

# 基于云模型的非一致性年径流预测

刘睿, 梁川

(四川大学 水利水电学院, 四川 成都 610065)

**摘要:** 采用云模型和非一致性水文算法对岷江流域年径流建立预测模型, 利用非一致性算法将水文序列分成相对一致的随机性成分和非一致的确定性成分, 再分别对确定性成分和随机性成分进行云模型的不确定性推理, 最后根据时间序列分析的分解与合成理论进行数值合成, 得到最终的预测值。通过与传统预测方法相比较, 非一致性云预测方法有较高的预测精度和推广应用价值。

**关键词:** 非一致性; 年径流; 云模型; 岷江流域; 不确定性

中图分类号: TV121+.1

文献标识码: A

文章编号: 1672-1144(2012)03-0106-05

## Prediction for Non-compliance Annual Runoff Based on Cloud Model

LIU Rui, LIANG Chuan

(College of Water Resources and Hydropower, Sichuan University, Chengdu, Sichuan 610065, China)

**Abstract:** The cloud model and non-compliance hydrologic calculation method are used for building the prediction model. Firstly, the non-compliance calculation method is used to divide the hydrologic sequence into the relatively consistent random elements and non-uniform certainty elements. Next, the cloud model is used for the uncertainty reasoning of the deterministic components and random elements respectively. Finally, the unmerical synthesis is made based on the decomposition of time series analysis and synthetic theory to get the final prediction value. Compared with the traditional forecasting methods, the method discussed here has the higher prediction accuracy and application value.

**Keywords:** non-compliance; annual runoff; cloud model; Minjiang River Basin; uncertainty

## 0 引言

未来年份的径流量预测对防汛抗旱起着十分重要的作用。准确地预测出来年的径流量, 对水量的分配调度进行预估, 进行合理的水量分配, 提前做好防汛抗旱的准备。因此, 研究年径流的预测问题对维护生态系统的安全和保障人类社会经济的可持续发展有着重要的意义。传统方法都没有体现不确定性推理, 以传统模糊集理论和概率统计为基础的云模型能更多的挖掘历史数据中的规律, 提供了一种新的解决途径。因此, 本文在阐述云预测模型原理的基础上, 以岷江流域控制站为例进行年径流预测研究。

## 1 云模型

### 1.1 云模型的概念

设  $U$  是一个精确数值量表示的论域,  $C$  是  $U$  上的定性概念, 对于任意一个论域中的元素  $x$ , 都存在一个有稳定倾向的随机数  $\mu(x) \in [0, 1]$ , 叫做  $x$  对  $C$  的隶属度, 则  $x$  在论域  $U$  上的分布称为云, 每个  $x$  称为一个云滴<sup>[1]</sup>。云的整体形状反映了定性概念的重要特性, 云滴则是对定性概念的定量描述, 云滴产生过程, 表示定性概念和定量值之间的不确定性映射<sup>[2]</sup>。

### 1.2 云的数字特征

云的数字特征用期望  $E_x$ 、熵  $E_n$  和超熵  $H_e$  来表示<sup>[2]</sup>。 $E_x$  是云滴在论域空间分布的期望; 熵  $E_n$  是定

收稿日期: 2012-02-10

修稿日期: 2012-03-10

基金项目: 中国长江上游三峡工程开发公司“长江上游变化条件下的流域水文预报方案研究”项目

作者简介: 刘睿(1987—), 男(汉族), 四川成都人, 硕士研究生, 研究方向为水文与水资源。

通讯作者: 梁川(1957—), 男(汉族), 四川成都人, 教授, 主要从事水文水资源水环境的教学与科研工作。

性概念的不确定性度量,反映了云滴的离散程度和取值范围;超熵  $H_e$  是熵的不确定性度量,即熵的熵,反映了不确定度的凝聚性。根据正态分布的特性,有意义的云滴主要落在区间  $[E_x - 3E_n, E_x + 3E_n]$ ,因此可以忽略  $[E_x - 3E_n, E_x + 3E_n]$  区间之外的云滴,即正态云的“ $3E_n$  规则”<sup>[2]</sup>。云模型的 3 个数字特征值把模糊性和随机性完全集成到一起,构成定性和定量相互间的映射,作为知识表达的基础<sup>[3]</sup>。

