

多维线性插值方法的构建及应用

刘少华^{1a, 1b}, 严登华^{1a, 1b}, 王刚^{1a, 1b}, 李传哲^{1a, 1b, 2}, 翁白莎^{1a, 1b}

(1. 中国水利水电科学研究院 a. 流域水循环模拟与调控国家重点实验室; b. 水资源研究所, 北京 100038;
2. 河海大学 水文水资源与水利工程科学国家重点实验室, 南京 210098)

摘要: 根据全国 2010 年 607 个气象站点的日降水量资料以及各站点平面分布与高程信息, 构建了多维线性空间插值方法。在不同插值控制站点数目($n=4, 8, 12$) 情景下, 分析了多维线性空间插值方法与反距离权重插值、样条插值和克里格插值法(球面模型和高斯模型)在我国年降水量空间插值中的应用情况。对比发现, 不同的控制站点情境下, 各种插值方法的插值结果差异显著, 但随着控制站点数目的增加, 各插值结果的差异性减小; 多维线性插值方法在多种情景下都表现出较好的插值精度。

关键词: 多维线性插值; 反距离权重插值; 样条插值; 克里格插值; 年降水量

中图分类号: P426 **文献标识码:** A **文章编号:** 1672-1683(2013)04-0009-04

Development and Application of Multidimensional Linear Interpolation

LIU Shao hua^{1a, 1b}, YAN Deng hua^{1a, 1b}, WANG Gang^{1a, 1b}, LI Chuan zhe^{1a, 1b, 2}, WENG Bai sha^{1a, 1b}

(1. China Institute of Water Resources and Hydropower Research, a. State Key Laboratory of Water Cycle Simulation and Regulation in River Basin; b. Department of Water Resources, Beijing 100038, China; 2. State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 210098, China)

Abstract: Spatial interpolation is an important approach to investigate the spatial distribution of meteorological and hydrological variables, and the suitable interpolation methods and reasonable control stations are beneficial to improve the precision of interpolation results. In this paper, a multidimensional linear spatial interpolation method was developed based on the daily precipitation at each of 607 stations in 2010 and the horizontal distribution and elevation of each station. Under different scenarios with varying control stations ($n=4, 8, \text{ and } 12$), The applications of multidimensional linear interpolation, inverse distance weighting interpolation (IDW), spline interpolation, and kriging interpolation (Spherical model and Gaussian model) to calculate the annual precipitation of 30 test stations were analyzed. The results showed that the interpolation results from different methods are significantly different, and the number of control stations has a significant impact on the precision of results with the existence of more control stations resulting in less difference between interpolation results. The multidimensional linear interpolation showed good interpolation precision under all scenarios and it can be an effective and reasonable solution for the future research of spatial interpolation.

Key words: multidimensional linear interpolation; inverse distance weighting interpolation; spline interpolation; kriging interpolation; annual precipitation

1 研究背景

降水是陆地生物圈模型的主要驱动变量^[1], 其空间化信息对于区域水文、水资源分析以及区域水资源管理、旱涝灾害管理、生态环境治理都具有重要意义^[2,3]。气象水文信息

空间化是研究气象水文变量空间分布的基本方法和重要内容^[4,5], 而空间插值是气象水文站点数据空间信息化的重要途径, 有效的插值方法和合理的控制站点数目对于空间信息化精度的提高具有重要意义^[6,7]。

空间插值的本质是通过建立已知站点与未知站点变量

收稿日期: 2013-03-13 修回日期: 2013-06-13 网络出版时间: 2013-07-28

网络出版地址: <http://www.cnki.net/kcms/detail/13.1334.TV.20130728.1310.037.html>

基金项目: 国家重点基础研究发展计划(973计划)项目(2010CB951102); 国家自然科学基金创新研究群体基金项目(51021066); 国家自然科学基金(51209225); 水文水资源与水利工程科学国家重点实验室开放研究基金(2011490511)

