

过程预测定性评估方法-POV

完成人：刘凑华（liucouhua@163.com）、牛若芸（niury@cma.gov.cn）

以基于目标的降水预测检验为例介绍过程预测定性评估方法 POV（Precipitation Object Verification）。基于目标的检验，其基本思想是将降水场看作天气系统相伴随的多个降水目标组成，它将观测场和预报场中的降水目标进行识别和配对后再进行检验，以提供具有天气学含义的定量检验信息。

首先，根据区域过程事件的识别算法从降水的预报场和观测场中识别出几个降水目标之后，需要将两个场的目标进行最优的配对。两个场中的目标可以一对一的配对，也将多个观测目标合并后再与一个预报目标配对，或者反过来，可以称之为一对多的配对，类似的还可以进行多对多的配对。Gilleland 等（2008）提出目标配对中应该主要考虑一对一和一对多的配对方式，并提出了有效的计算步骤。但其算法中采用距离的倒数作为对目标匹配度的定义，这种算子只能考虑到位置因素，容易出现不合理的配对结果，为此这里给出了一种更为合理的匹配度定义如下：

$$C_{A,B} = \frac{n_A + n_B}{N_A + N_B} \quad (1)$$

其中 N_A 和 N_B 分别是目标 A 和目标 B 的总像素个数， n_A 和 n_B 是目标 A 和 B 分别处于对方临近区域的像素个数。式中分子部分为两个目标重合或临近的像素，分母为两个目标的总像。两个目标位置、形状和尺度越匹配， $C_{A,B}$ 的取值会越大。

以示意图 1 中的四个例子来说明目标的位置、形状和尺度三个因素对匹配度的影响。图 1 中 A 和 B 分别代表预报场和观测场的两个目标。例 i 中 A 和 B 两个目标相距较近，互相完全处于对方的临近区域，因此匹配度为 1。例 ii 中 A 和 B 的情形和例 i 中相似，但两者相距更远，因此两者不完全处于对方的临近区域，匹配度比例 i 小。例 iii，A 同 B 的中心位置的距离与例 i 相同，但两者形状不一致，图中表现为主轴方向不一致，则两者也不能完全处于对方的临近区域，匹配度比例 i 小。例 iv，A 同 B 的形状和中心位置的关系和例 i 中相同，只是例 iv 中 B 的尺度明显比 A 小，虽然 B 完全处于 A 的临近区域，但 A 只有部分处于 B 的临近区域，因此匹配度也小于例 i 的情况。可以看出(1)式定义的匹配度能够综合反映两个目标的位置、形状和尺度的一致性。

