

降雨空间插值方法的比较与改进

石 朋, 芮孝芳

(河海大学水文水资源与水利工程科学国家重点实验室, 江苏 南京 210098)

摘要 对几种常用的降雨空间插值方法, 如泰森多边形法、距离平方倒数法、克里金方法的优缺点进行了分析和比较。考虑到高程对降雨量影响较大, 在协克里金方法的基础上将高程作为第 2 类影响因素引入降雨量的空间插值方法中, 并提出了引入高程信息的协克里金方法。将 3 种方法(距离平方倒数法、普通克里金方法、引入高程信息的协克里金方法)用于沿渡河流域降雨量的插值计算, 结果表明, 考虑高程信息的协克里金方法的插值效果明显优于其他 2 种方法。

关键词 降雨插值; 协克里金方法; 高程信息

中图分类号: P338

文献标识码: A

文章编号: 1000-1980(2005)04-0361-05

径流模拟是水文模拟研究中最基本、最重要的环节, 是研究其他水文问题的基础。而作为流域水文模型的主要输入项——降水, 其空间分布特征是影响径流模拟及一系列其他水文问题的主要控制因素。洪水预报不确定性主要来源于降雨的不确定性而非流域模型的不确定性, 对洪峰预测最敏感的因素就是降雨的空间变化^[1]。获取精确的降雨空间分布特征的一种方法就是建立密度极高的雨量站网。虽然我国已建立了相当多的遥测雨量站和气象站, 但这只是有限的观测点, 其资料密度远不能满足分布式流域水文模型的需求。因此, 有必要采用一些合适的空间插值方法。

1 常用空间插值方法

目前, 用于降雨资料空间插值的方法有多种^[2], 如泰森多边形法、距离平方倒数法、趋势面法及克里金(Kriging)方法^[3]。

1.1 泰森多边形法

该方法是一种极端的边界内插方法, 它只使用最近的单个点进行区域插值。即将采样点两两相连并作连线的中垂线, 中垂线相交形成若干个多边形, 从而将大流域分割成若干个子区域, 每个子区域中包含一个雨量站, 则区域内雨量值均用该雨量站实测值替代。泰森多边形法实际上是假设空间属性(降雨)在边界上发生突变, 在区域内均匀分布。这种假设是不符合实际的。

1.2 距离平方倒数法

这是 20 世纪 60 年代末提出的计算区域平均雨量的一种方法。实际上, 这是一种加权移动平均方法。该法将计算区域划分成若干矩形网格, 每个网格的宽度和长度分别为 ΔX 和 ΔY 。这样, 网格格点处的雨量 x_j 可用其周围邻近的雨量站实测资料按距离平方的倒数插值求得, 即

$$x_j = \frac{\sum_{i=1}^m \frac{p_i}{d_i^2}}{\sum_{i=1}^m \frac{1}{d_i^2}} \quad (1)$$

式中: x_j ——第 j 个格点处插得的降雨量; p_i ——第 j 个格点周围邻近的第 i 个雨量站的实测雨量; d_i ——第 j 个格点到其周围邻近的第 i 个雨量站的距离; m ——第 j 个格点周围邻近的雨量站的个数。距离平方倒数法的缺点是, 计算值易受数据点集的影响, 计算结果常出现一种孤立点数据明显高于周围数据点的情况。

1.3 克里金(Kriging)方法

克里金方法是建立在地质统计学基础上的一种插值方法.该方法最早由法国地理学家 Matheron 和南非矿山工程师 Krige 提出^[3],并用于矿山勘探.这种方法充分吸收了地理统计的思想,认为任何在空间连续变化的属性是非常不规则的,不能用简单的平滑函数进行模拟,只可以用随机表面函数给予恰当描述.克里金方法的关键在于权重系数的确定.该方法在插值过程中根据某种优化准则函数来动态地决定变量的数值,从而使内插函数处于最佳状态.

克里金方法的插值公式为

$$z(x_0) = \sum_{i=1}^n \lambda_i z(x_i) \quad (2)$$

式中: $z(x_0)$ —— x_0 处的估计值; $z(x_i)$ —— x_i 处的观测值; λ_i ——克里金权重系数; n ——观测点个数.

克里金方法包括普通克里金方法、泛克里金方法及协克里金方法等,最常用的是普通克里金插值方法.普通克里金方法认为,当空间变量的结构性成分确定后,剩余的差异变化属于同质变化,不同位置之间的差异仅是距离的函数.

2 引入高程信息的协克里金方法

Dirks 等^[4]发现,当站网密度较高时,普通克里金方法的插值效果与其他常用方法相比并无多大优势. Borga 等^[5]也曾得到过同样的结论.但是,随着高程的增加,降雨量有增加的趋势. Hevesi 等^[6]研究了年平均降雨量与高程的相关性,并得到了其相关系数达到 0.75 的结果.因此,本文采用协克里金方法,并将高程作为第 2 影响因素引入到对降雨量的空间插值中来.

