

用非线性多因子动态组合方法 作降水概率预报

邵明轩^{1,2} 刘凤辉³ 程维中⁴ 魏建明³ 窦以文³ 陈大刚³

(1 北京大学物理学院大气科学系,北京 100871; 2 中国气象局北京城市气象研究所,北京 100089;
3 北京市气象局,北京 100089; 4 苏州市气象局,苏州 215021)

摘要 将逐时段定点降水预报这个以往一般作为非连续量处理的预报问题转化为对一个连续的降水可能函数的预报问题。降水可能函数为一由实况值确定的函数。因为现在模式变化较快,且许多因子和降水可能函数为非线性关系,文章采用一种新的非线性多因子动态组合方法,通过对降水可能函数的预报,作客观分站降水概率预报。

关键词 降水概率预报 降水可能函数 动态组合方法

引言

以往在制作降水预报或降水概率预报时,因为降水量在时间分布上极不均衡,经常很长一段时间一点降水也没有,所以一般采用 0、1 化方法来制作预报方程^[1,2]。而如能使用连续量来建预报方程,就可比 0、1 化方法吸收更多的信息。为了解决这个问题,本文试验把对降水这个不连续量的预报转化为对一个连续量的预报,即对降水可能函数的预报。降水可能函数是个构造函数,由实况值确定,通过对它的预报可以达到对降水的预报。

现今数值预报模式发展很快,预报准确率也有很大提高,所以一般概率预报均是在数值预报模式基础上做出的。由于数值预报模式发展很快,也带来一个问题,模式变化很快,这给通常的 MOS 预报方法带来很大的困难^[3]。因此,预报需要采用新的方法,本文提出了一个新的非线性多因子动态组合预报方法。该方法的基本原理是各系数均为时间的函数,需用最新最近的资料确定,即利用最新最近的预报误差反馈信息来及时修正预报方程,以提高下一时刻的预报精度,克服模式变化带来的问题,同时在该方法中预报因子和预报函数的关系还可取其非

线性关系。

1 降水可能函数

逐时段降水量本质上是一个非连续量,因此逐时段降水预报是一种对非连续量的预报。能否通过转换预报对象使此问题变成一个对连续量的预报问题,这个问题的关键在于如何确定没有降水时的值。如果存在一个连续的降水可能函数,通过对此函数的预报达到对降水的预报,那么此问题就可解决。此函数应该遵守这样的条件:在有降水时,函数和此时段的降水量有一定的关系;在没有降水时,函数和此时段的降水可能性有一定的关系。

某一地点某一时段的降水量和这一点这一时段的水汽凝结量有关。大气中的水汽凝结过程就其产生的原因而言,主要可分为大尺度上升运动产生的水汽凝结和积云对流导致的水汽凝结^[4]。由大尺度上升运动引起的水汽凝结可用凝结函数法计算^[5,6],由积云对流产生的水汽凝结可用对流调整方案或对流参数化方案计算^[4]。由于大气中水汽主要集中在低层,所以低层水汽距离饱和的程度对于大尺度上升运动能否引起水汽凝结或降水具有很大作用;对于积云对流来说,积云单体起源于大气底

层并具有底层空气的属性,因此底层空气湿度对于能否产生对流凝结或对流性降水也具有非常大的作用。因而降水的可能性主要和两个条件有关:第一与大气中水汽特别是低层水汽距离凝结的程度即相对湿度大小有关,第二与垂直上升运动条件即是否有可使水汽上升冷却凝结的垂直运动有关。由于观测手段所限,各测站上空的垂直速度目前还不能直接测到,所能测到的只是某一时段的降水量、温度、湿度。考虑到降水量和垂直速度的密切关系,且如有能使水汽上升冷却凝结的垂直运动一定在低层相对湿度上有所反映,因此将某一时段内降水可能函数构造如下形式:

$$P_t = \Phi(R_t, f_t) \quad (1)$$

式中 R_t 为实际降水量, f_t 为低层相对湿度。经过反复试验,考虑到大部分测站只进行地面观测,对于 12h 有无降水可能函数 $P_{0,t}$, 定义如下:

$$P_{0,t} = \begin{cases} 50 & f_t < 50 \\ f_t & 0 \leq f_t \leq 100 \\ 100 + 5 R_t & 0 \leq R_t < 10 \\ 150 & 10 \leq R_t \end{cases} \begin{cases} \text{无降水时} \\ \\ \\ \text{有降水时} \end{cases} \quad (2)$$

f_t 为 12h 中地面最大相对湿度值(%), R_t 为 12h 的降水量(mm), $P_{0,t}$ 为无量纲指数。当没有降水时, $P_{0,t}$ 值取这一时段中最大相对湿度值;当有降水时, $P_{0,t}$ 值大于或等于 100。为了实际计算的方便, $P_{0,t}$ 最大值定义为 150, 最小值定义为 50。

