

基于混沌神经网络的径流预测模型

王义民¹, 张 珏²

(1 西安理工大学 西北水资源与环境生态教育部重点实验室, 陕西 西安 710048;

2 深圳市广汇源水利勘测设计有限公司 设计六室, 广东 深圳 518020)

【摘要】【目的】揭示径流时间序列变化规律并进行预测,为水库调度提供指导。【方法】针对径流时间序列的非线性特点,利用重构相空间的嵌入维数确定神经网络的结构,建立了基于混沌相空间重构的径流量预测 BP 网络模型,并利用该模型对位于陕西省汉江上游的石泉水文站的径流时间序列进行了预测。【结果】实例计算结果表明,石泉水文站月平均流量的时间序列具有混沌性,最大嵌入维数为 12,依此构建的 BP 神经网络收敛速度快、预测精度较好。【结论】利用重构相空间中的最佳嵌入维数,可合理确定 BP 神经网络的输入层节点数。

【关键词】 混沌径流时间序列;径流预测;神经网络

【中图分类号】 TV121

【文献标识码】 A

【文章编号】 1671-9387(2010)06-0200-05

Run-off prediction model based on the chaotic and BP network

WANG Yi-min¹, ZHANG Jue²

(1 Key Laboratory of Northwest Water Resources and Environment Ecology of MOE at Xi'an University of Technology,

Xi'an, Shaanxi 710048, China; 2 Design Office No. 6, Shenzhen Guanghuiyuan Survey & Design Co., Ltd.

for Water Resources, Shenzhen, Guangdong 518020, China)

Abstract: 【Objective】 It is important for reservoir operation to obtain and predict run-off variation law. 【Method】 According to the nonlinear characteristic of runoff time series, the network structure is determined by using embedding dimensions of reconstructed embedding phase space, and runoff prediction BP network model is established based on the chaotic phase space reconstruction. This method has been applied to predict the runoff time series of Shiquan hydrologic station. 【Result】 The calculation results show that the monthly average discharge time series of Shiquan hydrologic station is chaos, and its maximum embedding dimension is 12. The BP network is structured based on the dimension, and convergence speed is improved, calculation error is reduced. 【Conclusion】 Embedding dimensions of reconstructed embedding phase space can provide evidence for determining input layer node number.

Key words: chaotic runoff time series; runoff forecast; neural network

径流时间序列的预测对于水库运行调度和水资源配置具有极为重要的作用。由于水文现象的复杂性,单纯的物理方法很难对水文现象进行完全描述,因此需要借助其他不确定分析方法及非线性时间序列分析方法。混沌理论是近年来迅速发展起来的非线性时间序列分析方法之一,其研究始于 Packard 等^[1]于 1980 年提出的重构相空间理论。20 世纪 90

年代,人们开始了对混沌与神经网络相互交融的研究,即研究大脑中的混沌现象,并在各学科展开应用研究^[2-3]。人工神经网络是以人脑的生理研究成果为基础,其目的在于模拟人脑的某些机理与机制,实现某些方面的功能^[4]。研究表明,混沌理论可以用来解释大脑中某些不规则的活动,因此,混沌动力学为深入研究神经网络提供了新的方法^[5]。本研究以

* [收稿日期] 2009-11-30

[基金项目] 国家自然科学基金项目(50709027, 50779053);教育部重点研究项目(209125);陕西省教育厅科技项目(09JK664)

[作者简介] 王义民(1972—),男,山西新绛人,副教授,博士,主要从事水资源系统工程研究。E-mail: wang_yimin@163.com

陕西省汉江上游石泉水文站 1955—2006 年的径流序列为例,结合混沌相空间重构理论,通过混沌系统动力吸引子确定神经网络输入结构,建立基于混沌相空间重构的径流量预测模型,以期对径流时间序列预测提供支持。

1 径流序列的相空间重构

混沌理论认为,原本不可能预测的复杂事物具有可预测性,揭示了有序与无序、确定性与随机性的统一^[6]。相空间重构的目的,是在高维相空间中恢复混沌系统的主要特征之一——混沌吸引子,从而为具有混沌特性的时间序列预测奠定基础。

