

基于分形理论的需水量预测方法

陶涛, 刘遂庆

(同济大学环境科学与工程学院, 上海 200092)

摘要: 在研究有关分形理论的基础上, 提出了基于分形拼贴定理和分形插值函数迭代生成过程的需水量预测新方法. 该方法根据分形拼贴定理, 求取一个与历史记录相近的吸引子的迭代函数系统来求取分形插值函数; 通过分形插值函数的迭代生成, 建立预测模型. 并以实例证实该方法的合理性及可靠性.

关键词: 需水量; 预测; 分形理论; 吸引子; 迭代函数系统

中图分类号: TU 991.31

文献标识码: A

文章编号: 0253-374X(2004)12-1647-04

Water Demand Forecasting Method Based on Fractional Theory

TAO Tao, LIU Sui-qing

(College of Environmental Science and Engineering, Tongji University, Shanghai 200092, China)

Abstract: Based on the introduction of fractional theory, a new method of forecasting water demand is presented in this paper. It uses the iterated function system whose attractor is close to the historical data to establish the fractional function, and according to the iterative process of fractional function to set the forecasting model. The example shows the method is rational and reliable.

Key words: water demand; forecast; fractional theory; attractor; iterated function system

预测, 是根据以往的实际数据资料, 应用现代的科学理论和方法、丰富的经验和敏锐的判断力, 去探索人们所关心的事物在今后的发展趋势, 做出分析和估计, 以指导未来的行动, 把未来事件的不确定性最小化. 需水量预测是供需管理方面的基础工作之一, 正确地预测需水量, 使预测误差尽可能地小, 对区域的社会、经济和环境的协调发展, 水资源供需分析以及重大水资源工程(如引江工程等)的正确决策和实施, 乃至对市场经济条件下的用水管理, 都有着重大的意义^[1].

1977年以来, 随着著名数学家 B. B. Mandelbrot 几本专著的出版^[2], 分形理论即告诞生. 分形理

论可以有效地描述在自然界和科学研究中遇到的各种各样复杂的图形, 即便一些看似毫无规则的自然形状, 分形理论也可以通过严谨而有效的定量方法加以处理. 在数学中, 标准的分形可以和现实世界中的现象结合起来: 现实中许多事物的复杂性和随机性往往是由简单事物的大量重复产生出来的, 这与分形图形借助计算机以递归和迭代算法形成的过程相类似, 因此, 分形绘图方法可以模拟现实世界中的复杂和“随机”现象^[3].

目前的预测方法多以回归分析为主, 需要适当考虑影响因子的作用. 影响需水量变化的因子很多, 诸如气候变化、节假日等等, 为分析的准确性, 一般

收稿日期: 2003-11-28

基金项目: 国家自然科学基金青年基金资助项目(50409016); 水资源与水电工程科学国家重点实验室开放基金资助项目(2004B012); 中欧科技合作第五框架资助项目(ICA4-CT-2002-10004)

作者简介: 陶涛(1974-), 女, 重庆人, 讲师, 工学博士. E-mail: taotao@mail.tongji.edu.cn

总以为考虑的影响因子越多越好.事实上,因子的分析与评价本身就带有许多的不确定性,并且诸多的影响因子是难以用数学表达式表征的,这在一定程度上会对利用该法预测需水量带来不确定性.为此,本文提出基于分形理论的需水量预测方法:用插值于历史用水量记录点的分形插值函数来模拟需水量的动态随机演变过程,在分形插值函数的迭代生成中对未来的需水量做出预测.这为变化规律难以把握或只能给出一些不确定信息的需水预测提供了新的思路和方法.

