

区域水土流失研究中降雨量空间插值方法选取

汪翠英^{1,2}, 杨勤科^{1,2}, 郭伟玲^{1,2}, 姚志宏^{1,2}, 李文凤^{1,2}

(1. 中国科学院 水利部 水土保持研究所, 陕西 杨陵 712100; 2. 西北农林科技大学 资源环境学院, 陕西 杨陵 712100)

摘 要:在区域水土流失模型研究中, 空间插值可提供每个计算栅格的气象要素资料。考虑到研究区域降雨与高程相关性很弱, 不宜采用梯度距离反比法(GIDS), 故采用距离反比法(IDW)和普通克里格法(Kriging), 对延安示范区及其周围共50个站点2000—2003年的5—10月逐月降雨量进行插值。交叉验证结果表明: 对2种插值方法, 二者经过对数变换后平均相对误差(MRE)为8.30%和7.67%, 分别比原始数据插值后的MRE下降了23.17%和23.50%, 说明插值精度得到了提升, 对研究区域某一年逐月降水的插值Kriging方法比IDW方法更加精确。

关键词:普通克里格法; 距离反比法; 降雨

中图分类号:S157

文献标识码:A

文章编号:1005-3409(2008)02-0088-04

Selection of the Spatial Interpolation Methods for Precipitation in the Regional Soil Erosion

WANG Cui-ying², YANG Qin-ke^{1,2}, GUO Wei-ling^{1,2}, YAO Zhi-hong^{1,2}, LI Wen-feng^{1,2}

(1. Institute of Soil and Water Conservation, Chinese Academy of Sciences And Ministry of Water Resources, Yangling, Shaanxi 712100, China; 2. College of Resources and Environment, Northwestern A & F University, Yangling, Shaanxi 712100, China)

Abstract: Every cell's climatic information can be used by interpolation method when researching district soil erosion model, it is important to compare the results obtained using alternative methods applied to each data set. In this study, two methods for spatial interpolation of precipitation data were compared using 2003 year six months precipitation from the Yan'an demonstration area was interpolated using inverse distance weighted (IDW) and Universal Kriging methods. Corresponding cross-validation test show that a log-transformation improves the accuracy, by a log-transformation the value of MRE is 8.30% and 7.67% in the methods of IDW and Kriging respectively that is lower 23.17% and 23.50% than the result of original data. In addition, the method of GIDS can not be used in the study. Universal Kriging method is more accurate in predictions than IDW in this area.

Key words: inverse distance weighted; Universal Kriging; precipitation

水土流失的区域研究以多种水土流失因子综合分析运算为基础, 降雨量作为一个重要的输入数据, 其精确程度直接影响模型评价或其他评价方法的进行。在研究区域水土流失模型中^[1-2], 为反映降雨量及其空间分布对区域土壤侵蚀产沙的影响, 需要利用雨量站的资料进行网格化的空间插值分析。而实际中气象站点布设有很大的局限性, 分布也不一定均匀, 绝大多数空间位置上的数据无法获得, 所以需要通过试验选择合适空间插值的方法。区域尺度水土流失的研究中, 在降雨插值的时间尺度上要进行月雨量插值生成月降雨量表面, 作为模型的一个重要输入参数。所以要选择一个适合月降雨量插值的方法, 为模型的进一步研究奠定良好的基础。

目前, 进行降雨数据内插的方法有多种, 主要有泰森多边形法、梯度距离反比法(GIDS)、PRISM(Parameter-ele-

vation Regressions on Independent Slopes Model)、反距离权重法(IDW)、克里格法(Kriging)、样条插值法等^[3-4]。泰森多边形法是一种极端的边界内插方法, 其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。GIDS要考虑站点气象要素值与海拔高程之间的相关性^[5], PRISM插值方法以数字高程模型为基础, 要综合考虑高程、坡度、坡向、距离、垂直分层、距离海岸线的远近及有效地形等权重因子对气象要素的影响^[6], 在研究区气象要素与高程之间的相关性较弱则2种方法难以很好的应用^[7-8]。距离反比法由美国国家气象局于1972年提出, 其算法较为简单, 应用很广泛^[9-10]。克里格法由南非地质学家Krig提出, 其实质是最佳线性无偏估计, 应用广泛。该方法比较复杂, 受人为主观因素的影响^[11-14]。样条插值法是用多项式拟合的方法来产生平滑的插值曲线, 该方法

