

非方程灰色系统方法 在长期水文预报中的应用初探

李正最 吴雅琴

(湖南湘潭水文水资源勘测大队 湘潭 411100)

提 要 基于水文过程的复杂性和影响因子信息收集不完善这一基本事实,本文将流域水文系统视为含有灰元和灰信息的灰色系统,并初步实践了非方程灰色预报方法在长期水文预报中的应用。方法在一定意义上脱离了传统的以方程为中心的预报模式,它既考虑预报因子对预报量的不同影响程度,又不过分强调预报量与预报因子的具体相关函数形式,为长期水文预报提供了一种新思路。文中列出了实例,效果令人满意。

关键词 水文预报 非方程灰色系统 预报因子 流域水文系统

1 前 言

现行的各种长期水文预报方法已经形成了一个很大的方法库。但其中所有的定量预报方法的共同特点都是围绕建立预报量与预报因子或时间的某种函数关系而展开研究的,即以方程作为预报问题的核心。事实上,流域水文系统有较强的非线性动力学特征及其高度的复杂性。由于水文过程的不可人工控制,各个水文环节也不都是可观测或者还没有能够全部被观测到,并且样本系列不长,所以我们很难占有充分的信息,也往往不能彻底弄清预报量的运行机理。因而无法确切地确定预报方程的形式。退一步说,即使是知道了方程的形式,如果它的本质是非线性的,现有的计算技术也常常难以通过历史数据回归出可供预报作业的方程来,所以,迄今长期水文预报的难度很大,以致现行的长期水文预报无论在方法上还是在结果上都不同程度地带有主观性。这就要求我们离开以方程为中心的传统思维方式,研究非方程定量预报方法。

按照灰色系统理论的观点,部分信息已知,部分信息未知或非确知的系统称为灰色系统。基于水文过程的复杂性和影响因子信息收集不完善这一基本事实,将流域水文系统看成含有灰元和灰信息的“信息不完全”系统是合乎实际的。因此,运用灰色系统理论的学术思想,采取求实、宽容和综合分析的态度,尽可能少地验前强加约束和假定,可能是长期水文预报在现阶段比较实用和可行的途径和方法。本文既考虑水文过程形成的物理背景,又不强求预报量与预报因子之间具体的相关函数形式,通过对预报因子数据列的几何比较进行类比性预测。当收集到待预测状态的预报因子状态值后,借助关联度分析手段去寻找与这组状态值最贴近的一个历史数据类,再根据该历史数据类的预测量特征作出预报。经过实例检验,预报效果令人满意。

2 非方程灰色预报方法的基本原理

流域水文系统是一个有原型的物理系统。水文现象的发生不是孤立的,尽管水文现象形成

的物理机制和过程十分复杂,但是根据灰色系统白化的观点,采取多途径开发和获取,如物理成因考查和统计分析等手段,可以筛选出与流域特定时期水文要素预报量密切相关的前期影响因子。虽然有些信息特别是系统内部的物理环节没有完全观测到,已观测到的各种因子中有些因子对预报量的影响机制比较模糊,甚至具有某种意义上的不确定性。但是可以认为从“前期影响因子”到“预报量”是一个由“灰”到“白”的映射,水文过程的形成正是各种相关因子综合作用的效应。

设每个预报量的历史状态 $y_t(t = 1, 2, \dots, T)$ 对应一个相应的前期影响因子的数据列 x_t , x_t 由 m 个通过筛选的对 y_t , 影响较大的前期因子构成,即:

$$\begin{cases} x_1 = [x_1(1), x_1(2), x_1(3), \dots, x_1(m)] \\ x_2 = [x_2(1), x_2(2), x_2(3), \dots, x_2(m)] \\ \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \\ x_T = [x_T(1), x_T(2), x_T(3), \dots, x_T(m)] \end{cases} \quad (1)$$

其中 T 表示历史系列长度。

y 与 x 之间的关系显然是未知的,具有“灰”特征。灰色系统理论允许这种“灰”关系的存在,认为 y 是 x 的白化映射,且存在以下对应关系:

$$\begin{cases} y_1 \Leftrightarrow x_1 \\ y_2 \Leftrightarrow x_2 \\ \dots \quad \dots \quad \dots \\ y_T \Leftrightarrow x_T \end{cases} \quad (2)$$