例:没有降水时, f_t 为 90%, $P_{0,t}$ 的值为 90;有 3.0 mm 降水时, $P_{0,t} = 100 + 5 \times 3.0 = 115$;有 10 mm 降水或有 10 mm 以上降水时, $P_{0,t}$ 值均为 150。

不难看出,式(2)中 $P_{0,t}$ 的值在有降水时与降水量的大小正相关,在没有降水时与降水可能性的大小正相关,且降水可能函数为一连续函数。由此降水预报问题就转成预报降水可能函数问题。当预报 $P_{0,t}$ 值大于或等于 100 时,预报有降水,反之则预报无降水。同样对于有无某一数量降水的预报,也定义类似的降水可能函数,详见附录。

对于降水概率预报,则需要根据由历史资料计算出相应的降水可能函数值和降水概率函数值的关系,通过预报降水可能函数值,再计算出降水概率预报。

2 预报方法

客观降水预报目前面临的问题,既有数值预报产品本身的更新换代问题,还有如何考虑预报因子和降水可能函数的非线性问题,为此设计了一个新的动态的方法,非线性多因子动态组合方法,来预报降水可能函数。为此令降水可能函数 P_t 为以下形式:

$$P_t = A_t \cdot Y_t \quad (3)$$

式中 $A_t = [A_{1t}, A_{2t}, \dots, A_{mt}]$, 为权重向量, $Y_t = [Y_{1t}, Y_{2t}, \dots, Y_{mt}]^T$, 为预报向量, 每个预报分量 Y_{it} 为对 P_t 的一个预报。

权重向量 A_t 和预报向量 Y_t 都随时间变化而变化。预报向量 Y_t 是所选的因子 X_t 的函数:

$$Y_t = \Psi(X_t) \quad (4)$$

为了使方法更易实现,设:

$$Y_{it} = \Psi_i(X_{it}) \quad (5)$$

即 Y_{it} 只是单个预报因子 X_{it} 的函数。使用泰勒公式展开,并取其前二阶,可得:

$$Y_{it} = B_{0it} + B_{1it} X_{it} + B_{2it} X_{it}^2 \quad (6)$$

式中 $B_{0it}, B_{1it}, B_{2it}$ 也随时间变化而变化。 $B_t = [B_{1t}, B_{2t}, \dots, B_{mt}]^T$, $B_{it} = [B_{0it}, B_{1it}, B_{2it}]$, 由计算时间之前 n 个资料用最小二乘法动态确定,使得各个 Y_{it} 均为对 P_t 的一个最优化拟合。 B_{it} 也可由卡尔曼滤波方法推出^[7]。

式(3)中的权重向量因子 A_{it} 由预报分量 Y_{it} 和 P_t 的相关系数 R_{it} 及 CSI 评分动态决定。在预报方程中, A_{it} 是这样确定的,当用 Y_{it} 单因子预报 P_t , 其 CSI 值大于或等于某一临界值时:

$$A_{it}^1 = R_{it} \times R_{it} \quad (7a)$$

当其 CSI_{it} 值小于这一临界值时,

$$A_{it}^1 = 0 \quad (7b)$$

然后,

$$A_{it} = A_{it}^1 / \sum A_{it}^1 \quad (8)$$

这是为了所有权重之和为 1。资料由计算时间之前 n 个最新资料动态组成。 n 为一适当的常数,根据当地平均气候情况确定。CSI 评分由下式计算:

$$S_{CSI} = N_a / (N_a + N_b + N_c) \quad (9)$$

其中 N_a 为报对次数, N_b 为空报次数, N_c 为漏报次数^[8]。

据式(3) 计算出下一时段的 P_t 值,即降水可能函数的预报值,再据由历史资料确定的降水可能函数值和降水概率函数值的关系,作出降水概率预报。

3 试验情况

3.1 试验资料和试验时间

试验所用的预报因子取自 T106 产品和 MM5 预报产品。主要因子有 T106 的 36h 和 60h 降水预报量、850hPa 和 700hPa 相对湿度、850hPa 和 700hPa 水汽通量散度、850hPa 南风分量、500hPa 和 700hPa 垂直速度、850hPa 温度平流、K 指数、24h 变量、850hPa、700hPa 和 500hPa θ_{se} 、以及 850hPa 和 700hPa 比湿,MM5 的 12~36h 逐时段降水预报量。因子是根据积累的历史资料以及经验而确定的,选用一些物理意义明确、相关性较好的物理量预报量。所取 T106 产品为由 4 点内插和 9 点内插按一定权重确定的观象台上空插值,MM5 产品为插到各站的值。实况值用的是北京地区 15 站实测值。