作者简介: 刘少华(1989-), 男, 河南驻马店人, 硕士研究生, 主要从事水文学与水资源学方面研究工作。E-mail: shliuiwhr@hotmail.com

通讯作者: 严登华(1976-), 男, 安徽太湖人, 教授级高级工程师, 主要从事气候变化下水资源综合应对及 3S 技术的应用。E-mail: yandh@iwhr.com

之间的关系,来估计未知站点的变量值。常用的空间插值方法可以分为几何插值法、数学插值法以及地学插值方法^[8]。几何插值方法的基本原理是距离越近相似性越高,常用方法包括泰森多边形法和反距离权重插值法,适用于下垫面特征较为一致的地区,而在地形起伏较大的区域其插值误差普遍较大^[3,9]。数学插值方法是假定插值要素在研究区域呈一定的曲面分布,通过拟合曲面方程对未知站点进行插值,常见方法有趋势面法和样条插值法^[1,3]。地学插值法主要指克里格插值法,其原理是利用区域化变量数据和半变异函数的结构特点,对未知点的区域化变量进行最优无偏估计。数学插值方法和地学插值方法均是根据数学原理通过已知点的信息拟合区域曲面方程或者半变异函数对未知点变量进行无偏估计,因对区域下垫面差异性较大的区域具有一定的适用性,且通常已知站点密度越大插值结果精度越高^[10]。然而由于区域下垫面的差异性的分布特征,其空间变量场的分布是复杂多样的,如何确定合理的曲面方程和半变异函数来准确描述空间变量场的分布特征是十分困难的^[3]。实际上,区域变量场的分布在空间维度的上通常满足一定的递变规律,因此在空间任一维度上,均可通过线性插值对未知点变量值进行近似估计。然后,通过权衡空间未知点变量值在各个维度的权重,组合多维度线性插值结果,就能得到该变量值的非线性多维空间插值结果。

本文基于我国 2010 年 607 个气象站点的日降水量以及平面坐标、海拔高程等资料,构建多维线性插值模式,对我国的年降水量进行插值估计,并在不同的插值控制站点数目情境下与反距离权重插值、样条插值和普通克里格插值等传统插值结果进行对比,分析不同方法的插值精度。

2 多维线性插值方法构建

线性插值实质是假设未知变量分布在单一维度上满足线性递变特征,则根据线性函数斜率相同的特点对未知变量进行插值求解的方法。但是实际中真正满足线性递变特征的空间变量是极少的,原因是未知变量往往是多种要素综合影响的结果。因此,需要组合多维度线性插值结果,得到能够反映未知变量多维度的综合估计值。

以年降水量的多维线性插值为例,见图 1。求解未知点 C 处的未知变量值 m_c , 已知其邻近 A、B 两点的变量值 m_a 、 m_b , 同时已知 A、B 和 C 处的影响要素 x_a 、 x_b 、 x_c 值, 则根据线性插值公式(4)、(5)可得到 m_c 在要素 x 维度的插值结果 $m_{c,x}$ 。

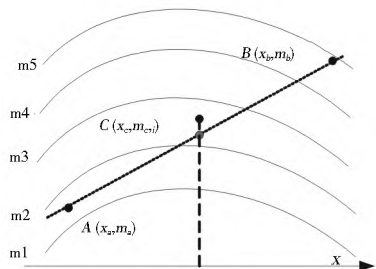


图 1 变量场在 x 维度线性插值示意图

Fig. 1 Schematic diagram of linear interpolation in x dimension of variable field

$$\frac{m_{c,x} - m_a}{x_c - x_a} = \frac{m_b - m_a}{x_b - x_a} \quad (1)$$

$$m_{c,x} = \frac{m_b - m_a}{x_b - x_a} \times (x_c - x_a) + m_a \quad (2)$$

假设在 A、B 和 C 三点处与其变量 m 相关的已知影响要素有 s 个, 则根据 s 个影响要素的线性插值可得到 s 个维度的插值结果 $m_{c,i} (i = 1, 2, \dots, s)$ 。根据已知点 A 和 B 的 s 个影响要素(维度)估计得到的 s 个未知变量估计值, 加权求和可得如下: $m_c = \sum_{i=1}^s (w_i, m_{c,i})$ (3)

