

杜钧, 邓国. 单一值预报向概率预报转变的价值: 谈谈概率预报的检验和应用[J]. 气象, 2010, 36(12): 10-18.

单一值预报向概率预报转变的价值: 谈谈概率预报的检验和应用^{*1}

杜 钧¹ 邓 国²

¹ 美国国家海洋大气局, 华盛顿

² 国家气象中心, 北京 100081

提 要: 基于大气系统的非线性和复杂性, 加上初值和模式等本身无法避免的一些不确定性, 天气预报从单一值的确定论向多值的概率论转变不但符合气象科学的实际也是更好地服务社会之必需。但是这种转变的理性思考对很多用户、甚至气象工作者自己都不是很清楚。为了帮助人们理解和加速这一转变, 本文就这种转变所面临的一些普遍困惑谈一些想法。作为促成这一转变的中心技术——集合预报——有关的问题已经在本文的姐妹篇中作了讨论。在本文中主要讨论另外几个问题, 即如何衡量一个概率预报的好坏? 如何应用概率信息来作决策? 以及为什么概率预报比单一值决定论预报更有经济价值? 为了易懂, 本文用浅显的说理结合举例的方式进行阐述, 以便读者特别是广大预报员和用户能在实际中运用。

关键词: 预报不确定性, 概率预报, 检验, 气象用户, 经济价值, 花费—损失比, 行业气象决策模式

The Utility of the Transition from Deterministic to Probabilistic Weather Forecasts—Verification and Application of Probabilistic Forecasts

DU Jun¹ DENG Guo²

¹ National Oceanic and Atmospheric Administration, Washington, DC, U. S. A.

² National Meteorological Center, Beijing 100081, China

Abstract: Given the nonlinear and complex nature of atmospheric system combined with the uncertainties associated with initial conditions and models, it is not only a requirement from science but also a necessity to better serve user community to transit weather prediction from deterministic to probabilistic format. However, the reason behind the transition is not clear but full of debate and confusion to many including public, end-users and even meteorologists themselves. To clarify some of these and to expedite the transition, a few issues were discussed. The ensemble forecasting which is a central technique to facilitate such a transition has been discussed in the companion paper of this article. In this paper, a few other issues related to ensemble-based probabilistic forecasts were discussed, which include (1) how to measure the merits of a probabilistic forecast? (2) how to use probabilistic information in decision-making process? and (3) why an ensemble-based probabilistic forecast has, in general, more utility or economical value than a single-value deterministic forecast does? To be easily understood and applicable to real world situations by readers especially forecasters and end-users, plain language illustrated with examples was used to explain the related concepts.

Key words: forecast uncertainty, probability forecast, verification, end-user, economic value, cost-loss ratio, decision-making model

* 国家自然科学基金 41075079; 公益性行业科研专项(GYHY200706001, GYHY200906007)资助
2010年4月28日收稿; 2010年7月29日收修定稿
第一作者: 杜钧, 主要从事数值集合预报业务和开发研究. Email: Jun. Du@noaa.gov

引言

基于大气系统的非线性和复杂性,加上初值和模式等自身无法避免的一些不确定性^[1-4],天气预报从单一值的决定论向多值的概率论转变不但符合气象科学的实际,也是更好地服务社会大众之必需^[5-6]。在本文的姐妹篇^[7],我们已经谈了在这一转变中扮演中心角色的集合预报技术。不同于传统的统计办法,集合预报是用动力学的手段定量地把每个个体预报(不是平均情况)相对应的、特定的预报不确定性给度量出来。但如何来表达不确定性这是集合预报研究中的重要子课题^[8],是目前国内外数值预报研究和业务预报的重要方法^[9-14]。概率是其中一个较常用的,也被大部分人所熟知的手段^[15]。虽然如此,但许多人仍对概率预报的真正含义不很清楚,对如何来衡量概率预报的优劣可能更是模糊。因此本文首先讨论了一下概率预报的含义及其相应的检验方法。

正如 Alan Murphy^[16]所指出,任何一个预报只有当被用户充分理解并付之于决策行动之中,才有其真正的价值。否则,哪怕是一个 100% 准确的预报也没有任何价值。所以,本文接着讨论了如何应用概率信息来作决策的基本原则和方法。并用两个例子来说明,在承认预报有不可避免的误差即存在预报不确定性的前提下,应用概率预报其实比用单一值的决定论预报来决策更具经济价值。最后,用相对价值的概念从理论上来说明为什么概率预报一般比单一值确定预报更有应用价值。希望本文能帮助人们理解应用预报不确定性的价值。

1 如何检验概率预报的准确性

不像单一值决定性预报,其黑白分明,从理论上来说它很容易被检验。但怎样来检验一个概率预报呢?譬如预报说 60% 会下雨,那么如果实况下雨了,该说它是对的还是错的呢?其实,对于某一个例来说,我们既不能说它是对的也不能说它是错的。要检验概率预报,必须要有很多的个例,有了足够的样本数,才可以来检验一个概率预报是不是可靠(注意,检验一个概率预报通常是说它是否“可靠”而不是说它是否“准确”)。一个完全可靠的概率预报,譬如说 60%。那么每 100 次“60%”的预报

