

# 变参数混沌时间序列的神经网络预测研究

王永生<sup>†</sup> 孙 瑾 王昌金 范洪达

(海军航空工程学院兵器科学与技术系 烟台 264001)

(2007 年 12 月 17 日收到 2008 年 5 月 10 日收到修改稿)

研究一类复杂变参数混沌系统时间序列的预测问题. 首先构造一个变参数 Logistic 映射, 分析变参数混沌系统的特点, 指出动力学特征不断变化的这类系统不存在恒定形状的吸引子. 结合 Takens 嵌入定理和神经网络理论, 阐述神经网络方法预测具有恒定吸引子形状的混沌系统可行的原因. 分析研究其用于预测变参数混沌系统的潜在问题. 变参数 Ikeda 系统的神经网络预测试验验证了理论分析结果. 试验还表明, 简单增大预测训练样本数可能降低泛化预测精度, 训练集的选择对这类系统的泛化预测效果影响极大. 指出混沌时间序列预测实用化必须研究解决这类变参数混沌系统的预测.

关键词: 混沌, 预测, 神经网络, 变参数系统

PACC: 0545

## 1. 引 言

混沌现象是自然界中广泛存在的一种不规则运动. 随着非线性混沌动力学的发展, 混沌时间序列预测已经成为一个非常重要的研究方向. 在国民经济、控制、生物医学信号的检测、复杂环境中的信号分离以及混沌保密通信的破译等领域都有重要的应用<sup>[1,2]</sup>. 自 Farmer 等人<sup>[3]</sup>开始研究混沌系统预测以来, 尤其是近 20 年来, 人们在低维混沌的预测理论、方法及应用方面已经做了大量研究工作. 对于非线性复杂系统的预测研究主要集中于基于 Takens 定理<sup>[4]</sup>考虑的各种预测模型的研究上, 建立了低维混沌时间序列的各种预测理论与方法, 如局域预测法、全局预测法、基于 Volterra 级数分解和基于支持向量机技术等非线性预测方法<sup>[5-9]</sup>. 其中全局预测法包括全局多项式预测器、神经网络预测方法. 其优点是用尽可能多的采样时间序列来近似产生序列的重构轨迹, 使其能够在预测模型中反映出原系统的动力学行为. 防止过拟合和提高泛化性能是其考虑重点. 目前很多研究仍致力于寻找有效的预测模型和算法<sup>[2,10-12]</sup>, 提高模型的预测精度.

由于内在的非线性动力学特性, 混沌系统不可长期预测, 但其确定性结构决定了系统运动轨迹短

期内发散较小, 人们认为利用观察资料进行短期预测是可能的<sup>[2]</sup>. 揭开了混沌研究序幕的、用于描述大气对流环境的 Lorenz 方程, 在 1963 年就有 Lorenz 研究天气预报问题时提出<sup>[13]</sup>, 然而至今, 也没有出现适合于所有时段的准确的短期天气预报系统. 与之情况类似的很多, 如金融系统中股指期货的预测研究<sup>[14]</sup>, 海杂波背景下目标的预测检测<sup>[15]</sup>等, 针对这些复杂性系统的预测理论的深入研究并不多见, 致使目前实际应用中很少出现实用并取得显著效益的预测系统. 复杂混沌系统的预测作为一类非常重要的问题, 当前基本上局限于参数不可变问题的预测研究, 而对实际问题中经常出现的更重要的参数可变混沌系统的预测, 由于构成几乎所有已有预测技术基础的 Takens 嵌入定理不再成立, 所以这方面的研究成果极少<sup>[16-18]</sup>. 文献 [7-12] 针对参数不变混沌系统的自适应预测做了研究, 文献 [16] 研究了不同时段参数不同的 Rossler 系统的时序预测, 也研究了参数微变 Henon 映射时序预测, 但其参数变化没有深刻影响系统运动特征. 文献 [17, 18] 虽然提到变参数混沌系统预测复杂性, 仍然作为参数不可变混沌系统预测处理, 没有深入分析.

因此, 开展针对复杂变参数混沌系统的预测研究不仅具有重要的学术意义, 也具有重要的工程和社会应用价值. 为此, 本文构造并研究变参数混沌系

<sup>†</sup> E-mail: hjhywys@sohu.com

统的动力学特性,分析变参数混沌系统预测特点,着重探讨基于神经网络模型预测方法的泛化性能,以及变参数混沌系统预测训练样本的合理选择等问题.研究表明,与具有恒定形状奇怪吸引子的混沌系统的预测不同,变参数混沌时间序列预测需要精心考虑训练样本集的选择,训练数据集对预测的泛化性能有显著影响.

## 2. 变参数混沌系统

1963 年 Lorenz 提出了用于描述大气对流环境的 Lorenz 方程<sup>[13]</sup>:

$$\begin{cases} \dot{x} = \sigma(y - x), \\ \dot{y} = \rho x - y - xz, \\ \dot{z} = xy - \beta z, \end{cases} \quad (1)$$

取参数  $\sigma = 16$ ,  $\rho = 45.92$ ,  $\beta = 4$ , 初始值  $x_0 = 1$ ,  $y_0 = 1$ ,  $z_0 = 1$ , Lorenz 方程三维时空轨迹构成的奇怪吸引子形状如图 1(a)所示.还有很多常见混沌系统,吸引子的形状简单的,如 Chen 系统<sup>[19]</sup>, Mackey-Glass 时滞微分系统<sup>[20]</sup>较为复杂的,如四维共存双翼吸引子<sup>[21]</sup>,多涡卷吸引子<sup>[22]</sup>等,在这些混沌系统的相空间中奇怪吸引子具有恒定形状,去除初始段后不随时间的延长而发生显著改变,如图 1(a)(b)(c)所示,这表明混沌系统全局动力学特性一直保持不变.

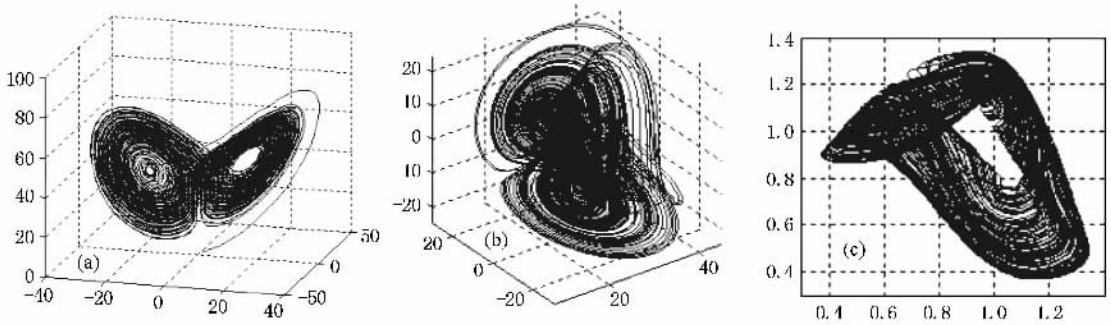


图 1 恒定形状的奇怪吸引子 (a) Lorenz 系统;(b)Chen 系统;(c)Mackey-Glass 系统

随着对气候问题复杂性的深入研究,人们逐渐认识到,采用类似 Lorenz 的这种具有全局恒定形状吸引子的混沌系统也难以描述诸如大气环境等复杂系统.在许多实际系统中,参数总是跟随时间不断变化,特别是在参数变化比较慢的情况下,经常会发生随时间周期性的分岔现象,而且不同于通常的周期性分岔.通常情况下,变参数混沌系统动态行为是充分发展的,暂态行为在每一个参数值会消亡<sup>[18]</sup>.这类系统随时间延长展现出复杂丰富的动力学特性.复杂的时空混沌系统<sup>[23]</sup>也可以根据具体耦合对象及强度参数分属于以上情况.下面构造并分析典型变参数混沌系统的特点.