式中: w_i 为第 i 个影响要素线性插值结果的权重。权重的确定方法要根据所选择的影响要素的差值而不同, 本文中应用中未知变量是降水量, 影响要素选择的是平面距离(水平方向 x 和垂直方向 y) 和海拔高程要素(h)。由于 3 个影响要素 x 、 y 和 h 均是长度单位, 本文中通过其距离差/高差确定其权重, 公式如下:

$$w_i = \frac{l_{a,i} + l_{b,i}}{\sum_{i=1}^s (l_{a,i} + l_{b,i})} \quad (i = x, y, h) \quad (4)$$

式中: $l_{a,i}$ 和 $l_{b,i}$ 分别为影响要素 i 在未知点 C 与点 A 和 B 间差值的绝对值。本文中平面距离 x 、 y 的单位 km, 海拔高程的单位是 m。

其次, 在空间插值时通常未知点 C 邻近的已知点有多个, 假设其邻近的已知点(控制点)数目为 n , 则可以形成已知直线数目为 $n(n-1)/2$, 相应的多维线性插值结果就有 $n(n-1)/2$ 个 $m_c^j (j = 1, 2, \dots, n(n-1)/2)$, 取其中间值作为点 C 处的未知变量值如下:

$$M_c = \text{median}(m_c^j, j = 1, 2, \dots, n(n-1)/2) \quad (5)$$

3 实例应用及检验

3.1 资料及检验方法

本文所用数据为 2010 年全国 607 个气象站点日降水量资料, 通过计算可得到各站点年降水量, 数据来源于中国气象科学数据共享服务网(<http://cdc.cma.gov.cn/>)。根据气象站点编号由小到大排序, 每隔 20 个站点选择一个气象站点作为检验站点, 共选择 30 个检验站点, 其余 577 个站点作为插值站点, 见图 2。根据上文插值方法分别运用插值站点数据对检验站点的年降水量进行插值检验。

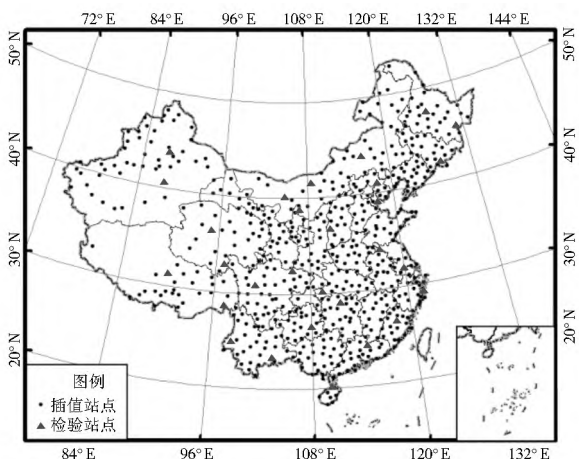


图 2 气象站点位置分布

Fig. 2 Distribution of weather stations

3.2 情景及方案设计

插值结果的精度与插值方法和控制站点数目直接相关。因此,本文在不同控制站点数目($n=4, 8, 12$)情景下,以反距离权重插值、样条插值和普通克里格插值(四种半变异函数模型:球面模型(Spherical)和高斯模型(Gaussian))方法作为对比方案,研究不同影响要素下两种多维线性插值方法的插值精度,具体方案设计见表1。