### 1.3 正向云发生器

正向云发生器是最基本的云算法,实现了从定性信息中获得定量的范围和分布规律。通过给定云的 3 个数字特征  $(E_x, E_n, H_e)$ ,产生正态云模型的若干云滴(如图 1)。

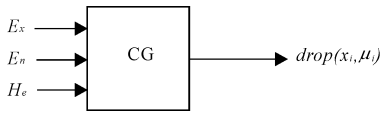


图 1 正向云发生器

具体步骤:

- (1) 生成以  $E_n$  为期望,  $H_e^2$  为方差的正态随机数  $E_{n_i}' = NORM(E_n, H_e^2)$ ,以  $E_x$  为期望,  $E_{n_i}'^2$  为方差的正态随机数  $x_i = NORM(E_x, E_{n_i}'^2)$ ;
- (2) 计算确定度  $\mu_i = e^{-(x_i - E_x)^2 / 2E_{n_i}'^2}$ ,具有确定度  $\mu_i$  的  $x_i$  成为论域中的一个云滴;
- (3) 重复步骤(1)–(2),直到产生要求的  $n$  个云滴为止。

### 1.4 逆向云发生器

逆向云发生器是实现从定量到定性概念的转换模型<sup>[2]</sup>,可以将一定数量的精确数据转换为以数字特征  $(E_x, E_n, H_e)$  表示的定性概念<sup>[4]</sup>(如图 2)。

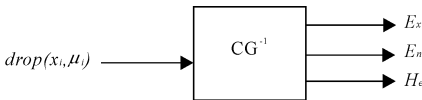


图 2 逆向云发生器

具体步骤如下:

- (1) 计算样本均值  $\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i$ ,一阶样本绝对中心矩  $\frac{1}{n} \sum_{i=1}^n |x_i - \bar{X}|$ , 样本方差  $S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{X})^2$
- (2)  $E_x = \bar{X}$ ,熵  $E_n = \sqrt{\frac{\pi}{2}} \times \frac{1}{n} \sum_{i=1}^n |x_i - E_x|$ ,超熵  $H_e = \sqrt{S^2 - E_n^2}$

## 2 云预测

### 2.1 云推理规则

云推理是一种不确定性的推理方法,基础是不确定性知识,通常使用 if  $A$  then  $B$  形式来表达<sup>[5]</sup>。 $A$  称为规则的前件,可以是一个或多个; $B$  是规则的后件,通常是一个结果。如果给定论域  $U_1$  中的一个特定点  $x$ ,通过正向云发生器可以生成  $x$  属于概念  $C_1$  的确定度  $y$ ,我们称其为前件云发生器。给定一确定度  $y$  且  $y \in [0, 1]$ ,通过正向云发生器可以生成论域  $U_2$  中的概念  $C_2$  上满足  $y$  的定量值  $x$ ,我们称其为后件云发生器。 $U_1$  和  $U_2$  可以是一维或多维。通常后件云发生器都是一维的,前件云发生器如果是多维的话,可以通过一维的组合实现<sup>[4]</sup>。一维前件云发生器如下:

输入:给定云的 3 个数字特征  $(E_{x_A}, E_{n_A}, H_{e_A})$  以及定量值  $x_A$ ,

输出:定量值  $x$  满足定性概念的确定度  $y = \exp(-\frac{(x - E_{x_A})^2}{2E_{n_A}'^2})$ 。

一维后件云发生器如下:

输入:给定云的 3 个数字特征  $(E_{x_B}, E_{n_B}, H_{e_B})$  以及确定度  $y$  且  $y \in [0, 1]$ ,

输出:满足确定度  $y$  的定量值  $x_B = E_{x_B} \pm E_{n_B}' \sqrt{-Ln y}$ 。

若输入值  $x_A$  处在前件云上升沿即:  $x_A \leq E_{x_A}$ ,后件云同取上升沿:  $x_B = E_{x_B} - E_{n_B}' \sqrt{-Ln y}$ ;反之则:  $x_B = E_{x_B} + E_{n_B}' \sqrt{-Ln y}$ 。将两者相连接,构造出单条件单规则发生器(如图 3 所示)。