### 2.1. 变参数 Logistic 系统

最简单的用于描述虫口模型的 Logistic 映射如下:

$$x_{n+1} = \mu x_n (1 - x_n), \quad (2)$$

取初始值  $x_1 = 0.22$ ,  $n = 1, 2, \dots$ . 已知当  $\mu > 3.5669$

以后系统由倍周期走向混沌态.随着参数  $\mu$  的增大,系统表现出动力学形态越来越复杂,图 2(a)为 Logistic 映射随参数  $\mu$  变化的 Feigenbom 图 (b)为 Logistic 映射处于混沌态时在二维相空间上重构的吸引子形状 (c)为对应的三维相空间中的吸引子形状.通过图 2(b)和(c)可知,只要参数  $\mu$  确定后吸引子形状不随迭代计算次数的增加而改变,只会使吸引子形状曲线更为密实.

考虑参数  $\mu$  发生慢变的 Logistic 映射

$$\begin{cases} x_{n+1} = \mu_n x_n (1 - x_n), \\ \mu_{n+1} = \mu_n + 0.00005(1 - 0.2 \sin(n)), \end{cases} \quad (3)$$

取初始值  $x_1 = 0.512$ ,  $\mu_1 = 3.4$ ,  $n = 1, 2, \dots$ . 分析变参数 Logistic 映射的动力学特性,如图 3 所示,其中图 3(a)为系统时序图,对比图 2(a)图,系统随着叠代次数增加系统状态也会逐渐经历倍周期分叉转入非常复杂的运动状态,图 3(b)和(c)为  $n = 12000$  迭代后二维和三维空间中的重构吸引子形状,放大可以观察到更为丰富的细节.

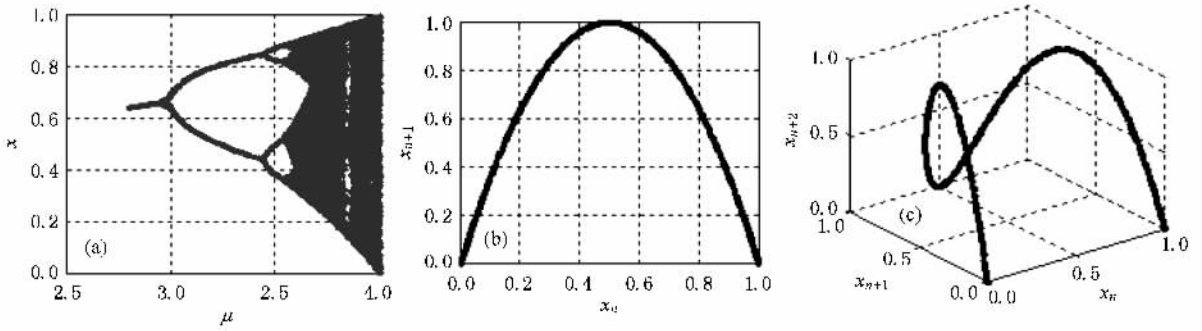


图2 Logistic 映射 (a)Feigenbom 图;(b)重构二维相平面图;(c)重构三维相平面图

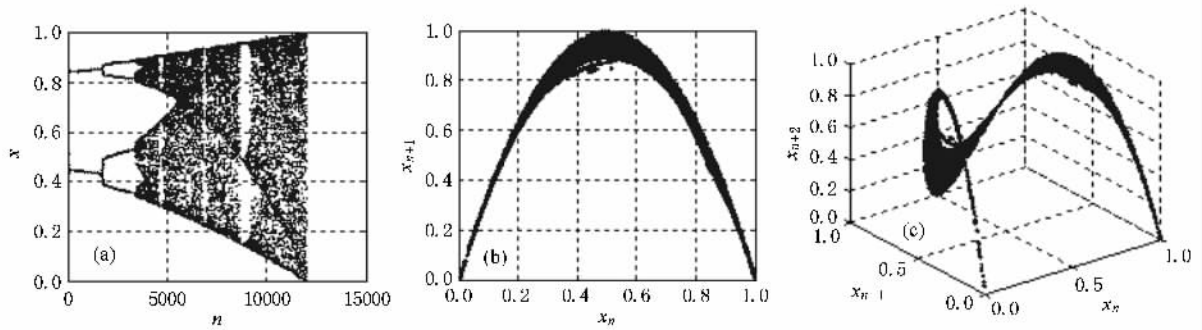


图3 变参数 Logistic 系统 (a)时序图;(b)重构二维相平面图;(c)重构三维相平面图

### 2.2. 变参数 Ikeda 系统

1979 年 Ikeda 在研究激光时给出了一个复数映射方程<sup>[24, 25]</sup>：

$$z_{n+1} = 1 + 0.9z_n \exp\left(0.4i - \frac{6i}{1 + |z_n|^2}\right), \quad (4)$$

初始值  $z_1 = 0.1$ ,  $n = 1, 2, \dots$  分别取方程(4)的实部和虚部 构成如下二阶系统,即为典型的混沌 Ikeda

系统

$$\begin{cases} x_{n+1} = 1 + \mu(x_n \cos(t) - y_n \sin(t)), \\ y_{n+1} = \mu(x_n \sin(t) + y_n \cos(t)), \end{cases} \quad (5)$$

其中  $t = 0.4 - 6(1 + x_n^2 + y_n^2)$ , 取初始值  $x_1 = y_1 = 0.1$ ,  $\mu = 0.9$ ,  $n = 1, 2, \dots$ . 图 4(a)为 Ikeda 系统吸引子的形状 (b)为由  $x$  分量重构二维相空间中吸引子形状. 只要参数  $\mu$  确定后吸引子的形状也是不随迭代计算次数的增加而改变.

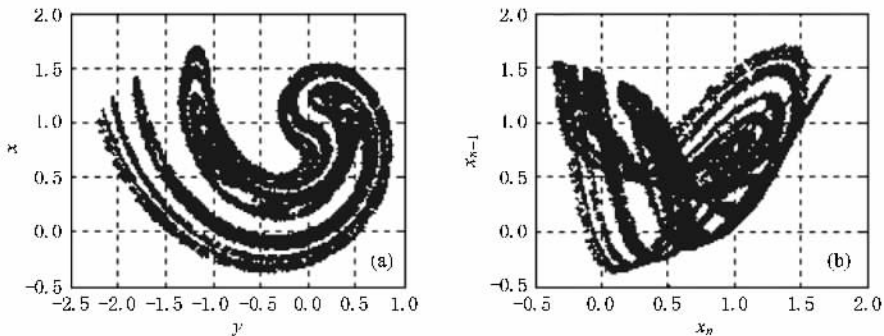


图4 混沌 Ikeda 系统 (a)二维吸引子;(b)x 分量重构二维相平面图

当方程(5)中参数  $\mu$  发生慢变时,得到如下变参数 Ikeda 系统<sup>[17, 18]</sup>：

$$\begin{cases} x_{n+1} = 1 + \mu_n(x_n \cos(t) - y_n \sin(t)), \\ y_{n+1} = \mu_n(x_n \sin(t) + y_n \cos(t)), \end{cases}$$