表1 插值方案设计

Table 1 Design of interpolation schemes

方案设计	插值方法
方案1	反距离权重插值
方案2	样条插值
方案3	普通克里格插值(半变异函数:球面模型)
方案4	普通克里格插值(半变异函数:高斯模型)
方案5	多维线性插值(选择影响要素:平面坐标 x, y)
方案6	多维线性插值(选择影响要素:平面坐标 x, y 和高度 h)

3.3 结果评价

对各插值方案(假定插值方案数目为 n)的插值结果中各检验站点平均相对误差、最大相对误差以及标准误差分别进行由大到小排序,误差最大方案得分为1,其次的方案得分为2,以此类推则误差最小的方案得分是 n 。然后合计各方案在平均相对误差、最大相对误差和标准误差评价中的综合得分,公式如下:

$$S_i = S_{i,1} + S_{i,2} + S_{i,3} \quad (i = 1, 2, \dots, n) \quad (6)$$

式中: S_i 为插值方案 i 的最终得分, $S_{i,1}$ 、 $S_{i,2}$ 和 $S_{i,3}$ 分别是插值

方案 i 在平均相对误差、最大相对误差和标准误差排序中的得分。

3.4 结果与讨论

各方案插值结果的误差见表2,可以看出:随着控制站点数目的增加、已知的有效信息增多,各插值结果之间的差距有所减小。当控制站点 $n=4$ 时,各插值结果差异性显著(各方案综合得分方差为22.7),其中方案4和5插值结果精度具有明显的优势;当增加控制站点数目 $n=8$ 时,各插值结果的差异性缩小(各方案综合得分方差为15.5),其中方案6的插值结果精度具有显著的优势;当继续增加控制站点数目 $n=12$ 时,各插值结果的差异性进一步缩小(各方案综合得分方差为9.1),且方案6的插值效果最优。

然而,方案1和5的插值结果精度随着控制站点的增加而降低,而其他方案的插值精度随控制站点数目的变化规律不明显。分析其原因:方案1和方案5均是采用平面就近原则,通过平面最近的控制站点进行插值,当控制站点较少且站点密度较大时该方法效果较好,但随着控制站点数目增加,导致远处的控制站点(地形地貌的差异导致距离越远相关性越差)被应用于插值,使得插值误差不断的增加。而方案2、3、4则都是通过控制站点信息来估计曲面方程和半变异函数模型进行插值,通过增加控制站点数目得到的曲面方程和半变异函数模型,未必更符合实际空间变量场的分布情况,因而其插值精度并不完全是随其控制站点的增加而提高。另一方面,与方案5相比,方案6插值过程中考虑了海拔高程信息,减小了地形差异性对插值结果的影响,在控制站点较丰富的地区对于插值结果精度的提高显著。

表2 插值结果误差及评价结果统计

Table 2 Statistics of the error and evaluation of interpolation results

检验方法	控制站点	方案1	方案2	方案3	方案4	方案5	方案6
平均相对误差(%)	$n=4$	21.72	25.05	24.51	20.48	19.63	23.79
	$n=8$	25.54	24.93	25.23	26.84	23.79	21.09
	$n=12$	27.42	24.76	24.51	27.00	24.86	22.79
标准误差/mm	$n=4$	273.33	314.86	230.97	228.16	253.95	296.46
	$n=8$	273.66	322.78	242.26	239.53	240.22	238.19
	$n=12$	271.25	323.70	230.97	224.48	243.40	256.28
最大相对误差(%)	$n=4$	108.02	110.76	129.60	95.47	68.82	147.72
	$n=8$	108.97	95.92	125.15	153.04	125.46	64.36
	$n=12$	126.69	97.30	129.60	162.52	163.77	75.55
综合得分 S	$n=4$	11	5	9	16*	16*	6
	$n=8$	8	10	9	7	11	18*
	$n=12$	7	10	13	10	8	15*
	合计	26	25	31	33	35	39*

注:带*表示该情境下评价结果中的最大值。

总之,综合对比分析各插值结果可知,两种多维线性插值方案(方案5和方案6)结果均表现较好,尤其是考虑高程修正的多维线性插值方法(方案6),在控制站点丰富的情境下,其插值结果具有明显优势。因此,多维线性插值对于我国年降水量的插值具有较好的适用性。