中,60 次实况应该出现而 40 次不会出现,至于具体哪一次会出现哪一次不会出现则无从得知。这 100 次预报可以是时间上的(即在同一地点出现在不同时间上的 100 次同一概率预报)也可以是空间上的(即同一时间出现在 100 个不同地点上的同一概率预报),当然也可以是时间和空间上混合的。概率预报的可靠与否通常用所谓的“可靠性曲线”来表征(如图 1)。当预报概率和观测概率(准确地说应该是频率,因为概率是基于无限多的样本上的,而频率是基于有限的样本数上的)完全一致时(即预报完全可靠)其曲线就是对角线;当预报概率小于观测概率时(即低估或预报太保守)其曲线则位于对角线之上,当预报概率大于观测概率时(即高估或预报太自信)其曲线则位于对角线以下。图 1 所示的是“0~30%”的预报过于保守(低估)，“30%~60%”的预报基本可靠,而“60%~100%”的预报则太自信(高估)。

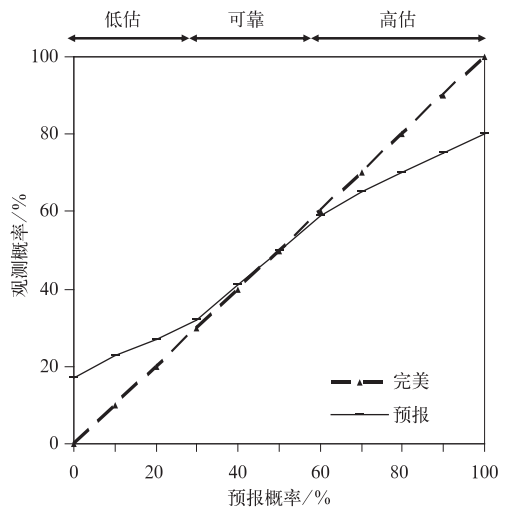


图 1 衡量概率预报是否可靠的“可靠性曲线”。水平轴(x轴)为预报的概率(频率),垂直轴(y轴)是观测到的概率(频率)。对角线(虚线)代表完全可靠的完美概率预报,而实线为一个不完美的实际预报的例子。

Fig. 1 An example of “reliability diagram” to measure how reliable a probabilistic forecast is. The x-axis is forecast probability (frequency) and y-axis observed probability (frequency). The diagonal dash line represents a completely reliable (perfect) forecast and the solid line an example of an imperfect forecast

我们知道一个比较“尖窄”的概率分布(图 2 中的实线)比一个“平宽”的概率分布(图 2 中的虚线)含有更多的信息,因为它的不确定度较小,因而也有

更高的应用价值。因此在保持概率预报的可靠性不变的前提下,这概率密度函数分布的尖窄或平宽也是衡量概率预报质量好坏的标志之一。例如,对一个有两种可能性(出现或者不出现)的事件,系统 1 的预报概率分布可能是(50%,50%),而系统 2 的可能是(80%,20%)的分布。显然,系统 2 优于系统 1,因为系统 2 比较确定。而相对于比较“尖窄”的概率分布,其对某一事件的预报也就较确定(譬如表的例子)。表 1 列出了两个预报系统对 5 次气温的概率预报。假定每个预报系统所给出的概率预报对该系统本身来说是可靠的,那么预报系统 B 就优于预报系统 A,因为从统计结果来看它对实况出现的事件的预报概率值(51%)一般大于预报系统 A 所提供的(42%),因此也就更有可用的信息。通常,我们把一个预报概率分布同其相应的气候概率分布(它是完全可靠的,因为它代表的就是实况)来相比较并称之为“分辨率”。如果它比气候概率分布还偏平的话,那么就没有分辨率可言,因为它没有比仅仅用气候概率来作预报提供更多的信息和更有价值。只有比气候分布包含更多信息的概率预报才有分辨率、才有更大的价值。所以,我们在追求概率预报的“可靠性”(不确定性的衡量)的同时也要追求它的高“分辨率”(确定性的衡量)。一般来说“可靠性”依赖于集合预报技术的好坏,而“分辨率”则依赖于模式和初值的质量。这儿我们又看到了文献[7]和[8]所强调的一个观点:即一个好的概率预报不仅

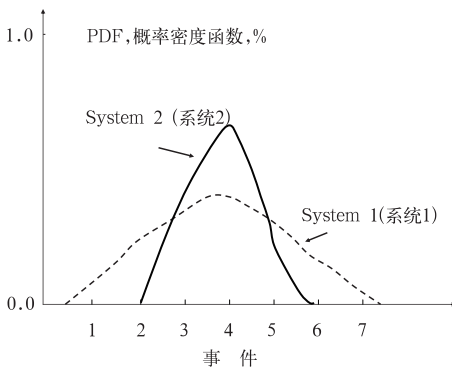


图 2 两个不同预报系统所产生的不同概率密度分布函数示意图。系统 2(实线)的比较“尖窄”,而系统 1(虚线)比较“平宽”,所以系统 2 的分辨率高于系统 1

Fig. 2 Schematic example of probability density distributions over events produced by two forecast systems where System 1 (dash line) is less sharp than System 2 (solid line)

表 1 两个独立的预报系统所给出的 5 种地面气温超过某一临界值的概率预报和相应的观测结果(超过同一临界值)
Table 1 Schematic example of probabilistic forecasts of surface air temperature exceeding a threshold over five occasions from two independent forecast systems as well as the corresponding observed events (exceeding the same thresholds)

预报系统	观测值					
	$\geq 3^{\circ}\text{C}$	$\geq 0^{\circ}\text{C}$	$\geq 10^{\circ}\text{C}$	$\geq 7^{\circ}\text{C}$	$\geq 20^{\circ}\text{C}$	平均
系统 A 预报	30%	20%	80%	30%	50%	42%
系统 B 概率	30%	40%	75%	50%	60%	51%