1 方法与预测模型

1.1 迭代函数系统及其吸引子定理

到目前为止,用一个数学系统去解析地构造、研究一类存在于人为或自然的具有“自相似”(比例自相似)结构的分形,最为成功的就是迭代函数系统(iterated function system, IFS).它既可包含确定性过程又可包含随机性过程.一个迭代函数系统是由一个完备距离空间 (X, d) 及其上的紧度量空间到自身的一组收缩映射 $w_i: X \rightarrow X$ 组成, w_i 的收缩因子为 S_i ,且 $0 \leq S_i < 1, i = 1, 2, \dots, N$.系统表示成 $\{X; w_i, i = 1, 2, \dots, N\}$,且系统的收缩因子为 $S = \max\{S_i, i = 1, 2, \dots, N\}$.一般用收缩仿射变换来表示这些映射,则 IFS 的形式为

$$\left\{ X \mid w_i \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix}, i = 1, 2, \dots, N \right\} \quad (1)$$

设 $W(B) = \bigcup_{i=1}^N w_i(B) = w_1(B) \cup w_2(B) \cup \dots \cup w_N(B)$ 对所有的 $B \in K(X)$ 也是一个收缩变换,并有收缩因子 S ,即 $h(w(B), w(C)) \leq Sh(B, C)$ 对所有 $B, C \in K(X)$ 成立,且在 $K(X)$ 上必存在唯一的不动点 A ,即满足 $A = W(A) = \bigcup_{i=1}^N w_i(A)$;并且 A 可由下式给出,即对 $\forall B \in K(X)$,有

$$A = \lim_{n \rightarrow \infty} W^{0n}(B) \quad (2)$$

式中, $W^{0n}(B)$ 表示变换 w 的 n 次复合,即 $W^{0n}(B) = w(w(\dots w(B)))$, A 即为该 IFS 的吸引子.

1.2 分形插值

所谓分形插值,就是构造一个 R^2 平面上的 IFS,使其吸引子恰为插值于给定数据集 $\{(x_i, F_i):$

$i = 1, 2, \dots, N\}$ 的连续函数 $f: [x_0, x_N] \rightarrow R$ 的图像.它的意义就是找出的插值函数在 H 距离下近似于数据,尤其可以保证分形插值函数的分数维在一定的比例范围内相应于数据.这是区别于传统插值方法的根本所在.分形插值说明,可以用 IFS 去表达一个连续函数 $f(x)$.这个 IFS 借助的是分形拼贴定理.业已证明,由此方法求取的迭代函数系统总有唯一吸引子,且该吸引子必定是某个连续函数的图形,并同时通过各插值点.这个连续函数即为分形插值函数.利用分形拼贴定理构造分形插值函数过程详见文献[4,5].

1.3 分形插值需水量预测方法

分形插值函数的构造仍然用拼贴定理解决,但此时 d 将在 $[0, 1]$ 区间任意取值,即 d 已不再取定.因此,模拟需水量演变过程的不是由某一个,而是由无数多个分形插值函数共同作用了.从操作行为看,分形插值有别于分形拼贴定理之处在于,拼贴定理是各窗口图形的连接,即按图形的连续性进行连接,而分形插值是局部点的递归迭代.应用分形插值函数模拟需水量随机演变过程预测需水量的步骤一般可分为3步:用水量预估,分形插值函数的求取,分形插值函数的迭代生成及未来时刻需水量预测.

1.3.1 用水量预估

为使所构造的分形插值函数能迭代收敛到未来时刻需水量,需事先给待预测时刻需水量一个预估值.在供水系统需水预测中,预测的主要是月需水量和中长期需水量.

(1) 月需水量 它与其相关时刻值存在相关性,是上一年同月需水量在一定范围内的波动值.当各时段间的联系较稳固时,这个范围将很小,相反则较大,从而难以准确定位.但是这个变换范围的上、下限值从历史用水量中可以轻松得到,用预测月上一同月的需水量加上变化区间的上、下限就可得到该月的2个预估值;分别赋予它们一定的权重值,参加分形插值函数生成的迭代运算.