水文系统的相空间重构方法关键在于如何确定时间延迟 τ 和嵌入维数 m ,这 2 个参数对于重构相空间影响极大^[7]。目前 τ 值选取的最常用方法是自相关函数法, m 值选取的算法包括 G-P 算法、伪邻近点法、奇异值分解法和填充因子法等^[8-10]。

1.1 时间延迟 τ 的确定

自相关函数法是推求时间延迟 τ 的常用方法,它主要是对时间序列间的线性相关性进行提取。对于时间序列 $x_t (t=1, 2, \dots, N)$,自相关函数计算式如下:

$$r_\tau = \frac{\sum_{t=\tau+1}^N (x_t - \bar{x})(x_{t-\tau} - \bar{x})}{\sum_{t=1}^N (x_t - \bar{x})^2} \quad (1)$$

式中: r_τ 表示时间延迟为 τ 时的自相关系数; \bar{x} 为序列均值; τ 为滞时,也称时间延迟。

对于有限长度时间序列,相空间的特征量依赖于 τ 的选择。自相关系数 r_τ 首次到达 0 时对应的 τ 值,即为时间延迟 τ ^[10]。

1.2 嵌入维数 m 的确定

从一个时间间隔一定的单变量时间序列 x_1, x_2, x_3, \dots 出发,构造一批 m 维的矢量,支起一个嵌入空间。只要嵌入维数足够高(通常要求 $m \geq 2D + 1$, D 为吸引子的维数,即关联维数),就可以在拓扑等价的意义下恢复原来的动力学性态。对于时间序列 $x(t_i) (i=1, 2, \dots, N)$,如果能适当选定嵌入维数 m 和时间延迟 τ ,便可以重构相空间如下^[10]:

$$Y(t_i) = [x(t_i), x(t_i + \tau), x(t_i + 2\tau), \dots, x(t_i + (m-1)\tau)] \quad (2)$$

关联维数 D 对于重构相空间嵌入维数 m 的选取具有重要作用。Grassberger 等^[11]提出了从时间序列计算关联维数的 G-P 算法,该算法的主要步骤如下:

设重构相空间中有 n 个矢量,计算其中有关联矢量(即两矢量间距离小于给定正数 r 的矢量)的对数,其在所有 n^2 个矢量对中所占比例称为关联积分,有下式:

$$C(r) = \frac{1}{n^2} \sum_{i,j=1}^n \theta(r - |y_i - y_j|) \quad (3)$$

式中: $C(r)$ 为关联积分; n 为重构相空间中的矢量个数; $\theta(x)$ 为 Heaviside 单位函数, $x \leq 0$ 则 $\theta(x) = 0$, $x > 0$ 则 $\theta(x) = 1$; r 为给定的较小正数; $|y_i - y_j|$ 为重构相空间中任意两矢量间的欧氏距离。 r 的适当选取可使其在某个区域且满足下式:

$$C(r) = r^D \quad (4)$$

对式(4)两边取对数可得:

$$D = \frac{\ln C(r)}{\ln r} \quad (5)$$

在实践中,给定一个 m 值,在一定范围内改变正数 r 值的大小,根据公式(3)可得到 $\ln C(r) - \ln r$ 的关系曲线,曲线中除去斜率为 0 或 ∞ 的直线外,考察其间的最佳拟合直线,那么该直线的斜率就是关联维数 D 。不断增加嵌入维数,可得到 $D-m$ 关系曲线。如果 D 值随 m 值的增大而增加,并不收敛于一个稳定值,则表明所考虑的系统是一个随机时间序列;如果 D 值不再随 m 增长且在一定误差范围内不变,则表明该序列为混沌系统,此时得到的 m 值即为所求重构相空间的最佳嵌入维数^[10,12]。

2 混沌神经网络的结构

混沌时间序列内部的规律产生于非线性,它表现出时间序列在时间延迟状态空间中的相关性,这种特性使得系统似乎具有某种记忆能力,同时,通常的解析方法又无法将这种规律很好地描述出来,而神经网络恰好具备这种信息处理方式^[13]。