式(1)、(2)综合反映预报量与相关因子相互关系中 T 个不同的发展变化态势,是非方程灰色预报的 T 个模式。

众所周知,水文过程作为地球物理过程的一部分,常常表现出一定程度的周期性特点,其可预报性正是和这种周期性变化相联系的。由于水文过程的周期性,使未来某一时刻的水文要素状态与历史上某一状态具有某种意义上的相似。换言之,某一历史水文状态可能在未来“重现”。所以当 T 期观测数据的信息量包含了流域可能出现的各种典型水文态势后,第 $s(s > T)$ 期的水文态势即可能是已观测到的 T 个态势之一的“重现”。

如果第 s 期的相关因子状态是一个确定的点

$$x_s = [x_s(1), x_s(2), x_s(3), \dots, x_s(m)] \quad (3)$$

只要能够确认 x_s 的发展态势与 T 期数列中某一数列 $x_k(k \in [1, T])$ 相类似,则认为 x_s 对应预报状态量 y_k (因尚待预报故标记为灰数 \otimes):

$$\otimes_s \Leftrightarrow x_s \quad (4)$$

是历史状态 Y_k 的“重现”,从而可以相应作出预测

$$\hat{\otimes}_s = Y_k \quad (5)$$

3 非方程灰色预报方法所涉若干技术手段

3.1 预报因子的挑选

预报因子的挑选主要通过物理考察和统计考察相结合途径实现。即首先分析预报量与预

报因子之间的物理联系,分析它们之间的因果关系,从大量因素中挑选出一批具有一定物理意义的可能因子,然后进步考察挑选出来的因子是否符合数理统计的有关原则,从而在可能因子中筛选出部分有一定预报价值,且符合统计学原则的有效因子。具体方法可参阅有关中长期水文预报专著,本文不拟详述。

3.2 预报因子的预处理

因经过筛选后构成数列 x_t 中的 m 个因子的物理量纲和数值范围是不同的,为了消除由于因子量纲和量级不同的影响,故需对预报因子数据进行预处理。有如下两种方法可供选用。

(1) 中心化预处理法。设

$$\bar{x}(j) = \frac{1}{T} \sum_{t=1}^T x_t(j) \quad , s(j) = \sqrt{\frac{1}{T} \sum_{t=1}^T [x_t(j) - \bar{x}(j)]^2}$$

则原始数据可据下式进行变换

$$x'_t(j) = [x_t(j) - \bar{x}(j)]/s(j) \quad j = 1, 2, \dots, m; t = 1, 2, \dots, T \quad (6)$$

(2) 极差化预处理法。令

$$x(j)_{\min} = \min_{1 \leq t \leq T} \{x_t(j)\}, x(j)_{\max} = \max_{1 \leq t \leq T} \{x_t(j)\}$$

极差化预处理公式为

$$x'_t(j) = \frac{[x_t(j) - x(j)_{\min}]}{[x(j)_{\max} - x(j)_{\min}]} \quad j = 1, 2, \dots, m; t = 1, 2, \dots, T \quad (7)$$

为行文方便,经预处理后的预报因子数据仍记为 $x_t(j)$ 。

3.3 模式归判的灰色关联度分析途径

鉴于系统信息的不完全,关系不明确,灰色系统理论提出了灰关联度分析途径,并通过关联度的大小比较反映事物在发展过程中的动态整体接近程度。由于式(1)、(2)综合反映了水文预报量及其相关因子相互间关系和变化态势的 T 个不同的模式,因此待预测状态 ($t=s$) 究竟与历史状态中的哪一个模式最贴近,可以通过因素数据列 x_t 与 x_1, x_2, \dots, x_T 的关联度分析来识别。关联度计算公式为:

$$r_t = \sum_{j=1}^m \left[\frac{A_{\min} + \rho \cdot A_{\max}}{A_t(j) + \rho A_{\max}} \cdot W(j) \right] \quad (8)$$

式中 r_t —— 因素数据列 x_t 与 x_t 的关联度;

$A_t(j)$ —— $x_t(j)$ 与 $x_t(j)$ 的绝对差, $A_t(j) = |x_t(j) - x_t(j)|$;

A_{\min} —— 各因子间的二级最小绝对差,

$$A_{\min} = \min_i \min_j |x_i(j) - x_t(j)|;$$

A_{\max} —— 各因子间的二级最大绝对差,

$$A_{\max} = \max_i \max_j |x_i(j) - x_t(j)|;$$

ρ —— 分辨系数, $\rho \in [0, 1]$, 一般取 $\rho = 0.5$;

$W(j)$ —— 预报因子 $x(j)$ 的权重, 满足 $\sum_{j=1}^m W(j) = 1$ 。

将关联度 r_1, r_2, \dots, r_T 按由大到小排队获得关联序。关联序说明了因素数据列 x_t 与历史数据列的接近程度的大小顺序。若 $r_k > r_p$

