

# 关联度分析及其在作降雨因子 等值线图中的应用

杨 艳 生

(中国科学院南京土壤研究所)

## 摘 要

灰色系统的关联分析法有较广泛的实用性。在水蚀区,为了反映土壤侵蚀能量的大小,需要绘制降雨因子等值线图,需确定降雨因子值同雨量值或地表径流量之间关系的密切程度。由于这些数值都是灰值,故采用了关联度分析法。结果表明,降雨因子值同地表径流量关系更为密切。

事物之间,同一事物的各影响因子之间都存在着复杂的关系。在一个复杂系统内,对某一研究对象有许多影响因子,每一个影响因子在不同条件或不同阶段,有不同的表现状态,这种状态有的可以采用一定的量测手段,获得不同的量值来表征它。这种量值可以有三类,一是准确的量测数值;二是一定的数值区间值;三是语言值。对于复杂系统,虽然能取得一定的量测数值,但是若用精确的量测值来表征某因子或某研究对象在不同条件下的表现状态时,常常就会同实际不符。比如说某一块地的氮素含量,假定取样很符合标准,分析误差也排除,得到的氮含量是0.15%,这样的量值就很可能得不到重复结果,为了使得同实际更加符合,人们就说这块地的氮素含量大概是0.15%,或0.15%左右。本来得到的是精确值,实际应用时常常加上“约”、“大概”、“左右”等词语,使原来的精确值加以模糊化,使得模糊化以后的量值更能符合实际情况。第二类是很难或不能得出准确的量值的,如作物产量的估测,有经验的人,估测某块地的单产为1200斤,估测本身的含义就是不那么准确,但由于估产人的经验所在,估产1200斤,依多年经验,实产总是在1100—1300斤的范围之内,这就是数值区间值。在实际生活中,区间值用得也是很多的。第三类是很难或得不出量值的,如不同生长阶段的作物生长情况可分为生长良好,生长正常,生长差。土壤肥力更是这样,肥力高,肥力中,肥力差等,这里的好、中、差,并不能用量测手段去得出某量值来表征,所以只能用语句来表述,或称它们为语言值。对于如同土壤那样的复杂研究对象,在进行描述时,语言值是用得很普遍的。

从上可见,对复杂的研究对象非精确量值较精确量值用得更普遍。经典数学是处理精确量值的,而非精确量值的处理问题要用模糊集理论和灰色系统理论来解决。本文就讨论应用灰色理论作各组数之间的关联分析及其应用举例。

## 一、数组之间的关系分析

由若干个数组成的一个数组Y与另一个数组X之间的关系,在经典数学里有比例关系;回

归关系和相关关系。相关关系实际上是反映两组精确数的线性趋势关系。如果这两组数不是精确值或不反映线性趋势,那它们间的关系就不能用相关系数去描述。灰色系统理论给出了关联度的概念和定义,为数组间的关联分析提供了方法<sup>[1]</sup>。

**1. 关联度** 关联度所考虑的对象是非精确的或区间数值数组间的关系。由于对研究对象表征并不是精确值,所以各关联度的大小顺序是主要的,而其数值大小并不重要<sup>[2]</sup>。所谓关联度,可以说是数组间关系密切的程度。例如现有某研究对象,采用不同方法进行量测,而每种方法各测出若干个值即若干个重复,这样不同方法的量测值同标准值之间就可用关联度大小来反映不同方法的精度。

**2. 关联度的数学定义** 假定现有一个母序列:  $\{X_0^{(0)}(i)\}, i=1,2,\dots,N_0$ , 并有  $N_k$  个子序列:  $\{X_k^{(0)}(j)\}, j=1,2,\dots,N, k=1,2,\dots,N_k$  这些母序列与子序列都是时间序列,因此可以将这些序列点划在座标图上(图1)。从图上可见,凡几何形状与母序列较接近者,关联度就更大。关联度可按如下定义:

实数  $r_{0k}$  满足如下条件:(1)  $r_{0k} \in \{0,1\}$ , (2) 将序列  $\{X_0^{(0)}(i)\}, \{X_k^{(0)}(j)\} i=1,2,\dots,N_0, j=1,2,\dots,N_k, N_0 \leq N_k$  在座标图上作成折线图(如图1),交于参考点  $O$ , 并将折线记为  $X_0, X_k \dots$  等,若  $X_0$  与  $X_k$  的形状越接近,  $r_{0k}$  的值就越大。(3)  $r_{0k}$  只同  $X_0$  及  $X_k$  的几何形