仅决定于集合预报技术,同时也依赖于模式和初值的质量。所以,我们在这两方面都不可倚靠。

从以前的讨论中知道(如文献[7]中的图 6),集合预报的主要目的就是定量估计预报不确定性的。所以,可预报性大的事件(容易报对,其对应的观测概率也大)其不确定性就应该小,即预报概率大;反之,可预报性小的事件(不易报对,其观测概率也小)其不确定性就大,预报概率就小。所以一个基于集合预报系统上的概率预报的可靠与否也就是反映了该集合预报系统预报不确定性或反映大气可预报性能力的好坏。注意,因为预报概率的极大或极小其实都表明了事件具有较小的不确定性,所以为了防止误解,必须注意这儿所说的“可预报性”和“预报概率大小”所对应的事件是指在实况中出现了的事件,而不是指没有出现的反事件。譬如,某次“暴雨出现”的预报概率是 80%,而另一次“暴雨出现”的预报概率为 20%。如果实况是两次都出现了暴雨,那么我们则说前一次暴雨过程较后一次暴雨过程有较大的可预报性。然而,对于其反事件“没有暴雨”而言,其可预报性的结论则刚好相反,即如果两次实况都没有出现暴雨,那么对于“没有暴雨”事件而言,后一次的预报(80%不会出现暴雨)所对应的形势则比前一次的预报(20%不会出现暴雨)所对应的形势有较大的可预报性。这就是这儿所指的“可预报性”与“预报概率”成正比的含义。另外,需指出的是,用统计的办法所产生的概率预报一般无法反映个别大气的可预报性,所以本文所说的概率预报一般都是指由集合预报产生的有动力学意义的概率预报。

2 如何应用概率预报来决策

很多人会问如果预报说 50%会下雨,那么该如何来决策呢?是信“下雨”还是信“不下雨”呢?其实

表 2 预报(决策)和实况的可能搭配情况及相应的经济损失
Table 2 Contingency table of deterministic (yes-no) forecast (and action) vs. observation and its associated economic loss resulted from the action

预报	实况	
	实况出现	实况不出现
预报“有” (采取防范措施)	命中率(HR) $Lu+C$ (较小损失)	空报率(FAR) C (花费)
预报“无” (不采取防范措施)	漏报率(MR) $L=Lu+Lp$ (较大损失)	正确否决率(CRR) 0 (无损失)

这一问题的答案取决于谁是用户,不同的用户会有截然不同的答案。对此问题过去已有一些研究^[17-21]。下面让我们来分析一下,表 2 是预报和实际可能会出现的所有搭配情况以及决策和经济损失情况。实况出现并有预防措施时(即命中,HR),经济损失有两部分:无法避免的财产损失(Lu)和采取的花费(C);而无预防措施时(漏报,MR),则有较大的财产损失 L, L 是 Lu 加上本可避免的财产损失(Lp)。当实况没有出现而采取了防范措施(空报, FAR)则有花费 C;而无防范措施(正确否决, CRR)则无任何经济损失和花费。这样当事件发生时,采取措施比不采取措施时可以减少的经济损失或说“利益”为 $(Lu + Lp) - (Lu + C) = Lp - C$ 。但采取措施所面临的可能“风险”为 C(如果实况不出现的话)。根据上面第 1 节的讨论,对于一个完全可靠的概率预报 P% 而言,采取措施而获得“利益”的可能性为 P%(实际会出现的可能性),而承担“风险”的可能性为 $(100 - P)%$ (实际不会出现的可能性),即

$$\text{利益} = (Lp - C)P\% \quad (1)$$

$$\text{风险} = C(100 - P)\% \quad (2)$$

有了这些定量的数据后,我们就可以按照经济价值(即“利益/风险”比)来得到如下的决策模式:

$$\text{利益 / 风险比} = \frac{(Lp - C)P}{C(100 - P)} \begin{cases} > 1 & \text{采取措施} \\ = 1 & \text{中性} \\ < 1 & \text{不采取措施} \end{cases} \quad (3)$$

在式(3)中, $0 \leq P \leq 100$ 。

有了决策模式后,概率 P% 就不再是“无意义”的数字了,而是对不同的用户而言有不同的实质决策含义(如图 3 所示)。图 3 表示用户 1 对气象条件依赖性较大($Lp \gg C$),当预报概率 $\geq 20\%$ 时就决定采取措施;而用户 3 对气象条件依赖性较小,所以待概率 $\geq 80\%$ 时才采取行动;用户 2 居于用户 1 和用户 3 之间,对气象条件依赖程度中等,待概率 \geq

50% 时才采取行动;而用户 4 几乎不依赖于气象条件 ($C > Lp, Lp - C < 0$),所以在任何预报概率下都不采取措施,这类用户不在本文讨论之中。可见,概率预报同单一值的决定论预报的一个根本不同处就是用户现在可以根据自身对气象条件的依赖程度做出最佳的决策而不用再听任预报员的善意“摆布”了(在单一值决定论预报中,预报员实质上替所有的用户统一作了决策而使用户无选择余地)。从决策模式(3)可以看到概率预报的“可靠性”很重要。如果预报概率不可靠,那么可能会导致决策的失误,所以集合预报系统的好坏是关键。

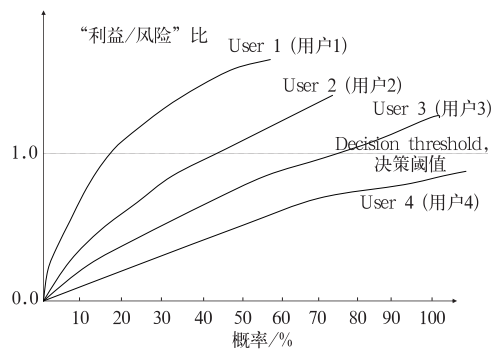


图 3 不同用户利用自身特有的“利益-风险”比来运用概率预报信息来决策的示意图。当“利益-风险”比大于 1.0(决策阈值)时,用户就采取有关的防范措施。对不同的用户而言,其相应的“临界决策概率”值是不同的