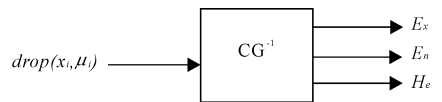


图 3 单条件单规则发生器

由于存在随机数  $E_{n_A}'$ ,确定度  $y$  带有不确定性;后件云发生器在确定度  $y$  的控制下,  $x_B$  具有不确定性。因此,规则发生器实现了推理过程中不确定性的传递,其预测值具有不确定性。

### 2.2 规则合并

根据物理学的惯性定理,任何事物都有保持其当前状态的趋势,同时事物的变化是渐进的。当前趋势只反映了未来的可能性,因此不能根据当前数据产生一个确定的数值作为预测结果,它应当表示为

定性趋势而不是定量数值<sup>[6]</sup>。云模型具备将定性和定量相互转换的能力,所以引入当前云  $C_1(E_{x_1}, E_{n_1}, H_{e_1})$  来表示当前趋势,期望值  $E_x$  作为期望预测值;熵  $E_n$  反映了可被接受的预测值;超熵  $H_e$  反映了当前趋势的模糊性的大小<sup>[7]</sup>。将时间序列数划分为历史数据集  $HD$  和当前数据集  $CD$ 。从  $HD$  中可挖掘规则  $\{A_1 \rightarrow B_1, A_2 \rightarrow B_2, \dots, A_m \rightarrow B_m\} | k \cdot T \leq a_i \leq a_i\}$ 。从  $CD$  中利用逆向云发生器可得到趋势云<sup>[6]</sup>。当两个规则的前件云或后件云相近时,可以进行云合并运算<sup>[7]</sup>。假定历史规则云为  $C_2(E_{x_2}, E_{n_2}, H_{e_2})$ , 则合并后的综合云为  $C(E_x, E_n, H_e)$ 。

$$E_x = \frac{E_{x_1} E_{n_1} + E_{x_2} E_{n_2}}{E_{n_1} + E_{n_2}} \quad (1)$$

$$E_n = E_{n_1} + E_{n_2} \quad (2)$$

$$H_e = \frac{H_{e_1} E_{n_1} + H_{e_2} E_{n_2}}{E_{n_1} + E_{n_2}} \quad (3)$$

综合云通常用于概念提升,即将两个想念的概念综合为更一般的概念,在这里我们将它实现惯性加权,综合两种不同的预测知识,综合历史云和当前云生成预测云。如图 4 所示。

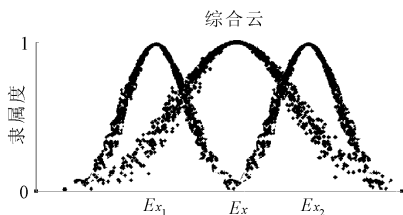


图 4 综合云、历史云与趋势云

### 3 径流非一致性研究

#### 3.1 年径流非一致性研究

岷江流域的水系和国家气象站点的分布情况如图 5。

从小波分析<sup>[8]</sup>中可以得出,降水和径流存在一定的相关性,可以得出在天然期内,对径流影响最大的是降水量的改变。为了得出从何时起人类活动开始对径流的突变产生显著影响。分析降水和径流的相关性随时间的变化,找出突变点即人类活动影响的介入点。选取岷江流域平均年水量和高场控制站径流量的相关性做为研究对象,以 12 年为一个周期做降水量和径流量相关系数递推变化(1977—1988,

