气象要素观测随机误差的定量估算

杨 贤 为

（北京气象中心）

提 要

本文证明由于观测随机误差的作用，采用观测资料计算的结构函数和相关函数，与要素真值的这两个函数之间存在一定的差异。通过对这种差异的理论推算并利用曲线外推法，可以估算出任何气象要素的观测随机误差和标准化观测随机误差的均方值。

一、引言

随着观测系统的不断完善和气象观测资料的广泛运用，在资料统计、数值预报、站网设计以及气候分析中，观测误差的作用愈来愈引起关注和重视[1]、[2]、[3]。众所周知，在测量某一点的气象要素时，不可避免地存在着系统误差与随机误差。当我们收集和使用气象要素的偏差时，系统误差可消去[4]，但随机误差依然存在。在具体工作中，当我们使用观测资料来进行线性内插和最佳内插时，由于观测随机误差（以下简称观测误差）的作用，往往导致内插误差的增大。同时，观测误差本身的大小又是确定气象观测代表性的重要判据之一[5]。因此，根据不同的地区、要素、季节，对观测误差进行定量分析，其意义是不言而喻的。由于观测误差无法直接计算[4]，本文介绍如何采用结构函数曲线外推法来估算观测误差的均方值（即平方的平均值，下同），并尝试以相关函数曲线外推法来估算标准化观测误差的均方值。

二、观测误差和标准化观测误差

如前所述，本文所说的观测误差，是一种随机误差，或正或负，可大可小，而且随着不同的要素、地区、季节而变。造成观测误差的因子很多，但大体上可归为这样三类：（1）观测仪器的不完善[6]；（2）人为的误差因素[6]；（3）小气候的不规则变化[7]。

设 $f'(A)$ 和 $D_f(A)$ 分别表示气象要素 f 真值在A点的偏差和方差，且以“——”表示时间序列的平均，即：

$$f'(A)=f(A)-\overline{f(A)} \tag{1}$$

$$D_f(A)=\overline{f'^2(A)} \tag{2}$$

对于任何气象要素的观测值来说，其偏差无不包含着两个分量，一个是要素 f 真值的偏差 $f'(A)$，另一个便是观测误差 $\delta_f(A)$，$\delta_f(A)$ 的量纲与 $f'(A)$ 一致，又称绝对观测误差。根据 δ_f 的随机性，可假设不同点之间的观测误差互不相关[8]：

$$\overline{\delta_f(A)\cdot\delta_f(B)}=0 \tag{3}$$

一点的观测误差与该点或其它点的偏差也不相关：

$$\overline{\delta_f(A)\cdot f'(B)}=0 \tag{4}$$

当两点重合时：

$$\overline{\delta_f(A)\cdot\delta_f(A)}=\sigma_f^2(A) \tag{5}$$

σ_f^2（A）是 A 点观测随机误差的均方值，$\sigma_f(A)$ 又称观测随机标准误差，下面要估计的就是 σ_f 的大小。

所谓标准化观测误差，系指某点观测误差与该点均方差之比，可表达为：

$$\mu_f(A)=\frac{\delta_f(A)}{\sqrt{D_f(A)}} \tag{6}$$

同样，该点真值的标准化偏差为真值偏差与均方差之比：

$$S_f(A)=\frac{f'(A)}{\sqrt{D_f(A)}} \tag{7}$$

显然，该点观测值的标准化偏差包含 $S_f(A)$ 和 $\mu_f(A)$ 两个分量。S_f 和 μ_f 都是无量纲值，μ_f 又称相对观测误差。

类似于（5）式，标准化观测误差的均方值为：

$$\overline{\mu_f(A)\cdot\mu_f(A)}=\eta_f(A)=\frac{\sigma_f^2(A)}{D_f(A)} \tag{8}$$

η_f 也是本文要估算的量值。

三、估算 σ_f^2 的结构函数外推法

某区域任意两点 f' 值之差平方的平均叫做要素 f 的结构函数[8]：

$$b_f(A,B)=\overline{[f'(A)-f'(B)]^2} \tag{9}$$

根据观测资料（其中包含观测误差）计算的结构函数为：

$$b_f'(A,B)=\overline{\{[f'(A)+\delta_f(A)]-[f'(B)+\delta_f(B)]\}^2} \tag{10}$$

展开（10）式并利用（3）、（4）、（5）式的结果，不难得到：

$$b_f'(A,B)=b_f(A,B)+\sigma_f^2(A)+\sigma_f^2(B) \tag{11}$$

如要素 f 的结构函数在该区满足均匀性和各向同性，则结构函数仅为距离的函数。所谓均匀性和各向同性，简单地说来，是指区内任何两点不论位于何处，也不论两点连线的走向如何，只要这两点的距离固定，其结构函数都相等。再假设区内各点的 σ_f^2 都相等，当A、B 的距离为d时，（11）式可变换成：

