

混沌时间序列可预报时间长度分析

王 文¹, 许武成²

(1. 河海大学水资源环境学院, 江苏 南京 210098; 2. 西华师范大学国土资源学院, 四川 南充 637002)

摘要:以经典混沌序列——Henon 映射序列为例, 对不同长度、不同噪声水平的序列进行了预报试验. 结果表明 (a) 纯净序列, 序列越长, 预报精度越高, 同时预报精度衰减速度随之减慢, 并且, 当序列足够长时, 即使较长预报步长的多步预报, 仍可取得相当高的预报精度 (b) 含噪声序列, 其预报精度随步长的增加迅速地呈指数衰减, 并且, 序列长度基本上不会对预报时间长度产生影响, 但噪声水平对可预报时间长度有一定影响, 噪声水平越高, 可预报时间长度越短. 因此, 不明确可预报的标准、不考虑序列长度、噪声水平等因素的影响而简单地以最大 Lyapunov 指数的倒数定义最大可预报时间长度, 是不可取的. 同样, 根据最大 Lyapunov 指数推断水文过程的可预报时间长度也是不合适的.

关键词:混沌时间序列; 可预报时间长度; Lyapunov 指数; 水文过程

中图分类号: P338 **文献标识码:** A **文章编号:** 1000-198X(2004)04-0367-05

1 水文过程可预报时间长度研究回顾

衡量动力系统对初始条件敏感依赖性的一个重要指标是李亚普诺夫(Lyapunov)指数. 该指数定义是: 假设状态空间中相近的两个相点为 x_n 和 x_m , 两点间的初始距离为 $\Delta_0 = \|x_n - x_m\|$, 在系统演化 k 时段后, 两点间的距离为 $\Delta_k = \|x_{n+k} - x_{m+k}\|$, 如果 $\Delta_k \approx \Delta_0 e^{\lambda k}$, 即两点间的距离呈指数分离, 则其中 λ 就是最大 Lyapunov 指数^[1]. 从定义可知, 混沌系统, 当系统从 t_0 时刻演进至 $t_0 + 1/\lambda$ 时刻时, 两个轨道初始距离为 Δ_0 的相邻相点的轨道距离 Δ_k 会增加到初始距离的 e 倍, 即有 $\Delta_k = \Delta_0 e$. 部分研究者据此将混沌序列的最大可预报时间长度定义为最大 Lyapunov 指数的倒数^[2~4], 或者以与 Lyapunov 指数密切相关的 Kolmogorov 熵、二阶 Renyi 熵的例数来定义混沌序列的平均可预报时间长度.

在水文过程研究方面, 部分研究者以上述定义为依据, 在确定所研究的水文过程为混沌序列后, 根据 Lyapunov 指数进行水文过程的可预报时间长度的推断. 例如: 傅军等^[5]算得金沙江屏山站日流量序列 Lyapunov 指数估计值为 0.075, Kolmogorov 熵为 0.085, 推断平均可预报时间长度为 11 d; 魏一鸣等^[6]算得九江年降水量序列的 Kolmogorov 熵为 0.452, 推断平均可预报时间长度为 2~3 a; 王宝灵等^[7]算得西北地区 30 个站月降水量序列的二阶 Renyi 熵, 推断平均可预报时间长度为 2~3 个月; 王德智等^[8]算得东北地区若干站月降雨过程的 Lyapunov 指数估计值为 0.13, 推断月降雨过程的最大有效预报时间长度为 8 个月; 陈云浩等^[9]算得上海市的旬平均降水量最大 Lyapunov 指数为 0.062, Kolmogorov 熵为 0.124 8, 推断该混沌系统的最大可预报时间长度为 160 d, 平均可预报时间长度为 80 d; 周寅康等^[10]在淮河流域洪涝灾害研究中, 根据二阶熵推算, 推断淮河流域洪涝灾害过程可预报时间长度为 5~6 a; 谢正栋等^[11]根据淮河流域蚌埠站的多年逐日平均流量资料建立了洪水时间序列, 计算出其 Lyapunov 指数和 Kolmogorov 熵, 同时推断淮河流域洪水的可预报时间长度为 4~7 a, 平均可预报时间长度为 5 a.