$\mu_{n+1} = \mu_n + 10^{-4}(1 - 0.5\sin(n))$ , (6) 其中  $t = 0.8 - 15(1 + x_n^2 + y_n^2)$  取初始值  $x_1 = 0.87$ ,  $y_1 = -0.4$ ,  $\mu_1 = -0.34$ ,  $n = 1, 2, \dots, 17400$ . 与固定参数 Ikeda 系统动力学特性完全不同,在一些关键

迭代计算时刻,系统的暂态行为在不同参数值会很快消亡.随迭代计算次数增多,变参数 Ikeda 系统在相空间中会逐渐展现出犹如由细到粗的观察星系一样的丰富现象.变参数 Ikeda 系统二维相空

间状态如图 5 所示.在 13900 点前后数据已经开始急剧增大,到 14000 后系统  $x, y$  分量已经是天文数字,17400 点后超出计算机计数范围,不再考虑.

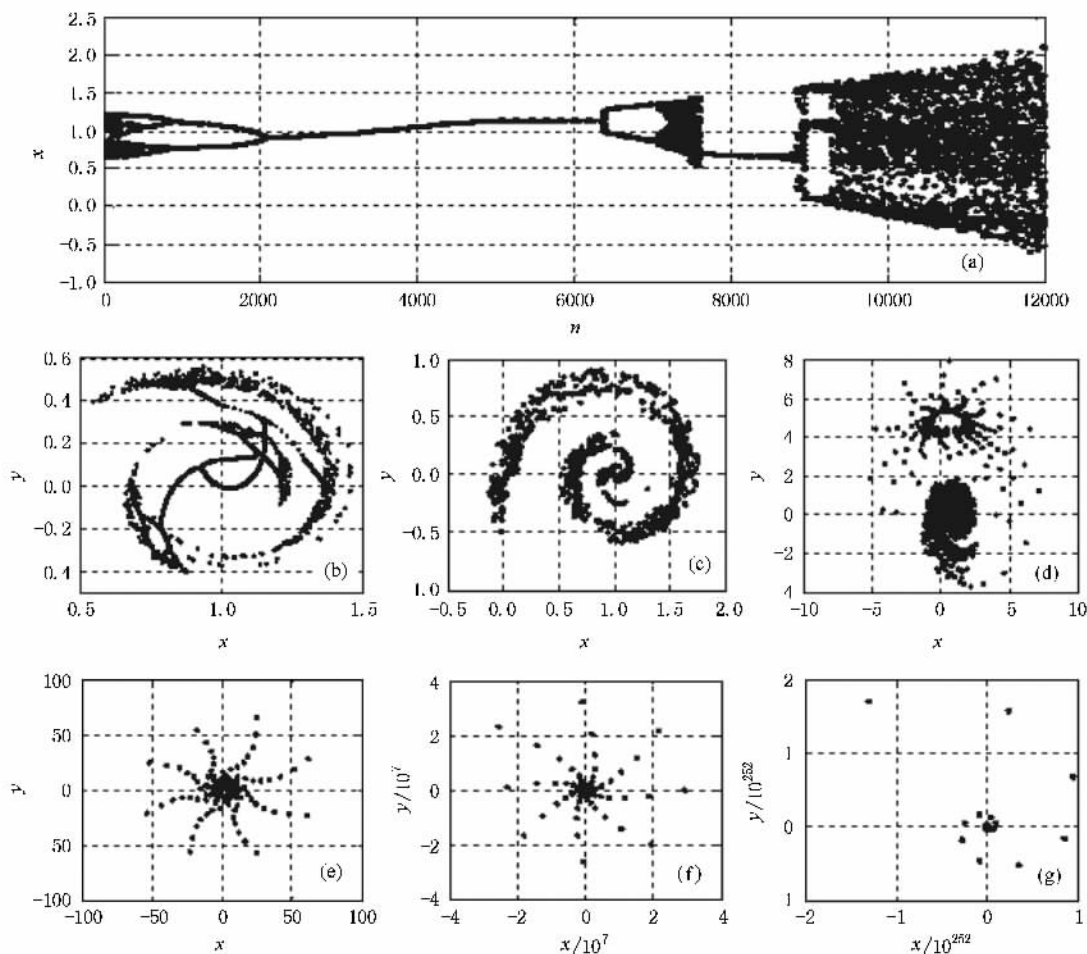


图 5 变参数 Ikeda 系统 (a)  $x$  分量时序图;(b) 1—8000 点二维相平面图 (c) 8000—10000 点二维相平面图 (d) 10000—13500 点二维相平面图 (e) 13500—13700 点二维相平面图 (f) 13700—14000 点二维相平面图 (g) 14000—17000 点二维相平面图

### 3. 变参数混沌系统的神经网络预测性能分析

分析表明,根据相空间中吸引子的时空形状,可以把混沌系统概略分为两类,一类混沌系统的奇怪吸引子具有全局恒定形状,另一类混沌系统则不具有恒定吸引子形状,属于后一类的变参数混沌系统表现出复杂多变的动力学特性.本节分析研究变参数混沌序列预测问题的特殊性,为便于说明以变参数 Ikeda 系统为例探讨.

#### 3.1. 变参数混沌系统的动力学特性

Lyapunov 指数是描述系统混沌特性的一个重要

特征值<sup>[26]</sup>,对于相空间中奇怪吸引子形状恒定的参数不变混沌系统,它的动力学特性在时间上保持不变,因此描述系统混沌程度的 Lyapunov 指数是常值.文献<sup>[27]</sup>给出 Lorenz 方程(1)的三个 Lyapunov 指数参考值分别为  $\lambda_1 = 1.497, \lambda_2 = 0.00, \lambda_3 = -22.46$ .采用 Eckmann 等人提出的算法<sup>[28]</sup>分别计算 Lorenz 方程(1)和 Ikeda 系统(5)的 Lyapunov 指数,结果如图 6 所示,图中每点 Lyapunov 值是该时刻之前数据累计计算结果.从图 6 可以看出经历初始段后描述系统混沌程度的 Lyapunov 指数稳定不变,说明了系统的动力学特性不随时间发生显著改变.