预报分两个时段:当日夜间(当日 20:00 至次日 08:00,北京时,下同),次日白天(次日 08:00 至 20:00)。

预报结果在每天下午 15:15 左右由计算机自动提供给预报员,提供当日夜间和次日白天的预报。

3.2 评分方法

采用概率天气预报中最常用的 Brier 评分法^[9],其评分公式如下:

$$B = \frac{1}{N} \sum (P_i - O_i)^2 \quad (i = 1, 2, \dots, N)$$

(10)

N 为预报次数, P_i 为预报概率, O_i 为实况(当天气事件出现时取 1,不出现时取 0)。 B 的得分在 0 与 1 之间,预报完全正确得分为 0,完全不正确为 1。

3.3 试验情况

预报试验是在北京气象台实时做的,预报方程在试验运行期间一直运行良好。以下评分用的是 2002 年 3 月 1 日至 5 月 19 日共 80 天的资料。15 站当日夜间预报和次日白天预报 Brier 评分见表 1。

从评分情况看,当日夜间预报好于次日白天预报。当日夜间各站 Brier 评分在 0.095 至 0.147 之间,观象台为 0.142,15 站平均为 0.115;次日白天各站 Brier 评分在 0.141 至 0.210 之间,观象台为 0.150,15 站平均为 0.170。

表 1 15 站降水概率预报 Brier 评分

站名	夜间	白天
顺义	0.101	0.193
海淀	0.095	0.180
延庆	0.104	0.170
密云	0.147	0.169
怀柔	0.131	0.160
平谷	0.102	0.141
通县	0.121	0.159
朝阳	0.101	0.190
昌平	0.146	0.210
门头沟	0.097	0.152
观象台	0.142	0.150
石景山	0.104	0.164
丰台	0.114	0.160
大兴	0.097	0.163
房山	0.121	0.181
平均	0.115	0.170

注:表内所用资料为 2000 年 3 月 1 日至 5 月 19 日

试验表明,利用降水可能函数,用非线性多因子动态组合方法作的降水概率预报是成功的,该方法具有一定的实用性。今后还需在预报因子的选用上,不同模式产品间如何更好的搭配使用上,以及方法本身的参数如何更合理地确定上作进一步的试验和研究,以提高该方法的实用性,达到提高预报准确性这一目的。

4 结语

降水量随时间变化序列在没有降水时不管任何情况其值都为零,但降水可能函数随时间变化序列在没有降水时和有降水时一样,每一时段都有一个据实测值确定的值。降水可能函数给降水量随时间变化序列在没有降水时补充确定了一个值。试验表明,将对降水的预报转换为对连续的降水可能函数的预报是可行的,设计的降水可能函数是成功的,大部分预报是合理的。由于降水可能函数是一个连续函数,这方法为降水概率预报提供了良好的基础。

试验表明,用非线性多因子动态组合方法作降水客观预报具有一定的实用性,虽然还需用更长时期的试验来增强此方法,但从已做的试验看,此方法是有前景的。非线性多因子动态组合方法克服了通常的 MOS 预报方法对模式变化极为敏感的问题,

也是一个较好的作定点降水预报的路径。在大城市中作分区定点预报是今后预报工作中一个重点。本方法如基于准确率更高的 MM5 或其他中尺度模式的输出产品,有望作出分辨率更高的指导预报。

在用非线性多因子动态组合方法作预报计算 A_i 和 B_i 时,需根据当地平均气候变化情况,确定一个适当的 n 。 n 不能太小,太小方程会不稳定; n 也不能太大,太大可能使方程不易反映出当时的情况。试验表明, n 取 30 ~ 60 较为适当。在确定权重向量 A_i 时,临界 CSI 值不能太大,太大将使大部分 A_{it} 值为 0;也不能太小,太小会使一部分近阶段预报意义不大的预报因子也入选预报方程,使动态组合方程不稳定,达不到应有效果。试验表明,临界 CSI 值取 0.20 ~ 0.35 较为适当。