此外, 有些研究者在确认流量过程为混沌序列的基础上进行了预报研究, 如进行了未来 1~10 步的日流量预报^[12, 13]、1~20 步的日流量预报^[14]、1~10 步的月流量预报^[15]等. 部分预报结果显示, 预报精度随步长的增加呈指数衰减. 有些预报结果则对很长预见期的多步预报给出了相当高的预报精度. 例如, Sivarkumar

等^[15]对未来1~10个月的流量预报结果的效率系数均在0.83以上,并且第3个月至第10个月的预报效率系数稳定在0.8318~0.8553之间,没有指数衰减迹象(见表1)。

表1 Sivarkumar等^[15]关于月流量的非线性预报精度

Table 1 Nonlinear forecast accuracy for monthly discharge made by Sivarkumar et al. ^[15]										
预见期/月	1	2	3	4	5	6	7	8	9	10
CE	0.9388	0.8692	0.8518	0.8318	0.8553	0.8326	0.8335	0.8484	0.8382	0.8435

注:CE为效率系数, $CE=1-(\text{预报值的均方误差}/\text{实测值的方差})$ 。CE即我国《水文情报预报规范》所称的确定性系数^[16]。如果 $CE\geq 0.8$,则可认为预报结果满意。其非线性预报的嵌入维数为3,预报方法为近邻多项式拟合。

这里有两大问题值得分析:(a)单纯以最大Lyapunov指数的倒数而不考虑序列长度、噪声水平等因素的影响来定义一个混沌序列的可预报时间长度,是否合适?根据最大Lyapunov指数的倒数来推断水文过程的可预报时间长度是否有意义?因为某些研究者以此推断的水文过程的可预报时间长度之长(如160d的降水、5~6a的洪涝灾害)在现实中几乎是不可能的。(b)随预报步长的增加,含噪声混沌序列的多步预报精度应具有怎样的衰减特征?一方面认为一个水文观测序列是混沌的,另一方面却又能如Sivarkumar等^[15]那样对月流量过程做出高精度的多步预报,这是否矛盾?

2 Henon映射序列预报试验

2.1 预报试验方法说明

本文以经典混沌序列——Henon映射序列为例来分析混沌序列可预报时间长度与序列长度、序列噪声水平之间的关系。

Henon映射是经典的具有混沌特性的动力系统,其数学表达式为

$$x_{n+1}=1-ax_n^2+by_n \quad y_{n+1}=x_n$$

试验中先模拟生成一个长度为10000的Henon映射序列($a=1.4$, $b=0.3$)然后取其最后的200个、500个、2000个、5000个以及全部10000个数据共5个序列分别进行预报试验。预报对象均为每个序列的倒数第51个至倒数第11个共41个数据点之后的10步的预报值。

首先基于混沌序列的相空间重构理论^[1]重构Henon序列的相空间(相空间重构参数为:嵌入维 $m=3$;时延量 $\tau=1$)然后以重构相空间为基础,采用最近邻算术平均法^[17]进行Henon序列的预报。最近邻算术平均法是混沌时间序列预报中最常用的方法之一。其基本思路是,先找到当前状态的 k 个最近邻状态(或称相点)然后将这些近邻相点下一时刻的观测值取平均,最后将平均值作为当前状态下一时刻的预测值。最近邻预报模型的近邻数 $k=m+1$,即 $k=4$ 。

下面以序列倒数第50点的预报来比较一下不同长度序列的近邻状态选择差异。待预报的倒数第50点的实际数值为0.73346,而在相空间中用于预报该点的当前相点由序列的倒数第53、52、51点构成,坐标为 $(-0.92243,-0.57602,0.25875)$ 。表2列出了该相点在长度为200、2000及10000的3个序列中所找到的最近的4个近邻状态及相应的近邻算术平均预测值。