收稿日期: 2007-10-30

基金项目: 国家重点基础研究发展计划“973计划”项目(2007CB407204), “十一五”国家科技支撑计划项目(2006BAD09B05); 黄委会治黄专项(2004SZ01-04)

作者简介: 汪翠英(1982-), 女, 在读硕士, 主要从事GIS在水土保持中的应用研究。

通信作者: 杨勤科(1962-), 男, 研究员, 博士, 主要从事区域水土流失和GIS应用研究。E-mail: qkyang@ms.iswc.ac.cn

适用于渐变的表面属性,如高程、水深、污染聚集度等,所以多用于气象要素的时间序列插值^[15]。对降雨量空间插值各学者以自己的研究需要有不同的选取方法,该文旨在根据逐月降雨的特点,基于 ARC GIS 软件中的地统计分析模块进行月降雨量的空间插值,从而得到一个比较精确的月降雨量表面。

1 资料与方法

1.1 数据来源与分析方法

以黄土高原延安示范区市内 53 个降雨量站测得的降雨数据为样本。该区位于 N36.85°—36.47°,E109.20°—109.58°(图 1),海拔 944~1 330 m。数据处理中对无降雨数据记录的站点予以剔除,最终选取 50 个站点的 2000—2003 年的 5—10 月的降雨数据进行插值。数字高程模型(DEM)选用国家基础地理信息中心制作的 1:5 万地形图生成的 25 m 分辨率 DEM。

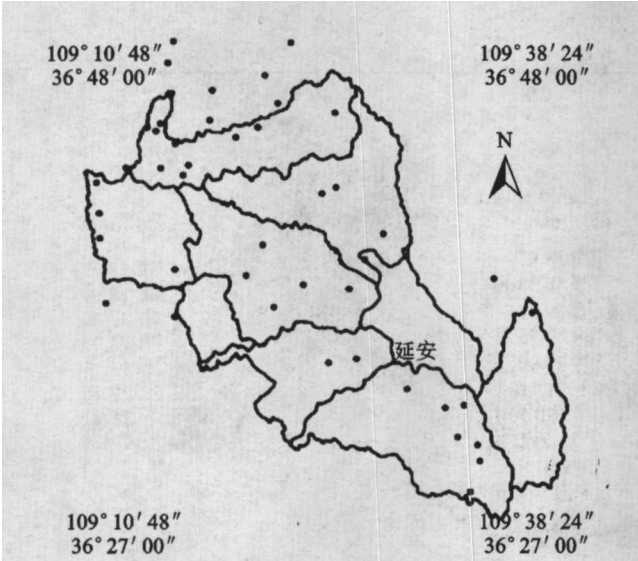


图 1 降雨站点分布

在大区域研究中如果气象要素与高程之间的相关性高就可以用 GIDS 方法,反之,此方法很难应用,由于本区雨量站点高程差异不大,降水与高程的关系较弱^[8](表 1),相关性最高的也只有 7 月,相关系数仅 0.07,因此在降雨插值时不宜采取此方法。

表 1 测站 2000—2003 年的 5—10 月平均降雨与测站高程相关系数

月份	5	6	7	8	9	10
相关系数	0.01	0.03	0.07	0.02	0.03	0.03

对降雨进行数据变换会使数据更趋于正态分布并且会提高预测的精度^[16-17]。研究中运用对数变换和立方根变换方式对数据进行预处理,用偏态(Skewness)来描述数据分布左右对称性,即对数据进行了变化方式的偏斜度优化选择,正偏斜度表示不对称部分的分布更趋向正值;负偏斜度表示不对称部分的分布更趋向负值。结果表明:原始数据 8 月、9 月和 10 月更趋于正态分布;5 月和 6 月数据经对数变换后