Fig. 3 Illustration of an economic value (Benefit/Risk ratio) based decision-making diagram for end-users using probabilistic weather information. An action should be taken when the ratio > 1.0 (decision threshold) to maximize user's economic value, where the corresponding "critical probability threshold for decision" varies from one user to another

3 为什么概率预报比单一值决定论预报更有价值

在承认预报有不确定性、有误差的前提下^[1],我们来举两个模拟的例子来说明多值的概率预报比单一值的决定论预报更有经济使用价值。假定结冰对某一工程的影响很大。如结冰,“可避免的损失”是 $Lp = 800$ 元,“不可避免的损失” $Lu = 200$ 元,而采取措施的“费用” $C = 100$ 元。所以“利益”是 $Lp - C = 700$ 元,其“风险”是 $C = 100$ 元。从“利益/风险”

比为 1 (公式 3), 即 $700P=100(100-P)$, 可求得其“临界决策概率”为 $P\% \approx 12.5\%$ 。假定有 10 次预报决策过程, 表 3 合理地比较了概率预报和单一值决定论预报所带来的经济损失(元)。可以想象, 单一值预报会有较多的“漏报”和“空报”; 而概率预报(尤其当临界决策概率较小时)一般不容易出现“漏报”, 但会有不少“空报”。这时利用概率预报来决策就会大大地减少经济损失。在这个例子中, 损失减

少了 30%。所以, 如果气象部门能够提供可靠的概率预报, 而所有的气象专业用户都能按照自己的行业决策模式来科学地、经济地决策, 那将对整体国民经济的保值起巨大的作用。与气象有关的自然灾害每年对人的生命和财产带来很大的损失^[22-24]。据统计目前中国每年由于气象灾害所造成的经济损失至少达整个国家国民总产值的 1%, 所以哪怕只减少 1% 的损失就是一笔可观的数字!

表 3 基于单一值的确定预报和概率预报的不同决策所导致的不同经济损失情况。

模拟的例子是针对 10 次“地面气温低于 0°C ”这一事件的预测

Table 3 Comparison of economic losses resulting from different actions taken based on deterministic and probabilistic forecasts, respectively, in the event of “surface air temperature $\leq 0^{\circ}\text{C}$ ” over 10 simulated cases

例子	预报		决策		实况与检验			经济损失/元	
	概率	单值/ $^{\circ}\text{C}$	概率($\leq 0^{\circ}\text{C}$)	单值	概率	观测/ $^{\circ}\text{C}$	单值	概率	单值
1	0	15%	防范	防范	0	命中	命中	300	300
2	-1	25%	防范	防范	0	命中	命中	300	300
3	2	20%	不防范	防范	1	正确否决	空报	0	100
4	-2	70%	防范	防范	-1	命中	命中	300	300
5	-4	100%	防范	防范	-5	命中	命中	300	300
6	1	20%	不防范	防范	0	漏报	命中	1000	300
7	1	30%	不防范	防范	2	正确否决	空报	0	100
8	3	5%	不防范	不防范	4	正确否决	正确否决	0	0
9	1	15%	不防范	防范	-1	漏报	命中	1000	300
10	0	15%	防范	防范	2	空报	空报	100	100
总计								3300	2100

下面再举一个稍为复杂一些的用户例子, 如利用风力发电。假如一个电力公司应用风的预报(可以转换为风能的预报)来计划调度电网中电源的分配。电力分配的主要任务是保证电的供与需的平衡, 否则就会停电(如供不应求)或造成浪费(如供大于求)。我们知道利用风力发电的最大挑战是风的无常性, 这使得风力发电的供电量经常随着时间变化。所以电力调度的计划, 如备用方案的选择和准备等就特显其经济价值。有些方案的准备是需要提前安排的, 如上游水库的放水到下游水库的发电可能有几天的时间差, 所以不提前采取行动是无法实施的。如果电力公司不知道风力预报的不确定性, 那么他们只能假定预报是 100% 准确的, 由此计算出来的风力发电也只能假定是对的, 而无法事先估计它会错成什么样——是高估了还是低估了? 如果高估了, 会高估多少? 如果低估了, 又会低估多少? 等等信息。所以也就无法事先实施备用方案(表 4a), 而只能待到实况出现时才没有选择余地的匆匆采取补救的应急措施: 如发现实际发电量低于计划电量时, 只能到市场上临时用高价买电; 或相反, 如

实际发电量多于计划发电量时, 就只能匆匆减少其他发电源如水力发电而使水库泄洪而浪费大量的水资源等等。但是如果电力公司事先定量的预知预报的不确定性, 知道有多少可能性计划发电被高估了或低估了, 这样他们就可以根据其出现概率的大小用上节中所讨论的“利益/风险”比(公式 3 或图 3)的行业决策模式来预选和提前启用具最佳经济价值的备用方案以作好准备(如表 4b 所示)。注意, 表 4b 所列的备案并不一定是真的反映实际情形中的最佳决策, 这里只是作为一个模拟例子来说明。类似地, 我们可以想象降水预报的不确定性对于水资源的管理和利用也是具有巨大的经济价值的。从这些例子可以看到, 对有些用户而言他们如能预先知道气象预报的不确定性对他们的风险管理是非常重要的。并且, 随着中国经济的高速发展, 气象风险管理工作就会愈显重要。

下面我们再进一步从理论检验上来更普遍地说明这一点。如有人用“相对经济价值”的方法和概念显示了集合预报所产生的概率预报比单一值预报更有价值^[25-26]。这儿先来介绍一下如何衡量一个预报

表 4 计划风力发电与备案的实施模拟例子:(a) 在不知道预报不确定性的情况下;(b) 在知道预报不确定性的情况下

Table 4 Schematic example of scheduled electrical power generated by wind and the execution of possible alternative plans with (a) no knowledge of uncertainties in wind forecasts over a 7-day period. (b) full knowledge of uncertainties in wind forecasts over a 7-day period.