表1 数 组 序 列

| 数组 | 数 号   |       |       |       |       |
|----|-------|-------|-------|-------|-------|
|    | 1     | 2     | 3     | 4     | t     |
| A  | 87.9  | 209.7 | 293.0 | 150.9 | Y     |
| B  | 147.8 | 375.9 | 88.8  | 39.9  | $X_1$ |
| C  | 188.5 | 129.9 | 149.3 | 74.6  | $X_2$ |

状有关,同它们在空间的相对位置无关。或有实数  $\rho$ , 则  $\{X_0(1)\}$  与  $\{X_k(1)\}$  的  $r_{0k}(1)$ , 和序列  $\{\rho X_0(1)\}$  与  $\{\rho X_k(1)\}$  的  $r_{0k}(\rho, 1)$  应彼此相等,即

$$r_{0k}(1) = r_{0k}(\rho, 1) \quad (4) \text{ 只有当序列 } \{X_0(1)\} \text{ 与 } \{X_k(1)\} \text{ 完全重合时, } r_{0k} \text{ 才为 } 1. \text{ 则称 } r_{0k} \text{ 为 } X_k \text{ 对于 } X_0 \text{ 关于参考点 } O \text{ 在区间 } [1, N_0] \text{ 的关联度。}$$

从关联度的定义可知,关联度实质上是反映各组折线间的相似程度。比如现有四个数的A、B、C数组,其中A为母数组,B、C为子数组(表1)。由于数组的个数少,若要作相关分析意义不大。若就此四位数字计算相关系数,则得到  $r_{AB} = 0.075, r_{AC} = -0.138$ 。可见,A与B,A与C数组的相关系数达不到显著水准。但作关联度计算时,  $r_{AB} = 0.772, r_{AC} = 0.752$ 。可见,同C数组相比,B数组同A数组关系更密切。

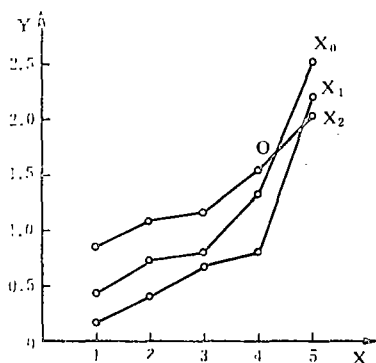


图1 折线形状与关联度的关系

## 二、关联度的求取方法

为了进行关联度分析处理的方便,把数号表示为  $t$ , A、B、C数组分别用  $Y, X_1, X_2$  表示(表1)。关联度计算步骤是:

1. 计算各数组各数值的变化速率。按下式计算:  $\Delta y(t) = y(t+1) - y(t), \Delta X_1(t) = X_1(t+1) - X_1(t)$ 。 计算结果如表2。

2. 计算  $X_1$  对于  $Y$  的关联数列  $\xi_1(t)$ 。计算公式:

$$\xi_1(t) = \frac{1}{1 + \left| \frac{\Delta X_1(t)}{X_1(t)} - \frac{\Delta Y(t)}{Y(t)} \right|}$$

计算结果列于表3。

3. 计算  $X_1$  对  $Y$  的关联度  $r$ 。计算公式:

$$r_1 = \frac{1}{N-1} \sum_{t=1}^{N-1} \xi_1(t) \quad \text{式中 } N \text{ 是 } Y$$

的数据个数。计算结果:  $r_{AB} = 0.755$ ,

$r_{AC} = 0.719$  结果说明数组  $B$  比数组  $C$  同数组  $A$  有更密切的关系。

表2 各数值的变化速率

| t                | 1      | 2      | 3      |
|------------------|--------|--------|--------|
| $\Delta Y/Y$     | 1.386  | 0.397  | -0.485 |
| $\Delta X_1/X_1$ | 1.543  | -0.764 | -0.551 |
| $\Delta X_2/X_2$ | -0.311 | 0.149  | -0.500 |

表3  $X_1$  对  $Y$  的关联数列

| t          | 1     | 2     | 3     |
|------------|-------|-------|-------|
| $\xi_1(t)$ | 0.864 | 0.463 | 0.938 |
| $\xi_2(t)$ | 0.371 | 0.801 | 0.985 |