则说明 x_t 与 x_k 在发展态势上的接近程度较之 x_t 与 x_p 的接近程度更大。当第一关联度

$$r_k = \max_{1 \leq i \leq T} \{x_i\} \quad (9)$$

即可推断在发展态势上 x_i 与 x_k 最贴近, 即形成 y_i 的物理成因可归判为历史数据 T 个模式中的

$$y_i \leftrightarrow x_k = [x_k(1), x_k(2), x_k(3), \dots, x_k(m)]$$

模式。从而可以利用式(5)作出类比性推理预测。

3.4 预报因子权重的确定

前已指出, 预报量 y_i 是因子数据列 x_i 中诸元素综合作用的效应, 但各因子对预报量的影响程度并不相同, 也就是说预报因子具有不同的权重。因此, 在运用灰色关联度分析进行模式识别决策时如何合理地考虑各个预报因子的权重是十分重要的问题。

通常情况下, 预报量与预报因子的关系越密切, 则该因子的权重越大。如近似地认为预报量与预报因子之间显线性相关, 则可以建立以下相关模型。

$$y = B(0) + \sum_{j=1}^m B(j) \cdot x(j) \quad (10)$$

式中 $B(0)$ —— 期望值;

$B(j)$ —— 第 j 个预报因子的偏回归系数。

当 $x(j)$ 取预处理后的数据时, $B(j)$ 的大小实际上反映了相应因子对预报量的贡献的大小。故预报因子的归一化权重可取为:

$$W(j) = |B(j)|/K \quad (11)$$

其中 $K = \sum_{j=1}^m |B(j)|$

其实当 $x(j)$ 为中心化预处理后数值时, 可以证明 $B(j)$ 就是 y 与 $x(j)$ 的相关系数。因此权重计算还可直接通过相关系数求得。即

$$\rho_j = \frac{\sum_{i=1}^n [x_i(j) - \bar{x}(j)][y_i - \bar{y}]}{\sqrt{\sum_{i=1}^n [x_i(j) - \bar{x}(j)]^2 \cdot \sum_{i=1}^n [y_i - \bar{y}]^2}} \quad (12)$$

$$W(j) = |\rho_j| \sum_{j=1}^m |\rho_j| \quad (13)$$

4 非方程灰色系统方法在长期水文预报中的应用

我们用此方法做了年最高水位(最大流量)和月径流量的长期预报, 限于篇幅现仅列两例。

4.1 石镇街站年最高水位预报

以该站 1960~1985 年为样本, 经过物理成因分析和统计考察后筛选以下四个因子作为预报因子(因子挑选过程从略):

$x(1)$ —— 当地当年 1 月平均气温;

$x(2)$ —— 前一年 10 月至当年 1 月太平洋 —— 欧洲 W 型出现日数之和;

$x(3)$ —— 石镇街站前一年 10 月至当年 1 月降水量之和;

$x(4)$ —— 前一年乌拉尔山 1 月 500 hPa 平均图上 9 时平均高度。

然后试报 1986、1987、1988 共三年的年最高水位并与实测值比较, 据以检验预报效果。

该站 1960~1988 年各年最高水位及相应预报因子值(已经预处理)详见表 1。为求得各因

子的权重,以 1960~1985 共 26 年的资料建立年最高水位 y 与各预报因子的多元线性回归方程后,进而求得:

$$W(1) = 0.149; W(2) = 0.236; W(3) = 0.214; W(4) = 0.272。$$

表 1 石镇街站历年最高水位及预报因子序列

年份	年最高水位 $y(m)$	$x(1)$	$x(2)$	$x(3)$	$x(4)$	年份	年最高水位 $y(m)$	$x(1)$	$x(2)$	$x(3)$	$x(4)$
1960	17.75	0.618	0.096	0.451	0.976	1975	21.48	0.891	0.385	0.558	0.440
1961	19.01	0.618	0.173	0.440	0.420	1976	21.32	0.564	0.115	0.670	0.504
1962	20.88	0.382	0.519	0.352	0.808	1977	21.10	0.000	0.250	0.412	0.532
1963	18.12	0.218	0.000	0.176	0.624	1978	18.71	0.745	0.480	0.385	0.628
1964	19.48	0.891	0.346	0.731	0.808	1979	20.52	0.818	1.000	0.190	0.704
1965	20.77	1.000	0.481	0.000	0.000	1980	20.57	0.655	0.346	0.209	0.644
1966	21.68	0.945	0.904	1.000	0.796	1981	18.94	0.418	0.343	0.203	0.536
1967	23.22	0.491	0.673	0.368	0.396	1982	20.59	0.800	0.481	0.464	0.936
1968	20.55	0.636	0.615	0.275	0.504	1983	22.70	0.600	0.846	0.478	0.408
1969	21.42	0.655	0.577	0.775	0.228	1984	19.56	0.091	0.231	0.272	0.480
1970	21.52	0.327	0.519	0.195	0.340	1985	18.52	0.691	0.481	0.225	1.000
1971	22.74	0.509	0.769	0.505	0.568	1986	18.84	0.636	0.231	0.429	0.404
1972	19.50	0.873	0.558	0.137	0.752	1987	21.22	0.982	0.212	0.184	0.752
1973	22.76	0.709	0.596	0.681	0.048	1988	20.55	0.709	0.615	0.275	0.296
1974	22.70	0.436	0.615	0.319	0.136						