表2 不同长度Henon序列的近邻状态选择差异比较

Table 2 Neighborhood selection differences for Henon series of different sizes						
序列长度	近邻状态	近邻下一时刻值	与待测相点距离	平均距离	预测值	预测误差
200	(-1.16199,-0.52334,0.26797)	0.74247	0.24546	0.36273	0.57786	0.15560
	(-1.10389,-0.34510,0.50210)	0.54352	0.38141			
	(-1.09906,-0.33075,0.51712)	0.52639	0.39763			
	(-1.09422,-0.30607,0.54058)	0.49906	0.42640			
2000	(-0.92584,-0.58519,0.24282)	0.74190	0.01869	0.01477	0.73889	-0.00542
	(-0.92145,-0.57339,0.26327)	0.73094	0.00532			
	(-0.92682,-0.58781,0.23822)	0.74421	0.02408			
	(-0.92443,-0.58142,0.24941)	0.73849	0.01097			
10000	(-0.92278,-0.57696,0.25713)	0.73435	0.00190	0.00295	0.73376	-0.00030
	(-0.92286,-0.57723,0.25667)	0.73460	0.00244			
	(-0.92156,-0.57377,0.26264)	0.73130	0.00458			
	(-0.92293,-0.57744,0.25631)	0.73479	0.00286			

2.2 预报结果分析

从表 2 可以看出 ,随着序列长度的增加 ,当前相点的近邻状态选择越来越准确 ,近邻状态与当前相点的平均距离由 0.36273 缩小为 0.00295.相应地 ,一步预报的预测误差也由 0.15560 降为 - 0.00030.比较不同长度序列对未来 1 ~ 10 步的预测值的效率系数 CE (见图 1)可知 ,随着序列长度的增加 ,多步预报的精度会逐渐提高 ,对长度为 10000 的序列来说 ,即使 10 步预报 ,精度仍相当高.

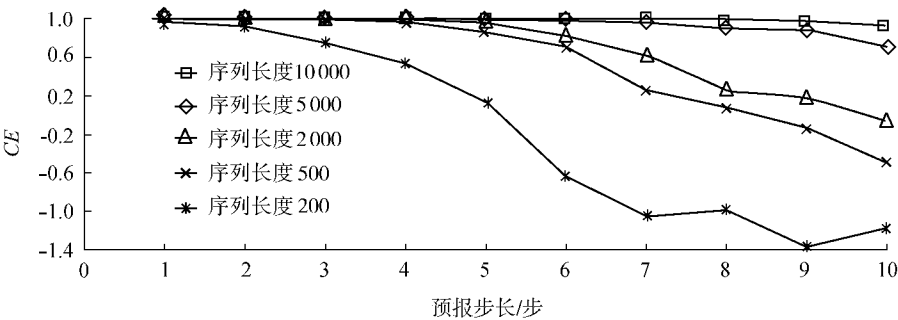


图 1 不同长度纯净 Henon 序列的未来 1 ~ 10 步的预测效率系数比较

Fig.1 Accuracy of prediction of 1 ~ 10 steps for pure Henon series of different sizes

Henon 序列的最大 Lyapunov 指数为 0.4162 ~ 0.418^[18] ,如将最大 Lyapunov 指数的倒数作为最大可预报时间长度 ,则可预报步长还不到 3 步.但上述预报试验表明 ,以预测效率系数 CE 大于 0.8(即预报精度满意)作为可预报标准 ,对纯净序列来说 ,当序列长度超过 500 时 ,可预报步长远大于 3 步 ,并且 ,随着序列长度的增加 ,Henon 序列的可预报长度显著增加 ;如果以 CE 大于 α (即预报值优于序列均值)作为可预报标准 ,则即使对于长度为 200 的 Henon 序列 ,可预报步长均超过 3 步.