人工神经网络是通过模拟人的大脑神经处理信息的方式,进行信息并行处理和非线性转换的复杂网络系统,其具有强大的学习能力和非线性映射能力。BP 网络是一种多层前馈神经网络,在网络训练中,调整网络权值的训练算法主要是误差反向传播学习算法。网络结构包括输入层、隐含层和输出层^[14-16]。输入层节点数为重构相空间的最佳嵌入维数,隐含层节点数由试凑法确定且遵循隐层节点数要大于或等于输入层节点数的原则。试凑法初始值由下式确定:

$$m = \sqrt{nl} \quad (6)$$

式中: m 为隐含层节点数, n 为输入层节点数, l 为输

出层节点数。

其中输入层到隐层的激活函数采用 sigmoid 型函数, 隐含层到输出层的激活函数采用线性函数。网络的工作形式为:

$$z = f\left(\sum_{i=1}^n w[i]x[i] - \theta\right). \quad (7)$$

式中: z 为网络输出; $f(x)$ 为非线性函数; $x[i]$ 、 $w[i]$ 分别为网络输入及相应权值; i 为输入神经元个数, $i=1, 2, \dots, n$; θ 为节点阈值。

混沌神经网络结构就是利用混沌理论判断系统是否具有混沌特性, 并采用 G-P 算法确定时间序列的最小嵌入维数 m , 将其作为神经网络的输入层神经单元的数目, 通过构建神经网络模型进行时间序列预测。

3 实例应用

汉江是长江的一级支流, 发源于陕西省宁强县潘家山, 流域面积 159 000 km², 是南水北调的重要水源。石泉水库是汉江上游具有较好调节能力的水库, 对于汉江水资源综合利用起着重要作用。径流补给主要以降雨形成的地表径流为主, 且径流分布不均, 年内变化和年际变化明显。7—9 月为主汛期, 年最大流量多出现在 7 月或 9 月。流域洪水主

要来源于暴雨, 洪水涨落极快, 汛期具有明显的双峰型特点。本例建模采用石泉水文站 52 年(1955—2006 年)的径流资料进行计算。

3.1 重构相空间

石泉水库 1955—2006 年的月平均入库流量时间序列如图 1 所示。由图 1 可知, 从水文径流时间历程图中很难直接分析出其系统的动力学特征, 因此需将其入库径流时间序列进行相空间重构。

首先用自相关函数计算石泉水文站月平均流量自相关系数, 根据公式(1), 得到石泉站月平均流量自相关函数变化曲线(如图 2 所示)。 r_τ 第一次过 0 点时所对应的 τ 值即为所求的时间延迟 τ , 由图 2 得 $\tau=4$ 。

确定时间延迟 τ 值后, 根据前文提到的 G-P 法求解嵌入维数 m 值, 可得到 $\ln C(r) - \ln r$ 、 $D - m$ 关系曲线, 分别如图 3、4 所示。从图 3 可以看出, 不同嵌入维数 m 对应的 $\ln C(r) - \ln r$ 关系存在直线相关的部分, 每条曲线中直线段部分的拟合斜率即为不同 m 值所对应的关联维数; 从图 4 可以看出, 当嵌入维数 $m=12$ 时, 关联维数 D 达到饱和值 $D=4.42$ 。这表明重构相空间的最大嵌入维数为 12, 且系统是混沌的。

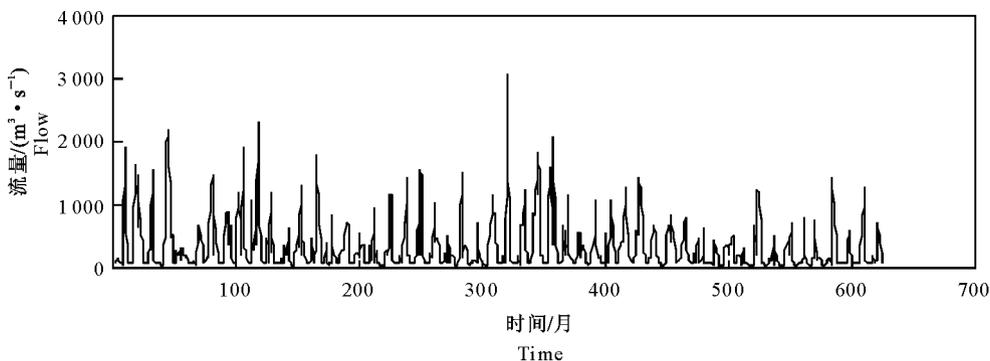


图 1 石泉水库月平均流量的时间序列(1955—2006 年)