4 结语

本文基于线性插值构建了多维线性插值方法,并在不同

的控制站点数目情境下对比研究了该方法与反距离权重插值、样条插值和克里格插值方法在我国降水量插值中实际应用效果,结果表明:不同的控制站点情境下,各种插值方法的插值结果具有略有差异,整体上多维线性插值方法对我国的降水量年值具有较好的估计能力。多维线性插值提供了一种开放式的插值模式,它不仅可以根据空间要素信息,也可以考虑其他相关影响要素,进行多维度线性组合插值,以提高插值精度。作为一种思路清晰、操作简单的开放式插值模

式, 多维线性插值方法可为今后的空间插值提供一个简洁有效的手段。

参考文献(References):

- [1] 卢毅敏, 岳天祥, 陈传法, 等. 中国区域年降水空间分布高精度曲面建模[J]. 自然资源学报, 2010, 25(7): 1194-1205. (LU Yi min, YUE Tian xiang, CHEN Chuan fa, et al. Surface Modeling of Annual Precipitation in China. Journal of Natural Resources [J]. Journal of Natural Resources, 2010, 25(7): 1194-1205. (in Chinese))
- [2] 蔡福, 于慧波, 矫玲玲, 等. 降水要素空间插值精度的比较—以东北地区为例[J]. 资源科学, 2006, 28(6): 73-89. (CAI Fu, YU Hui bo, JIAO Ling ling, et al. Comparison of Precision of Spatial Interpolation of Precipitation Factors: A Case Study in Northeastern China [J]. Resources Science, 2006, 28(6): 73-89. (in Chinese))
- [3] 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): 34-39. (ZHU Hui yi, JIA Zhao feng. Uncertainty in the Spatial Interpolation of Rainfall Data [J]. Progress in Geography, 2004, 23(2): 34-39. (in Chinese))
- [4] Houghton J T, Meira Filho L G, Callander B A, et al. Climate Change 1995: The Science of Climate Change [M]. Cambridge: Cambridge University Press, 1996.
- [5] 姜晓剑, 刘小军, 黄芬, 等. 逐日气象要素空间插值方法的比较[J]. 应用生态学报, 2010, 21(3): 624-630. (JIANG Xiao jian, LIU Xiao jun, HU AN G Fen, et al. Comparison of Spatial Interpolation Methods for Daily Meteorological Elements [J]. Chinese Journal of Applied Ecology, 2010, 21(3): 624-630. (in Chinese))
- [6] 林忠辉, 莫兴国, 李宏轩, 等. 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. (LIN Zhong hui, MO Xing guo, LI Hong xuan, et al. Comparison of Three Spatial Interpolation Methods for Climate Variables in China [J]. Acta geographica sinica, 2002, 57(1): 47-56. (in Chinese))
- [7] 王舒, 严登华, 秦天玲, 等. 基于 PER-Kriging 插值方法的降水空间展布[J]. 水科学进展, 2011, 22(6): 757-764. (WANG Shu, YAN Deng hua, QIN Tian ling, et al. Spatial Interpolation of Precipitation Using the PER-Kriging Method [J]. Advances in Water Science, 2011, 22(6): 757-764. (in Chinese))
- [8] 李新, 程国栋, 卢玲. 空间内插方法比较[J]. 地球科学进展, 2000, 15(3): 260-265. (LI Xin, CHENG Guo dong, LU Ling. Comparison of Spatial Interpolation Methods [J]. Advance in Earth Sciences, 2000, 15(3): 260-265. (in Chinese))
- [9] Watson D F, Philip G M. A Refinement of Inverse Distance Weighted Interpolation [J]. Geoprocessing, 1985, (2): 315-327.