趋于正态分布;7 月降水数据经立方根变换后更趋于正态分布。在整个分析过程中,对降雨均采用了原始数据,对数变换和立方根变换三套数据,并对插值结果进行了对比分析。

表 2 降雨数据变换偏斜性分析

月份	5	6	7	8	9	10
原数据	1.864	1.346	1.664	-0.504	-0.428	0.258
对数变换	0.844	-0.242	-0.850	-1.661	-1.387	-1.183
立方根变换	1.263	0.366	0.001	-1.214	-2.885	-1.866

1.2 空间插值方法

1.2.1 反距离权重法

IDW 是一种常用而简便的空间插值方法,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。它以插值点与降雨量站点的距离为权重进行加权平均,离插值点越近的样本点赋予的权重越大。

$$z(s) = \left[\sum_{i=1}^n \frac{z_i}{d_i^2} \right] / \left[\sum_{i=1}^n \frac{1}{d_i^2} \right]$$
 (1)

式中: $z(s)$ —— s 处插得的降雨量; n ——预测计算过程中使用的雨量站个数; z_i ——第 i 个格点周围邻近的第 i 个雨量站实测雨量; d_i ——插值点到已知雨量站点的距离。

1.2.2 克里格法

Kriging 方法是一种地统计学方法,是一种最优、线性、无偏估计方法,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。

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

式中: $z(x_0)$ —— x_0 处的降雨量估计值; $z(x_i)$ —— x_i 处的观测值; λ_i ——克里金权重系数; n ——观测点个数。

其中 λ_i 是与降雨观测点有关的权重

$$\begin{cases} \sum_{i=1}^n \lambda_i \gamma(x_i, x_j) + \psi = \gamma(x_i, x_0) \\ \sum_{i=1}^n \lambda_i = 1 \end{cases}$$
 (3)

式中: $\gamma(x_i, x_j)$ ——两测点间的半方差; $\gamma(x_i, x_0)$ ——测点与未测点间的半方差; ψ ——拉格朗日乘子。

气象要素空间结构特性由半变异函数描述,其表达式为

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(x_i) - z(x_i + h)]^2$$
 (4)

式中: $N(h)$ ——被距离区段 h 分割的试验数据对数目; $z(x_i), z(x_i + h)$ ——区域化随机变量 Z 在空间位置 x_i 和 $x_i + h$ 的值; $\gamma(h)$ ——变异函数值(又称为半方差, Semivariance)。

在讨论中,根据试验半变异函数得到的试验变异函数图,其理论模型选用球形函数模型。

1.3 检验数据标准

采用交叉验证法(cross-validation)来验证插值的效果。首先假设每一雨量站点的气象要素值未知,都用周围站点的值来估算,然后计算所有站点的实际观测值与估算值的误差,以此来评判估值方法的优劣。研究过程中对 50 个站点的插值结果进行全交叉验证。运用相对平均误差(MRE)及均方差(RMSE)作为检验的标准。相对平均误差可以反映