Fig. 1 Monthly average flow series of Shiquan reservoir (1955—2006)

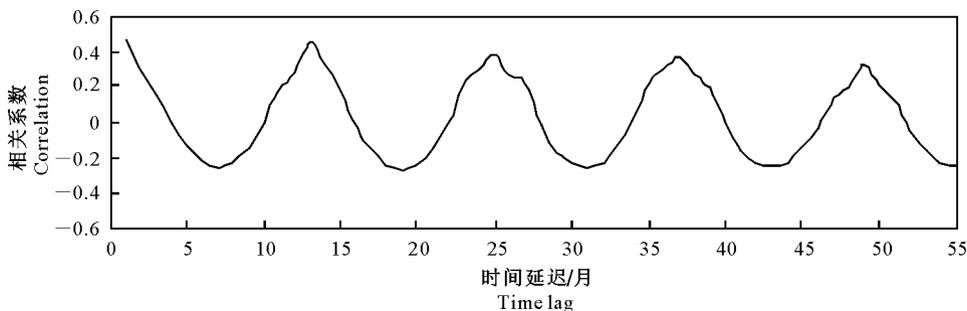


图 2 石泉站月平均流量自相关函数的变化曲线

Fig. 2 Autocorrelation function curve of monthly average flow at Shiquan station

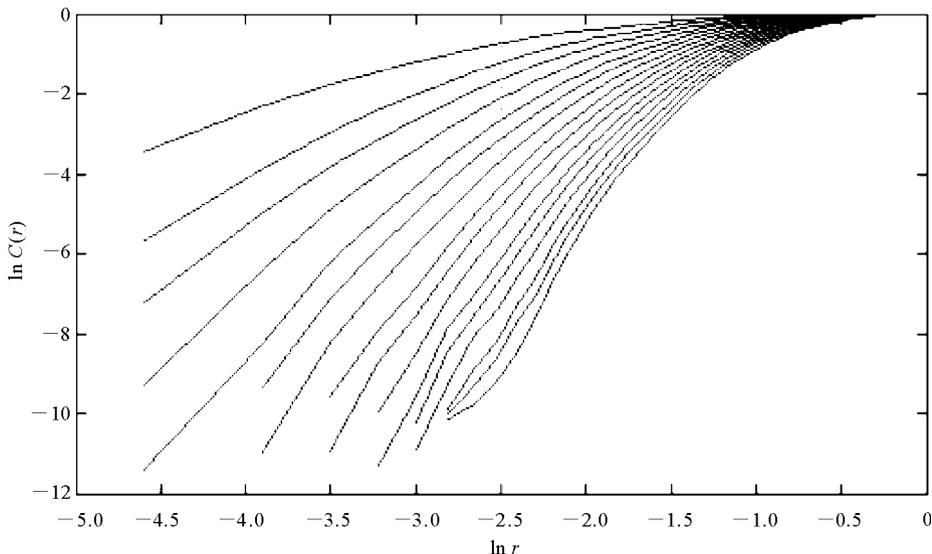


图 3 石泉站月平均流量的 $\ln C(r) - \ln r$ 关系曲线
Fig.3 Relation curve of monthly average flow $\ln C(r) - \ln r$