除序列长度对预报精度有一定影响外 ,噪声水平对预报精度也有相当大的影响.为说明噪声水平对预报精度的影响 ,对上述不同长度的 Henon 序列分别叠加 5% 的正态分布噪声(即噪声的标准差为原序列标准差的 5%) ,生成相应长度的新的序列 ,再采用近邻算术平均法 ,对倒数第 51 个至倒数第 11 个共 41 个数据点之后的 10 步数值进行了预报 ,结果如图 2 所示.

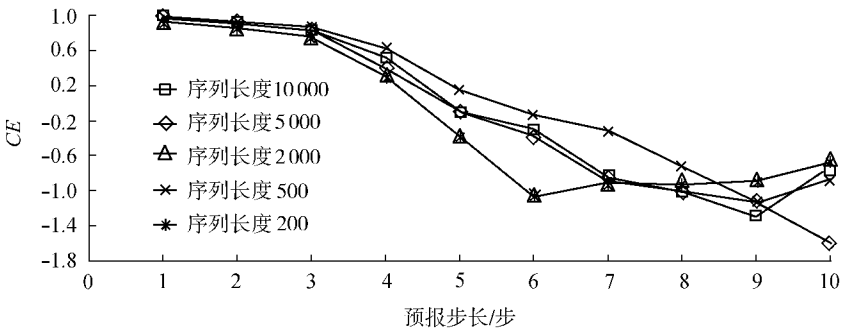


图 2 含 5% 噪声的不同长度 Henon 序列的未来 1 ~ 10 步的预测精度比较

Fig.2 Accuracy of prediction of 1 ~ 10 steps for Henon series of different sizes with 5% noise

从图 2 可以看出 ,当序列存在噪声时 ,预报精度均有不同程度的降低 ,较长序列的多步预报精度降低尤为显著.此时 ,Henon 序列的可预报长度几乎与序列长度无关.即不论序列长度如何 ,如果以预测效率系数 CE 大于 0.8 作为可预报标准 ,则可预报长度为 2 ~ 3 步 ;如果以 CE 大于 0 作为可预报标准 ,则可预报长度为 4 ~ 5 步.但可预报长度与序列噪声水平有较显著的关系.由不同噪声水平(2% ,5% ,10%)预报精度的比较(见图 3)可知 ,噪声水平越高 ,可预报长度越短.

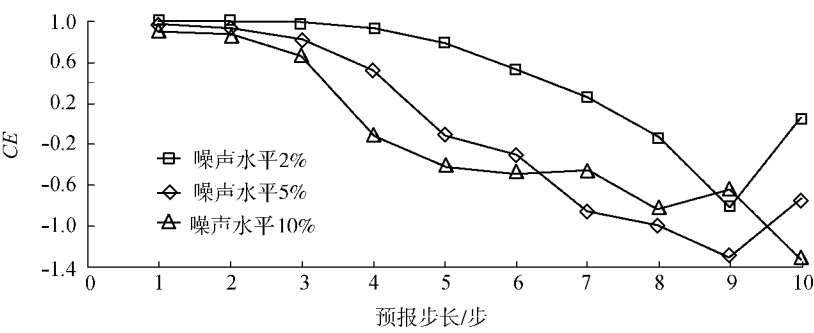


图 3 不同噪声水平时长度为 10 000 的 Henon 序列的未来 1~10 步预测精度比较

Fig.3 Accuracy of prediction of 1~10 steps for 10 000-point Henon series with 2% ~10% noise