插值的相对精确性,均方差可以反映利用数据估值灵敏度和极值效应。

2 结果分析

2.1 数据变换结果分析

由国内外的插值经验,对样本点数据进行变换通常可以提高插值精度,但从数据变换结果看出,这并不是绝对的,不同季节、不同插值方法,数据变换的插值效果是不一样的。

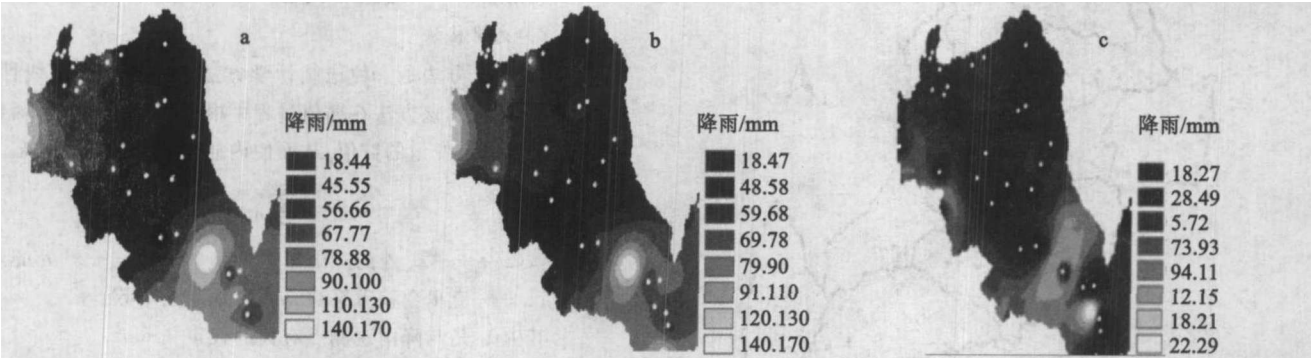
在研究过程中,对 2000—2003 年的 5—10 月的降雨数据进行了对数与立方根变换,并在整个插值过程中与原始数据插值的结果做了对比分析,交叉验证结果表明:对于距离平方反比法,对数变换和立方根变换后的 MRE 和 RMSE 的值都小于原始数据插值的结果,对数变换后的 MRE 的值下降 23.17%,立方根变换后下降 19.33%。MRE 经对数变换后的值比立方根变换后的要小;RMSE 的值也有下降趋势。其中采用普通克里格插值方法,对数变换和立方根变换后 MRE 的值分别下降了 23.50%和 18.67%。总之,交叉验证结果表明:数据变换后插值精度得到了提高,用对数变换的方式对月降雨进行插值可以提高插值精度。

表 3 IDW 插值中降雨数据变换前后交叉验证结果

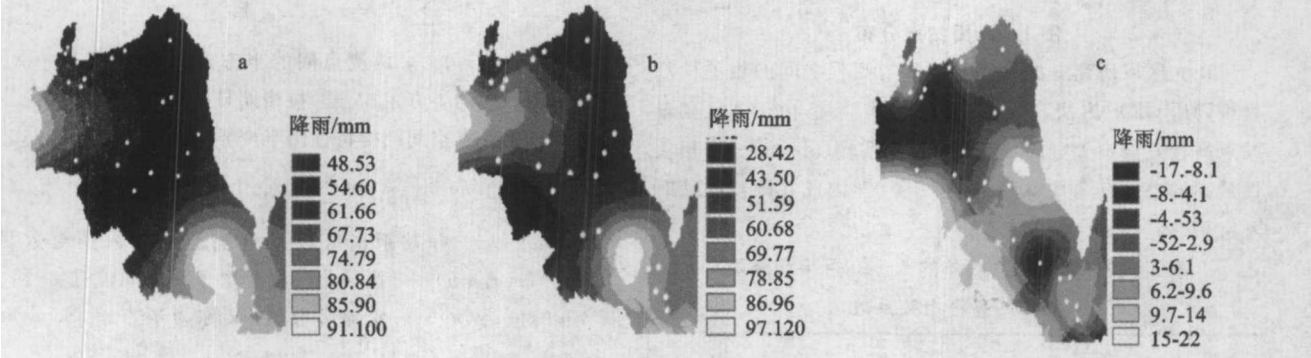
月份	5	6	7	8	9	10
相对平均误差(MRE)						
原数据	0.21	0.28	0.35	0.38	0.26	0.41
对数变化	0.05	0.08	0.09	0.11	0.07	0.10
立方根变换	0.06	0.10	0.12	0.15	0.13	0.17
均方根误差(RMSE)						
原数据	16.24	24.78	34.62	58.95	41.50	29.13
对数变换	4.64	7.80	9.38	13.43	8.80	10.24
立方根变换	5.76	9.90	12.32	18.20	15.68	15.39