现在结合 Lyapunov 指数分析变参数 Ikeda 系统的动力学特性.利用文献<sup>[28]</sup>中算法计算变参数 Ikeda 系统的 Lyapunov 指数,结果如图 7(a)和(b)所

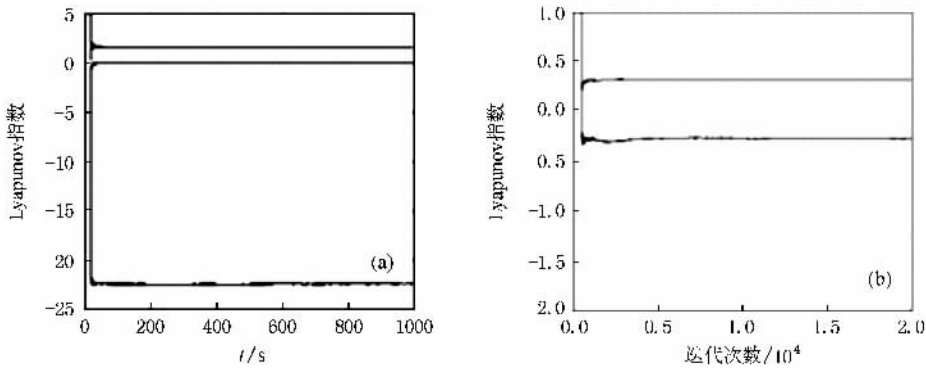


图 6 参数不变混沌系统的 Lyapunov 指数图 (a)Lorenz 方程 ;(b)Ikeda 系统

示.图 7(a)是从  $n = 7101$  点开始累计计算 500 点的 Lyapunov 指数,最后给出的最大 Lyapunov 指数为 0.355976 (b)是从  $n = 1$  开始计算 Lyapunov 指数的结果.结合图 5 和图 6 不难理解 Lyapunov 指数变化情况,当系统进入非混沌态时计算值都逐渐减小,累计变为负值,再随混沌程度的增加累计变为正值.最后用 Rosenstein 提出小数量法<sup>[29]</sup>计算变参数 Ikeda 系统 7101 点至 7600 点  $x$  分量的最大 Lyapunov 指

数.如图 7(c)所示,取嵌入维为 6,拟合斜率得到最大 Lyapunov 指数约为 0.3199;1 点至 200 点  $x$  分量的最大 Lyapunov 指数约为 0.2563;8501 点至 9000 点  $x$  分量的最大 Lyapunov 指数约为 0.1491;12001 点至 12500 点  $x$  分量的最大 Lyapunov 指数约为 0.3192.以上 Lyapunov 指数分析充分说明了变参数 Ikeda 系统前后时段处于不同混沌态和非混沌态,系统的动力学特性在不断的发生改变.

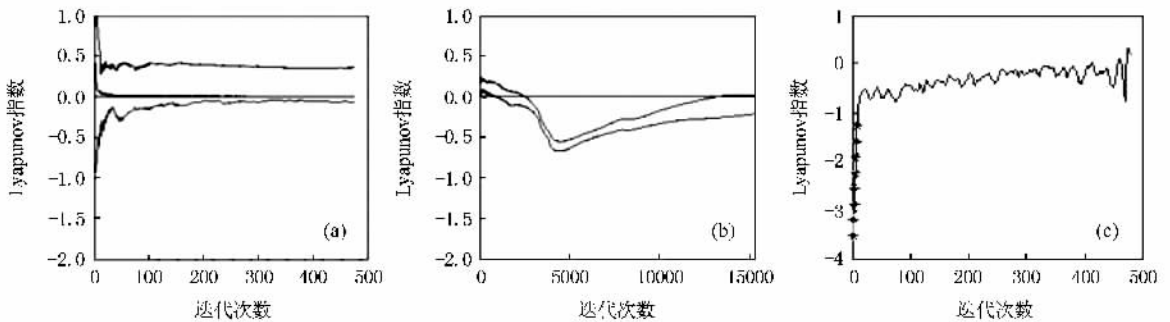


图 7 变参数 Ikeda 系统 Lyapunov 指数 (a)7100 点至 7600 点累计结果 ;(b)1 至 15300 点累计结果 ;(c) $x$  分量最大 Lyapunov 指数 (“- \* -”为拟合)

### 3.2. 变参数混沌系统神经网络预测

#### 1) 参数不变混沌系统的神经网络预测

与定义域上的函数逼近略为不同,混沌系统预测研究主要是利用已获得的数据对将来进行预测.在采用神经网络模型预测混沌系统时,通常通过已有数据训练神经网络,使得神经网络能够“学习”具备系统的特性,从而实现预测.首先回顾两个定理<sup>[30-31]</sup>:

**定理 1 (Kolmogorov 定理)** 给定任一连续函数  $f : U^n \rightarrow R^m$ ,  $f(X) = Y$ , 其中  $U$  是闭单位区间  $[0, 1]$ ,  $f$  可以精确地用一个三层前向网络实现.

**定理 2 (BP 定理)** 给定任意  $\epsilon > 0$  和任意  $L^2$  函数  $f [0, 1] \rightarrow R^m$ , 存在一个三层 BP 网络,它可在任意  $\epsilon$  平方误差精度内逼近  $f$ .

Kolmogorov 定理即映射网络存在定理. Kolmogorov 定理保证任一连续函数可由一个三层前向网络来实现, BP 定理说明只要用三层的 BP 网络就可以实现  $L^2$  函数.基于 Takens 定理,在吸引子相似的意义下可以利用从混沌系统中观测到的单变量时间序列实现混沌系统的嵌入空间重构.那么对于具有恒定形状奇怪吸引子的参数不变混沌系统,利用 Takens 嵌入定理容易在有限维空间中重构出吸引子形状,由定理 1 和定理 2 不难利用神经网络学

习这类系统的动力学行为,同时又由于系统的动力学特性始终不发生变化,在已有时段数据上训练后的神经网络模型可以适用于将来时间段上数据的预测,也就是说训练后的网络能够保证对在将来时段上也有较好泛化性能.对连续方程混沌系统,在由方程产生混沌数据时只要计算步长足够小,或者是对采样数据的采样频率和采样精度足够高,则都能够精细的刻画混沌吸引子的形状,那么采用的训练数据越多则就能够得到更多对该混沌系统的经验知识,所以能够解释训练样本越多能够获得越好的训练效果这一现象.对迭代映射混沌系统,只要迭代计算的次数越多,那么获得的混沌数据也越多就能够细致的刻画混沌吸引子,在相空间中获得一个更连续的吸引子形状曲线,正如前文分析数据越多吸引子形状曲线更为密实,所以同样能够解释训练样本越多能够获得越好的训练效果.如 Logistic 为例,迭代次数越多,则图 2(b)中吸引子的越密实,当已知一个数据预测下一个数据时,很容易根据该图估算出下一个点的数值.基于经验风险最小化原则的神经网络学习理论,决定了训练样本越多则泛化预测效果也越好.

用神经网络预测这类混沌系统时,如何选取网络的隐层数和节点数,目前还没有确切的方法和理论,要达到一定的预测精度,具体算法上的计算复杂度也无法事先确定,这往往与待学习的问题相关.从时序图来看,一维 Logistic 混沌映射非常类似于随机过程,但 Logistic 映射容易用一个具有较少隐节点的三层网络实现较为精确的预测,而 Mackey-glass 时滞微分系统的准确预测相对较难,后者被很多学者作为预测研究的基准问题.可见,用神经网络精确预测这类具有恒定形状吸引子的混沌系统的实现难易,与混沌系统的复杂度<sup>[19,32]</sup>关系密切.

## 2) 参数可变混沌系统的神经网络预测

Takens 定理仅仅说明了在已知嵌入维空间里可以把有规律的轨迹恢复出来,具体实现可能需要较多的时间序列先验知识.对于实际中性质未知的复杂时间序列,基于 Takens 定理的预测方法往往无能为力.由于参数的变化,变参数混沌系统随着时间变化,系统的动力学特性可能在不停的发生改变,如变参数 Ikeda 系统,该系统由混沌到周期态,再进入混沌态,在混沌态和周期态中反复变化,在系统的整个相空间中不存在一个有规律的奇怪吸引子形状,当  $n$  增加到很大时,系统的动态特性会趋于无穷大,就

不存在吸引子了,因此得到的混沌时间序列具有很强的暂态特性.在这种情况下,Takens 定理至少在原理上是无效的,然而即使这样,在相对短的时间内仍然表现出很强的规律性,所以仍希望可以找到其存在的近似函数关系,从而实现短期预测.