2.1 协克里金方法

对多个具有空间相关性的空间变量进行估计的克里金方法可以归类为协克里金方法.借助这类方法,可以利用几个空间变量之间的相关性,对其中的一个变量或多个变量进行空间估计,以提高估计的精度和合理性.

设有 m 个协同的随机函数 $z_1(x), z_2(x), \dots, z_m(x)$, 则其中任意 2 个随机函数的交叉变异函数可以定义为

$$r_{ij}(x_k, x_{k'}) = \frac{1}{2} E\{ [z_i(x_k) - z_i(x_{k'})][z_j(x_k) - z_j(x_{k'})] \} \quad (3)$$

如果这些函数是二阶平稳的,则相应的交叉变异函数的表达式为

$$\begin{aligned} r_{ij}(x+h, x) &= \frac{1}{2} E\{ [z_i(x+h) - z_i(x)][z_j(x+h) - z_j(x)] \} = \\ &= \frac{1}{2} E\{ z_i(x+h)z_j(x+h) + z_i(x)z_j(x) - z_i(x+h)z_j(x) - z_i(x)z_j(x+h) \} = \\ &= \frac{1}{2} [C_{ij}(0) + C_{ij}(0) - C_{ij}(h) - C_{ij}(-h)] = \\ &= C_{ij}(0) - \frac{1}{2} [C_{ij}(h) + C_{ij}(-h)] \end{aligned} \quad (4)$$

可见, $r_{ij}(x+h, x)$ 只是 h 的函数,而与 x 无关,并且同单变量普通克里金方法的情况一样,交叉协方差函数和交叉变异函数可以相互表示.式中 $\bar{\alpha}(\ast)$ 表示交叉协方差函数.

协克里金方程组用矩阵形式可表示为

$$\begin{bmatrix} \bar{\alpha}(x_1 - x_1) & \dots & \bar{\alpha}(x_1 - x_n) & I_m \\ \bar{\alpha}(x_2 - x_1) & \dots & \bar{\alpha}(x_2 - x_n) & I_m \\ \vdots & \vdots & \vdots & \vdots \\ \bar{\alpha}(x_n - x_1) & \dots & \bar{\alpha}(x_n - x_n) & I_m \\ I_m & & & 0_m \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \bar{\mu} \end{bmatrix} = \begin{bmatrix} \bar{\alpha}(x_1 - x) \\ \bar{\alpha}(x_2 - x) \\ \vdots \\ \bar{\alpha}(x_n - x) \\ I_m \end{bmatrix} \quad (5)$$

式中 $\bar{\mu}$ 为考虑权重系数约束条件的拉格朗日算子.

虽然协克里金方法的基本原理早已为人们所熟知,但至今它仍是一个不经常使用的方法.其主要原因就

在于其符号和算法比较复杂,交叉协方差函数和交叉变异函数的求取比较困难^[7,8]。

2.2 引入高程信息的协克里金方法

当研究区域内的第 2 类信息(如高程)随处可知且变化平稳时,可将这类信息作为第 2 类影响因素引入协克里金方法。

考虑高程影响的协克里金估值方法可表示为

$$z^*(x) = \sum_{i=1}^n \lambda_i z_{ui} + \lambda [y(x) - m_Y + m_Z] \quad (6)$$

式中: $z^*(x)$ —— x 点处插得的降雨量; z_{ui} ——第 i 站的实测降雨量; $y(x)$ —— x 点处的高程; n ——实测雨量站个数; m_Y, m_Z ——高程和降雨的全局平均值; λ_i, λ ——协克里金插值权重系数。

式(6)中的插值权重系数可通过求解下列 $(n+2)$ 个线性方程组获得。

$$\begin{cases} \sum_{i=1}^n \lambda_i r_{ZZ}(x_i - x_j) + \lambda r_{ZY}(x_j - x) + \mu(x) = r_{ZZ}(x_j - x) & (j = 1, 2, \dots, n) \\ \sum_{i=1}^n \lambda_i r_{YZ}(x - x_i) + \lambda r_{YY}(0) + \mu(u) = r_{ZY}(0) \\ \sum_{i=1}^n \lambda_i + \lambda = 1 \end{cases} \quad (7)$$

式中: $\mu(x), \mu(u)$ ——考虑权重系数约束条件的拉格朗日算子; $r_{ZY}(x - x_i)$ ——位置 x 与 x_i 处第 1 类与第 2 类信息的交叉变异函数值,其计算公式为

$$\hat{r}_{ZY}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(x_i) - z(x_i+h)] [y(x_i) - y(x_i+h)] \quad (8)$$

