a)极可能发生,概率在90%以上;(b)相当可能发生,概率为75%左右;(c)可能、不可能发生机会均等,概率为50%左右;(d)不太可能发生,概率为25%左右;(e)发生的可能性极小,概率小于9%。本文所说的小概率事件,是指概率小于9%的事件。

1.2 强降雪是一种小概率事件

1.2.1 呼和浩特地区强降雪日出现的概率
借用文献[2]的资料,我们给出了5~9月呼和浩特地区强降雪日出现的概率(表1)。

表1 呼和浩特地区5~9月强降雪日出现概率

月份	5	6	7	8	9
26年强降雪日总次数	10	42	58	46	39
强降雪日出现概率/%	1.2	5.4	7.2	5.7	5.0

表1中的强降雪日,是指大范围降雪(一日内降雪点有3处以上)或(与)雪块直径 $D \geq 1.5\text{cm}$ 的降雪或(与)造成重灾的降雪日。这里所说的冰雹,应按广义去理解,即无论其体积大小、质地软硬都视为冰雹,全部参加统计,当呼和浩特地区内的气象台站(4个)、防雹炮点(30个)、雨情点(10个)及灾情调查报告一日内有一处以上(包括一处)记录有降雪时就定义为一个降雪日^[2]。由表1可知,呼和浩特地区的强降雪是一种小概率事件。

1.2.2 青藏高原以东地区强降雪是一种小

① “内蒙防雹减灾新技术和冰雹专项预报研究”、“冰雹落区预报逐级指导技术研究”、“新一代气象服务体系研究”共同资助。

概率事件

由全国年平均降雹日数分布图^[3,4]看出:除青藏高原是一片多雹区外,在高原以东有南北两支多雹带。北支从青藏高原北部出祁连山、六盘山,经黄土高原和内蒙古高原连接,延伸到包括冀北和东北三省的一些地区,是我国最宽、最长的一条多雹地带,其雹日明显比南支多。呼和浩特地区即位于北支多雹带内,且在一个多雹中心附近。

如上所述,文献[2]统计的是呼和浩特地区的降雹日,这无疑比呼和浩特市降雹日多些。另外,文献[2]不仅考虑气象台站记录到的降雹,还考虑了防雹炮点、雨情点及灾情报告中记录的降雹,这无疑也增多了降雹日。即便这样,强降雹日仍是小概率事件,这显然支持我们欲得的结论。

由1.2.1与1.2.2的讨论,我们可以得出以下初步看法:对于我国,在青藏高原以东地区,强降雹是一种小概率事件。其它地区,本文暂不讨论。

2 评估小概率事件预报质量的参数

对于检验、评估小概率事件预报质量用的参数,近年虽然有了更深入研究并提出了一些较新者(例如见文献[5]),但考虑到目前我国天气预报质量检验和评估的有关规定^[6],本文暂只考虑文献[7]引入的临界成功指数 CSI (即文献[6]中的 TS 评分)、探测概率 POD 与虚假报警率 FAR 。

根据表2及 CSI 的定义^[7],有:

$$CSI = \frac{x}{x + y + z} \quad (1)$$

由表2可见,本文说的预报属于二分类预报,即只考虑事件的出现与否。在评价这类预报的成功率时,不仅要考虑预报该事件发生时发生者占的比率,还要考虑预报该事件不发生时不发生者占的比率;两者比率越大(小),其预报成功率越高(低)。照此,可以认为选用现有的 CSI 作为评价小概率事件预报成功率的一个参数是有道理的。

根据表2,我们还可以给出探测概率 POD 与虚假报警率 FAR 的表达式^[7]:

$$POD = \frac{x}{x + y} \quad (2)$$

$$FAR = \frac{z}{x + z} \quad (3)$$

上面引入的 CSI 、 POD 、 FAR 是衡量预报质量、预报指标或预报方法优劣的很有用的参数。需要说明的是,在本文中 POD 的含义改为:报对次数与事件出现总次数的比值。

表2 定义 CSI 、 POD 、 FAR 用表

预报(强降雹)	观测(强降雹)		总计
	有	无	
有	x	z	$x + z$
无	y	w	$y + w$
总计	$x + y$	$z + w$	N

3 提高预报成功率的途径

由式(1)看出,只有当 x 较大以及(或) y 与 z 较小时, CSI 的数值才较大。这就是说, CSI 的大小,可以代表预报成功率的大小;若找出增加 CSI 的途径,就等于找到了提高预报成功率的途径。

3.1 CSI 表达式的变换

假定在预报质量评估时段内发布过强降雹预报,亦即 $(x + y) \neq 0$, 则对式(1)进行变换,得:

$$CSI = \frac{\frac{x}{x + y}}{\frac{x + y}{x + y} + \frac{z}{x + y}} \quad (4)$$

假定在预报质量评估时段内发布过正确的无强降雹预报 w 与(或)出现过“空报” z , 亦即 $(z + w) \neq 0$, 则对式(4)中分母 $z/(x + y)$ 进行变换,得:

$$\begin{aligned} \frac{z}{x + y} &= \frac{z(z + w)}{(x + y)(z + w)} \\ &= \frac{z[(z + w)/N]}{(z + w)[(x + y)/N]} \end{aligned} \quad (5)$$