### 三、关联度在作降雨因子等值线图中的应用

在长江三峡土壤侵蚀水土保持调查中,需了解和绘制决定土壤侵蚀力的降雨因子 ( $R$ ) 值。但是由于  $R$  值的计算是按月算出,然后再累加得出年  $R$  值。每个县的雨量站有数十年的降雨资料,各县又有上十个乃至十多个雨量站,整个三峡区又有一、二十个县,所以就不太可能利用全部降雨资料计算  $R$  值。如果能利用现有降雨量或地表径流等值线图,参考绘制  $R$  等值线图,不仅可以弥补现有资料的不足,而且还可以提高  $R$  因子等值线图的精度。但  $R$  因子值同降雨量、地表径流量都有一定关系,因此就有必要采用关联分析法来确定  $R$  因子同降雨量、地表径流量的关系密切程度差别。

1. 降雨量、地表径流量观测值的灰色性。降雨量和地表径流量都是可量测值,降雨量还可以采用自动量测仪器测出,但是由于降雨有显著区域性差别,所以一定点上取得的雨量值,即使排除观测误差的可能性,也不可能作为完全反映整个面的精确量值。地表径流量亦有类似情况。地表径流量不仅随雨量而变化,雨量是灰值,当然径流量亦为灰值,而且即使雨量是白值,由于地表径流同土壤、地形、植被等关系密切,这些因素发生变化时,也就使得地表径流产生变化,所以地表径流也是灰值。 $R$  因子值是直接根据降雨量算出的,当然也都是灰值,所以作这些量值的关系分析时,用关联度分析法就比较合理。

2. 降雨因子 ( $R$ ) 值的计算方法。 $R$  值的大小是反映降雨时不同的雨滴大小和降落速度对地面打击的动能大小,它同降雨过程中连续30分钟最大降雨强度有关,一般可按每次降雨的雨量过程线求出。这种求取方法较麻烦,尤其要求取各年的  $R$  值时,缺少各次雨量过程线或每小时的雨量记录,  $R$  值就无法计算。因此就采用陕西省农林科学院主编的联合国粮农组织专家组讲授教材《土地资源利用讲义》中介绍的魏斯曼经验公式:

$$R = \sum_{i=1}^{12} 1.735 \times 10 \exp(1.5 \times \log \frac{p_i^2}{p} - 0.8188), \text{ 其中 } p \text{ 为全年降雨量(毫米); } p_i \text{ 为月平均降雨量(毫米)。}$$

按各月雨量代入计算,累计12个月得出全年  $R$  值。

3. 以四川省万县地区云阳县龙角水文站22年的雨量(毫米)、年地表径流量(亿方),进行  $R$  值计算,一并列入表4。以表4中的  $R$  值为母序列,以其中的年雨量和年径流量为子序列,计算出的  $R$  值与年雨量的关联度  $r_1 = 0.692$ ,  $R$  值与年径流量的关联度  $r_2 = 0.925$ 。

4. 从关联度分析结果可见, R值同地表径流关系更密切, 这一结果同普通情况相符。因为R值是反映降雨时产生对地面打击的动能大小, R值较大说明雨滴对地表打击力量较大, 易

表4 云阳龙角水文站历年雨量、径流资料(毫米)

