文章编号:1674-7070(2010)02-0162-04

空间插值法在降水分布中的应用

秦伟良[1] 刘悦[1]

摘要

空间降水信息需求日益增加,应用于空间降水的插值方法也日益完善,不同的插值方法因为所应用的地区不同会产生不同的结果.对于不同的地区,应采用不同的方法进行比较,才能得出最佳的插值方法.采用江苏省的南京市、无锡市等12个市自1999—2008年间的1月的降雨量为数据源,分别用普通克里格插值法(OK)和反距离加权法(IDW)对同时期江苏省扬州市的降雨量进行插值,通过分析发现:普通克里格插值法的结果整体上优于反距离加权平均法.

关键词

空间插值;普通克里格法(OK);反距离加权法(IDW);江苏省

中图分类号 O241.3;P426.6

文献标志码 A

收稿日期 2010-01-10

作者简介

秦伟良,男,博士,副教授,主要研究概率论与数理统计.qinwl@nuist.edu.cn

刘悦(通讯作者),女,硕士生,主要研究方向为应用统计.yueliu20080808@163.com

1 南京信息工程大学 数理学院,南京,210044

0 引言

Introduction

经济的发展、人口的膨胀及生态环境的变化,使得社会对水资源的开发利用等一系列的问题更加关注,我国是水资源大国,对于水资源的研究显得尤为重要,降雨是水资源的重要来源之一,对工业和农业的生产建设都有着重大的意义.所谓降雨量,即从天空降落到地面上未经蒸发、渗透及流失的降水在水平面上聚集的水层深度[1].降雨具有间断性和空间不连续性的特点,而且小时降雨量零值较多,因此,与其他要素相比,降雨的空间插值具有更大的困难.国内外学者一般采用反距离加权平均法、普通克里格法、样条插值等对降水分布进行研究.李朝奎等[2]采用反距离加权平均法、普通克立格法、规则样条函数法及趋势面法等对美国爱达荷州105个气象站点及其30 a年平均降水量数据进行插值,分析了不同的插值方法中站点数量化、像元尺度变化对降雨数据空间插值结果的影响,指出要得到最理想的插值结果,须对不同研究区的实测样本数据进行分析,反复进行试验比较,从而选择最佳插值方法.鲁振宇等[3]使用反距离加权法、张力样条函数法和普通克里格法对黄河源及其周围地区65个气象站点,1990—2001年的12月平均降水数据进行空间插值,并对模拟结果进行了交叉检验和站点检验,指出普通克里格法优于其他2种方法.本文主要侧重于运用普通克里格法用江苏省的南京市、无锡市、苏州市等12个市自1999—2008年间的1月降雨量对同时期江苏省扬州市的降雨量进行插值,并与反距离加权平均法(IDW方法)进行比较,发现普通克里格插值法整体上优于反距离加权平均法.本文所选取的数据来源于2000—2009年江苏省统计年鉴.

1 反距离加权法(IDW)

Inverse Distance Weighting

反距离加权法又名空间滑动平均法[4],它是根据近邻点的平均值估计未知点的方法[5],该方法基于地理学第一定律——相似相近原理,即根据样本点周围数值随着其到样本点距离的变化而变化,并且呈现反相关,距离样本点越近,其数值和样本点的数值越近.可表示为

$$Z^*(x_0) = \sum_{i=1}^{n} \frac{1}{(D_i)^k} Z(x_i) \Big/ \sum_{i=1}^{n} \frac{1}{(D_i)^k}. \quad (1)$$

其中:$z^*(x_0)$为待估值,x_0 为观测的待估值点;$z(x_i)$为区域内位于 x_i的观测值;D_i是样本点之间的距离;n 为参与插值的样本点的个数;k 为距离的幂,显著影响着插值的结果,国内外学者一般取 $k=1$ 或$k=2$ 进行插值.

$k=2$ 时,该方法称为反距离加权平均法,简单易行,但其存在着明显的缺点:需要多大的局部邻域内的样本点个数对未知点数据进行估计是未知的;当要素场存在空间异质性或各向异性时,邻域的大小、方向和形状都会对估计产生影响;对权重系数的估计过分依赖于经验,缺乏相应的理论支持;对未知点的数据估计不能超过观测数值的值域,其结果的好坏依赖于样本点的布局[6].