结合上述分析,不能简单依据定理 1 和定理 2 建立一个结构和参数都不变的神经网络模型,对变参数混沌系统进行全时段的精确预测.系统参数的不断改变,导致系统状态发生变化,致使预测问题的复杂化.从抽取的单个混沌序列去学习、辨识和跟踪一个变参数系统一般都要求有较多采样数据或样本数目,然而在积累观测数据的过程中系统的动力学状态又发生了深刻的改变.换句话说,利用一段时间内的数据训练神经网络,如果这个神经网络学习获得了这段时间的系统的动力学特性,那么它能够在这一段时间范围内实现很好的预测结果,在延伸的一段时间内也可能具有很好的泛化性能.但是迁移到相距较长时段的数据上,再采用原来训练后的网络进行预测,显然不会获得较好的预测结果,泛化性能很难保证.还有可能所建神经网络预测模型在最近一段时间内堪用,过一段时间再用于该系统的预测,则泛化性能下降很大.

影响神经网络预测泛化性能的主要因素有训练样本、网络结构和学习算法,这三个方面是相互关联的.这里仅分析在样本数据采样充分并且足够多情况下训练样本对网络泛化性能的影响.当前神经网络学习问题研究基于如下假设开展:神经网络学习问题是根据训练样本确定输入  $x$  和输出  $y$  之间映射关系,其中  $(x, y)$  服从某未知的联合分布  $P$ ,而有  $N$  个样本组成的训练样本集  $D = \{(x_1, y_1), \dots, (x_N, y_N)\}$  则是  $(x, y)$  的一组实现,于是神经网络对训练样本集的学习可以描述为一个非线性回归问题,就是根据训练样本集  $D$  构造一个函数  $f(x; D)$ ,使得  $f(x; D)$  在未知点  $x$  处能够较好的逼近实际输出.但对于变参数混沌时间序列,显然这个基础并不严格成立,因为系统参数在变,系统动力学特性跟着发生改变,那么训练样本  $(x, y)$  所服从的某未知联合分布  $P$  是不稳定的,这正是造成神经网络方法预测变参数混沌时间序列研究复杂化的根本原因所在.

总之,由于内在的非线性动力学特性,混沌系统的确定性结构决定了系统运动轨迹短期内发散较小,所以利用观察资料进行短期预测仍然是有可能



的神经网络预测方法需要用尽可能多的观察时间序列来近似产生序列的重构轨迹,使其能够在预测模型中反映出原系统的动力学行为,但变参数混沌系统的动力学特性在不断变化,要取得较好的短期预测效果,训练样本集的合理选择以及训练后网络的适用范围需要考虑。接下来通过试验分析训练数据集的选取对变参数混沌时间序列预测的精度和泛化性能产生的影响。

## 4. 仿真试验

为了深入研究神经网络预测变参数混沌系统遇到的问题并验证上述理论分析结果,本文将变参数的 Ikeda 混沌时间序列预测作为仿真实例。采用典型的三层前向神经网络,隐层神经元传递函数为正切 S 型函数(tansig),输出神经元传递函数为对数 S 型函数(logsig)。为了便于检验试验结果,利用 Matlab 神经网络工具箱,采用其中 Levenberg-Marquardt 算法训练神经网络。Levenberg-Marquardt 算法是牛顿法的变形,实践研究表明该算法收敛速度快,训练精度高,对具有恒定形状吸引子混沌系统预测精度高于很多文献所用方法,且不存在振荡效应,受初始参数的影响较小,只要初始梯度矩阵行列式值不是很小,算法一般都能收敛,主要的缺点是受制于该算法提出时计算机性能的影响,参数数量很多时难以实现操作很大的梯度矩阵<sup>[33]</sup>。为了衡量对比网络训练效果,采用如下误差平方和(SSE)和均方误差(MSE)作为评价指标。

$$SSE = \sum_{n=1}^N (x_n - \hat{x}_n)^2, \quad (7)$$

$$MSE = \frac{1}{N} \sum_{n=1}^N (x_n - \hat{x}_n)^2, \quad (8)$$

选取变参数 Ikeda 混沌系统的  $x$  分量进行预测研究。试验中选择如下数据集,总共取三组数据,每组含有两段连续数据,数据分段:第一组(1—200), (201—1000);第二组(7101—7600)(7601—8100);第三组(8551—8950),(8951—9550),从图 5 中时序图上观察,每组两段数据分别是系统处于不同运动状态的数据采样,若不作特殊说明每组中第一段数据为训练集,第二段数据为测试集,着重分析前两组数据试验结果。分组归一化,由于每组中两段数据分布范围比较接近,因此不会因归一化问题影响本文分析结果可信性。构造神经网络预测用数据结构:  
 $[x(t-6), x(t-5), x(t-4), x(t-3), x(t-2), x$

$(t-1), x(t)]$  即选择嵌入维为 6,由过去六个数据点实施一步预测。

### 4.1. 第一组数据分析

从时序图上观察,第一组数据选取在系统由混沌状态逐渐转入周期分叉过程阶段。

先取(1—200)点数据为训练数据(201—1000)点数据为测试数据。试验中选择网络学习速率为 0.0001,隐层节点 20 个,进行 20000 次批处理训练,结果如图 8 所示,其中在训练集上的均方误差  $MSE = 4.8855 \times 10^{-8}$ ,误差平方和  $SSE = 9.7709 \times 10^{-6}$ 。保持网络结构和训练参数不变,改变学习速率,多次试验结果相近,均方误差量级保持在  $10^{-8}$ — $10^{-9}$  左右。

从图 8 可以看到尽管在训练集上获得了很好的预测效果,在测试集上的泛化性能较差。当选择 10 个隐节点时训练预测精度降低(如  $MSE = 3.1561 \times 10^{-6}$ ),泛化性能也较差;增加隐节点为 40 个,能在训练集上获得了更好的训练效(如  $MSE = 4.9824 \times 10^{-13}$ ),但在测试集上的泛化性表现更差;选择嵌入维为 3 时上述现象更明显。分析原因在于由于混沌系统参数改变,系统的动力学特性发生改变,前段数据主要处于混沌态,后段数据主要处于周期态,在训练集上获得的神经网络会逐渐不适用于周期态的数据。

现在分析在训练集上是否存在过拟合现象,保持隐节点 20 个的网络结构不变,从训练集 200 样本点中随机挑选出 160 个和 40 个样本,前一部分用作训练,后一部分用作测试,采用同样仿真参数,结果如图 9 所示。虽然有少数点的预测有一定偏差,考虑到该段数据上系统动力学的前后变化,可以认为不存在明显的过拟合。但随隐节点数的逐渐增多,在此训练集上的预测精度提高很快,过拟合会慢慢显现。