式中 $N(h)$ 为相距 h 的采样点的配对数。

3 实例应用与比较

本文在长江三峡区间沿渡河流域上分别采用 3 种方法(距离平方倒数法、普通克里金方法、考虑高程影响的协克里金方法)进行降雨量的插值计算。三峡区间沿渡河流域面积为 601 km^2 , 共有板桥、下古、堆子、送子园、万碑、界河及沿渡河 7 个雨量站,其基本信息如表 1 所示。

分别利用距离平方倒数法、普通克里金(Kriging)方法、考虑高程的协克里金(Co-Kriging)方法对各站 1987 年各月的降雨量及 1980~1987 年的年降雨量进行了插值估算,误差统计如表 2、表 3 所示。图 1 所示为 3 种方法用于 1987 年年降雨量的插值效果。

表 1 雨量站点基本信息

Table 1 Basic information of precipitation stations

站点	位置		高程/ m
	经度	纬度	
板桥	110°09'	31°24'	1100
下古	110°14'	31°22'	730
堆子	110°16'	31°18'	200
送子园	110°24'	31°20'	1280
万碑	110°22'	31°18'	510
界河	110°11'	31°15'	780
沿渡河	110°18'	31°12'	145

表 2 3 种方法月平均降雨插值结果(选用 1987 年资料)

Table 2 Results of average monthly rainfall by three methods(1987)

月份	平均相对误差 $E_{MA}/\%$			平均绝对误差 $E_{MR}/\%$			月份	平均相对误差 $E_{MA}/\%$			平均绝对误差 $E_{MR}/\%$		
	距离法	Kriging 法	Co-Kriging 法	距离法	Kriging 法	Co-Kriging 法		距离法	Kriging 法	Co-Kriging 法	距离法	Kriging 法	Co-Kriging 法
1	13.4	9.9	9.6	2.99	2.27	1.99	7	16.3	18.3	9.8	52.20	58.70	27.93
2	29.8	30.4	16.7	1.81	1.97	1.11	8	47.6	43.0	17.5	54.03	59.63	27.27
3	10.1	11.7	9.5	4.67	5.34	4.26	9	8.8	9.2	5.4	20.60	21.94	11.81
4	10.8	10.1	10.0	5.97	5.69	5.30	10	18.8	17.5	9.6	15.94	15.07	10.96
5	9.4	8.9	3.8	12.03	11.37	4.86	11	36.4	44.1	10.9	18.36	22.44	5.34
6	4.6	4.6	3.5	13.90	12.80	10.49	12	13.8	13.8	5.8	5.91	5.56	2.77

注: $E_{MA} = \frac{1}{n} \sum_{i=1}^n \left| \frac{z(u_i) - z^*(u_i)}{z^*(u_i)} \right|$; $E_{MR} = \frac{1}{n} \sum_{i=1}^n |z(u_i) - z^*(u_i)|$ 。

表3 3种方法年平均降雨插值结果

Table 3 Statistics of average results of annual rainfall by three methods

年份	平均相对误差 $E_{MA}/\%$			平均绝对误差 $E_{MR}/\%$		
	距离法	Kriging 法	Co-Kriging 法	距离法	Kriging 法	Co-Kriging 法
1980	6.7	7.7	2.2	105.0	120.0	34.8
1981	8.1	9.6	1.8	122.6	145.6	28.0
1982	12.6	13.3	5.2	206.7	218.0	89.9
1983	8.2	8.2	1.9	158.6	162.6	38.8
1984	8.7	9.5	3.7	130.7	142.2	55.8
1985	6.7	7.9	2.6	100.1	118.3	38.4
1986	10.2	11.3	1.6	138.1	153.7	23.6
1987	8.6	9.0	1.9	129.3	131.4	29.0



图1 3种方法对于年(1987年)降雨量的空间插值效果

Fig.1 Effect of spatial interpolation of annual rainfall(1987) by three methods

为了更加直观地看到各种插值方法的效果,将板桥站的实测资料作为真值,利用其余各站资料插值推求板桥站降雨过程.图2给出了板桥站不同方法得到的1987年月平均降雨过程和1987年9月10日的一次降雨过程的插值结果.