(2) 中长期需水量 表征中长期需水量增长特性最直观的参数就是需水量增长率,预测年度的需水量预估值可以以当年用水量乘以增长率而求得.如认为将来需水量有一个快速增长期,增长率可取历史上较大值,否则取较小值;如无法估计,则可取平均值.获得预估值后即能参加分形插值函数的生成迭代运算.在预估未来时刻需水量时,有时会因为多种考虑而得到多个可能值,则可以根据预估人员的经验对多个预估值分别赋予不同的权重来参加随

机迭代运算.

1.3.2 分形插值函数的求取

用需水量预估值构造分形插值函数,即构造一个能够迭代收敛到插值点的迭代函数系统

$$w_i \begin{bmatrix} x \\ F \end{bmatrix} = \begin{bmatrix} a_i & 0 \\ c_i & d_i \end{bmatrix} \begin{bmatrix} x \\ F \end{bmatrix} + \begin{bmatrix} e_i \\ f_i \end{bmatrix} \quad (3)$$

求出 IFS 中的仿射变换参数,即确定一个分形插值函数.参数求取过程中, d 值为 $[0,1]$ 区间随机数,对应 d 的每一个随机取值存在一个分形插值函数.

1.3.3 分形插值函数的生成及需水量预测

为模拟需水量在满足收敛于历史用水量插值点约束条件之外的随机性,对每个分形插值函数仅取一个点,每一个点在 IFS 作用下生成的过程,就是分形插值函数生成的过程.这个过程即用于模拟需水的动态随机变化形态.选取一个以未来时刻点为中心的邻域,经迭代进入此邻域范围内的点所对应的需水量即可认为是未来时刻最可能的需水量;当有多个点通过时,可通过求均值或取大、小值或中值等运算来确定未来需水量.

2 实例验证

现将所建立的模型应用于河北省 1994 年各行政区域的需水量预测,来验证模型的合理性.采用本文所建立的基于分形理论的需水量预测模型,以 1980~1993 年的用水量及 1994 年的用水量预估值为基础资料,对河北省各行政区域 1994 年用水量进行预测,与实际用水量进行比较.预测结果见表 1 和图 1.

表 1 河北省各行政区 1994 年用水量预测
Tab.1 Water demand prediction of administrative area, Hebei province in 1994

市地	用水量/ 10^6 m^3			相对误差/%
	实际值	预测值	差值	
邯郸	2 247	2 282	-35	-1.55
邢台	1 908	1 945	-37	-1.94
石家庄	3 905	3 782	123	3.15
保定	3 256	3 135	121	3.72
衡水	1 353	1 372	-19	-1.40
沧州	1 391	1 306	85	6.11
唐山	3 217	2 976	241	7.51
廊坊	940	898	42	4.47
秦皇岛	990	1 004	-14	-1.42
张家口	1 199	1 174	25	2.09
承德	942	894	48	5.10
全省	21 348	20 768	580	2.72

从表 1 可以看出,用分形插值函数来预测需水量,预测值与实际值的差值基本上都在合理范围之内.但不可否认的是,预估值也直接影响到预测值,而本文为验证模型,所取预估值较为接近实际值,因此预测结果比较好,误差最大只达到 7.51%.在实际应用过程中,预估值的选取同样为随机状况的出现提供了处理的手段,即可以通过预估值的选取来表现预测时段可能出现的各种情况.这也显示了分形从无序中寻找有序的特点.