目前数值预报产品越来越丰富,如何使用好这个资源,进一步提高各气象要素的客观预报水平,是当前的一个重要课题。在确定预报因子时,非线性多因子动态组合方法克服了以往多元回归方程等常有的因子不能取太多的问题,或因子间需要严格独立不相关的要求,可以选用任意多个因子。

对于多个数值模式的集合预报问题,只要其初始场中的扰动不完全是随机产生的,非线性多因子动态组合方法的思想也可应用于其中:某个预报在最近的预报中准确率较高,其在最终预报中的权重相应较大;某个预报在最近的预报中准确率较低,其在最终预报中的权重相应较小,甚至权重为零;亦即每个预报在最终预报中的权重应随时间而动态变化。

附录

为了进一步发展用降水可能函数和非线性多因子动态组合方法作降水预报这个方法,还设计了一个对于有量降水的预报方法。

对于某一量级降水的预报,可以将降水可能函数大体分成三段,即没有降水时一段,有降水时从微量降水到该量级以下为一段,该量级以上为一段。对于 12h 有无 10 mm 以上降水的预报,有无 25 mm 以上降水的预报,也作了类似式(2)的定义:

$$P_{10,t} = \begin{cases} 40 & 0 \leq f_t < 50 \\ 0.8 f_t & 50 \leq f_t \leq 100 \\ 80 + 2 R_t & 0 \leq R_t < 10 \\ 100 + \frac{10}{3} (R_t - 10) & 10 \leq R_t < 25 \\ 150 & 25 \leq R_t \end{cases} \begin{matrix} \text{无降水时} \\ \\ \\ \text{有降水时} \end{matrix} \quad (A1)$$

$$P_{25,t} = \begin{cases} 35 & 0 \leq f_t < 50 \\ 0.7 f_t & 50 \leq f_t \leq 100 \\ 70 + R_t & 0 \leq R_t < 15 \\ 85 + 1.5 (R_t - 15) & 15 \leq R_t < 25 \\ 100 + 2 (R_t - 25) & 25 \leq R_t < 50 \\ 150 & 50 \leq R_t \end{cases} \begin{matrix} \text{无降水时} \\ \\ \\ \text{有降水时} \end{matrix} \quad (A2)$$

$P_{10,t}$ 为有无 10 mm 以上降水可能函数, $P_{25,t}$ 为有无 25 mm 以上降水可能函数。当 $R_t < 10$ mm 时, $P_{10,t} < 100$, $R_t \geq 10$ mm 时, $P_{10,t} \geq 100$; 当 $R_t < 25$ mm 时, $P_{25,t} < 100$, $R_t \geq 25$ mm 时, $P_{25,t} \geq 100$ 。

这样对于某一量级降水的预报问题也转成预报降水可能函数问题。当预报 $P_{10,t} \geq 100$ 时,预报有 10 mm 以上降水,反之则预报没有;同样,当预报 $P_{25,t} \geq 100$ 时,预报有 25 mm 以上降水,反之则预报没有。

在作有无降水概率预报同时,根据历史资料确定的降水可能函数值和降水概率函数值的关系,试验制作了北京地区 15 站有无 10 mm 以上降水、有无 25 mm 以上降水这二级降水的概率预报,试验期同 3.3 节。从试验情况看,也是当日夜间的预报好于次日白天的预报。由于试验期(3 ~ 5 月)北京地区中到大雨很少,所以对于中到大雨的预报能力还有待于进一步试验。这里主要介绍如何用降水可能函数和非线性多因子动态组合方法作降水概率预报,以及如何评分。

对于分量级的降水概率预报,仅用 Brier 评分法进行评分是不够的。例如对于实际出现了大于 25 mm 的降水,一个预报无雨概率 0.3,有 0.0 ~ 9.9 mm 降水概率 0.7,有 10 mm 以上降水概率 0.0;另一个预报无雨概率 0.0,有 0 ~ 9.9 mm 降水概率 0.4,有 10 ~ 24.9 mm 降水概率 0.6,有 25 mm 以上降水概率 0.0。按 Brier 评分这两个预报全错,得分一样,而其实明显应该是后一个预报较好。为了充分考虑这种情况,可采用有序概率得分 RPS(Ranked Probability Scores)评分法^[10]。其评分公式如下:

$$M_{RPS} = \frac{3}{2} - \frac{1}{2(K-1)} \sum_{i=1}^{K-1} \left[\left(\sum_{n=1}^i P_n \right)^2 + \left(\sum_{n=i+1}^K P_n \right)^2 \right] - \frac{1}{K-1} \sum_{i=1}^K |i-j| P_j \quad (i=1,2,\dots,K) \quad (A3)$$