1978—1989……1999—2010)如图 6。

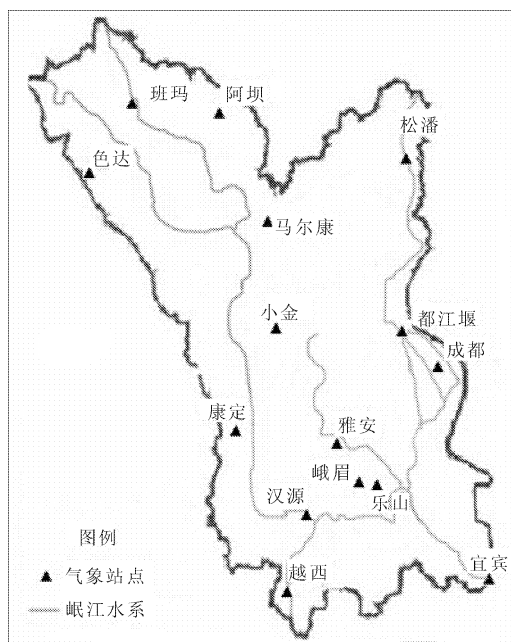


图 5 岷江流域水系图

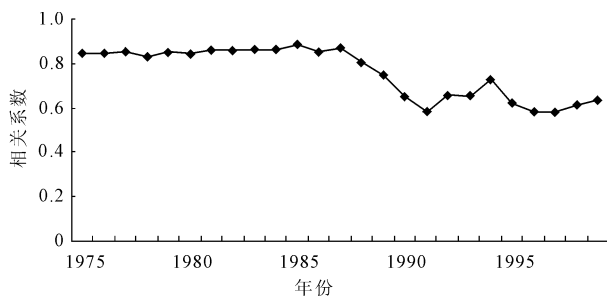


图 6 降水径流相关系数图

如图 6 中所示年降水和径流的相关性在 1998 年发生了显著的突变,另计算出变异前 1951—1997 时段的年径流均值为 636.5 mm,变异后 1998—2010 时段为 565.2 mm,两时段均值的绝对变化为 71.3 mm,说明岷江流域的年径流量变化比较显著。

对岷江流域分别绘制出变异点 1998 年前后的降雨径流关系线如图 7 所示。

从图 7 可以看出,变异前后的降水径流曲线存在着很大的差距,降水和径流的关系发生的较大的改变。经过分析降雨径流关系线为线性关系,则计算出变异前后的分别为相关方程为:

$$R_t = \begin{cases} 2.6382P_t + 102.69 & 1951 \leq t \leq 1997 \\ 3.6975P_t - 1006.7 & 1998 \leq t \leq 2010 \end{cases} \quad (4)$$

式(4)中径流  $R_t$  的单位是( $\text{m}^3/\text{s}$ ),降水的单位是(mm)。

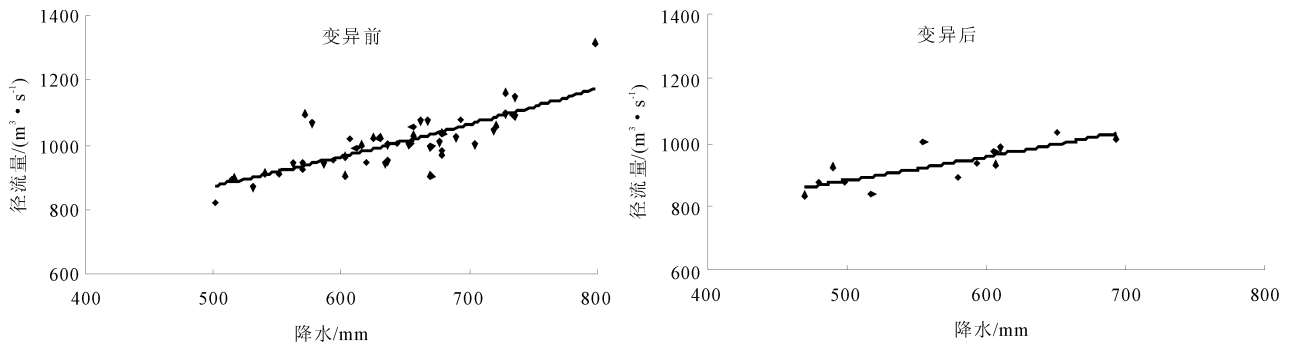


图 7 变异前后年降水径流关系曲线

### 3.2 确定性成份的提取

在岷江流域年径流序列变异点 1998 以前,径流时间序列是稳定的随机项,其确定性成份为零;对于变异点 1998 以后某年  $t$  的径流量,先通过变化前降水径流量方程,得出径流量  $R_1$ ,在查到其变化后的径流量  $R_2$ ,两者之差  $(R_2 - R_1)$  反映了水文序列变异以后确定性成分  $Y_t$ 。确定性成分  $Y_t$  的表达式为