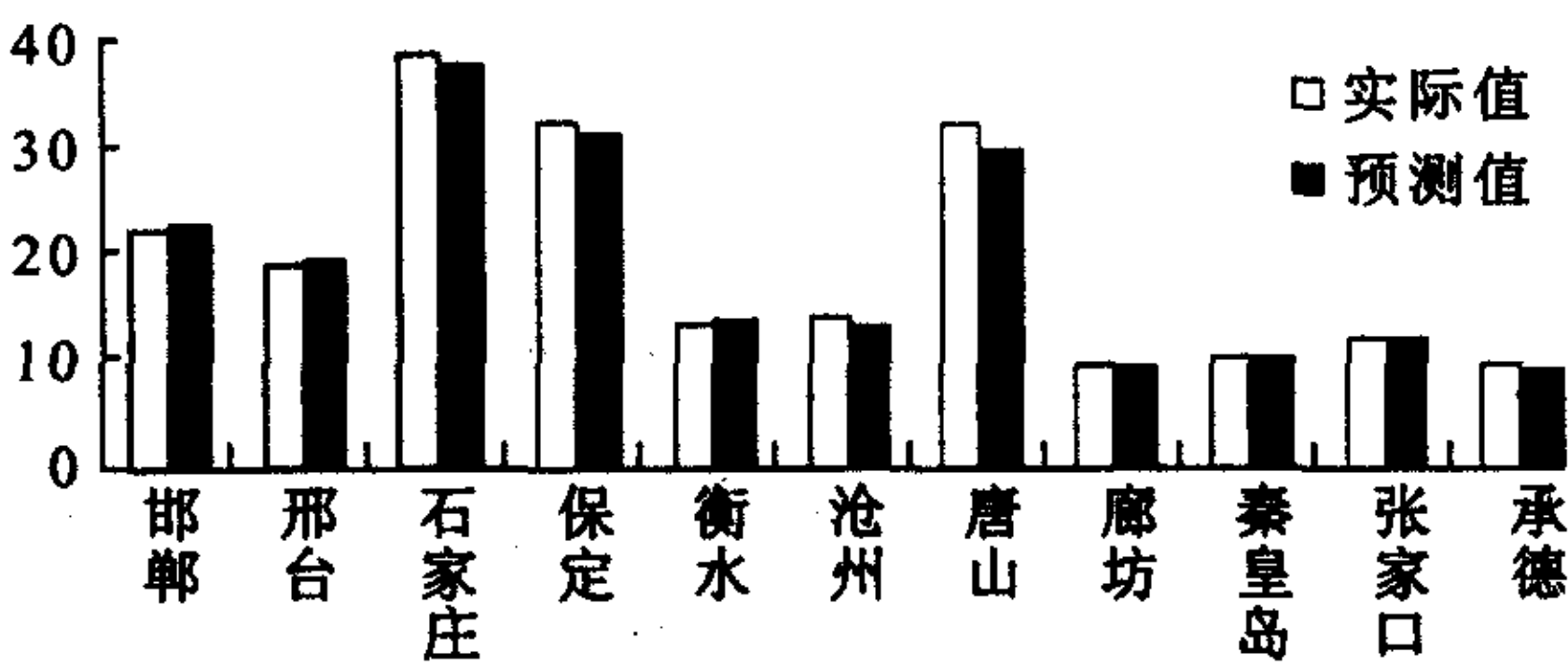


图 1 实际值与预测值的对照(单位:亿 m^3)

Fig.1 Comparison of real data and prediction

为此,再对河北省各行政区域各分区及各用水部门远景(2010 年、2020 年)需水量进行预测.预测结果如表 2 所示.

从目前世界需水的发展情况看,农村用水(特别是缺水地区)占总用水的比例一般都呈下降趋势,而工业用水和城镇生活用水增长迅速.这主要是因为城镇和工业用水的重要性要大于农业用水,往往会挤占农村用水.从历史资料得知,1990 年农业用水占总用水量的 83.3%,1993 年为 79.5%.而从表 2 可以算出,2010 年农业用水占总用水量的比例为 73.7%,2020 年为 69.2%.另外从区域发展的角度来看,工业越发达、生活水平越高的地区,工业和生活用水比例就越高,反之亦然.1990 年,石家庄城镇生活用水量占总用水量的 2.78%,而沧州为 1.72%;预计到 2010 年,将分别达到 13.42% 和 8.44%.