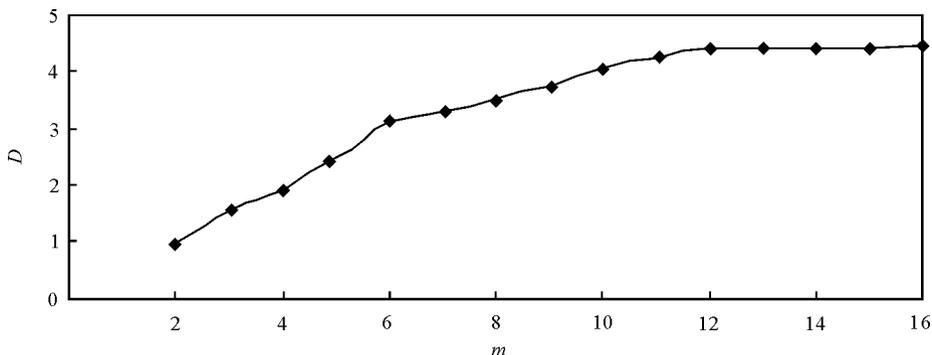


图 4 石泉站月平均流量的 $D - m$ 关系曲线
Fig.4 Relation curve of monthly average flow $D - m$

3.2 BP 网络设计

本研究中采用单隐层的 BP 网络进行预测。重构相空间的最佳嵌入维数 m 为 12, 所以网络输入层神经元有 12 个, 输出层神经元为 1 个, 通过试凑法确定网络隐含层神经元为 22 个。对于输入输出数据, 需对其进行尺度变换即归一化处理, 将数据变换为 $[0, 1]$ 区间的值, 这里采用如下归一化公式:

$$\hat{x}_i = (x_i - x_{\min}) / (x_{\max} - x_{\min}) \quad (8)$$

表 1 基于相空间重构的 BP 神经网络的训练参数

Table 1 Training parameter of BP network based on the chaotic phase space reconstruction

训练次数 Training times	训练目标 Training objective	学习速率 Learning rate	动量因子 Momentum factor
1 000	0.01	0.03	0.7

选取石泉站 1955—2005 年的径流资料作为网络的学习训练样本, 以 2006 年的径流资料作为预测样本, 利用 matlab 神经网络工具箱进行编程, 可得到预测月平均流量的 BP 网络模型。预测结果如表

式中: \hat{x}_i 为归一化后的数据, x_i 为原始的输入或输出数据, x_{\max} 、 x_{\min} 分别为原数据中的最大、最小值。

网络中间层神经元传递函数采用 S 型正切函数 tansig, 输出层神经元传递函数采用 S 型对数函数 logsig。网络的训练函数为 trainlm 函数, 采用 L-M (Levenberg-Marquardt) 算法进行网络学习。网络训练的参数设定如表 1 所示。

2 所示。由表 2 可知, 与实测值相比较, 模型预测值的最大相对误差为 7.62%, 平均相对误差为 2.81%, 预测结果较为准确, 能够满足实际应用的需求。

表 2 石泉站 2006 年月平均流量的预测结果

Table 2 Prediction results of monthly average flow in 2006 at Shiquan station

月份 Month	实测值/ ($\text{m}^3 \cdot \text{s}^{-1}$) Measure value	预测值/ ($\text{m}^3 \cdot \text{s}^{-1}$) Predicted value	相对误差/% Relative error	月份 Month	实测值/ ($\text{m}^3 \cdot \text{s}^{-1}$) Measure value	预测值/ ($\text{m}^3 \cdot \text{s}^{-1}$) Predicted value	相对误差/% Relative error
1	60.57	61.30	1.22	7	138.67	138.87	0.14
2	70.87	69.03	2.59	8	128.87	120.95	6.15
3	88.23	87.08	1.31	9	716.67	721.66	0.70
4	131.33	131.30	0.02	10	415.00	410.42	1.10
5	176.33	181.08	2.69	11	135.73	126.86	6.54
6	67.23	64.77	3.66	12	64.07	68.95	7.62

4 结 论

本研究将混沌理论和 BP 神经网络相结合,建立了基于混沌相空间重构理论的径流量预测 BP 网络模型,确定了重构相空间中最佳嵌入维数,为 BP 神经网络输入层节点数的合理选择提供了依据。实例研究结果表明,石泉水文站月平均流量时间序列具有混沌性,可对时间序列中所反映出的内在确定性进行短期预测。模型预测结果表明,本研究提出的模型合理可靠,能够有效地进行短期预测。

[参考文献]