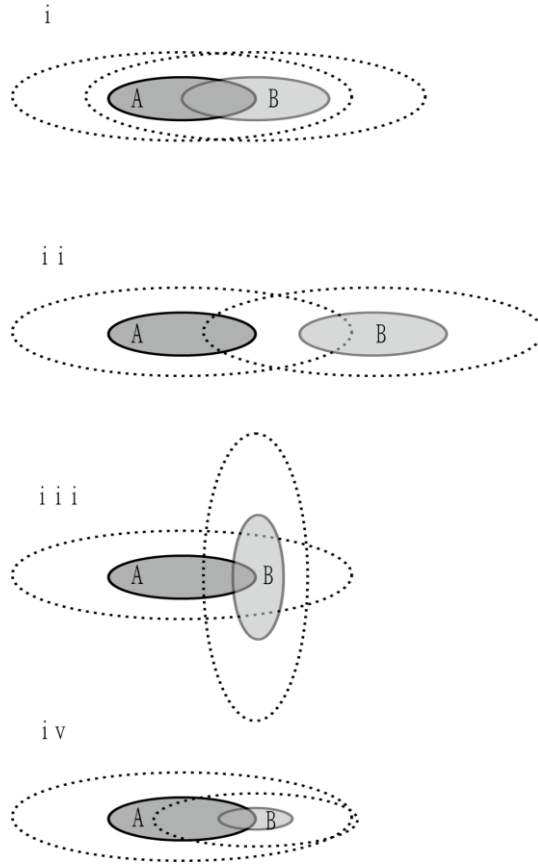


图 1 目标的位置、面积和形状对匹配度的影响示意图

在完成目标匹配之后，就可以针对每一对已经配对好的目标进行检验。一个降水目标最基本的属性包括中心位置、主轴方向、主轴的特征长度、次轴的特征长度和覆盖面积等特征。

(1) 目标的中心位置

采用矩阵 $X_{2 \times N} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_N)$ 表示预报或观测目标的点集，其中矢量 $\bar{x}_k$ 表示目标中第 k 个点的经纬度坐标。目标的中心位置为

$$\bar{\bar{x}} = \frac{1}{N} \sum_{k=1}^N \bar{x}_k \quad (2)$$

(2) 目标的主和次轴

目标中每个点相对于中心的坐标为 $\bar{d}_k = \bar{x}_k - \bar{\bar{x}}$ ，这些相对坐标构成矩阵 $D_{2 \times N} = (\bar{d}_1, \bar{d}_2, \dots, \bar{d}_N)$ ，对矩阵 $D_{2 \times N}$ 进行主分量分析，方法是计算协

方差矩阵 $S = DD^T$ 的特征值和特征向量。 S 的特征向量和特征值($\vec{v}_1, \vec{v}_2$ 和 λ_1, λ_2)即为目标的主轴和次轴的方向和特征长度。因为主轴和次轴的方向是垂直的,所以只需要计算主轴方向就足够了。实际应用中主轴方向可采用 $\vec{v}_1$ 和正东方向的夹角表示。

(3) 目标的覆盖面积

对于格点形式的模式降水目标,其面积可以通过目标的格点数来确定,考虑到不同纬度上每个格点代表的面积不同,可以用如下方式计算

$$s_f = \sum_{i=1}^N d_f^2 \cos(\varphi_i) \quad (3)$$

其中 N 为降水目标的像素点数, d_f 为格距(度), φ_i 为第 i 个像素的纬度(弧度)。

对于分布散乱的观测降水目标的覆盖面积采用如下方法计算:首先用等经纬度网格将整个区域划分成 $n \times m$ 个小区域,网格的格距设为 d_o (度),再通过如下公式计算降水目标的覆盖面积

$$s_o = \sum_{i=1}^n \sum_{j=1}^m d_o^2 \cos(\varphi_{i,j}) r_{i,j} \quad (4)$$

其中 $\varphi_{i,j}$ 为第 i 行,第 j 列网格位置的纬度,方程(4)中 $r_{i,j} = \frac{N_{i,j}^r}{N_{i,j}^t}$, $N_{i,j}^r$ 和 $N_{i,j}^t$ 分别为第 i 行,第 j 列个网格区域内目标的像素点数和总站点数。由于在一个网格内站点分布的密度可以看作是稳定的,一个目标站点和一个非目标站点代表的面积相同,因此 $r_{i,j}$ 的计算并不会受站点密度的影响,采用(4)式计算降水目标的覆盖面积能够自动适应整个区域站点分布的不均匀性,得到稳定的计算结果。在上述计算方法中,如

果一个网格区域内总站点数($N_{i,j}^t$)为 0, 则该区域 $r_{i,j}$ 的取值就无法判断, 实际计算取为 0。这样做会导致总覆盖面积的计算产生误差, 因此 d_0 不能取得过小。但如果 d_0 取得过大, 则一个网格区域站点分布可能很不均匀 $r_{i,j}$ 也就不能较为准确的反映一个网格区域内目标覆盖的面积比例。试验表明, 当格距 d_0 取为 1 到 2 度时可得到稳定的计算结果。

参考文献:

- Gilleland, E., T. C. M. Lee, J. Halley Gotway, R. G. Bullock, and B. G. Brown.
Computationally efficient spatial forecast verification using Baddeley's delta
image metric. Mon. Wea. Rev, 2008, 136: 1747-1757.
- 刘凑华, 牛若芸. 2013, 基于目标的降水检验方法及应用. 气象, 39(6): 681-690.