3 结语

本文对分形理论在供水系统需水量预测应用中的方法构造和运用作了一定的探索,证实了分形拼贴定理和分形插值函数的切实可行性.分形拼贴过程能使一些随时间序列具有一定演变特性的变量的基本关系固定下来,提取来对未来做出预测.分形插值函数通过插值于历史记录,保存演变过程中一些确定性的东西,由迭代生成过程得出在这种确定性基础之上的充分的随机行为,给预测带来了一定的

表 2 河北各行政区各部门远景需水量预测

Tab.2 Long-term prediction of water demand in administrative area, Hebei province

10⁶ m³

市地	2010 年				2020 年			
	工业	城镇	农业	总用水量	工业	城镇	农业	总用水量
唐山	735.42	281.76	2 921.52	3 938.70	848.13	589.56	2 664.24	4 101.93
秦皇岛	113.50	145.65	885.47	1 144.62	135.24	153.85	958.09	1 247.18
廊坊	373.41	255.96	1 082.95	1 712.32	350.42	339.08	1 196.35	1 885.85
邯郸	542.91	439.65	1 826.54	2 809.10	700.84	595.81	2 685.19	3 981.84
邢台	518.52	283.74	1 915.70	2 717.96	642.79	379.52	1 884.56	2 906.87
石家庄	597.40	588.41	3 198.61	4 384.42	600.46	688.14	3 156.18	4 444.78
保定	629.15	388.86	2 625.25	3 643.26	798.16	497.30	2 756.98	4 052.44
衡水	321.05	206.54	1 262.59	1 790.18	281.09	193.54	1 327.26	1 801.89
沧州	136.52	173.25	1 742.58	2 052.35	438.35	235.36	2 074.32	2 748.03
承德	302.52	220.09	1 085.64	1 608.25	415.32	346.45	1 223.58	1 985.35
张家口	208.02	123.06	1 585.24	1 916.32	241.53	234.19	1 878.41	2 354.12
合计	4 478.42	2 806.70	20 432.30	27 717.48	5 452.32	4 252.80	21 805.10	31 510.28

灵活性.从预测结果可以看出,该法用于需水量预测是可行的,且计算速度快,精度高,无收敛性问题.但是,该方法完全是以历史资料为基础,因此对于趋势的预测难免会存在一些缺陷.而对于用水量的总体发展趋势,诸如节水工程的大力发展、废水回用技术的发展可能会使用水量渐趋稳定而非逐年增长等问题,简单的预测方法是难以达到满意的结果的,必须结合发达国家的实际经验及国内发展现状进行全面分析.

参考文献:

[1] 聂相田,邱林,朱普生,等.水资源可持续利用管理不确定性分析方法及应用[M].郑州:黄河水利出版社,1999.

NIE Xiang-tian, QIU Lin, ZHU Pu-shang, et al. The uncertain

analysis method to sustainable water resources utilization and application[M]. Zhengzhou: Huang River Water Conservancy Press, 1999.

[2] Benoit B M. Fractals: Form, chance and dimension[M]. San Francisco: W H Freeman, 1977.

[3] 敖力布,林鸿溢.分形学导论[M].呼和浩特:内蒙古人民出版社,1996.

AO Li-bu, LIN Hong-yi. Fractal theory [M]. Huhehot: Inner Mongol Press, 1996.

[4] 齐东旭.分形及其计算机生成[M].北京:科学出版社,1994.

QI Dong-xu. Fractal and computer application [M]. Beijing: Science Press, 1994.

[5] 王东生,曹磊.混沌、分形及其应用[M].北京:中国科学技术出版社,1995.

WANG Dong-sheng, CAO Lei. Chaos, fractal and application [M]. Beijing: Chinese Science Technology Press, 1995.

(编辑:陶文文)

· 声 明 ·

自 2004 年 1 月起,本刊加入华艺中文电子期刊服务(CEPS).今后凡向本刊投稿者,均视为其文稿刊登后可供 CEPS 数据库收录、转载并上网发行,其文章著作权使用费包含在稿酬中.如作者不接受此项服务,请在投稿时声明.CEPS 的访问地址:www.ceps.com.tw.