

Using Monte Carlo Method for Predicting Rainfall -- Taking Xi'an Area, Shaanxi Province, China, as an Example

Wang Haike, Xu Panpan, Qian Hui

College of Environmental Science and Engineering, Chang'an University, Xi'an, China

Emial address:

Wanghaike321@163.com (Wang Haike), xupanpan0212@163.com (Xu Panpan), 55480918@qq.com (Qian Hui)

To cite this article:

Wang Haike, Xu Panpan, Qian Hui. Using Monte Carlo Method for Predicting Rainfall -- Taking Xi'an Area as an Example. *Earth Sciences*. Vol. 4, No. 5, 2015, pp. 201-204. doi: 10.11648/j.earth.20150405.16

Abstract: Rainfall forecasting plays a vital role in the national economy, social development and human life. Based on Monte Carlo method, this paper uses P- III distribution function to fitting precipitation data in the past 63 years so as to forecast precipitation. Using this model to forecast the rainfall for the past ten years (2003 ~2013) in Xi'an city, Shaanxi province, China, based on the past 63 years data. The predicted results indicate that the prediction has a high accuracy in normal rainfall year, but in extremely in dry condition and high rainfall year, the relative error is huge. So that, the method is more suitable for the prediction of rainfall in the flat water.

Keywords: Monte Carlo, P- III Distribution Curve, Precipitation Forecasting, Curve Fitting Method

蒙特卡洛方法预测降雨量——以中国陕西省西安地区为例

王海科, 徐盼盼, 钱会

长安大学环境科学与工程学院, 西安, 陕西, 中国

邮箱

Wanghaike321@163.com (王海科), xupanpan0212@163.com (徐盼盼), 55480918@qq.com (钱会)

摘要: 降雨量预测在国民经济、社会发展和人类生活方面有着至关重要的作用。以蒙特卡洛方法为基础, 用P-III分布曲线来拟合降雨数据, 以此来对降雨进行预测。基于中国陕西省西安市1951—2013年63年降水资料对西安市2004—2013年降雨量进行了预测。预测结果表明: 该方法在对平水年降雨量预测时精度较高, 对丰水年和枯水年的预测精度却有所降低, 因此该方法更适合对平水年降雨量进行预测。

关键字: 蒙特卡洛, P-III分布曲线, 降雨量预测, 曲线拟合

1. 引言

降水是陆地水资源的直接补给来源, 它的丰欠程度直接影响着内陆局部地区的水资源总量, 降水资源已经成为影响人类生活和社会经济发展的重要因素, 因此降雨量的预测受到越来越多的学者重视。目前, 预测降雨量的方法有很多, 有学者将预测方法分为定性、时间序列和因果关系三类[1]。其中定性预测主要是依据人们对系统过去和

现在的经验、直觉和主观判断, 通过现场调查、专家打分、主观评价等方法对系统进行预测, 常用的方法有特尔菲法[2]、专家会议等; 时间序列和因果关系主要是借助历史的观测数据, 通过采用一定的数据分析方法对系统进行定量预测, 其精度主要取决于数据的可靠性和分析方法的科学性、稳定性。目前定量的预测方法有很多, 具有代表性的研究方法主要有多元逐步分析回归法[3]、多元线性回归法[4]、自回归移动平均模型[5]、人工神经网络模型[6]、灰色模型[7]、加权马尔科夫链[8]、多目标遗传算法[9]、

遥感观测[10]等。近年来,利用统计实验技术进行随机过程预测的方法开始应用到水文预测中来,韦庆等利用蒙特卡洛方法模拟了中国辽宁省吉林白城地区年降水量的分布特征,对该地区未来20年的年降水量进行了预测[11]。杨金玲等[12]依据蒙特卡洛理论,采用P-III型分布函数对中国黑龙江省嫩江流域汛期降雨量做了预测,预测结果表明当实际汛期为平水年时预测精度较高。

西安市位于中国陕西省中部,属于黄河流域中部的关中盆地(107°40'E—109°49'E, 33°39'N—34°45'N),北部为冲积平原,南部为剥蚀山地。近年来随着经济的迅速发展和人口的急剧增加以及气候条件的变幻莫测,西安地区需水量逐年增加,可利用水资源量短缺,缺水严重影响着西安地区国民经济的发展[13]。近50年来,西安市年降雨量多年评价值为507.7—719.8mm,最大值为903.2mm(1983年),最小值为312.2mm(1995年),最大值约为最小值的3倍,年际变化较大。本研究采用蒙特卡洛法模拟西安市降雨的随机过程,对该地区降雨量进行预测,并且验证该方法在该地区的适应性。