### 4.2. 第二组数据分析

从时序图上观察,第二组数据选取在系统由混沌状态突然转入非周期运动过程阶段。

先取(7101—7600)点数据为训练数据(7601—8100)点数据为测试数据。试验中选择网络学习速率 0.0001,隐节点 40 个,进行 20000 次批处理训练,结果如图 10 所示,其中在训练集上的均方误差  $MSE = 7.9914 \times 10^{-9}$ ,误差平方和  $SSE = 3.9957 \times 10^{-6}$ 。保持网络结构和训练参数不变,学习速率可调,多次试

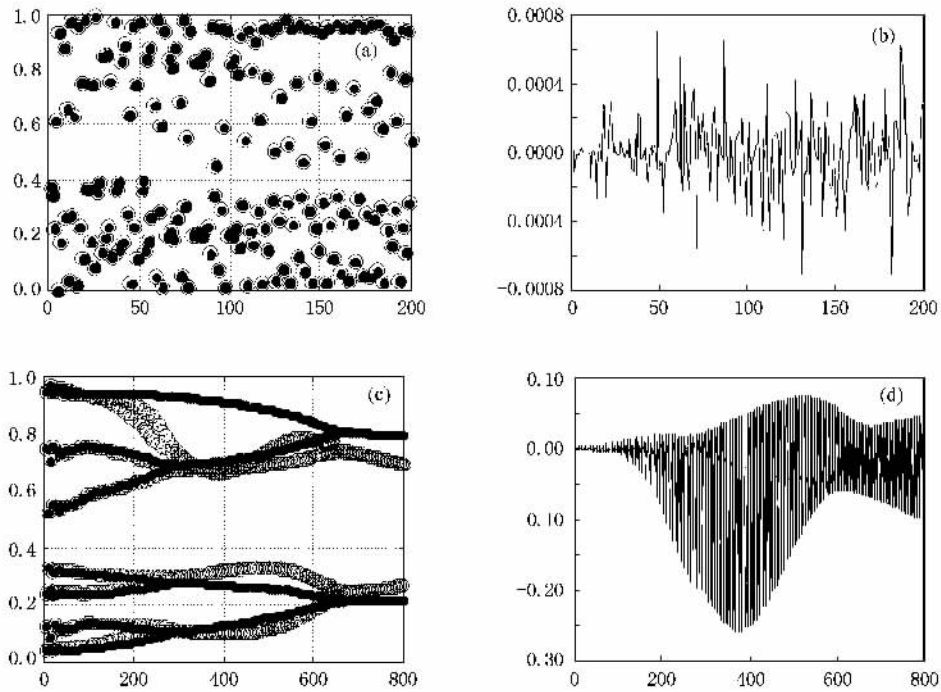


图 8 第一组前 200 点训练后 800 点测试试验 (a) 为训练后网络在训练集上的预测结果(“·”真实值,“o”预测值,下同);(b)为训练集上预测误差;(c)为在测试集的预测结果;(d)为测试集上预测误差

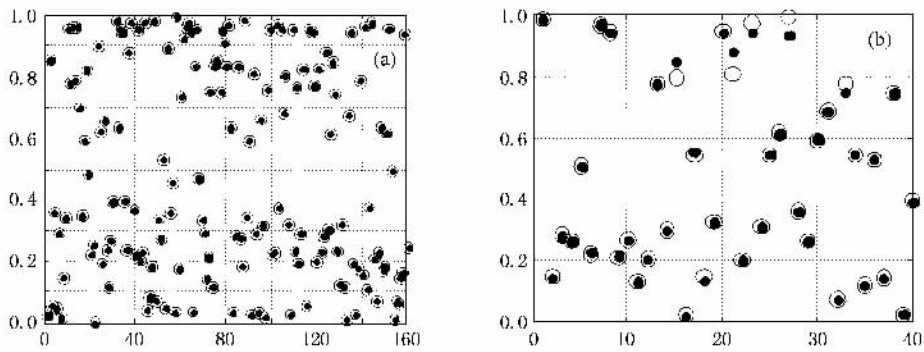


图 9 过拟合检验 (a) 过拟合训练预测结果;(b) 过拟合测试预测结果

验结果相近 均方误差量级保持在  $10^{-8}$ — $10^{-9}$  左右.

从图 10 可以看到尽管在训练集上获得了很好的预测效果,但在测试集上的根本不具有泛化性,增加隐节点也只能在训练集上获得了更好的预测效果.原因在于由于混沌系统参数改变,在该组两段数据上系统的动力学特性发生了深刻变化,即由混沌状态突然转入非周期态,所以在训练集上获得的神经网络不适用于测试集上的数据预测.过拟合检验的结果表明在训练集上不存在明显的过拟合.

若选择 20 个隐节点,调节训练参数反复试验,得到训练集上均方误差 MSE 量级约  $10^{-6}$  左右,在该组数据上的训练效果远远达不到 40 个隐节点的训

练效果,在训练集上预测结果相差很大,但都在测试集上不具有泛化性.已知变参数 Ikeda 混沌时间序列  $x$  分量的 1 点至 200 点最大 Lyapunov 指数约为 0.2563,7101 点至 7600 点  $x$  分量的最大 Lyapunov 指数约为 0.3199,两段数据上的动力学特性明显不同,所以适用于第一组中(1—200)点数据的具有 20 个隐节点的神经网络已经不再适用于(7101—7600)点数据段.

虽然以(7101—7600)点数据为训练样本不能实现对(7601—8100)点数据的预测,并不是所有(7601—8100)点数据都无法预测.选择(7601—7650)的 50 个点为训练集用较为简单的网络结构可以实



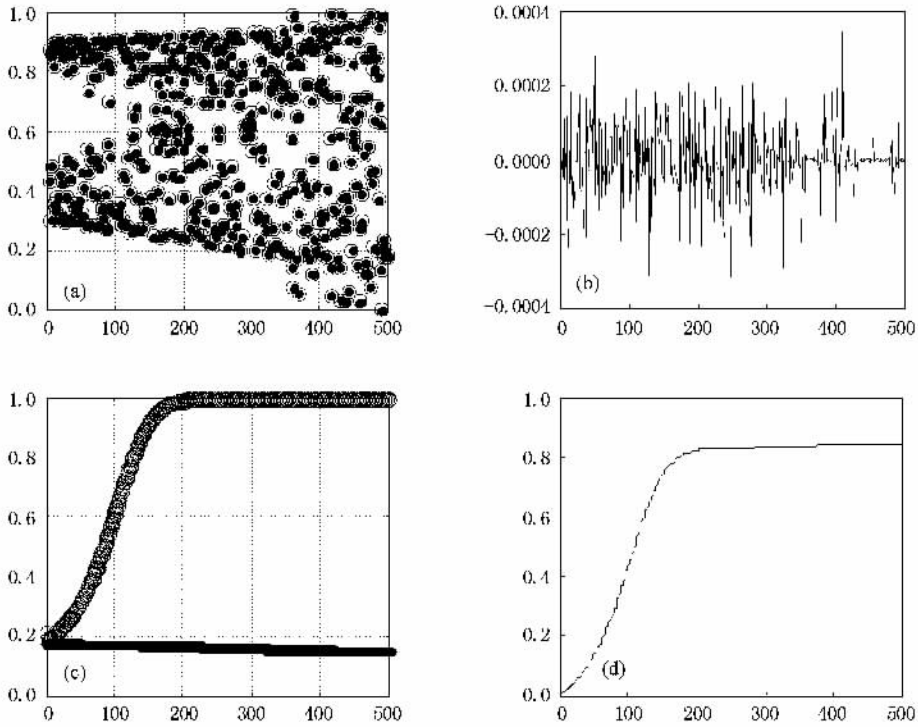


图 10 第二组前 500 点训练后 500 点测试试验 (a)为训练后网络在训练集上的预测结果;(b)为训练集上预测误差(c)为在测试集的预测结果;(d)为测试集上预测误差