表 4 Kriging 插值中降雨数据变换前后交叉验证结果

月份	5	6	7	8	9	10
相对平均误差(MRE)						
原数据	0.22	0.28	0.33	0.39	0.27	0.38
对数变化	0.05	0.07	0.08	0.10	0.06	0.10
立方根变换	0.07	0.10	0.12	0.15	0.14	0.17
均方根误差(RMSE)						
原数据	16.5	25.15	34.53	57.53	42.78	26.43
对数变换	4.90	7.50	9.23	12.37	9.27	11.40
立方根变换	6.16	7.55	7.69	18.59	16.19	16.03



a. 原数据降雨; b. 对数变换降雨; c. 原数据和对数变换降雨差值
图 2 延安示范区 2003 年 6 月 IDW 法插值降雨表面图



a. 原数据降雨; b. 对数变换降雨; c. 原数据和对数变换降雨差值
图 3 延安示范区 2003 年 6 月 Kriging 法插值降雨表面图

2.2 IDW 与 Kriging 的比较分析

在对研究区样本降雨数据进行对数变换、立方根变换的基础上,对降雨资料进行了 IDW 与 Kriging 两种方法插值比较。通过对比分析得出:对于两种插值方法,经过对数变换后平均相对误差分别为 8.33%和 7.67%,立方根变换后平均相对误差为 12.17%和 12.50%。6—9 月采用 Kriging 插值经对数变换后 MRE 小于反距离插值法的值,说明 Kriging

方法较优,6 月、7 月、8 月和 10 月经立方根转换后两种方法插值精度一样。图 2 为延安示范区 6 月反距离权重法插值的降雨空间分布,从图中可以看出,对于原始数据降雨表面(a)和对数变换后的降雨表面(b),二者没有很大的区别,进行研究区域现有数据插值时精度也比较高,但降雨分布在部分站点界限过于清晰,在极值分布区形成明显的“牛眼”,从原数据与对数变换差值图(c)可以看到,在站点周围差值较

小,差值以站点为圆心向外逐渐增大。图3为Kriging法插值的降雨空间分布,对数变换降雨表面(b)更为细致平滑些,正确反映了降雨的实际分布情况。从原数据与对数变换差值图(c)可以看出,在无降雨站点地区两者的差值很大,站点密集区差值较小,这就说明Kriging法对站点分布的密集程度更为依赖,站点多的区域插值出的栅格图比较平滑。总之,对月降雨数据经过对数变换后插值精度得到了提高,如果经对数变换后采用Kriging插值方法可以进一步提高插值精度。说明普通Kriging方法比IDW略胜一筹。

3 结论与讨论

经过对数变换可以提高研究区内月降雨插值精确程度,在此过程中有必要对其他数据变换采用不同的插值方法进行分析。对月降雨数据进行两种数据变换的基础上采用了IDW和Kriging进行插值,得到如下的结果:对IDW法对数变换后的MRE的值比原始数据平均下降23.17%,Kriging方法下降了23.50%。说明对数变换可以提高插值精度。对于两种插值方法,IDW法和Kriging方法二者经过对数变换后平均相对误差分别为8.33%和7.67%,说明Kriging方法比较适合本研究区月降雨插值。本研究降雨数据与海拔高度相关性很弱,所以不宜采用GIDS法。

目前插值方法较多,在进行具体的降雨插值时采用不同方法来说明插值的精确程度,目前用途比较广的是反距离权重法、克里格等方法,但考虑到高程及其他气象因素与降雨相关性较强时采用协克里金法、梯度距离平方反比法等,采取合适的插值方法对插值结果精度有很大的影响^[18]。对原始数据进行变换可以提高插值精度,但这不是绝对的,不同地区、不同季节、不同插值要素,数据变换的插值效果都是不同的。由于本文高程与降雨相关性不大,固采用反距离权重法和克里格法进行插值,对原始数据进行对数变换和立方根变换,以ARC GIS为实现工具,统计分析模块为Geostatistical Analyst^[19],实现了对延安示范区2000—2003年月降雨插值,结果表明,采用对数变换后的Kriging法为优选方法。