(a)	时间	1 天	2 天	3 天	4 天	5 天	6 天	7 天
计划发电量/kW		1000	1500	2000	3000	3000	2500	1200
备用方案实施		无	无	无	无	无	无	无
(b)	时间	1 天	2 天	3 天	4 天	5 天	6 天	7 天
高估可能性(≥500 kW)		5%	0%	0%	0%	0%	0%	5%
高估可能性(≥100 kW)		15%	0%	5%	0%	2%	0%	25%
计划发电量/kW		1000	1500	2000	3000	3000	2500	1200
低估可能性(≥100 kW)		0%	0%	15%	5%	0%	40%	0%
低估可能性(≥500 kW)		0%	0%	7%	2%	0%	20%	0%
备用方案实施		泄洪分流	无	增加核能发电	从市场上买入 100 kW	无	调上游的水增加水力发电	在市场上卖掉 100 kW

的价值。为了简化讨论和明确意义,假定表 2 中 $Lu = 0$,即表示所有的损失都是可以避免的。假定某一特定的天气事件发生的气候概率为 \bar{o} ,那么从平均来看如果你永远都不采取防范措施的话,根据表 2 你的平均经济损失将是 $E_{\text{不防范}} = \bar{o}L$;而相反,如果永远都采取保护措施的话,那么平均经济的损失(花费)是 $E_{\text{防范}} = C$ 。所以,如果只用气候概率当作预报,用户的经济损失应该是

$$E_{\text{气候}} = \min\{C, \bar{o}L\} \tag{4}$$

即防范和不防范的决策取决于 $E_{\text{防范}}$ 与 $E_{\text{不防范}}$ 二者之中较小者。如果用户有一个 100% 准确的完美预报(注意:完美预报报“有”的概率同该事件实际发生的气候概率应是一样的),那么他只在预报“有”时才防范,而“无”时不防范,这样平均的经济损失是

$$E_{\text{完美}} = \bar{o}C \tag{5}$$

当然,实际预报并不完美,它有时对有时错。对于一个特定的预报系统而言,它的平均表现好坏是固定的,即表 2 中的“命中率” HR ,”空报率” FAR ,”漏报率” MR ,和“正确否决率” CRR 是已知的。这样对照表 2 中的预报表现率和经济损失情况,对于一个不完美的单一值确定论预报其平均经济损失是

$$E_{\text{预报}} = HR \times C + FAR \times C + MR \times L \tag{6}$$

所以 $E_{\text{气候}} - E_{\text{预报}}$ 就表示了一个实际预报系统相对于“仅仅用气候概率作预报”的价值。如果 $E_{\text{气候}} - E_{\text{预报}} < 0$ 就表示没有任何价值,即还不如用气候概率当作预报为好。而 $E_{\text{气候}} - E_{\text{完美}}$ 是一个预报系统有可能实现的最大价值。因此人们一般把一个预报系统的“相对经济价值”定义为

$$V = \frac{E_{\text{气候}} - E_{\text{预报}}}{E_{\text{气候}} - E_{\text{完美}}} \tag{7}$$

这样,一个完美预报的 $V = 1$;只用气候概率作

预报的 $V = 0$;只有当 $V > 0$ 时,预报才比“仅仅用气候值作预报”更有意义。如果把式(4)~式(6)代入式(7)的话,就得到

$$V = \frac{\min\{C, \bar{o}L\} - HR \times C - FAR \times C - MR \times L}{\min\{C, \bar{o}L\} - \bar{o}C} = \frac{\min\{\frac{C}{L}, \bar{o}\} - HR \times (\frac{C}{L}) - FAR \times (\frac{C}{L}) - MR}{\min\{\frac{C}{L}, \bar{o}\} - \bar{o} \frac{C}{L}} \tag{8}$$

如果注意到表 2 中的 $HR + MR$ 实际上就是某一事件实际发生的频率(即气候概率) \bar{o} ,那么式(8)可进一步改写为

$$V = \frac{\min\{\frac{C}{L}, \bar{o}\} - HR \times (\frac{C}{L} - 1) - FAR \times (\frac{C}{L}) - \bar{o}}{\min\{\frac{C}{L}, \bar{o}\} - \bar{o} \frac{C}{L}} \tag{9}$$

式(9)就是“相对经济价值”的通常表达方式。可见一个预报系统的相对价值同三方面有关:预报系统的好坏(命中率 HR ,空报率 FAR),所预报的对象(气候概率 \bar{o} ,即大概率事件还是小概率事件等),以及用户对气象条件的依赖程度(“花费/损失”比 C/L ,低风险用户 C/L 大,而高风险用户 C/L 小)。所以,对于一个特定的预报系统如国家气象中心 T213 模式预报系统来说,虽然预报同一天气现象如大风,那么对不同的用户有不一样的价值;而对同一用户而言,对不同的天气现象其价值又不一样。既使对于同一用户和同一种天气现象,如果预报系统变化了,那其预报价值也就变了。甚至当改变某一预报系统时,对一部分用户而言可能变得更有价值了;而对另一部分用户却可能变得有较小价值了。因此,天气预报对用户的价值是一个很复杂