2 普通克里格法(OK)
Ordinary Kriging

克里格法由南非采矿工程师 D. G. 克里格(D. G. Krige)于 1951 年首次提出,并且运用到地矿评估. 克里格法是地质统计学的核心,是一种求最优、线性、无偏内插估计量(Best Linear Unbiased Estimator,BLUE)的方法. 法国地理学家 Matheron 于 1962 年给出了 Kriging 的一般公式[7],形式如下:

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

其中:$z^*(x_0)$为估计值,是 n 个观测点的观测值的属性加权之和;$z(x_i)$为区域内位于 x_i的观测值;x_0为观测的待估值点;λ_i为待求加权系数,表示各空间样本点 x_i处的观测值 $z(x_i)$对估计值 $z^*(x_0)$的贡献值;n 为已知的观测点的总数. 选取适当的 λ_i是克里格插值法中的关键.

克里格方法有简单克里格(Simple Kriging, SK)、普通克里格(Ordinary Kriging,OK)、泛克里格(Universal Kriging,UK)等方法. 克里格法于 20 世纪 70 年代引入我国,但是由于客观条件的制约,这个方法目前在我国的应用仍不是非常广泛.

普通克里格插值法是常用的空间插值方法之一. 具体步骤如下:首先对空间场进行结构分析,建立空间变量的协方差函数,提出变异函数模型,然后再在该模型基础上进行克里格计算. 由于普通克里格插值法是根据无偏估计和方差最小原则确定 λ_i,故选取的 λ_i 要求使得 $z^*(x_0)$的估计无偏且方差 $(\sigma_\varepsilon^*)^2$ 小于任意观测值线性组合的方差.

若 $z(x)$满足二阶平稳假设,即满足以下 2 个条件:1) $z(x)$的数学期望存在且等于常数,即 $E(z(x)) = m$(常数);2) $z(x)$的协方差 $\mathrm{Cov}(x_i,x_j)$只与 2 个观测点之间的位置相关. 由于要求 $z^*(x_0)$的估计无偏,即 $E(z^*(x_0)) = E(z(x_0))$,可以推导出 $\sum_{i=1}^{n} \lambda_i = 1$. 在此条件下使估计方差最小,即:

$$\mathrm{Min}(\mathrm{var}(z^*(x_0) - z(x_0)) - 2\mu \sum (\lambda_i - 1)).$$

由此可得求解加权系数 λ_i 的方程组为

$$\begin{cases} \sum_{i=1}^{n} \lambda_i \mathrm{Cov}(x_i,x_j) - \mu = \mathrm{Cov}(x_i,x_0), \\ \sum_{i=1}^{n} \lambda_i = 1. \end{cases} \quad (3)$$

其中:$\mathrm{Cov}(x_i,x_j)$为样本之间的协方差;$\mathrm{Cov}(x_i,x_0)$为样本点与待估值点之间的协方差;μ 为极小化处理时的拉格朗日乘子.

联合式(2)、(3)求解,即可得出加权系数 λ_i,从而可以算出估计值 $z^*(x_0)$.

若用一维条件下的半变异函数 $\gamma(x_i,x_j)$表示协方差 $\mathrm{Cov}(x_i,x_j)$,则方程组(3)形式变为

$$\begin{cases} \sum_{i=1}^{n} \lambda_i \gamma(x_i,x_j) - \mu = \gamma(x_i,x_0), \\ \sum_{i=1}^{n} \lambda_i = 1. \end{cases} \quad (4)$$

其中,$\gamma(x_i,x_j)$定义为

$$\gamma(x_i,x_j) = \gamma(x_i - x_j) = \frac{1}{2} E[z(x_i) - z(x_j)]^2. \quad (5)$$

通常选取的样本是有限的,这些有限个实验样本所构成的变异函数称为试验变异函数,定义如下:

$$\gamma(x_i,x_j) = \frac{1}{2n} \sum_{i=1}^{n} \sum_{j=1}^{n} (z(x_i) - z(x_j))^2.$$

其中 n 为选取的样本点的个数.

由式(5)可以看出,半变异函数依赖于所取 2 个样本点的位置以及它们之间的距离,当半变异函数仅仅依赖于 2 样本点之间的距离,而与样本点的位置无关,即 $h = x_i - x_j$时,式(5)可改写为[5]

$$\gamma(h) = \frac{1}{2} E[z(x_i) - z(x_i + h)]^2.$$

相应的试验变异函数可改写为

$$\gamma^*(h) = \frac{1}{2n} \sum_{i=1}^{n} \sum_{j=1}^{n} (z(x_i) - z(x_i + h))^2.$$

$\gamma(h)$称为半方差图,其中半方差的上界称为梁,半方差从最低值增加到梁的距离范围称为变程,变程描述了与空间有关的差异怎样随距离而变化.