$$b_f'(d)=b_f(d)+2\sigma_f^2 \tag{12}$$

上式表明由于观测误差的作用，同一距离下的 b_f' 大于 b_f。当 $d=0$ 时，$b_f(0)=0$，代入（12）式可得：

$$\sigma_f^2=\frac{1}{2}b_f'(0) \tag{13}$$

式中 $b_f'(0)$ 表示以观测资料计算的结构函数与距离的关系曲线外推到零距离的数值。采用这种方法，便可获得 σ_f^2 的估计值了。

四、估算 η_f 的相关函数外推法

相关函数可定义为任意两点 S_f 之积的平均，即：

$$r_f(A,B)=\overline{S_f(A)\cdot S_f(B)}=\frac{\overline{f'(A)\cdot f'(B)}}{\sqrt{D_f(A)\cdot D_f(B)}} \tag{14}$$

根据观测资料计算的相关函数为：

$$r_f'(A,B)=\frac{\overline{[f'(A)+\delta_f(A)][f'(B)+\delta_f(B)]}}{\sqrt{\overline{[f'(A)+\delta_f(A)]^2}}\sqrt{\overline{[f'(B)+\delta_f(B)]^2}}} \tag{15}$$

展开上式且利用（2）—（7）式的结果：

$$r_f'(A,B)=\frac{\overline{f'(A)\cdot f'(B)}}{\sqrt{D_f(A)+\sigma_f^2(A)}\ \sqrt{D_f(B)+\sigma_f(B)}}$$

$$=\frac{r_f(A,B)}{\sqrt{1+\eta_f(B)+\eta_f(A)+\eta_f(A)\eta_f(B)}} \tag{16}$$

若要素 f 的相关函数在某区满足均匀性和各向同性的条件，则相关函数仅为距离的函数[9]。假设区内各点的 η_f 都相等，当 A、B 的距离为 d 时，(16) 式可简化成：

$$r_f'(d)=r_f(d)\frac{1}{\sqrt{1+2\eta_f+\eta_f^2}}$$

$$=\frac{r_f(d)}{1+\eta_f} \qquad (17)$$

由于 η_f 始终为正值，(17) 式表明同距离下 $r_f'(d)$ 必小于 $r_f(d)$。当 $d=0$ 时，$r_f(0)=1$，代入 (17) 式可得：

$$\eta_f=\frac{1}{r_f'(0)}-1 \qquad (18)$$

其中 $r_f'(0)$ 表示以观测资料计算的相关函数与距离的关系曲线外延到零距离的数值。据此不难获得 η_f 的估计值。由于 η_f 只是一个相对值，通过 (8) 式，还可将 η_f 变换成 σ_f^2。

五、实例

现以四川盆地的雨量资料为例，采用结构函数法和相关函数法，分别估算 7 月雨量场的观测误差与标准化观测误差的均方值。

本试验区大体位于 104—107°E，28—32°N 之间，面积约 128,000km²，7 月份的平均雨量在 150—250mm 之间。我们利用区内 49 个台站 1961—1985 年逐年 7 月的雨量资料作为计算样本。为了获得该区 7 月雨量场的结构函数与距离的关系曲线，先据 (1) 式计算出各站逐年 7 月雨量的偏差，然后以 (9) 式算得区内可能组合的每对台站间的结构函数值，并根据台站的经纬度计算出每对台站的间距，公式如下：

$$d=\rho[2-2\sin\phi_A\cdot\sin\phi_B-2\cos\phi_A\cdot\cos\phi_B\cdot\cos(\lambda_A-\lambda_B)]^{\frac{1}{2}} \qquad (19)$$

式中 d 为 A、B 两站间的距离，ϕ_A、ϕ_B 和 λ_A、λ_B 分别为 A、B 两站的纬度和经度，ρ 为地球的平均半径 (6371.11km)。严格地说，据上式算得的 d 是曲率半径为 ρ 的弧长，由于所选的区域不大，可以把 d 看作两站的直线距离。不同等级的台站间距与相应的结构函数平均值绘制的曲线，可表示该区的结构函数与距离的关系（见图1）。从中