- Packard N H, Crutchfield J P, Farmer J D, et al. Shaw geometry from a time series [J]. *Phys Rev Lett*, 1980, 45(9): 712-716.
- 冯国章, 李佩成. 论水文系统混沌特征的研究方向 [J]. *西北农业大学学报*, 1997, 25(4): 97-101.
Feng G Z, Li P C. On the study of chaotic characteristics in hydrological systems [J]. *Acta Universitatis Agriculturae Boreali-occidentalis*, 1997, 25(4): 97-101. (in Chinese)
- 王红瑞, 宋宇, 刘昌明, 等. 混沌理论及在水科学中的应用与存在的问题 [J]. *水科学进展*, 2004, 15(3): 400-407.
Wang H R, Song Y, Liu C M, et al. Application and issues of chaos theory in hydrosience [J]. *Advance in Water Science*, 2004, 15(3): 400-407. (in Chinese)
- 高隼. 人工神经网络原理及仿真实例 [M]. 北京: 机械工业出版社, 2007.
Gao J. The principles and sample of ANN [M]. Beijing: China Machine Press, 2007. (in Chinese)
- 黄润生. 混沌及其应用 [M]. 武汉: 武汉大学出版社, 2000.
Huang R S. Chaos and its apply [M]. Wuhan: Wuhan University Press, 2000. (in Chinese)
- Breadford P W, Mark S S, Thor H M. Searching for chaotic dynamics in snow melt runoff [J]. *Water Resources Research*, 1991, 27(6): 1005-1010.
- 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.
- 马千里, 郑启伦, 彭宏, 等. 基于模糊边界模块化神经网络的混沌时间序列预测 [J]. *物理学报*, 2009, 58(3): 1410-1419.

Ma Q L, Zheng Q L, Peng H, et al. Chaotic time series prediction based on fuzzy bounda modular neural networks [J]. *Acta Physica Sinica*, 2009, 58(3): 1410-1419. (in Chinese)

- 王秀杰, 练继建, 费守明, 等. 基于小波消噪的混沌多元回归日径流预测模型 [J]. *系统仿真学报*, 2007, 19(5): 3605-3608.
Wang X J, Lian J J, Fei S M, et al. Chaotic multivariate autoregressive model of daily runoff prediction based on wavelet denoising [J]. *Journal of System Simulation*, 2007, 19(5): 3605-3608. (in Chinese)
- 王义民, 徐小元, 燕爱玲, 等. 最近邻点法预测模型及其在河川径流预测中的应用 [J]. *系统工程*, 2008, 26(6): 111-115.
Wang Y M, Xu X Y, Yan A L, et al. Stream flow forecasting based on the nearest adjacent [J]. *Systems Engineering*, 2008, 26(6): 111-115. (in Chinese)
- Grassberger P, Procaccia L. Measuring the strangeness of strange attractors [J]. *Physica D*, 1983, 9: 189-208.
- 刘国, 黄胜, 许模, 等. 金沙江流域月径流时间序列的混沌分析 [J]. *成都理工大学学报: 自然科学版*, 2007, 34(4): 390-393.
Liu G, Huang S, Xu M, et al. Chaos analysis of the monthly runoff time series in Jinsha River, China [J]. *Journal of Chengdu University of Technology: Science & Technology Edition*, 2007, 34(4): 390-393. (in Chinese)
- 吕金虎, 陆君安, 陈士华. 混沌时间序列分析及其应用 [M]. 武汉: 武汉大学出版社, 2002.
Lü J H, Lu J A, Chen S H. Analysis and application of chaos time series [M]. Wuhan: Wuhan University Press, 2002. (in Chinese)
- Zhang B, Govindaraju R S. Prediction of watershed runoff using Bayesian concepts and modular neural networks [J]. *Journal of Water Resources Research*, 2000, 36: 753-762.
- 王义民, 于兴杰, 畅建霞, 等. 基于 BP 神经网络马尔科夫模型的径流量预测 [J]. *武汉大学学报: 工学版*, 2008, 41(5): 14-18.
Wang Y M, Yu X J, Chang J X, et al. Prediction of runoff based on BP neural network and Markov model [J]. *Journal of Wuhan University: Engineering Edition*, 2008, 41(5): 14-18. (in Chinese)
- Ghiassi M, Saidane H, Zimbra D K. A dynamic artificial neural network model for forecasting time series events [J]. *Int J Fore-cast*, 2005, 21(2): 341-362.