$$Y_t = \begin{cases} 0 & 1951 \leq t \leq 2010 \\ 1.0507P_t - 1109.39 & 1998 \leq t \leq 2010 \end{cases} \quad (5)$$

### 3.3 随机性成分的提取

由于 1998 年以后的水文序列含有确定性成份,对于含有确定性成分的非一致性时间序列不能直接进行水文计算,须对其进行时间序列的分解,以求出相对一致的随机性成分和非一致的确定性成分。根据变化环境下非一致性年径流序列的水文频率计算原理,年径流序列,可以分解成确定性成分与随机性成分之和,因此根据式(4)和式(5)可得随机性成分为:

$$X_t = 2.6382P_t + 102.69 \quad 1951 \leq t \leq 2010 \quad (6)$$

### 3.4 非一致性水文序列的合成计算

根据研究问题的需要,对于非一致性水文序列,可以进行数值合成、参数合成以及分布合成<sup>[9]</sup>。而本文研究的是非一致性水文序列的预测问题,因此采用数值合成方法。可以分为三个步骤:一是根据确定性规律预测某个具体时刻的确定性成分;二是根据随机性规律预测某个时刻的随机性成分;最后进行数值合成,得到预测值。

选取岷江流域 1951—2007 年降水序列作为模型样本,用 2008—2010 年作为模型精度验证样本。令 1951—2007 年的年径流随机性成份数据为  $E = \{e_1, e_2, \dots, e_{57}\}$ 。将 1951—2006 年的年径流数据作为输入样本  $X$ ,1952—2007 年的数据作为输出样本  $Y$ 。即:

$$\begin{aligned} X &= \{x_1, \dots, x_{56}\} = \{e_1, e_2, \dots, e_{56}\} \\ &= \{2891.33, \dots, 2024.13\} \\ Y &= \{y_1, \dots, y_{56}\} = \{e_2, e_3, \dots, e_{57}\} \end{aligned} \quad (7)$$

$$= \{2826.92, \dots, 2232.22\}$$

利用  $(x_i, y_i)$  推导相邻年份之间的不确定性规则。将  $X$  中的数据从小到大排列,  $Y$  中的数据位置做相应改变。通过已有资料画出年径流 P - III 型曲线,根据频率将数据分为三类,  $Y$  中的数据做相应于  $X$  的分类,得到输入项分别为枯水年,平水年,丰水年的三组规则。

$$\begin{cases} (2161.08, 2218.75, \dots, 2490.92, 2531.33) \\ (2490.92, 2456.08, \dots, 2743.08, 2828.50) \\ (2563.92, 2600.83, \dots, 2891.33, 2918.42) \\ (2928.17, 2674.58, \dots, 3142.83, 2636.42) \\ (2924.67, 2928.17, \dots, 3173.50, 3444.08) \\ (2218.75, 2735.08, \dots, 2563.92, 3037.92) \end{cases} \quad (8)$$

令 1998—2007 年的年径流确定性成份数据为  $E = \{e_1, e_2, \dots, e_{10}\}$ 。将 1998—2006 年的年径流数据作为输入样本  $X$ ,1999—2007 年的数据作为输出样本  $Y$ 。即:

$$\begin{aligned} X &= \{x_1, \dots, x_{56}\} = \{e_1, e_2, \dots, e_9\} \\ &= \{-133.00, -45.07, \dots, -231.78\} \\ Y &= \{y_1, \dots, y_{56}\} = \{e_2, e_3, \dots, e_{10}\} \\ &= \{-45.07, -174.02, \dots, -229.97\} \end{aligned} \quad (9)$$