2. 蒙特卡洛预测降雨模型

2.1. 蒙特卡洛法简介

1777年,法国数学家布丰提出用投针实验的方法来求解圆周率 π ,这是目前被大家所公认的蒙特卡洛方法的起源。20世纪40年代,美国数学家S. M. 乌拉姆和J. 冯·诺依曼首次提出了该方法。其基本思想为:当所求解的问题是某个事件的概率,是某个随机变量的数学期望,或是与概率、数学期望有关的数值时,可以通过试验得出该事件发生的频率,凭借它来得出问题的解。从数学的角度来看,试验次数越大,试验的频率越接近事件发生的概率。因此,蒙特卡洛方法虽然很早就被人们提出却很少被使用。在本世纪四十年代,由于电子计算机的出现,使得人们可以通过计算机来模拟随机试验,把巨大的随机试验交给计算机来完成,使得蒙特卡洛方法可以广泛的应用。蒙特卡洛方法相比于其他方法能够比较逼真地描述具有随机过程性质的事物及物理过程,受几何条件限制小、收敛速度不受问题维数的影响,并且数据误差很容易确定,实验思路简单、程序易于实现,因此该方法在现代化的科学技术中得到了非常广泛的应用。

2.2. 伪随机数的选取

蒙特卡洛方法的成功与否取决于其所选取的随机模型的构造,但它很大程度上取决于模型计算中所产生的随机数的性质。由于在现实生活中不可能有真正的随机数,借助计算机所产生的随机数也是通过一定的算法及几个参数的计算来得出的伪随机数。无论用什么方法产生的随机数,它所产生的随机数总是具有一定的局限性,其所产生的随机数总是在一个有限长的循环数据中,而且序列偏差的上确界达到了最大值。所以人们希望能够得到低偏差的确定性序列,这样任意长的子序列都能均匀地充填函数空间。基于这些要求,学者们研究

出了许多序列,如Van der Corput序列, Halton序列, Sobol序列[14]和Niederreitre的(t, s)序列[15]等。目前我们可以借助一些软件来获取低偏差的伪随机数,如MATLAB中的rand函数可以用来获取在(0, 1)上均匀分布的随机变量;randa函数可以获取服从标准正态分布的随机数;类似的函数还有frnd函数、chi2rnd函数、normrnd函数、unifrnd函数等。

2.3. 分布函数选取

区域降雨受大气环流、海陆位置、地表地形、太阳辐射及人类活动等众多因素的影响,各影响因素的作用及其机理极其复杂,且每项因子都不起绝对的支配作用,因此可以将降雨在局部范围内视为时间上的随机过程。在对降雨分布函数选取时,选取的分布函数应该符合水文现象的基本物理性质,曲线不应出现负值,且概率密度函数的数学性质简单,计算起来方便,同时又具有一定的弹性,以便可以广泛的适用。20世纪50年代,陈志凯[16]对P-III型和K-M型分布函数进行了比较,得出这两种曲线对中国洪水拟合都具有很强的适应性;20世纪70年代,美国水资源委员会研究表明不同线型和不同的拟合方法对美国河流的适应性没有明显差异,他们建议采用P-III型分布函数[17]。中国从20世纪60年代就开始就采用P-III型分布函数对水文资料进行拟合,尤其是在推算洪水系列资料。因此,选取P-III型分布函数来拟合西安市降雨数据。

3. 研究区资料选取

3.1. 研究区数据选取及预处理

使用蒙特卡洛方法除了尽可能的要求预测值序列与实际值序列保持相同的概率特性外,还需整个预测序列与实测值序列的平均值近似相等,这就需要对原始数据进行分析。选取中国陕西省西安地区1951—2013年63年的数据,从1951年开始每30年为一组,然后逐年向后取平均值,采用U检验法在一定的置信区间上估计年均值的变化范围,要求预测值与原始收据的均值要落在置信区间内,否则重新生成预测值。

3.2. 方法具体步骤

预报方法的过程如下:

第一步:将样本数据从大到小排列,计算样本容量的平均值和方差,求其平均值的置信区间。

第二步:进行理论频率曲线计算,确定最佳拟合的P-III型分布函数。

第三步:采用蒙特卡洛方法进行预测,使得预测值与原始数据的平均值在置信区间内,否则重新生成预测值。

4. 计算结果与讨论

4.1. 降雨P-III分布曲线拟合

在用蒙特卡洛方法进行降雨量预测时,每一次预测应有一定的随机抽样次数。由理论分析可知,试验次数越大,

模拟的数据越准确。可是实验次数过多，会使得运算量过多，将会影响计算速度。采用中国陕西省西安市1951年—2013年实测降水数据拟合P-III分布曲线，以此模拟不同抽样次数下预测值的变化情况。如图1所示：