其中 K 为有序概率预报类别数(例如这里的预报 $K=4$:即无雨、有 0.0 ~ 9.9 mm 降水、有 10 ~ 24.9 mm 降水、有 25 mm 以上降水共 4 类), P_i 为第 i 类天气预报概率, j 为实际出现的天气类别(例如无雨时 $j=1$)。 M_{RPS} 的得分在 0 与 1 之间,当预报完全正确得分为 1,完全不正确得分为 0。

为了综合评述试验期间各量级降水预报情况,使用有序概率得分 RPS 评分法,各站评分情况见表 A1。

表 A1 各站降水概率预报 RPS 评分

站名	夜间	白天
顺义	0.963	0.919
海淀	0.966	0.925
延庆	0.962	0.924
密云	0.948	0.924
怀柔	0.953	0.927
平谷	0.962	0.939
通县	0.957	0.932
朝阳	0.963	0.920
昌平	0.948	0.914
门头沟	0.965	0.934
观象台	0.950	0.935
石景山	0.962	0.932
丰台	0.959	0.931
大兴	0.965	0.929
房山	0.956	0.921
平均	0.959	0.927

注:表内所用资料时间同表1

当日夜间各站 RPS 评分在 0.948 ~ 0.966 之间,平均为 0.959;次日白天对于有无降水预报各站 RPS 评分在 0.914 ~ 0.939 之间,平均为 0.927。这里需要指出的是这个评分是在中至大雨很少的情况下取得的,只代表试验期间的预报结果。

致谢:感谢陶祖钰老师、吴宝俊老师和吴正华老师的热情指导和帮助

参考文献

1 周兵,雷振发.北京降水概率预报业务试验.大气科学,1998,22(3):336-342

2 Kumar A, Maini P and Singh S V. An operational model for forecasting probability of precipitation and yes/no forecast. Wea. Forecasting., 1999,14(1):38-48

3 Mao Q, McNider R T and Mueller S F, et al. An optimal model output calibration algorithm suitable for objective temperature forecasting. Wea. Forecasting., 1999,14(2):190-202

4 欧阳首承,魏绍远.形势数值天气预报及其应用.北京:气象出版社,1993.180-194

5 王德瀚,吴宝俊.暴雨分析方法.北京:气象出版社,1985.159-163

6 陶祖钰,谢安.天气过程诊断分析和实践.北京:北京大学出版社,1989.140-141

7 陆如华,徐传玉,张玲,等.卡尔曼滤波的初值计算方法及其应用.应用气象学报,1997,8(1):34-39

8 Schaefer J T. The critical success index as an indicator of warning skill. Wea. Forecasting., 1990,5(4):570-575

9 Brier G W. Verification of forecasts expressed in terms of probability. Mon. Wea. Rev., 1950,78(1):1-3

10 Murphy A H. A note on the ranked probability score. J. Appl. Meteor., 1971,10(1):155-156

Using Dynamically Combined Nonlinear Multi-Factor Equation to Forecast Probability of Precipitation

Shao Mingxuan^{1, 2} Liu Fenghui³ Cheng Weizhong⁴
Wei Jianming³ Dou Yiwen³ Chen Dagang³

(1 Department of Atmospheric Sciences, Physics College, Peking University, Beijing 100871;
2 Institute of Urban Meteorology, CMA, Beijing 100089;
3 Beijing Meteorological Bureau, Beijing 100089; 4 Suzhou Meteorological Bureau, Suzhou 215021, China)

Abstract: As the forecast of precipitation is actually considered as the forecast of a discontinuous variable, the forecast is translated into the forecast of a continuous variable—possibility function of precipitation (PFP). PFP is determined by observational values. Because the model on whose output the objective forecast was made updates frequently now, and relation between some forecast factors and PFP is nonlinear, a new method, dynamically combined equation of nonlinear multi-factors (DCENMF), is presented to forecast the probability of precipitation at stations by forecasting PFP.

Key words: probability of precipitation, possibility function of precipitation, dynamically combined equation