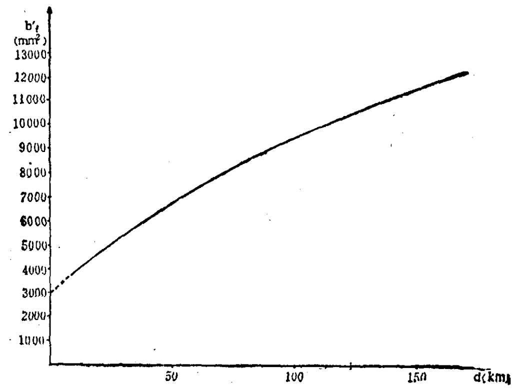

图 1　四川盆地 7 月雨量场的结构函数(b_f')与距离(d)的关系

不难看出，该曲线按趋势外延到零距离的纵坐标值为 3000mm²。也就是说，当 $d=0$ 时，$b_f'(0)=3000\text{mm}^2$，代入 (13) 式可得：

$$\sigma_f^2=\frac{1}{2}b_f'(0)=1500\text{mm}^2$$

或

$$\sigma_f\approx 39\text{mm}$$

于是可知四川盆地 7 月雨量场的观测随机标准误差是 39mm。

同样，在各站 7 月雨量偏差及方差的基础上，用 (14) 式可算得区内可能组合的每对台站之间的相关函数值，并据此绘制该区相关函数与距离的关系曲线（见图 2）。此曲线外延到零距离的相关函数值 $[r_f'(0)]$ 是 0.87，代入 (18) 式可得：

确的方程式来描述此类关系曲线，尚有待于今后进一步探讨。

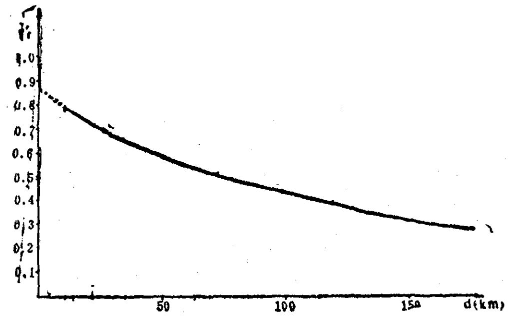

图2 四川盆地7月雨量场的相关函数(r_f')与距离(d)的关系

$$\eta_f = \frac{1}{r_f'(0)} - 1 = \frac{1}{0.87} - 1 = 0.15$$

上式表明四川盆地7月雨量场的标准化观测随机误差的均方值为0.15。

六、结语

根据以上分析可以认为，尽管观测误差不可能直接测算，却可以通过一定的数学方法估算。本文所介绍的结构函数外推法和相关函数外推法，都是在函数曲线外推到零距离的基础上进行演算的，因此，在观测站点相当密集的地区（台站间距趋近于零），可以获得比较客观的观测误差估计值。然而，曲线外推法毕竟难免含有一定的主观成分，在站点稀疏的地区尤其如此。国外的研究工作者曾试图以方程式来表达上述函数与距离的关系[10]，但效果不甚理想。至于能否用比较精确的方程式来描述此类关系曲线，尚有待于今后进一步探讨。

参 考 文 献

[1] Bergman. K. H., Role of Observational, Errors in Optimum Interpolation Analysis, Bulletin American Meteorological Society, Vol. 59, No. 12, p. 1603—1617, 1978.

[2] 廖洞贤，最优测站距离、最优垂直分层和最优观测时间间隔的决定，气象学报，43卷，1985年2期。

[3] Каган. Р. Л., Хлебникова. Е. И., Облиян иигустоты сетистанций на характеристики изменчивости интерполи рованных значений, МЕТЕОРОЛОГИЯ И ГИДРОЛОГИЯ, №. 5, 1981.

[4] Mooley. D. A., and P. M. Mohamed Ismail, Structure Functions of Rainfall Field and Their Application to Network Design in the Tropics, Arch. Met. Geoph. Biokl., Ser. B. 30, 95—105, 1982.

[5] Nappo. C. J., and Caneill. J. Y., 等，关于气象观测代表性的专题讨论会，周朝辅译，气象科技，1985年5期。

[6] 朱炳海等，气象学词典，p. 652，上海辞书出版社，1985年。

[7] Gandin. L. S., The Planning of Meteorological Station Networks, Tec. Note. III WMO 1970.

[8] 杨贤为等，江淮平原二类气象站网的设计，气象学报，45卷，1987年1期。

[9] Mooley. D. A., and P. M. Mohamed Ismail, Correlation Functions of Rainfall Field and Their Application in Network Design in the Tropics, PAGEOPH, Vol. 120, p. 249—260, 1982.

[10] Hendrick. R. L., and G. H. Comer., Space Variation and Implication of Raingauge Network Design, J. Hydrol. 10, p. 151—163, 1970.

The quantitative estimation of observational random errors of meteorological elements

Yang Xianwei

(Beijing Meteorological Center)

Abstract

In this paper, it is proved that there are some differences between the magnitudes of the structure functions or correlation functions measured from observed data and that derived from true values because of the rule of observational random errors, and the mean square of the observational random errors or the standardized observational random errors of any meteorological element can be estimated by extrapolating the curves of the two functions to the zero distances respectively.