| 月<br>年 | 1    | 2    | 3     | 4     | 5     | 6     | 7     | 8     | 9     | 10    | 11    | 12   | 年雨量    | R值    | 年径流量 |
|--------|------|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|------|--------|-------|------|
| 1959   | 20.9 | 19.3 | 96.6  | 80.9  | 162.5 | 190.9 | 11.9  | 62.5  | 118.5 | 103.8 | 51.2  | 8.0  | 927.0  | 148.0 | 28.8 |
| 1960   | 3.0  | 52.0 | 53.5  | 33.3  | 153.9 | 164.9 | 147.9 | 23.3  | 123.0 | 85.5  | 51.7  | 3.6  | 895.7  | 140.8 | 33.4 |
| 1961   | 10.1 | 15.0 | 62.4  | 90.9  | 123.8 | 68.2  | 166.5 | 151.8 | 74.5  | 69.4  | 41.5  | 12.3 | 886.4  | 121.2 | 27.3 |
| 1962   | 15.1 | 16.9 | 69.2  | 128.3 | 197.3 | 108.8 | 150.5 | 209.7 | 78.1  | 124.7 | 20.6  | 33.1 | 1152.3 | 178.8 | 41.9 |
| 1963   | 0.9  | 4.5  | 84.0  | 133.5 | 249.1 | 153.8 | 189.9 | 169.4 | 238.1 | 35.7  | 114.6 | 33.8 | 1407.3 | 243.7 | 46.7 |
| 1964   | 2.2  | 6.9  | 45.0  | 97.0  | 293.8 | 103.7 | 75.1  | 100.9 | 141.8 | 115.2 | 12.6  | 16.7 | 1010.9 | 273.0 | 47.7 |
| 1965   | 25.0 | 6.8  | 25.3  | 42.2  | 98.5  | 190.1 | 221.7 | 101.5 | 182.3 | 77.8  | 42.0  | 44.7 | 1057.9 | 203.3 | 40.9 |
| 1966   | 0.6  | 20.6 | 31.2  | 80.6  | 101.3 | 47.7  | 93.3  | 177.2 | 87.6  | 150.6 | 36.9  | 10.3 | 837.9  | 132.7 | 25.8 |
| 1967   | 1.9  | 33.8 | 75.2  | 80.3  | 270.3 | 182.4 | 261.0 | 56.0  | 106.4 | 77.1  | 47.5  | 0.2  | 1192.1 | 297.7 | 54.5 |
| 1968   | 7.9  | 1.8  | 78.5  | 112.2 | 101.8 | 56.8  | 135.4 | 232.1 | 244.7 | 61.1  | 51.8  | 16.9 | 1101.0 | 238.9 | 48.7 |
| 1969   | 4.4  | 2.4  | 38.8  | 96.3  | 157.6 | 138.0 | 195.5 | 157.0 | 141.1 | 57.3  | 28.9  | 3.6  | 1020.9 | 176.5 | 45.5 |
| 1970   | 8.7  | 19.9 | 29.9  | 117.1 | 189.7 | 113.9 | 162.1 | 58.5  | 219.8 | 71.7  | 19.6  | 5.0  | 1015.9 | 206.6 | 45.3 |
| 1971   | 33.0 | 16.0 | 17.3  | 130.1 | 132.9 | 275.9 | 55.3  | 88.6  | 300.2 | 83.5  | 57.8  | 11.0 | 1201.6 | 343.2 | 40.3 |
| 1972   | 2.7  | 14.0 | 120.8 | 61.5  | 161.2 | 133.4 | 167.4 | 44.1  | 192.3 | 103.8 | 74.6  | 8.6  | 1084.4 | 162.1 | 41.3 |
| 1973   | 6.1  | 9.4  | 21.2  | 172.3 | 124.6 | 173.1 | 278.5 | 86.6  | 332.9 | 22.1  | 15.3  | 14.9 | 1257.0 | 421.9 | 53.6 |
| 1974   | 12.8 | 7.8  | 53.9  | 124.1 | 195.6 | 139.7 | 228.6 | 243.6 | 201.0 | 49.7  | 23.6  | 22.8 | 1303.2 | 262.8 | 53.8 |
| 1975   | 4.4  | 11.0 | 17.6  | 118.5 | 106.0 | 316.1 | 200.7 | 115.3 | 87.0  | 131.5 | 20.7  | 8.6  | 1137.4 | 322.7 | 50.0 |
| 1976   | 6.8  | 29.8 | 33.4  | 63.5  | 171.3 | 227.2 | 174.8 | 59.9  | 17.8  | 111.4 | 42.9  | 8.9  | 947.7  | 217.5 | 32.4 |
| 1977   | 4.6  | 24.8 | 115.7 | 169.8 | 164.6 | 197.9 | 209.7 | 95.2  | 50.4  | 116.4 | 27.6  | 37.7 | 1214.4 | 190.0 | 46.3 |
| 1978   | 19.0 | 1.6  | 39.7  | 57.3  | 144.4 | 193.1 | 126.7 | 82.0  | 122.7 | 110.1 | 52.7  | 17.8 | 967.1  | 143.5 | 32.5 |
| 1979   | 12.7 | 5.2  | 10.8  | 136.5 | 156.7 | 151.2 | 304.0 | 192.5 | 366.6 | 21.7  | 5.9   | 46.6 | 1410.4 | 469.6 | 51.9 |
| 1980   | 4.8  | 20.9 | 51.8  | 61.6  | 198.4 | 278.3 | 176.4 | 231.8 | 84.7  | 89.7  | 39.5  | 0.6  | 1238.5 | 296.5 | 55.5 |