据表 1,1960~1985 年所构成的 26 个基本预报模式为:

$$[y_1 = 17.75] \Rightarrow x_1 = [0.618, 0.096, 0.451, 0.976]$$

$$[y_2 = 19.01] \Rightarrow x_2 = [0.618, 0.173, 0.440, 0.420]$$

... ..

$$[y_{26} = 18.52] \Rightarrow x_{26} = [0.691, 0.481, 0.225, 1.000]$$

以灰数 \otimes_{27} 表示 1986 年的年最高水位,则 1986 年可构成模式

$$[y_{27} = \otimes_{27}] \Rightarrow x_{27} = [0.636, 0.231, 0.429, 0.404]$$

分别计算 x_{27} 对 x_1, x_2, \dots, x_{26} 的关联度。表 2 所列为关联度计算结果

表 2 数据列 x_{27} 与 x_i 的关联度计算结果

序号 i	1	2	3	4	5	6	7	8	9	10	11	12	13
关联度 r_i	0.724	0.928	0.609	0.602	0.624	0.532	0.436	0.727	0.699	0.633	0.657	0.631	0.554
序号 i	14	15	16	17	18	19	20	21	22	23	24	25	26
关联度 r_i	0.582	0.606	0.759	0.754	0.814	0.700	0.506	0.727	0.714	0.635	0.733	0.807	0.598

因此有:

$$r_2 = \max_{1 \leq i \leq 26} (r_i) = 0.928$$

从而可知以作出预报: $\otimes_{27} = y_2 = 19.01 (m)$

实况为 18.84m,与预报值相差 0.17m。根据部颁《水文情报预报规范》预报误差小于历史变幅的 20%即为合格。由表 1 中 1960~1985 年的年最高水位变化范围 17.75~23.22m 得历年变幅为 5.47m,允许预报误差 $\pm 1.09m$ 。故 1986 年预报结果合格。

同理对 1987,1988 年进行试报,试报结果一并列入表 3,预报合格。

4.2 龙羊峡水库入库站——贵德站 10 月平均流量预报

经挑选从以下 4 个因子作为预报因子,即

$x(1)$ —— 贵德站 7 月份平均流量系列;

$x(2)$ —— 该站以上流域 8 月份平均降水量;

$x(3)$ —— 亚洲地区 ($45^{\circ} \sim 60^{\circ}N, 60^{\circ} \sim 150^{\circ}E$)500hPa, 8 月份纬向环流指数;

$x(4)$ —— $40^{\circ}N, 130^{\circ}E$ 500hPa 3 月份平均位势高。

并以 1959 ~ 1982 共 24 年 10 月份平均流量和 $x(1), x(2), x(3), x(4)$ 等 4 个主要预报因子(经预处理后数值,具体资料详见文[7])构成 24 个基本预报模式,然后对 1983、1984、1985 年各年 10 月平均流量进行预报。预报结果详见表 4。可见预报误差均小于 20%,预报全部合格。

表 3 石镇街站年最高水位预报检验

年份	实测年最高水位(m)	最大关联度	关联年	预报值(m)	误差(m)	检验结论
1986	18.84	0.928	1961	19.01	0.17	合格
1987	21.22	0.804	1980	20.57	-0.65	合格
1988	20.55	0.872	1968	20.55	0	合格

表 4 贵德站 10 月份平均流量预报效果检验

年份	实测流量(m ³ /s)	最大关联度	关联年	预报值(m ³ /s)	误差(%)	检验结论
1983	1960	0.7532	1975	1957	-0.2	合格
1984	893	0.7675	1965	824	-7.7	合格
1985	1170	0.8728	1979	968	-17.3	合格