现对(7651—7700)点数据很精确的预测,因为它们都处于近似线形段上,本文不再图示.选择(7101—

7650)点数据为训练样本对(7651—7700)点数据进行预测,这时选择 20 个隐节点,其他网络参数不变,

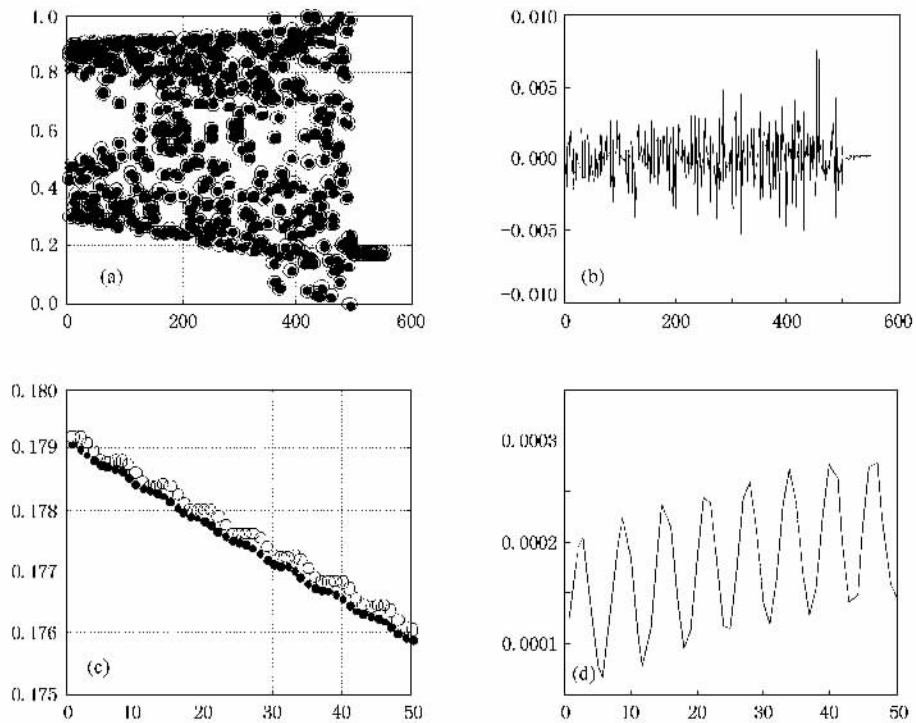


图 11 训练集选择试验 (a)为训练后网络在训练集上的预测结果;(b)为训练集上预测误差(MSE =  $2.5942 \times 10^{-6}$ );(c)为在测试集的预测结果;(d)为测试集上预测误差

20000 次批处理训练后的结果如图 11 所示. 可见选择(7101—7650)点数据为样本不能实现对(7651—7700)点数据非常准确的预测,而且对7700点后数据的预测逐渐变的很差. 选择不同的隐节点数和训练次数多次试验结果表明预测效果没有显著改进. 从中可以看出,对于变参数混沌系统进行预测选择合理的训练集对获得好的预测效果非常重要.

对于动力学特性保持不变的具有恒定形状吸引子的参数不可变混沌系统,训练数据越多预测效果越好,而对于变参数混沌系统并非训练数据越多泛化预测性能越好. 可见实际复杂系统预测中通常认为除增加计算量外训练样本数据越多则预测效果也越好并不都成立.

一个较为困难的问题是,要想利用7601点之前的数据(... ,7600)对(7601—7650)点数据准确预测(这里准确预测指 $\forall x_n \in [x_{7601}, x_{7650}], |x_n - \hat{x}_n| \leq 10^{-3}$ ),采用上述非自适应神经网络,多次选择不同训练集、隐节点数目以及网络训练参数,都未能实现较好预测效果,成为单变量预测中的“盲点”. 采用本文的预测数据结构,对这段数据实施精确的短期预测很困难,这种情况下不是缺少前期数据,而是系统的动力学特性发生了质的改变,并且预测时缺少与 $x$ 分量紧密相关的 $y$ 分量的信息,而在对实际复杂系统预测时往往也只能获得有限的单方面的数据信息.

最后采用上述方法可以分析在第三组数据上的试验情况. 从时序图上观察,第三组数据选取在周期分岔进入混沌再由周期分岔进入混沌的连续阶段. 试验结果仍然表明尽管在训练集上获得了很好的训练效果,在测试集上的根本不具有泛化性. 这三组训练数据包含了不同类型的数据,有混沌态,有拟周期态,甚至交叉状态,多次试验都在三组数据的训练数据集上获得了很好的训练预测效果,这说明了变参数 Ikeda 系统时间序列具有很强的规律性,不是随

机的时间序列,因为随机序列是无法准确学习预测的. 试验还表明适当选择的训练集能够实现对很多数据的精确预测,但也存在一些数据段很难通过以前数据来预测. 同样结构的神经网络模型对于变参数 Ikeda 混沌系统不同数据段预测能力表现出很大差异,这也使得一般的基于参数调整的自适应预测方法很难适用.

## 5. 结 论

根据混沌相空间时空轨迹所示奇怪吸引子形状,将混沌系统分为两类:具有恒定形状吸引子的混沌系和不具有恒定形状吸引子的混沌系统,复杂变参数混沌系统多属于后者,构造了一个变参数 Logistic 映射,分析了典型变参数混沌系统的动力学特性. 致力于将混沌系统的动力学特性与神经网络学习能力联系起来分析变参数混沌系统的预测,结合变参数混沌系统的动力学特点,分析指出神经网络预测这类系统可能存在的问题,进行了初步探讨. 以变参数 Ikeda 系统为例,通过仿真试验,研究分析神经网络预测变参数混沌系统遇到的难题. 试验结果表明训练集的选择对泛化预测效果影响极大,训练样本也不再是越多泛化效果越好,甚至系统存在难以预测的“盲点”. 分析原因在于,对变参数 Ikeda 系统的吸引子不再保持恒定形状,Takens 定理不再全程成立,系统的动力学行为不断发生改变,利用离线神经网络学习掌握系统的内在规律十分困难,学习后的网络难以保证在后续系统动力学特性已改变的数据上有较好的泛化性能.

虽然变参数混沌时间序列的预测问题已经引起了人们的注意,但是深入的分析还未见报道. 以上研究也表明,对于更具有实际意义的复杂系统的预测问题在认知上有待深化,预测理论和方法需要进一步完善.