的关系,我们应该多进行这方面的评估和研究,多了解接触用户以提高天气预报的总体经济效益,为更广泛的用户服务。而单看某一特定的预报检验指标是不够的,而应该用一个综合性的衡量指标如类似 UK-Index 来评价,这样才会比较全面一些。

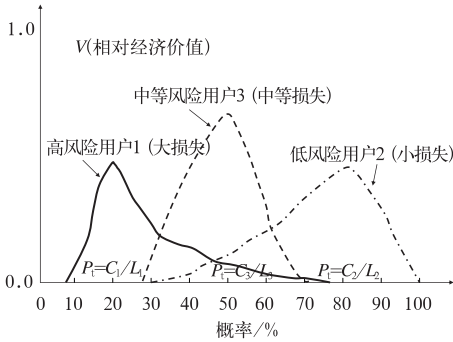


图 4 对不同的用户(对应不同的“花费-损失”比),其“相对经济价值” V (垂直轴)随概率(水平轴)变化的情况:对高(中,低)风险用户而言,其 V 值在小(中,大)概率时达到最大。通常,在概率值等于“花费-损失”比(即 $P=C/L$)时, V 达最大值

Fig. 4 Variation of relative economic value (y-axis) with probability(x-axis) for various users; the value is maximized at low (mediate, high) probability for high- (mediate-, low-) risk users. Normally, the maximum V corresponds to the probability value which is equal to that particular user's cost-loss ratio C/L (i. e., $P=C/L$)

下面再回到单一值预报和概率预报的优劣问题上。对于单一值确定论预报而言(如上面的二个例子),它没有选择余地而只能听预报的,如果预报说“有”就防范,否则就不防范。而对于概率预报,不同的用户就可以根据自己独特的花费-损失比 C/L 来选择(变化)某一事件的预报概率值(从0%到100%)来作为自己的确定预报(从而 HR 和 FAR 就相应地变化)来看相对经济价值 V 的变化。当 V 达最大时的概率值就是最适合自己的临界决策概率 P_i (图4)。一般来说当概率 $P=C/L$ 时, V 达最大值(图4)。譬如对高风险(潜在损失大)或低花费的用户(用户1)其花费/损失比 $P_i=C_1/L_1$ 较小,这样可能当概率 $P \geq 20\%$ 就当作“有”的预报;而对低风险(潜在损失小)或高花费用户(用户2)其花费/损失比 $P_i=C_2/L_2$ 很大,可能当概率 $P \geq 80\%$ 才当

作“有”的预报;对中等风险用户(用户3),其花费/损失比 $P_i=C_3/L_3$ 居中,待概率 $P \geq 50\%$ 当作“有”的预报。所以,同图3类似,用户也是根据自己的“花费/损失”比 C/L 来计算最大的相对经济价值 V 来决策。对于一个来自同一预报系统、可靠的概率预报而言,这两种方法(图3和图4)应该是一致的,它们只是从不同的角度来表达。正是由于不同的用户可以根据自己特有的 C/L 在概率分布中来挑选最适合自己的预报以达到最大的价值,所以基于集合预报的概率预报一般比单一值预报更有经济价值。图5是David Richardson根据欧洲中期数值天气预报中心(ECMWF)的集合降水预报所给出的检验结果^[25]。从该结果可以看到二点:第一,由于上述所分析的原因,基于集合预报系统上的概率预报比单一值预报(控制预报和集合平均预报)更有价值(垂直轴,其曲线位于单一值预报的曲线之上)。第二,基于集合预报系统上的概率预报相对单一值预报而言对更多的用户群有价值(水平轴,其曲线比单一值预报包含更宽的 C/L 范围)。这是因为小概率预报对高风险或低花费的用户具有价值,但这些预报却因为概率值小而往往不被单一值预报选中而被仍掉。由此可见,单一值预报对某些用户有一定价值(但不一定达到最大的价值),而对另外许多用户却没有用处。这儿又看到了我们在文献[7]中强调的观点:集合预报的主要目的是定量描述预报的不确定性、即给出预报的概率密度函数(分布)而不是给出一个更精确的单一值确定预报。因为即使一个更好更精确的预报也常常有误差,因而达不到最好的价值,且更重要的是满足不了广大的用户群特别是高风险的用户。这样就无法充分发挥集合预报的潜在优势。这儿顺便再讨论一下集合平均预报的价值。在图5中我们看到一个有趣的现象,就是虽然集合平均预报(EM)相对控制预报而言对更多的用户(主要是低风险用户)有价值,但对不少风险比较高的用户而言其价值反而不如控制预报。出现这一似乎“反常”现象的原因是由于集合平均的光滑作用。这一光滑作用对降水预报尤其明显,如文献[27]指出集合平均会使“小雨”区扩大而使“大雨”区缩小。根据上面的讨论可推知,这种平滑了的预报(缺少了峰值)通常会使低风险的用户受益而使高风险的用户受损。