- [10] 孙然好, 刘清丽, 陈利顶. 基于地统计学方法的降水空间插值研究[J]. 水文, 2010, 30(1): 14-18. (SUN Ran hao, LIU Qing li, CHEN Li ding. Study on Precipitation Interpolation Based on the Geostatistical Analyst Method [J]. Hydrology, 2010, 30(1): 14-18. (in Chinese))
- [5] 余烜. 气候变化对海河流域水资源影响研究 [D]. 北京: 中国水利水电科学研究院, 2009. (YU Xuan. Studies on Impacts of Climate Change on Water Resources in Haihe River Basin [D]. Beijing: China Institute of Water Resource and Hydropower Research, 2009. (in Chinese))
- [6] 刘健, 张奇, 许崇育. 近 50 年鄱阳湖流域径流变化特征研究 [J]. 热带地理, 2009, 29(3): 213-218. (LIU Jian, ZHANG Qi, XU Chong yu. Characteristics of Runoff Variation of Poyang Lake Watershed in the Past 50 Years [J]. Tropical Geography, 2009, 29(3): 213-218. (in Chinese))
- [7] 李滨江, 郭姚生. 佳木斯市水资源精确评价中需水系统若干参数的确定 [J]. 水文, 1999, (2): 23-29. (LI Bin jiang, GU O Yao sheng. Determine Several Parameters of Need Water System in Water Resources Precise and Accurate Evaluation in Jiamusi City [J]. Hydrological, 1999, (2): 23-29. (in Chinese))
- [8] 孙同洁, 纪万军. 佳木斯市水资源可利用量的分析 [J]. 黑龙江水利科技, 2006, 34(4): 164. (SU N Tong jie, JI Wan jun. Analysis of Water Resources Available Amount in Jiamusi City [J]. Heilongjiang Science and Technology of Water Conservancy, 2006, 34(4): 164. (in Chinese).)
- [9] 张戈平. 城市水资源承载力评价指标体系研究 [D]. 哈尔滨: 东北农业大学, 2003. (ZHANG Ge ping. Study on Evaluation Index System of Urban Water Resources Carrying Capacity [D]. Harbin: Northeast Agricultural University, 2003. (in Chinese))
- [10] 裴源生, 刘建刚, 赵勇, 等. 水资源用水总量控制与定额管理协调保障技术研究 [J]. 水利水电技术, 2009, 40(3): 8-11. (PEI Yuan sheng, LIU Jian gang, ZHAO Yong, et al. Study on Support Technique for Coordination Between Total Amount Control and Quota Management of Water Consumption for Water Resources [J]. Water Resources and Hydropower Engineering, 2009, 40(3): 8-11. (in Chinese))
- [11] 唐力, 赵勇, 肖伟华, 等. 水资源总量控制和定额管理制度实施进展 [J]. 人民黄河, 2008, 30(3): 43-44. (TANG Li, ZHAO Yong, XIAO Wei hua, Water Resources Total Control and Quota Management System Progress in the Implementation [J]. Yellow River, 2008, 30(3): 43-44. (in Chinese))
- [12] 张少文, 王文圣, 丁晶, 等. 黄河天然年径流超长期变化特性研究 [J]. 人民黄河, 2004, 26(8): 10-12. (ZHANG Shao wen, WANG Wen sheng, DING Jing, The Research Natural Annual Runoff Ultra long term Changes in Characteristics in Yellow River [J]. 2004, 26(8): 10-12. (in Chinese))
- [13] 许月卿, 李双成, 蔡运龙. 基于小波分析的河北平原降水变化规律研究 [J]. 中国科学 (D 辑), 2004, 34(12): 1176-1183. (XU Yue qing, LI Shuang cheng, CAI Yun long. The Research of Rainfall Variation Based on Wavelet Analysis in Hebei Plain [J]. Science China Ser. D Earth Sciences, 2004, 34(12): 1176-1183. (in Chinese))

(上接第 8 页)