对于众多的空间插值方法而言,没有绝对最优的方法,只有在特定条件下的最优方法,针对不同研究区的实际情况,对实测数据进行分析、变换、比较验证,才能得到较为理想的插值效果。不论采用那种方法降雨站点的分布和数据本身对插值精度结果有很大的影响,所以在充分利用现有站点数据的情况下根据月降雨数据的特点引入采样密度高的其它地理要素数据,如地形数据、卫星云图、大气环流等资料与插值模型整合将是进一步提高月降雨量空间插值精度的关键。

参考文献:

- [1] 姚志红,杨勤科,吴喆,等.区域尺度降雨径流估算方法研究工:算法设计[J].水土保持研究,2006,13(5):306-308.
- [2] 杨勤科,李锐,徐涛,等.区域水土流失过程及其定量描述的初步研究[J].亚热带水土保持,2006,18(2):20-23.
- [3] 李海滨,林忠辉,刘苏峡.Kriging方法在区域土壤水分估值中的应用[J].地理研究,2001(9):446-452.
- [4] Dubrule O. Two methods with different objectives; Splines and Kriging[J]. Mathematical Geology, 1983, 15(2): 245-257.
- [5] Nalder I A, Wei R W. Spatial interpolation of climate normal: test of new method in the Canadian boreal forest[J]. Agricultural and Forest Meteorology, 1998, 92: 211-225.
- [6] 赵登忠,张万昌,刘三超.基于DEM的地理要素PRISM空间内查研究[J].地理科学,2004,24(2):211-225.
- [7] 郭艺歌,康玲,王学立.小流域时段降雨量空间差值方法研究[J].中国农村水利水电,2006(2):41-43.
- [8] 朱求安,张万昌.新安江模型在汉江江口流域的应用及适应性分析[J].水资源与水工程学报,2004,15(3):19-23.
- [9] 李新,程国栋,卢玲.空间内插方法比较[J].地理科学进展,2000,15(3):260-263.
- [10] 刘登伟,封志明,杨艳昭.海河流域降水空间插值方法的选取[J].地理信息科学,2006,8(4):75-83.
- [11] 秦耀东,李保国.应用析取克里格方法估计区域地下水埋深分布[J].水利学报,1998(8):28-33.
- [12] 封志明,杨艳昭,丁晓强,等.气象要素空间插值方法优化[J].地理研究,2004,23(3):357-364.
- [13] Amani A, Lebel T. Lagrangian Kriging for the estimation of Sahelian rainfall at small time steps[J]. Journal of Hydrology, 1997, 192: 125-157.
- [14] Peter M, Chris D. Mapping precipitation in Switzerland with ordinary and indicator Kriging[J]. Journal of Geographic Information and Decision Analysis, 1998, 2(2): 65-76.
- [15] 林忠辉,莫兴国,高静.中国陆地区域气象要素的空间插值[J].地理学报,2002,57(1):47-56.
- [16] Michael F Hutchinson. Interpolation of rainoff data with thin plate smoothing splines-Part 1: two dimensional smootning of data with short range correlation [J]. Journal of Geographic Information and Decision Analysis, 1998, 2(2): 139-151.
- [17] Genton M G. Analysis of rainfall data by robust spatial statistic using S + SPATIALSTATS[J]. Journal of Geographic Information and Decision Analysis, 1998, 2(2): 116-126.
- [18] 石朋,芮孝芳.降雨空间插值方法的比较与改进[J].河海大学学报:自然科学版,2005,33(4):361-365.
- [19] 汤国安,杨昕. ARCGIS地理信息系统空间分析实验教程[M].北京:科学出版社,2006.