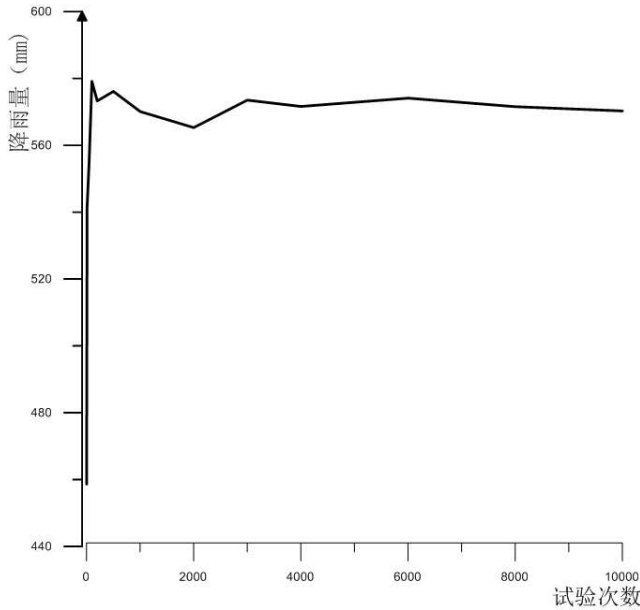


图1 不同试验次数模型预测结果。

如图1可知，随着试验次数的增加，预测值将趋于稳定。当试验次数超过3000以后，预测值变化范围很小，因此选取3000为每一次模拟的随机数生成次数。

表1 2003—2013年预测结果表。

| 年份 | 实测值 (mm) | 拟合曲线参数值 | | 理论频率 (%) | 预测值 (mm) | 相对误差 |
|------|----------|---------|--------|----------|----------|----------|
| | | Cv | Cs | | | |
| 2004 | 512.7 | 0.2277 | 0.4543 | 0.296875 | 541.4823 | 0.056139 |
| 2005 | 541.4 | 0.2264 | 0.4804 | 0.4375 | 574.1331 | 0.06046 |
| 2006 | 561.5 | 0.2246 | 0.4968 | 0.546875 | 569.9312 | 0.015015 |
| 2007 | 698.5 | 0.2227 | 0.5051 | 0.84375 | 571.7962 | -0.18139 |
| 2008 | 525.2 | 0.2218 | 0.4597 | 0.34375 | 544.1396 | 0.036062 |
| 2009 | 660.3 | 0.2204 | 0.4811 | 0.78125 | 583.1179 | -0.11689 |
| 2010 | 504.4 | 0.2188 | 0.4482 | 0.265625 | 524.4874 | 0.039824 |
| 2011 | 723.6 | 0.2180 | 0.474 | 0.875 | 667.1403 | -0.07803 |
| 2012 | 426.7 | 0.2178 | 0.4296 | 0.140625 | 575.58 | 0.34891 |
| 2013 | 497.5 | 0.2194 | 0.4496 | 0.234375 | 519.5992 | 0.044421 |

由表1可以看出，随着试验年份的增加，拟合曲线的Cv参数以微小的趋势在减小，表明拟合曲线在逐渐趋于平缓。从数学角度分析，说明试验数据越多，拟合曲线的端点差在减小。我们在水文频率计算中，往往将设计频率值在25%—75%之间的年份称为平水年，而小于25%的年份称为丰水年，大于75%的年份称为枯水年。由表1可以看出，预测值在平水年可以取得较高的预报精度，而在枯水年，尤其是在频率大于85%（如2012年）以后，预测精度将急剧下降；在丰水年，当频率小于20%（如2009年）后，预测精度也会有所下降。

利用中国陕西省西安市1951年—2003年的降水数据进行P-III分布曲线拟合（图2），以此来对2004年降雨数量进行预测，计算其误差。然后将2004年的实测值增加到拟合数据序列中，重新进行P-III分布曲线拟合，由此预测2005年的降雨量，计算其误差。以此类推，对西安市2004年—2013年10年的降雨量进行预测，求其误差。