考虑到

$$\frac{z}{z + w} = \frac{(z + w - w)}{z + w} = 1 - \frac{w}{z + w} \quad (6)$$

将式(6)代入式(5)得:

$$\frac{z}{x+y} = (1 - \frac{w}{z+w}) \left[\frac{(z+w)/N}{(x+y)/N} \right] \quad (7)$$

现引入以下定义:强降雹概率 $P_h = (x+y)/N$, 无强降雹概率 $P_n = (z+w)/N$, 强降雹探测概率 $(POD)_h = x/(x+y)$, 无强降雹探测概率 $(POD)_n = w/(z+w)$ 。考虑到以上定义及 $P_h + P_n = 1$, 得:

$$\begin{aligned} \frac{z}{x+y} &= [1 - (POD)_n] \frac{P_n}{P_h} \\ &= [1 - (POD)_n] \left[\frac{1 - P_h}{P_h} \right] \\ &= \left[\frac{1}{P_h} - 1 \right] [1 - (POD)_n] \quad (8) \end{aligned}$$

将式(8)代入式(4), 得:

$$CSI = \frac{(POD)_h}{1 + \left(\frac{1}{P_h} - 1 \right) [1 - (POD)_n]} \quad (9)$$

3.2 由公式(9)得出的几个推论

由式(9)看出:

当 P_h 与 $(POD)_n$ 一定时, CSI 与 $(POD)_h$ 成正比关系。

当 P_h 与 $(POD)_h$ 一定时, CSI 与 $(POD)_n$ 成正比关系。亦即当 $(POD)_n$ 增大(减小)时, CSI 亦增大(减小), 但两者不成线性关系。

当 $(POD)_h$ 与 $(POD)_n$ 一定时, 强降雹概率 P_h 愈小(大), 预报临界成功指数 CSI 愈低(高), 参见表3。

表3 说明 CSI 与 P_h 关系用表

		$(POD)_h =$ $(POD)_n = 0.6$	$(POD)_h =$ $(POD)_n = 0.7$
CSI	$P_h = 0.1$	0.13	0.19
CSI	$P_h = 0.2$	0.23	0.32

当然, 由式(9)还可以得出另外的一些推论。但考虑到本文的任务, 我们不再涉及。

综上推论, 要提高强降雹预报的 CSI , 除提高 $(POD)_h$ 与 $(POD)_n$ 外, 还要设法加大 P_h 。但实际上, 对于某一地区某一季节, 当做多年统计时, P_h 应当不变。因此, 严格地讲,

当设法加大 P_h 或说“变小概率事件为条件概率下的大概率事件”时, 是指在某些条件下减小分母 N 来提高 P_h 的数值, 即增大条件概率值。

4 小 结

对于我国, 在青藏高原以东地区, 强降雹是一种小概率事件。由目前我国天气预报质量评定中使用的 TS 评分公式(与本文式(1)相同)推出, 变小概率事件为条件概率下的大概率事件, 是提高小概率事件预报成功率的一条途径。与式(9)类似的公式虽然早已有人提出, 例如文献[9], 但考虑到该观点在提高灾害性天气预报成功率方面的重要性, 我们仍从不同角度做了阐述。至于日常业务中使小概率事件变为大概率事件的具体方法, 例如起报条件法、消空法等, 拟另文讨论。

参考文献

- 金 哲等. 当代新方法. 上海: 上海人民出版社, 1990, 103~106.
- 巴特尔, 巩 迪, 李一平. 呼和浩特地区冰雹天气的气候分析. 内蒙古气象, 1996, (3): 23~31.
- 刘全根, 汤懋苍. 中国降雹的气候特征. 地理学报, 1966, 32: 30~65.
- 雷雨顺, 吴宝俊, 吴正华. 冰雹概论. 北京: 科学出版社, 1978, 112~117.
- Marzban C. Scalar measures of performance in rare-event situations, *Weather and Forecasting*. 1998, 13: 753~763.
- 中国气象局业务发展与天气司. 天气预报业务规定(试行). 1998年6月.
- Donaldson, R J Jr, R M Dyer and M. J. Kraus. An objective evaluator of techniques for predicting severe weather events. Preprints, Ninth Conf. on Severe Local Storms, Norman, Oklahoma, Amer. Meteor. Soc., 1975, 321~326.
- 施 能. 气象科研与预报中的多元分析方法. 北京: 气象出版社, 1995, 18~21.
- 张承福. 人工神经网络在天气预报中的应用研究. 气象, 1994, 20(6): 43~47.

(下转第10页)

A Method to Improve the Forecast Accuracy of Rare Events

Peng Zhiban¹ Wu Baojun¹ Jiang Jianmin² Wang Shujing³

¹(Chinese Academy of Meteorological Science, Beijing 100081)

²(The Training Centre, China Administration of Meteorology, Beijing 100081)

³(National Meteorological Centre, Beijing 100081)

Abstract

Based on the definition of rare events to be given and the assessment of the successful parameters forecasting rare events, an equation by paramant transformation was suggested. It can be seen that a method to improve the forecast accuracy of rare events is to turn the rare eveuts from very little probability into large probability by conditional probability.

Key Words: rare event critical success index severe hail