3 问题讨论

a. 混沌系统的最大可预报时间长度应与相点初始距离直接相关,如找不到非常相近的相点,或找错了近邻相点,则在很短的时间步距后,就可能出现很大的轨道偏差,从而出现无法预报的情况;否则,在时间步距 = $1/\lambda$ 后仍可以得到较高的预报精度.在表 2 所描述的各纯净 Henon 序列的倒数第 50 点的预报试验中,当系统从 t_0 时刻演进至 $t_0 + 1/\lambda \approx t_0 + 2.39$ 时刻时(Henon 序列的最大 Lyapunov 指数 $\lambda \approx 0.418$)相邻相点轨道距离 Δ_k 应为 $\Delta_k = \Delta_0 e^k$ (其中 Δ_0 为相点初始距离),对长度为 200 的序列来说, $\Delta_0 \approx 0.36273$, $\Delta_k \approx 0.98599$, Δ_k 大于序列标准差(序列标准差为 0.72218),而对长度为 10 000 的序列来说, $\Delta_0 \approx 0.00295$, $\Delta_k \approx 0.00802$, Δ_k 远小于序列标准差.因此,在 $t_0 + 1/\lambda$ 时刻,对长度为 200 的序列来说,预报精度已很低,而对长度为 10 000 的序列来说,预报精度仍很高.这说明,对纯净 Henon 序列来说,序列越长,可预报时间越长;如序列足够长,多步预报仍具有相当高的精度.而另一方面,含噪声序列由于受到噪声的影响,在采用最近邻方法进行非线性预报时,无法准确选择近邻,因此预报精度衰减很快,即使是对于长序列,多步预报也无法做到;并且,噪声水平越高,近邻选择误差越大,预报精度衰减也越快.所以,如果不考虑序列长度、序列噪声水平、序列方差等因素对实际预报精度的影响而简单地以最大 Lyapunov 指数的倒数定义最大可预报时间长度,是很不严格的.

b. 根据上述分析,含噪声混沌序列的多步预测精度应具有显著的指数衰减特征.如果实测序列(如水文过程序列)为混沌序列,由于受到噪声的影响,则其预报精度也应具有指数衰减特征.这与表 1 中显示的对未来 1~10 个月流量过程的高精度多步预报情况完全不同.也就是说,一方面肯定一个流量序列为混沌序列,另一方面却能在其受到噪声影响的情况下仅依靠流量序列自身的非线性特征对其做出高精度的多步预报,是不符合逻辑的.

4 结 论

a. 对经典混沌序列——Henon 映射序列的预报试验表明 (a)纯净序列,其序列越长,则各步的预报精度越高,并且多步预报精度的衰减速度随序列长度的增加而减慢,序列的可预报时间长度随序列长度的增加而延长.当序列足够长时,即使较长步长的多步预报,仍可取得相当高的预报精度.(b)含噪声序列,不同长度序列的预报精度随步长的增加均会迅速地呈指数衰减,且序列长度对可预报时间长度基本上没有影响.但噪声水平对可预报长度有一定影响,噪声水平越高,可预报长度越短.因此,在不考虑序列长度、噪声水平等因素对可预报时间长度的影响,不明确达到怎样的精度标准算是“可预报”的情况下,简单地以最大 Lyapunov 指数的倒数定义最大可预报时间长度,是不可取的.

b. 水文过程的可预报性分析是一个比较复杂的问题.由于其预报精度受多方面的影响,如序列自身受到的随机噪声的干扰程度、预报时是否考虑外部因素(如大气环流、地球物理因素等)的影响,所以,即使不考虑水文过程是否为混沌序列这一有争议的问题,即将水文过程作为混沌序列,如何在考虑诸多预报因子的前提下定义可预报时间长度,也是有待讨论的问题.单纯地以最大 Lyapunov 指数(或相关的 Kolmogorov 熵、Renyi 熵)的倒数定义水文过程的可预报时间长度,从理论研究和应用角度来看,是不合适的.