- [1] Haykin S, Li X B 1995 *Proceedings of the IEEE* **83** 95
- [2] He T, Zhou Z O 2007 *Acta Phys. Sin.* **56** 693 (in Chinese) [贺涛、周正欧 2007 物理学报 **56** 693]
- [3] Farmer J D, Sidorowich J J 1987 *Phys. Rev. Lett.* **59** 845
- [4] Takens F 1981 *Lecture Notes in Mathematics* (Berlin: Springer-Verlag) **898** 366
- [5] Li J, Liu J H 2005 *Acta Phys. Sin.* **54** 4569 (in Chinese) [李军、刘君华 2005 物理学报 **54** 4569]
- [6] Tan W, Wang Y N, Zhou S W, Liu Z R 2003 *Acta Phys. Sin.* **52** 795 (in Chinese) [谭文、王耀南、周少武、刘祖润 2003 物理学报 **52** 795]
- [7] Zhang J S, Xiao X C 2000 *Acta Phys. Sin.* **49** 403 (in Chinese) [张家树、肖先赐 2000 物理学报 **49** 403]
- [8] Ren R, Xu J, Zhu S H 2006 *Acta Phys. Sin.* **55** 555 (in Chinese) [任韧、徐进、朱世华 2006 物理学报 **55** 555]
- [9] Zhang J S, Xiao X C 2000 *Acta Phys. Sin.* **49** 1221 (in Chinese)

- [ 张家树、肖先赐 2000 物理学报 **49** 1221 ]
- [ 10 ] Gan J C , Xiao X C 2003 *Acta Phys. Sin.* **52** 1096 ( in Chinese )  
[ 甘建超、肖先赐 2003 物理学报 **52** 1096 ]
- [ 11 ] Gan J C , Xiao X C 2003 *Acta Phys. Sin.* **52** 2995 ( in Chinese )  
[ 甘建超、肖先赐 2003 物理学报 **52** 2995 ]
- [ 12 ] Yan H , Wei P , Xiao X C 2007 *Acta Phys. Sin.* **56** 5111 ( in Chinese ) [ 闫 华、魏 平、肖先赐 2007 物理学报 **56** 5111 ]
- [ 13 ] Lorenz E N 1963 *J. Atmos. Sci.* **20** 130
- [ 14 ] Yao H X , Sheng Z H , Chen H X 2002 *Sys. Engi. Theo. Prac.* **22** 33 ( in Chinese ) [ 姚洪兴、盛昭瀚、陈洪香 2002 系统工程理论与实践 **22** 33 ]
- [ 15 ] Haykin S , Puthusserypady S 1997 *Chaos* **7** 777
- [ 16 ] Ye M Y , Wang X D , Zhang H R 2005 *Acta Phys. Sin.* **54** 2568 ( in Chinese ) [ 叶美盈、汪晓东、张浩然 2005 物理学报 **54** 2568 ]
- [ 17 ] Xiao F , Gao X P 2006 *J. of Soft.* **17** 1042 ( in Chinese ) [ 肖芬、高协平 2006 软件学报 **17** 1042 ]
- [ 18 ] Cui W Z , Zhu C C , Bao W X , Liu J H 2004 *Acta Phys. Sin.* **53** 3303 ( in Chinese ) [ 崔万照、朱长纯、保文星、刘君华 2004 物理学报 **53** 3303 ]
- [ 19 ] Lü J H , Lü J A , Chen S H 2002 *Analysis and Application of Chaotic Time Series* ( Wuhan : Wuhan University Press ) p114—131 ( in Chinese ) [ 吕金虎、陆君安、陈士华 2002 混沌时间序列分析及其应用( 武汉 : 武汉大学出版社 ) 第 114—131 页 ]
- [ 20 ] Mackey M C , Glass L 1977 *Science* **197** 287
- [ 21 ] Zhang Y H , Qi G Y , Liu W L , Yan Y 2006 *Acta Phys. Sin.* **55** 3307 ( in Chinese ) [ 张宇辉、齐国元、刘文良、阎 彦 2006 物理学报 **55** 3307 ]
- [ 22 ] Wang F Q , Liu C X 2007 *Acta Phys. Sin.* **56** 1983 ( in Chinese )  
[ 王发强、刘崇新 2007 物理学报 **56** 1983 ]
- [ 23 ] Zhang J S , Dang J L , Li H C 2007 *Acta Phys. Sin.* **56** 67 ( in Chinese ) [ 张家树、党建亮、李恒超 2007 物理学报 **56** 67 ]
- [ 24 ] Ikeda K 1979 *Opt. Commun* **30** 257
- [ 25 ] Kantz H , Schreiber T 1997 *Nonlinear Time Series Analysis* ( Cambridge : Cambridge University Press ) p90
- [ 26 ] Yang S Q , Zhang X H , Zhao C A 2000 *Acta Phys. Sin.* **49** 636 ( in Chinese ) [ 杨绍清、章新华、赵长安 2000 物理学报 **49** 636 ]
- [ 27 ] Briggs K 1990 *Phys. Lett. A* **151** 27
- [ 28 ] Eckmann J P , Ruelle D 1985 *Rev. Mod. Phys.* **57** 617
- [ 29 ] Rosenstein M T , Collins J J , Deluca C J 1993 *Phys. D* **65** 117
- [ 30 ] Hu S R 1993 *Neural Networks Rudiment* ( Changsha : National Defence Technology University Press ) ( in Chinese ) [ 胡守仁 1993 神经网络导论( 长沙 : 国防科技大学出版社 ) ]
- [ 31 ] Yan P F , Zhang C S 2005 *Artificial Neural Networks and Evolutionary Computing* [ Second Edition ] ( Beijing : Tsinghua University Press ) ( in Chinese ) [ 阎平凡、张长水 2005 人工神经网络与模拟进化计算( 第 2 版 ) ( 北京 : 清华大学出版社 ) ]
- [ 32 ] Liu B Z , Peng J H 2004 *Nonlinear Dynamics* ( Beijing : Higher Education Press ) p430—454 ( in Chinese ) [ 刘秉正、彭建华 2004 非线性动力学( 北京 : 高等教育出版社 ) 第 430—454 页 ]
- [ 33 ] Martin T H , Howard B D , Mark H B 2002 *Neural Network Design* ( Beijing : China Machine Press , CITIC Publishing House ) p12—27

# Prediction of the chaotic time series from parameter-varying systems using artificial neural networks

Wang Yong-Sheng<sup>†</sup> Sun Jin Wang Chang-Jin Fan Hong-Da

( *Department of Armament Science and Technology , Naval Aeronautical Engineering University , Yantai 264001 , China* )

( Received 17 December 2007 ; revised manuscript received 10 May 2008 )

## Abstract

Prediction of the chaotic time series generated by the complex parameter-varying systems is researched in this paper. The parameter-varying Logistic system is constructed firstly , and the properties of this kind of system are analyzed. These systems , whose parameter values change with time , do not have attractor shape invariable with time evolution because of their continually changing dynamical property. Combining the Takens ' embedding theorem and the artificial neural networks ( ANN ) theory , we interpret the feasible reason that ANN method can be used to predict the chaos systems with the invariable attractor shape , and then discuss the potential problem that will be met when using ANN to predict the parameter-varying system. Experiments of forecasting the chaotic time series from parameter-varying Ikeda system using neural networks have been performed. The previous theoretical analyses are validated by the experiment results. The results also show that if only simply increasing the training data , the neural networks ' predicting generalization ability may be reduced , the generalized predicting result on the parameter-varying system is especially seriously affected by the selected training data set. So prediction of the parameter-varying systems must be well resolved before the chaotic time series prediction can be made practical.

**Keywords** : chaos , prediction , artificial neural networks , parameter-varying dynamical system

**PACC** : 0545

---

<sup>†</sup> E-mail : hjhywys@sohu.com