使表层土壤结构破坏, 降低水分下渗速度, 使地表径流量迅速增加。所以如果降雨量相同, 雨强不同时, 降雨强度较大, R值也较大, 地表径流量同样较大。因此, 同降雨量相比, R值同地表径流量关系更密切是合理的。

5. 对长江三峡及其周围共15个县的范围, 共选取25个气象台站, 计算了十年至二十多年的R值, 并求出R平均值, 作为绘制R等值线图的控制点, 此外, 参考万县地区多年平均径流深度等值线图, 绘制出三峡及其周围地区的R等值线图<sup>[3]</sup>。根据各县站降雨资料, 算出的R值如表5。

应该同时指出在三峡地区, 各地R值的年际变化是很大的, 以忠县为例, 26年资料中, R最大值为554.6(1975年), 最小值为108.4(1961年)相差四倍之多。降雨量最小为913.7毫米, 最大为1390.7毫米。可见就忠县论降雨量的年差异性并不大, 但相同月份雨量的年际变化却较大。以雨季各月份为例, 最大最小月雨量年际变化为4—5倍(表6), 个别月份相差40倍之多。所以三峡地区, 月、旬、候、天降雨量变化的显著差异性, 就是造成该地区土壤侵蚀严重和危害甚大的主要原因之一。

从R值空间变化可以看出, 沿长江主航道两侧, R值为200—230, 随距长江渐远, 在长江北侧, R值稍有增加, 长寿、垫江、万县、云阳一线, R值约为250—260。在开县、巫溪一线以北, 及巫山、巴东一线以南, 有两个多雨中心, 也有两个R高值区, R值约达到280—300。

R值、降雨量和径流量之间都有密切的关系, 如果将它们都看作是精确量值时, 也可以

作相关分析,例如降雨量同R值间的相关系数为0.739;R值同径流量间的相关系数为0.721,这两数值虽然很接近,但它们间关系密切的顺序与关联分析得出的结果相反。根据分析,R值同径流量的关联程度更高更为合理。同时也能看到,尽管R值同径流量有更高的关联程度,但R值同降雨量的关联度值也很接近,从而也说明R值同降雨量亦有较高的关联性。从这里也可看出关联分析的特点和优点。由于土壤是非常复杂的系统,因而在以往的研究中,许多采用相关分析的地方,或无法采用相关分析的许多地方,当采用关联分析时,将会得到更合乎实际的结果。

表5 主要县级台站的R值

| 气象台站    | 平均R值  | 平均降雨量<br>(毫米) | 计算年数 |
|---------|-------|---------------|------|
| 开县      | 288.9 | 1230.3        | 25   |
| 垫江      | 247.9 | 1152.9        | 24   |
| 涪陵      | 214.6 | 1128.3        | 25   |
| 丰都      | 212.9 | 1084.5        | 21   |
| 石柱(二)   | 208.0 | 1069.8        | 28   |
| 秭归(陕西营) | 207.6 | 959.0         | 23   |
| 梁平      | 272.3 | 1207.1        | 38   |
| 万县(龙宝)  | 263.8 | 1188.4        | 26   |
| 云阳      | 251.3 | 1106.7        | 22   |
| 奉节      | 230.9 | 1080.9        | 29   |
| 巴东(野三关) | 297.3 | 1290.3        | 28   |
| 巫溪      | 243.4 | 1062.7        | 25   |
| 巫山      | 224.7 | 1034.6        | 29   |
| 忠县(两河)  | 233.6 | 1191.3        | 22   |
| 长寿      | 250.9 | 1214.3        | 18   |

表6 忠县26年间雨季各月最大最小雨量(毫米)

| 月 份   | 4     | 5     | 6     | 7     | 8     | 9     |
|-------|-------|-------|-------|-------|-------|-------|
| 最 大   | 265.4 | 260.9 | 430.5 | 270.1 | 301.4 | 364.5 |
| 最 小   | 53.0  | 65.5  | 68.5  | 55.1  | 7.4   | 69.1  |
| 最大/最小 | 5.0   | 4.0   | 6.3   | 4.9   | 40.7  | 5.3   |

## 参 考 文 献

- 〔1〕 邓聚龙:灰色控制系统,348-374页,华中工学院出版社,1985。
- 〔2〕 王亨:灰函数及其关联度,模糊数学,第2期,1985。
- 〔3〕 中国科学院三峡工程生态与环境科研项目领导小组编:长江三峡工程对生态与环境的影响及其对策研究论文集,498-521页,科学出版社,1987。