半方差函数模型中比较常见的有球状模型、指数模型、高斯模型等,这其中,以球状模型的使用最为广泛. 球状模型经过修正后公式如下[6]:

$$\begin{cases}\gamma(0)=0, & h=0;\\ \gamma(h)=c_0+c_1\left[\dfrac{3h}{2a}-(h/a)^3/2\right], & 0<h<a;\\ \gamma(h)=c_0+c, & h\geqslant a.\end{cases}$$

高斯模型经过修正后公式如下[6]:

$$\gamma(h)=c_0+c_1[1-\exp(h/a)^2].$$

指数模型经过修正后的公式为[6]

$$\gamma(h)=c_0+c_1[1-\exp(h/a)].$$

其中,块金常数 $c_0=\lim\limits_{h\to 0}\gamma(h)$,基台 $c(0)=\mathrm{var}[z(x)]=\gamma(\infty)$,拱高 $c=c(0)-c_0$.

当明显的变程和梁存在,而核方差的数值不大的时候,一般采用球状模型;当核方差很小,变程接近于钟形曲线的形状时,一般采用高斯模型;当变程和梁明显存在,而没有渐变的过程时,一般采用指数模型.

3 实证分析
Empirical analysis

在本次实例中,选取的研究区域为江苏省的南京市、无锡市、苏州市等 13 个市. 利用南京市、苏州市等 12 个城市自 1999—2008 年的 1 月降雨量对扬州市自 1999—2008 年 1 月降雨量进行空间插值分析.

首先用反距离加权法对扬州市 10 年间的 1 月降雨量进行空间插值,为使计算简便,取式(1)中的幂指数 k 为 2 进行计算.

其次,在普通克里格插值方法中,分别采用球状模型、高斯模型以及指数模型为变异函数模型进行插值,并且进行对比,对比结果见表 1.

由表 1 可以看出,对于江苏省扬州市 10 年来的降雨量插值,高斯模型的效果要优于球状模型与指数模型,而就球状模型与指数模型的比较而言,球状模型的预测效果优于指数模型.

根据表 1 的结论,采用克里格插值中效果最优的高斯模型与反距离加权平均法所得出的相对误差与绝对误差进行比较,结果如表 2.

由表 2 可以看出,在本次实例研究中,除了 1999 年和 2005 年的插值效果显示反距离加权平均法略优于普通克里格插值法中的高斯模型,对于其余 8 a 的降雨量插值,普通克里格插值的效果要明显优于反距离加权平均法.

4 结论
Conclusion

空间插值是研究区域变量空间分布的基本方法,不同的插值方法有着各自的特定假设、研究范围以及优缺点,由于所选区域的不同,各种插值方法的效果也不相同. 对于空间插值,没有绝对的最优方法,只有在特定条件下的最优方法.

通过本次研究表明,对于江苏省自 1999—2008

表 1 球状模型、高斯模型与指数模型的误差比较

Table 1 Error comparison of Spherical Model, Gauss Model and Exponential Model

年份	实测值/mm	球状模型			高斯模型			指数模型		
		预测值/mm	相对误差	绝对误差/mm	预测值/mm	相对误差	绝对误差/mm	预测值/mm	相对误差	绝对误差/mm
1999	18.4	18.938	0.029	0.538	18.031	0.020	0.369	18.996	0.032	0.596
2000	75.3	70.635	0.062	4.665	68.092	0.096	7.208	70.726	0.061	4.574
2001	117.2	107.750	0.081	9.450	107.790	0.080	9.410	107.350	0.084	9.850
2002	25.4	28.390	0.118	2.990	27.409	0.079	2.009	28.558	0.124	3.158
2003	45.8	44.002	0.039	1.798	45.093	0.015	0.707	43.778	0.044	2.022
2004	59.7	60.647	0.016	0.947	54.978	0.079	4.722	60.571	0.014	0.871
2005	27.0	31.253	0.158	4.253	30.342	0.124	3.342	31.544	0.168	4.544
2006	109.8	123.030	0.120	13.230	116.280	0.059	6.480	122.520	0.116	12.720
2007	7.6	8.906	0.172	1.306	8.936	0.176	1.336	9.306	0.224	1.706
2008	85.8	92.266	0.075	6.466	92.716	0.081	6.916	92.899	0.083	7.099