参考文献：

- [1] KANTZ H ,SCHREIBER T. Nonlinear time series analysis[M]. Cambridge :Cambridge University Press ,1997. 60—61.
- [2] NICOLIS C. Climate predictability and dynamical systems[A]. In : NICOLIS C ,NICOLIS G ,eds. Proceedings of the NATO Advanced System Study Institute on Irreversible Phenomena and Dynamical System Analysis in Geophysical Sciences[C]. Dordrecht : D Reidel , 1986. 321—354.
- [3] 刘式达 ,刘式适. 分形和分维引论[M]. 北京 :气象出版社 ,1993. 93—95.
- [4] 林振山. 长期预报的相空间理论和模式[M]. 北京 :气象出版社 ,1993. 39—40.
- [5] 傅军 ,丁晶 ,邓育仁. 洪水混沌特性初步研究[J]. 水科学进展 ,1996 ,7(3): 226—230.
- [6] 魏一鸣 ,孙国栋 ,胡平昭 等. 九江年降水时间序列的混沌特性[J]. 江西科学 ,1998 ,6(3): 141—145.
- [7] 王宝灵 ,谢金南 ,俞亚勋. 西北地区降水混沌特性的初步分析[J]. 高原气象 ,2000 ,19(1): 25—31.
- [8] 王德智 ,夏军 ,张利平. 东北地区月降雨时间序列的混沌特性研究[J]. 水电能源科学 ,2002 ,20(3): 32—34.
- [9] 陈云浩 ,史培军 ,李晓兵. 不同热力背景对城市降雨(暴雨) 的影响(1)——降雨时序的混沌分析[J]. 自然灾害学报 ,2001 ,10(3): 20—25.
- [10] 周寅康 ,王腊春 ,张捷. 淮河流域洪涝变化的混沌特征[J]. 自然灾害学报 ,1999 ,8(1): 42—47.
- [11] 谢正栋 ,张永勤 ,周寅康 等. 淮河流域洪水的分形特征及可预报时间研究[J]. 南京大学学报(自然科学版) ,2003 ,39(1): 113—119.
- [12] JAYAWARDENA A W ,LAI F. Analysis and prediction of chaos in rainfall and stream flow time series[J]. Journal of Hydrology ,1994 ,153 : 23—52.
- [13] 丁晶 ,邓育仁 ,傅军. 洪水变化的混沌特性和相空间预测[A]. 见 :刘昌明主编. 第六次全国水文学术会议论文集[C]. 北京 :科学出版社 ,1997. 151—157.
- [14] LIU Q ,ISLAM S ,RODRIGUEZ-ITURBE I ,et al. Phase-space analysis of daily streamflow : characterization and prediction[J]. Adv Water Resour ,1998 ,21 : 463—475.
- [15] SIVARKUMAR B ,BERNDTSSON R ,PERSSON M. Monthly runoff prediction using phase space reconstruction[J]. Hydrol Sci ,2001 ,46(3): 377—387.
- [16] SL 250—2000 水文情报预报规范[S].
- [17] FARMER J D ,SIDOROWICH J J. Predicting chaotic time series[J]. Phys Rev Lett ,1987 ,59(8): 845—848.
- [18] 李冬梅 ,王正欧. 基于 RBF 网络的混沌动力系统辨识[J]. 天津大学学报(自然科学版) ,2002 ,35(2): 191—195.

Analysis of length of predictable time of chaotic time series

WANG Wen¹ , XU Wu-cheng²

(1. College of Water Resources and Environment , Hohai Univ. , Nanjing 210098 , China ;

2. College of Land and Resources , West China Normal University , Nanchong 637002 , China)

Abstract : Prediction experiments were performed on a classic chaotic time series , the Henon mapping series , with different lengths and noise levels. It is shown that : (a) for the noise-free series , the accuracy increases with the size of the series , and the longer the series , the slower the accuracy of prediction decays , besides , the accuracy of multi-step prediction is fairly high if the series is long enough ; (b) for the noisy series , the accuracy decays exponentially with the increase of the step length , and the size of the series almost has no influence on the length of the predictable time , while the noise level is of certain influence on the length of the predictable time : the higher the noise level , the shorter the length of predictable time. Therefore it is unreasonable to take the reciprocal of the maximum Lyapunov exponent as the maximum length of predictable time without clear definition of the criterion of prediction and without consideration of the influences of the size and noise level of the series , and it is meaningless to deduce the length of predictable time of hydrological processes according to the maximum Lyapunov exponent.

Key words : chaotic time series ; length of predictable time ; Lyapunov exponent ; hydrological process