文章编号：1001-0920(2009)10-1526-05

基于储备池主成分分析的多元时间序列预测研究

韩 敏，王亚楠
（大连理工大学 电信学院，辽宁 大连 116024）

摘 要：提出一种基于回声状态网络储备池的非线性 PCA 方法，并将其应用于多元时间序列的预测中. 由于多维输入变量间的相关性会影响建模效果，通过储备池将输入在原空间的非线性特征转化成高维空间的线性特征. 在其中运用线性 PCA 技术寻找输入在储备池空间的最大方差方向，提取有效的多元变量综合信息. 经储备池主成分分析处理后的输入与预测点呈动态线性映射，可使用线性方法建模. 仿真结果表明了该方法的有效性.

关键词：储备池主成分分析；回声状态网络；多元时间序列；预测
中图分类号：TP183 **文献标识码**：A

Prediction of multivariate time series based on reservoir principal component analysis

HAN Min, *WANG Ya-nan*
(School of Electronic and Information Engineering, Dalian University of Technology, Dalian 116024, China. Correspondent: HAN Min, E-mail: minhan@dlut.edu.cn)

Abstract: A nonlinear principal component analysis (PCA) method based on echo state network (ESN) reservoir is proposed for multivariate time series prediction. As the correlations among multivariate inputs have adverse effect on modeling, the reservoir is utilized for translating the inputs' nonlinear features in the original input space to linear features in the high-dimension reservoir state space. Then the linear PCA is performed on mapped data to find the uncorrelated directions of maximum variance and thus to extract the joint information of multiple variables. Furthermore, dynamic linear mapping is obtained from the RPCA outputs to the predictor points, so linear algorithms are investigated for modeling. The simulation results show the effectiveness of the proposed method.

Key words: Reservoir principal component analysis; Echo state network; Multivariate time series; Prediction

1 引 言

在自然界和人类社会中，复杂系统是普遍存在的，很多复杂系统包含了非线性的混沌特征. 对复杂的混沌非线性系统研究的最主要任务之一是如何从多变量时间序列中提取有用的信息来刻画复杂系统的动力学动态特性[1].

处理混沌时间序列的基本方法是利用延迟向量重构相空间，在此基础上进行非线性预测[2]. 根据 Takens 嵌入理论[3]，只要嵌入维数和延迟时间选择得合理，单变量时间序列即可较好地重构相空间，并获得较理想的预测效果. 然而，实际中获得的有限单时间序列所包含的信息通常具有不完备性与不确定性，无法保证重构的相空间能十分准确地描述动力系统状态变量的演化轨迹. 多变量时间序列包含有比单变量时间序列更丰富的信息，能重构出更为准确的相空间. 但当多个输入变量所代表的状态间相关性较强时，会影响建模效果. 主成分分析(PCA)[4,5]是一种利用变量间线性相关关系对多维信息进行统计压缩的方法，能够有效解决输入相关性问题. 但对于复杂非线性问题，一般的线性 PCA 方法无法对数据信息进行有效提取. 核主成分分析(KPCA)[6,7]能有效捕捉数据的非线性特征. 但在求解过程中，需特征分解一个 $N\times N$(N 为训练样本数)的核矩阵，计算代价大，测试样本特征提取过程繁琐. 尽管很多相关改进算法的出现使其运算效率提高，但总体上 KPCA 及相关算法都是依托核函数

收稿日期：2008-11-11；**修回日期**：2009-01-20.
基金项目：国家自然科学基金项目(60674073)；国家 863 计划项目(2007AA04Z158)；国家科技支撑计划项目(2006BAB14B05)；国家 973 计划项目(2006CB403405).
作者简介：韩敏(1959—)，女，吉林延吉人，教授，博士生导师，从事神经网络、专家系统等研究；王亚楠(1985—)，女，河南南阳人，硕士生，从事神经网络、多元时间序列预测的研究.

进行，具有运算复杂、数据处理欠直观的缺陷[8].

回声状态网络(ESN)是 Jaeger[9] 提出的一种新型递归网络. 它具有独特的储备池回声状态机制，将非线性的部分分离出去由储备池处理，网络的训练仅需要确定网络的输出权值，可通过线性算法实现. 文献[10]利用储备池搭建高维特征空间，在其中运用线性支持向量技术，对非线性问题进行求解.

本文在已有研究的基础上，提出储备池主成分分析(RPCA)方法，并将其用于复杂系统建模. 利用储备池状态空间映射替代 KPCA 的未知 Φ 映射，然后运用线性 PCA 技术. 由于储备池状态可求，在 RPCA 中只需特征分解一个 $d\times d$ (d 为初始储备池维数，一般远小于 N)的储备池状态相关矩阵，并能直接对测试样本进行特征提取. 同时，经 RPCA 处理后的输入至预测向量间的映射关系，可用线性方法逼近.