5 讨 论

非方程灰色系统预报方法是建立在历史资料基础上的类比性预测,因此它要求历史系列资料具有代表性和典型性,备选样本应能比较全面地概括预报流域可能出现的各种水文态势。如果某种水文态势在备选样本中未能包含进去就可能产生误报、错报。

同现行的各种方程预报方法一样,预报因子的挑选也是非方程灰色系统预报方法成功预报的关键。本文对预报因子的挑选论述不多。但在实际应用中应充分注意各因子与预报量的物理联系。由于各种预报因子对预报量的影响程度是不同的,因此预报因子的权重计算也是值得深入研究的问题。实践证明本文在考虑权重时接受了方程预报方法中的有关信息,将方程预报方法与非方程预报方法相结合,对于提高预报精度起了一定的作用。但是,由于文中的计算公式是基于预报量与预报因子呈线性相关的假定提出来的。当预报量与预报因子之间呈非线性关系时,采用线性相关的权重计算公式必然对模式识别产生影响,从而影响预报效果。

6 结 语

1 非方程灰色系统预报方法改变了长期水文预报中传统的以方程为核心的预报模式,在预报中它通过权重考虑不同的预报因子对预报量的不同影响程度,又不苛求预报量与预报因子的具体相关函数形式,处理方式灵活,应用方便。其最大特点是既符合水文过程的物理特性,又可得到一定精度的预报结果。实例表明这种处理途径是基本成功的。

2 灰色系统理论及其在长期水文预报中的应用,在我国已经过了近十年的实践。但目前主要仍局限于灰色 GM(1,1)模型,而且应用效果时好时坏。而本文提出的基于灰色系统理论关联度分析原理的非方程灰色预测,不仅为长期水文预报提供了新的思路,而且在一定意义拓展了灰色系统理论关联度分析原理的应用范围。它不仅适用于长期水文预报,对于类似其它领域,诸如气象、地震的中长期预报也有一定的参考意义。

3 作为对长期水文预报方法的探讨,有许多问题尚待作更进一步的工作,这些问题包括

如何对备选样本中尚未包含的小概率事件(如特枯、特丰水情)进行预报,如何根据预报量与预报因子的不同关系合理地赋予权重等等。

参 考 文 献

- 1 邓聚龙. 灰色系统基本方法. 华中理工大学出版社, 1987
- 2 邓聚龙. 灰色系统理论教程. 华中理工大学出版社, 1990
- 3 夏 军. 水文尺度问题及灰色系统研究思想. 2000年中国水文展望. 河海大学出版社, 1991
- 4 李正最. 灰色系统理论与水文水资源研究. 2000年中国水文展望. 河海大学出版社, 1991
- 5 李正最. 长期径流预报的灰色关联决策模型. 水电能源科学, 1991年第4期
- 6 冯耀煌. 优化的非线性预报方法在中长期水文预报中的应用. 水文, 1990年第2期
- 7 杨百银. 龙羊峡电站下闸蓄水水情长期预报的模糊聚类方法. 水文, 1989年第2期
- 8 水利电力部. 水文情报预报规范. 水利电力出版社, 1985

Primary Study on Application of Non-equation Gray System Method in The Longterm Hydrologic Forecasting

Li Zhengzui Wu Yaqin

(Hydrology and Water Resources Team at Xiangtan, Hunan)

Abstract Because the complex of hydrological process and unperfected information of influence factor, hydrologic system in the basin is regarded as the Gray System with gray element and gray information. The non-equation gray forecast method is primarily applied in longterm hydrologic forecasting. The method separates itself from the traditional forecasting mode which takes equation as the center. The different effects of forecast factors on forecast estimation are considered, but the concrete form relative functions for forecast estimation and forecast factors are not stressed strongly. It is a new concept for longterm hydrologic forecasting. The examples are set in the paper. The good results are achieved.

Key Words hydrologic forecasting, non-equation Gray System, forecast factor, hydrologic system in a basin.

规划专委会全委会议召开

水能规划及动能经济专业委员会于1993年12月21日在成都召开了全委会议。会议由专委会主任委员晏志勇同志主持, 14位委员出席了会议。

会上, 张登仕同志对专委会1993年的工作进行了总结。他介绍了专委会目前的组织发展情况, 1993年进行的八项学术活动内容。晏志勇同志向会议介绍1993年12月在广州举行的中国水力发电工程学会水能规划及动能经济专委会年会的情况、“西电东送”科研课题的进展情况。会议学习了三届三次常务理事(扩大)会议纪要。根据纪要精神, 商讨了专委会1994年的学术活动计划。委员们决心在1994年中为四川省水电建设做出更大的贡献。

(成勘院 张登仕)