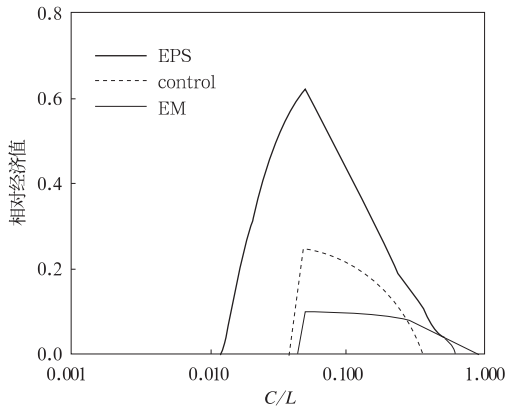


图 5 不同的预报对不同用户(不同的“花费-损失”比)的相对经济价值

粗实线(EPS)是基于集合预报的概率预报,虚线(Control)为单一值的控制预报,细实线(EM)为单一值的集合平均预报。结果是基于欧洲中期数值天气预报中心(ECMWF)全球集合预报系统所做的1999—2000年冬季欧洲的5天降水预报(超过 $10 \text{ mm} \cdot \text{d}^{-1}$)。水平轴(x轴)是“花费-损失”比(C/L),垂直坐标轴(y轴)是预报的相对经济价值(本图承蒙欧洲中心的David Richardson博士提供^[25])

Fig. 5 Relative economic values for different users (C/L) of the deterministic control (“control”, dashed line), ensemble (“EM”, thin solid line) and the ensemble forecast-based probabilistic forecasts (“EPS”, thick solid line) for precipitation event (exceeding $10 \text{ mm} \cdot \text{d}^{-1}$) over Europe at day 5 for winter 1999–2000 based on the ECMWF’s global ensemble prediction system (EPS). x-axis is C/L and y-axis the relative economic value [Courtesy of Dr. David Richardson of ECMWF^[25]]

这样,当气象用户认识到了天气预报有不可避免的不确定性^[1],并了解了概率预报实际上比单一值预报更有经济价值时,他们就会科学地制定适合自己的概率决策指标如表5。表5表示了某行业对其行为有影响的天气事件种类及相关的临界决策概率。当一个行业有了适用于自己的类似表5的决策指标以后,他们不但不会反对概率预报反而会主动地要求气象部门向其提供有关天气事件的概率预报。所以气象界需要向广大的用户进行广泛的科普宣传并同时跟他们深入合作。如何同各行各业合作建立科学的行业气象决策模式将是一项艰巨而有重要意义的工作。建议能有更多的优秀人才投身到这一创新的前沿阵地为社会作贡献。

表 5 某行业的影响天气事件种类及相应的临界决策概率
Table 5 Impacting weather types and their critical probability thresholds for decision-making yielding best possible economic value for a particular type of industry (end-user)

影响天气种类	临界决策概率/%
气温 $\leq 0 \text{ }^{\circ}\text{C}$	15
24小时累计雨量 $\geq 100 \text{ mm}$	30
地面风速 $\geq 20 \text{ m} \cdot \text{s}^{-1}$	20

虽然,上面用“花费-损失”比的经济模式可以说明并应用到许多实际情形中去,但在现实世界中并不是所有的情况都可以用数字来精确表达他们的价值。譬如人的生命就无法这样计算(除非把它当作无穷大的损失,这样无论多么小的概率都被当作是“有”的预报),以及有些人为制定的操作规则等。所以在这些情况下,就要做些变通来运用上述的原理。另外,也可用其他不同的方法来利用概率预报的信息。从上分析知道,由于用户的决策风险不一样,对预报的要求就不一样。对于高风险的专业用户,定量的概率分布非常重要,即使是位于概率分布两尾端的小概率事件也极具重大的经济价值。但对于风险小的用户和一般日常生活起居所需的广大民众,他们也许不需要非常精确的概率数字,而只需了解预报不确定性的分级或分类的语言或文字表达,如最大的可能性是A,同时也有较小的可能性会出现B等。电视天气预报也可用不同的颜色如红色、橙色、黄色、绿色、蓝色或图画等来表达预报的不确定性等级。电视气象播报员也可用语调和动作来表达预报的不确定性。如何简单易懂而有效地来表达预报的不确定性这需要大量的社会调查并同有关的社会科学工作者合作一起来制定。对于上级领导,可以坐下来面对面地向他们分析讲解预报的结果有哪几种可能性以及各种可能性的大小程度如何、并作出最好的决策(如最有可能的方案和备用方案是什么等等)。总之,对于不同的用户,要用最佳的不同方法来表达天气预报的不确定性以达到最佳的服务效果。

4 总结

本文试图用浅显易懂和举例的表达方式谈谈从单一值决定论预报向多值的概率预报转变的价值。这包括如何检验一个概率预报的好坏,如何使用概率预报来决策和为什么概率预报在实现世界中比单一值预报更有价值等。我们谈到一个好的概率预报首先应是“可靠的”,即能准确描述预报的不确定性,因为这是决策对错的基础。与此同时,也要追求高的“分辨率”(确定性的衡量)因为它会含有更多的信息,从而更有应用价值。根据用户对气象条件的依赖程度不一样,同一概率预报对不同的用户而言可能意味着截然不同的“确定预报”。譬如一个小概率预报对高风险用户而言,它就表示“有”,而对低风险用户就是“没有”。也就是说“确定预报”其实是用户的函数。具体而言,就是要运用“花费-损失”比

等经济价值的计算方法来应用概率信息来决策。正是由于用户可以根据自身对气象条件的依赖程度来选择“预报”使自身的气象决策最优化, 概率预报比单一值预报更有价值并适用于更广大的用户群, 其价值对于高风险用户尤其明显。这儿也进一步看到了我们在本文的姐妹篇^[1]中强调的, 就是集合预报的主要任务是定量描述预报的不确定性以给出可靠的概率密度函数(分布)而不是提供一个更好的单一值确定预报。譬如, 由于集合平均预报(从统计意义上来说它比较准确)的光滑作用, 反而使它对较高风险的用户有害而无利。