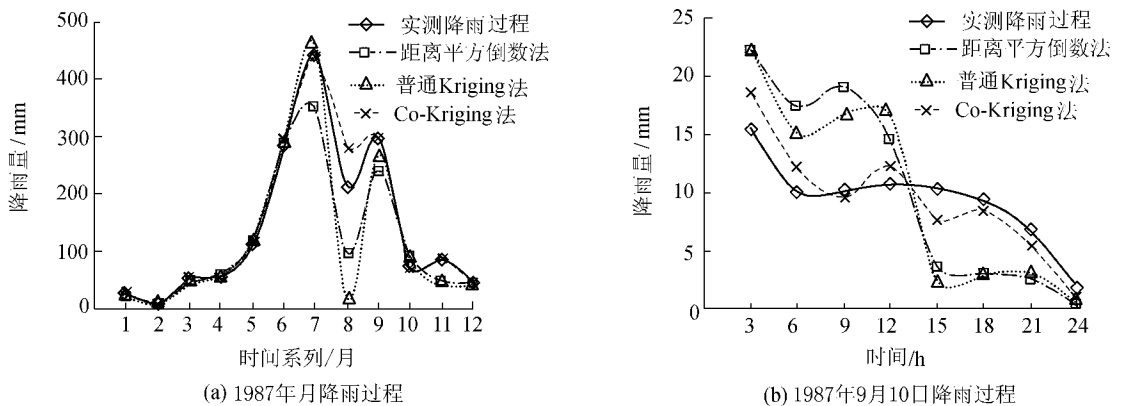


图2 3种方法对板桥站月降雨插值结果的比较

Fig.2 Comparison of results of monthly rainfall of Banqiao Station by three methods

4 结 论

- a. 当不考虑高程影响时,无论是比较传统的距离平方倒数法还是较新的普通克里金方法,都出现了较大的误差,而考虑高程影响后,用协克里金方法插值,效果明显优于前2种方法.
- b. 从表2和表3可以看出,对于年雨量和月雨量的插值,考虑高程的协克里金方法都取得了比较好的效果.从图2(b)可以看出,对于时间尺度较小的场次降雨,考虑高程的协克里金方法同样适用.这就为分布式水文模型的应用提供了较好的输入条件.
- c. 从图2(a)可以看出,对于降雨量相对较小的1月~6月,3种方法的插值效果差异不很明显,但对于

降雨量相对较大的7月~10月,距离平方倒数法出现了较大误差.2种克里金方法相比,考虑高程的协克里金方法又明显优于普通克里金方法.

d. 当采样点(雨量站)个数很少时,要想获得克里金方法中所必须的半变差函数是相当困难的.这也会在一定程度上影响插值效果.因此,有必要保证一定密度的站网分布.

e. 作为模型的输入条件,降雨资料空间分布的准确性对模型效果的影响尤其显著,因此探讨适当的降雨资料处理方法是很有必要的.还应对是否还有其他因素(如站点与海洋之间的距离,站点所在处的坡度、坡向,实时的气象条件(如风速)等)对降雨资料的空间分布造成影响,以及如何将这些可能的影响因素引入到降雨资料的处理方法中来这些问题进行探索.

参考文献:

- [1] MOHAMED A S. Reliability estimation of rainfall-runoff models[D]. New York: State University of New York, 1999.
- [2] LAM N. Spatial interpolation methods: a review[J]. The American Cartographer, 1983, 10(2): 129—149.
- [3] 侯景儒, 黄竞先. 地质统计学的理论与方法[M]. 北京: 地质出版社, 1990. 69—78.
- [4] DIRKS K N, HAY J E, STOW C D, et al. High-resolution studies of rainfall on Norfolk Island. Part II: interpolation of rainfall data[J]. J Hydrol, 1998, 208(3A): 187—193.
- [5] BORGA M, VIZZACCARO A. On the interpolation of hydrologic variables: formal equivalence of multiquadratic surface fitting and Kriging[J]. J Hydrol, 1997, 195(1—4): 160—171.
- [6] HEVESI J A, FLINT A L, ISTO J D. Precipitation estimation in mountainous terrain using multivariate geostatistics. part I: structural analysis[J]. J Appl Meteor, 1992, 31: 661—676.
- [7] 王家华, 高海余, 周叶. 克里金地质绘图技术[M]. 北京: 石油工业出版社, 1999. 157—179.
- [8] 朱会义, 贾绍凤. 降雨信息空间插值的不确定性分析[J]. 地理科学进展, 2004, 23(2): 34—42.

Comparison and improvement of spatial rainfall interpolation methods

SHI Peng, RUI Xiao-fang

(Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 210098, China)

Abstract: A comparison was made on several spatial rainfall interpolation methods commonly used, including the Thiessen polygon method, the reverse distance weighting method and the ordinary Kriging method. In consideration of great influences of the elevation information, it was taken as an influencing factor of the second class, and introduced into the Co-Kriging method for spatial rainfall interpolation, and a new Co-Kriging method was put forward. The application of the three methods, the reverse distance weighting method, the ordinary Kriging method and the new Co-Kriging method, to rainfall interpolation calculation for the Yanduhe River Basin shows that the effect of the last one is obviously superior to the other two methods.

Key words: rainfall spatial interpolation; Co-Kriging method; elevation information