表 2　反距离加权平均法与普通克里格法误差比较

Table 2　Error comparison of Inverse distance weighting and ordinary Kriging

年份	实测值/mm	反距离加权平均法(IDW)			普通克里格法(OK)		
		预测值/mm	相对误差	绝对误差/mm	预测值/mm	相对误差	绝对误差/mm
1999	18.4	18.725	0.018	0.325	18.996	0.032	0.596
2000	75.3	66.108	0.122	9.192	70.726	0.061	4.574
2001	117.2	104.500	0.108	12.700	107.350	0.084	9.850
2002	25.4	21.800	0.142	3.600	28.558	0.124	3.158
2003	45.8	34.258	0.252	11.542	43.778	0.044	2.022
2004	59.7	47.250	0.208	12.450	60.571	0.014	0.871
2005	27.0	30.749	0.139	3.749	31.544	0.168	4.544
2006	109.8	86.017	0.217	23.783	122.520	0.116	12.720
2007	7.6	18.975	1.497	11.375	9.306	0.224	1.706
2008	85.8	76.167	0.112	9.633	92.899	0.083	7.099

年 1 月的降水分布,普通克里格插值法要整体优于反距离加权平均法,而对于特定的年份以及所选取模型的差异,反距离加权平均法是优于普通克里格插值的.本文采用的空间插值方法,只考虑了观测点的位置以及降水信息,这是不够的,还应该考虑到地形、地理因素等对降水分布的影响,从而找出影响因子,得到更精确的插值结果.

参考文献
References

[1] 杨振东.降雨量的观测计算[J].甘肃农业科技,1980(2):28
YANG Zhendong. Observation calculation of precipitation[J]. Gansu Agricultural Science and Technology,1980(2):28

[2] 李朝奎,陈良,王勇.降雨量分布的空间插值方法研究——以美国爱达荷州为例[J].矿产与地质,2007,21(6):684-687
LI Chaokui,CHEN Liang,WANG Yong. Research on spatial interpolation of rainfall distribution:A case study of Idaho State in the USA[J]. Mineral Resources and Geology,2007,21(6):684-687

[3] 鲁振宇,杨太保,郭万钦.降水空间插值方法应用研究——以黄河源区为例[J].兰州大学学报:自然科学版,2006,42(4):11-14
LU Zhenyu,YANG Taibao,GUO Wanqin. Application of the spatial interpolation of rainfall:A case study of the headstream region of the Yellow River[J]. Journal of Lanzhou University:Natural Sciences,2006,42(4):11-14

[4] Richard F. Scattered data interpolation:Tests of some methods[J]. Mathematics of Computation,1982,38:181-200

[5] Zoubeida K B,Afef C. Comparison of two kriging interpolation methods applied to spatiotemporal rainfall[J]. Journal of Hydrology,2009,365(1/2):56-73

[6] 王远飞,何洪林.空间数据分析方法[M].北京:科学出版社,2007
WANG Yuanfei,HE Honglin. Analysis of spatial data[M]. Beijing:Science Press,2007

[7] Oliver M A. Kriging:A method of interpolation for geographical information systems[J]. International Journal of Geographical Information Science,1990,49(4):313-332

Application of spatial interpolation in rainfall distribution analysis

QIN Weiliang[1]　LIU Yue[1]

1　College of Mathmatics and Physics,Nanjing University of Information Science & Technology,Nanjing　210044

Abstract　With the increasing need of spatial precipitation information,spatial interpolation method of precipitation is widely used. Different interpolation methods applied in different areas will produce different results. The best interpolation method should be obtained by comparing different interpolation methods in various areas. Using precipitation data of January from 1999 to 2008 of 12 cities of Jiangsu Province,like Nanjing,Wuxi etc.,we interpolate the precipitation data of Yangzhou of the same period with Ordinary Kriging and Inverse Distance Weighting. The result showed that Ordinary Kriging method is more suitable than Inverse Distance Weighting in the whole analysis.

Key words　spatial interpolation;Ordinary Kriging(OK);Inverse Distance Weighting(IDW);Jiangsu Province