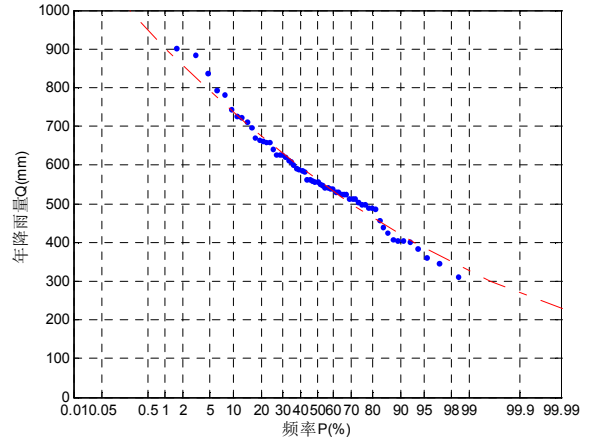


图2 西安1951-2003年降雨量P-III频率曲线。

4.2. 结果分析

对2004年—2013年实测值的预测结果见表1。

5. 结论

以P-III分布曲线来对中国陕西省西安市63年降雨量数据进行拟合，采用蒙特卡洛方法对2004年—2013年降雨量数据进行预测，在2004—2006年、2008年、2010年、2013年中取得了不错的预测精度，而在2007年、2009年、2011年和2012年这4年中预测结果存在较大误差。预测结果表明该方法在对平水年预测时有较高的精度，而对枯水年和丰水年预测精度较低。因此，使用该方法来预测平水年的

降雨量时, 会取得较高的预测精度, 但用于枯水年和丰水年时, 却存在较大误差, 预测精度也往往不尽人意。

蒙特卡洛方法和P-III分布曲线都在数学上有着明确的意义和严格的推导, 但其和水文气象等物理作用却没有着必然的联系, 在预测过程中并没有考虑各物理要素之间的联系, 只是对数据进行了数学上的处理分析。由于实验次数的有限性以及伪随机数的循环性, 必然会对试验结果产生一定的影响。因此, 预测结果在枯水年和丰水年的精度有所下降。但是, 由于降雨是一个十分复杂的物理过程, 它的发生受到各种作用的影响, 因此, 在预测平水年降雨量时, 只有该地区降雨资料, 且对该地区降雨的物理成因不是很明确的时候, 采用蒙特卡洛方法将是一种十分简单迅速且准确的方法。

致谢

中国水利部公益性行业科研专项经费项目(201301084)。

参考文献

- [1] 朱聪. 径流中长期预测模型研究[M]. 四川大学. 2005. 04.
- [2] 冯文权. 预测·头脑风暴法·特尔菲法[J]. 科学决策, 1997(2): 39-41.
- [3] Munot A, A., Kumar, K. K., . Long range prediction of Indian summer monsoon rainfall [J]. Journal of Earth System Science 118(I), 2007: 73-79.
- [4] Nayagam, L. R., Janardanan, R., Mohan, H. S. R.. An empirical model for the seasonal prediction of southwest monsoon rainfall over Kerala, a meteorological subdivision of India[J]. International Journal of Climatology 28(6), 2008: 823-831.
- [5] Toth, E., Brath, A., Montanari, A.. Comparison of short-term rainfall prediction models for real-time flood forecasting[J]. Journal of Hydrology 2000. 239, 132-147.
- [6] 张德同, 王学敏. 径向基函数神经网络在白城市年降雨量预测中的应用[J]. 技术应用. 2013, 17-0122-01, 118-122.
- [7] 侯祥龙, 阳辉. 基于滑动无偏灰色马尔科夫模型的水库年降水量预测[J]. 水土保持通报. 2014, 34(3), 181-184.
- [8] 钱会, 李培月, 王涛. 基于平均滑动一加权马尔科夫链的宁夏石嘴山市年降雨量预测[J]. 华北水利水电学院学报. 2010, 30(1), 6-9.
- [9] 采用遗传算法的小波神经网络在降雨量预测中的应用[J]. 河南工程学院学报(自然科学版). 2015, 27(1), 53-57.
- [10] Ganguly, A. R., Bras, R. L.. Distributed quantitative precipitation forecasting (DQPF) using information from radar and numerical weather prediction models [J]. Journal of Hydrometeorology. 2003. 1168-1180.
- [11] 韦庆, 卢文喜, 田文君. 运用蒙特卡罗方法预报年降雨量研究[J]. 干旱区资源与环境. 2004, 18(4), 144-146.
- [12] 杨金玲, 吴亚楠, 谢淼等. 蒙特卡罗法在嫩江流域汛期降雨量预测中的应用[J]. 南水北调与水利科技. 2011, 9(3), 28-32.
- [13] 吴建华, 李培月, 钱会. 西安市气象要素变化特征及可利用降雨量预测模型[J]. 南水北调与水利科技. 2013, 11(1), 50-54.
- [14] I. M. Sobol'. The distribution of points in a cube and the approximate evaluation of integrals. USSR Comp [J]. Math. and Math. Phys., 7: 86-112, 1967.
- [15] H. Niederreiter. Point sets and sequences with small discrepancy [J]. Monatsh. Math. 104: 273-337, 1987.
- [16] 陈志凯, 暴雨及洪水频率计算方法的研究[R], 北京: 中国水利水电科学研究院水文研究所, 1957.
- [17] Water Resources Council, Hydrology committee, A uniform technique for determining flood flow frequencies, Bulletin 15, Washington, D. C., 1967.