文章进一步指出要提高气象预报的经济效益, 建立各行各业的行业气象决策模式是关键。气象工作者应该花大力气同用户建立关系, 了解他们的需要, 同他们合作以帮助他们建立合理的行业气象决策模式。同时对于低风险用户和广大普通民众而言, 也应发展一套合适的表达方式(如语言、图表、文字、颜色等)来清楚地传达预报的不确定性, 这也是今后有待于调查研究和发展的课题。总之, 由单一值预报向多值的概率预报转变不但是气象科学本身的要求而且是提高气象预报经济价值的需要, 并有许多事需要跟上, 特别是同用户的关系。现在是该理性地思考和设计未来天气、气候和水文等预报及服务体系的时候了^[28]。气象部门应该率先向公众和用户发布有关预报不确定性的信息, 使他们有机会逐步熟悉和习惯预报的不确定性并最终在他们的决策过程中应用这些不确定性的信息。我们将在另文^[29]讨论具体的步骤和措施如何来实现这一宏伟的目标。

参考文献

- [1] Lorenz E N. 混沌的本质[M]. 刘式达, 刘式适, 严中伟译, 北京, 气象出版社, 1997.
- [2] 丑纪范. 大气科学中的非线性与复杂性[M]. 北京, 气象出版社, 2002.
- [3] 柳崇健. 复杂大气系统的性质和问题. 东亚季风和我国暴雨[M], 465—469, 北京, 气象出版社, 1998.
- [4] 薛纪善. 和预报员谈数值预报[J]. 气象, 2007, 33(8): 3-11.
- [5] 杜钧, 陈静. 天气预报的公众评价与发布形式的变革[J]. 气象, 2010, 36(1): 1-6.
- [6] U. S. National Research Council. Completing the forecast: characterizing and communicating uncertainty for better decision using weather and climate forecasts [M]. Washington DC: The National Academy Press, 2006.
- [7] 杜钧, 陈静. 单一值预报向概率预报转变的基础: 谈谈集合预报及其带来的变革[J]. 气象, 2010, 36(11): 1-11.
- [8] 杜钧. 集合预报的现状和前景[J]. 应用气象学报, 2002, 13(1): 16-28.
- [9] Kalnay E. Atmospheric modeling, data assimilation and predictability[M]. Cambridge, Cambridge University Press, 2003.
- [10] 邓国, 龚建东, 邓莲堂, 等. 国家级区域集合预报系统研发和性能检验[J]. 应用气象学报, 2010, 21(5): 513-523.
- [11] 马清, 龚建东, 李莉, 等. 超级集合预报的误差订正与集成研究[J]. 气象, 2008, 34(3): 42-48.
- [12] 谭燕, 陈德辉. 河南“75.8”大暴雨的中尺度集合预报试验[J]. 气象, 2008, 34(9): 10-21.
- [13] 杨学胜. 业务集合预报系统的现状及展望[J]. 气象, 2001, 27(6): 3-9.
- [14] 刘家峻, 张立凤, 关吉平, 等. 集合预报产品在台风麦莎预报中的应用[J]. 气象, 2010, 36(5): 21-31.
- [15] Wilks D S. Statistical methods in the atmospheric sciences [M]. New York: Academic Press, 2005, 627pp.
- [16] Murphy A H. Decision making and the value of forecasts in generalized model of the cost-loss ratio situation [J]. Mon Wea Rev, 1985, 113: 362-369.
- [17] Epstein E S and A H Murphy. Use and value of multiple-period forecasts in a dynamic model of the cost-loss ratio [J]. Mon Wea Rev, 1988, 116: 746-761.
- [18] Murphy A H. Decision-making models in the cost-loss ratio situation and measures of the value of probability forecasts [J]. Mon Wea Rev, 1976, 104: 1058-1065.
- [19] Murphy A H. The value of climatological, categorical, and probabilistic forecasts in the cost-loss situation [J]. Mon Wea Rev, 1977, 105: 803-816.
- [20] Keith R and S M Leyton. An experiment to measure the value of statistical probability forecasts for airports [J]. Weather and Forecasting, 2007, 22: 928-935.
- [21] Eckel F A, J G Cunningham, and D E Hetke. Weather and the calculated risk: exploring forecast uncertainty for operational risk management [J]. Air & Space Power Journal, 2008, 12: 71-82.
- [22] 邓国, 陈怀亮, 周玉淑. 集合预报技术在暴雨灾害风险分析中的应用 [J]. 自然灾害学报, 2006, 15(1): 115-122.
- [23] 章国材. 防御和减轻气象灾害 [J]. 气象, 2006, 32(3): 3-5.
- [24] 陈云峰, 高歌. 近20年我国气象灾害损失的初步分析[J]. 气象, 2010, 36(2): 76-80.
- [25] Richardson D S. Applications of cost-loss models [C]. Proceedings of 7th ECMWF Workshop on Meteorological Operational Systems, 2000, Reading, U. K., 209-213.
- [26] Richardson D S. Skill and relative economic value of the ECMWF ensemble prediction system [J]. Quart J Royal Meteor Soc, 2000, 126: 649-668.
- [27] Du J, S L Mullen and F Sanders. Short-range ensemble forecasting of quantitative precipitation [J]. Mon Wea Rev, 1997, 125: 2427-2459.
- [28] 叶笃正, 严中伟, 戴新刚, 等. 未来的天气气候预测体系[J]. 气象, 2006, 32(4): 3-8.
- [29] 王东海, 杜钧, 柳崇健. 正确认识和对待天气预报的不确定性[J]. 气象, 2011, 待发表