由于确定性成份数据长度较短,不进行分类得到一组云规则。整个预测模型如图 8 所示。

其中  $x_s$  代表随机性成份,  $x_y$  代表确定性成份,  $X$  代表最终的年径流预测值。

## 4 方法对比

本文还利用神经网络、灰色预测与云模型进行比较,各种方法的计算结果如表 1 所示。

以平均相对误差和作为评价算法优劣的标准。即:

$$E_{rr} = \frac{1}{3} \sum_{i=1}^3 \frac{|\hat{e}_i - e_i|}{e_i} \quad (10)$$

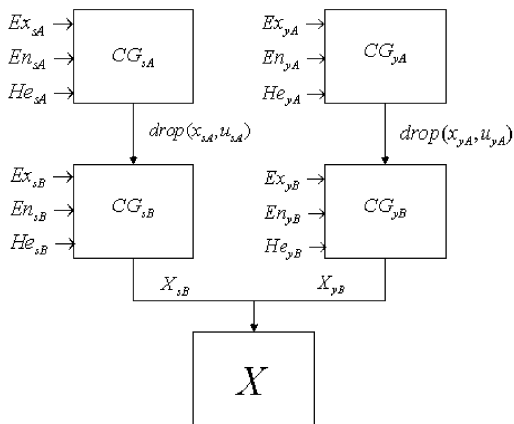


图 8 非一致性径流云预测模型

表 1 不同预测方法结果对比

模型	年份		
	2008	2009	2010
实测值	2113	2149	2393
非一致性云预测	2179	2020	2128
灰色预测	2203	1952	1983
BP 神经网络	2378	2365	2579

## 5 结 论

神经网络、灰色预测的误差分别为 10.1%、10.2%。由于云推理是一种不确定的推理，因此每次计算的精度不尽相同，但其结果必在某一范围内。进行多次计算得出期望值的误差为 6.7%。BP 神经网络算法存在局部极小点问题，得到的计算结果可能不是全局最优解，而且受神经网络学习、训练的随机性影响，数据最终不收敛或者只是达到局部

收敛；灰色预测通过对数据进行累加处理，发现数据整体趋势，但它并未体现当前的变化趋势。

本文建立的岷江流域年径流量非一致性云预测模型。不但综合考虑了当前变化的趋势以及相邻数据之间的关系，而且分别考虑历史年份径流量随机性成分和确定性成份各自的关系，最大的减少了历史数据中不同成份相互之间的干扰，挖掘了历史数据的规律。结果表明此方法得出的预测值比起传统方法预测模型，具有较高的精度，满足了预测需要。

## 参考文献：

- [1] 王思鹏. 基于云模型的识别技术[J]. 计算机技术与发展, 2006, 16(6): 187-190.
- [2] 李德毅, 杜 鹞. 不确定性人工智能[M]. 北京: 国防工业出版社, 2005.
- [3] 刘德地, 陈晓宏, 楼章华. 基于云模型的降雨时空分布特性分析[J]. 水利学报, 2009, 40(7): 850-857.
- [4] 高文荣, 段春青, 陈晓楠, 等. 基于云推理的年降水预测模型[J]. 中国农村水利水电, 2009, (3): 32-35.
- [5] 陈 昊, 李 兵. 云推理方法及其在预测中的应用[J]. 计算机科学, 2011, 38(7): 209-211, 224.
- [6] 于少伟, 曹 凯, 赵 模. 基于一维云模型的交通信息预测算法[J]. 山东大学学报(工学版), 2007, 37(2): 121-126.
- [7] 蒋 嵘, 李德毅, 陈 晖. 基于云模型的时间序列预测[J]. 解放军理工大学学报, 2000, 1(5): 13-17.
- [8] 王文圣, 丁 晶, 衡 彤, 等. 水文序列周期成分和突变特征识别的小波分析法[J]. 工程勘察, 2003, (1): 32-35.
- [9] 谢 平, 陈广才, 夏 军. 变化环境下非一致性年径流序列的水文频率计算原理[J]. 武汉大学学报(工学版), 2005, 38(6): 6-9, 15.