2　基于 RPCA 的多元时间序列预测建模

本节将对 RPCA 方法进行较为具体的介绍，并将其用于多变量嵌入延迟向量相空间重构中.

混沌时间序列非线性分析的第 1 步就是相空间重构，延迟坐标状态空间重构是较常用的方法. 设多元离散时间序列为

$$\{x_1(t),x_2(t),\cdots,x_n(t)\},\ t=1,2,\cdots,N,$$

其中 N 表示时间序列的长度. 选取 m_i 和 τ_i($i=1,2,\cdots,n$) 为第 i 个混沌时间序列的嵌入维数和时间延迟，则第 i 个时间序列的嵌入延迟窗为

$$T_{im}=(m_i-1)\tau_i. \tag{1}$$

得到多元变量嵌入延迟向量为

$$\boldsymbol{X}(k)=[\boldsymbol{x}_1^{\mathrm{T}}(k),\boldsymbol{x}_2^{\mathrm{T}}(k),\cdots,\boldsymbol{x}_n^{\mathrm{T}}(k)], \tag{2}$$

其中

$$\boldsymbol{x}_i(k)=[x_i(k),x_i(k-\tau_i),\cdots,x_i(k-(m_i-1)\tau_i)]^{\mathrm{T}}. \tag{3}$$

取 $\hat{L}=N-\max\limits_i(T_{im}),k=1,2,\cdots,\hat{L}$. 多元嵌入延迟矩阵 $\boldsymbol{X}$ 为 $\hat{L}\times M$ 矩阵. 对于 m_i 和 τ_i 的选取，目前仍没有统一的方法. 本文采用 C-C 方法进行选择，它利用关联积分原理并结合统计学思想，具有计算量小、需要数据量少、容易操作和抗噪能力强的优点[11]. 对于重构后的相点，选用线性回归方法逼近其与预测向量直接的映射关系，建立多元变量预测模型为

$$\boldsymbol{y}_i(t+\eta)=F(\hat{x}_i). \tag{4}$$

其中：$F(\cdot)$ 表示线性模型，$\boldsymbol{y}_i(t+\eta)$ 为未来 $t+\eta$ 时刻 $\boldsymbol{x}_i$ 的预测值，$\hat{\boldsymbol{x}}_i$ 为 t 时刻储备池主元向量.

2.1　储备池主成分分析

回声状态网络由输入层、中间层和输出层构成，各层间通过权值连接. 网络输入层和输出层间的部分即称为储备池，是一种大且稀疏的递归结构，其内部节点状态为输入信号的高维显现，本质上表征了输入信号在高维空间中线性特性. 网络储备池的状态方程为

$$\boldsymbol{x}(k+1)=\mathrm{tansig}(W_x\boldsymbol{x}(k)+W_{\mathrm{in}}\boldsymbol{u}(k)). \tag{5}$$

其中：$\boldsymbol{x}(k)$，$\boldsymbol{u}(k)$ 分别代表 k 时刻 ESN 的状态变量、输入变量；W_x 和 W_{in} 分别为储备池内部的连接矩阵及外部输入与储备池神经元间连接矩阵. 通常，状态变量 $\boldsymbol{x}$ 的维数很高(一般取 100 ～ 1000 之间)，矩阵 W_x 保持 1% ～ 5% 的稀疏连接. W_x 和 W_{in} 经初始化后保持不变，当 W_x 的谱半径(矩阵所有特征值模的最大值，下文表示为 $\rho(W_x)$) 小于 1 时，可保证网络的稳定运行.

正是由于储备池独特的回声特性，ESN 较传统递归网络在网络结构和网络训练方面均有极大的突破. 根据 ESN 的储备池机制，非线性处理部分完全分离出去由储备池完成，网络学习过程仅需要确定储备池至输出节点间的权值矩阵. 而输入数据激发得到储备池状态的过程，相似于核方法中将待处理的数据通过未知函数 Φ 映射至高维 Hilbert 空间的过程[10].

由于“核方法”和“储备池方法”之间具有共通性，而 KPCA 是将待处理的数据先通过未知函数 Φ 映射至高维 Hilbert 空间中，再利用线性 PCA 对 Hilbert 空间中的数据进行处理. 将 KPCA 中的非线性 Φ 映射替换为 ESN 中的储备池映射就构成了本文的储备池主成分分析方法. 然而不同于 Hilbert 空间中的未知 Φ 映射，由于储备池的节点状态可通过式(5) 计算得到，本文的 RPCA 较 KPCA 方法具有较大的性能优势. 首先，KPCA 中的 Φ 映射是一种静态函数映射，而 RPCA 储备池内部状态的递归特性使得其具有映射动态特性的能力. 其次，在 KPCA 中 Φ 映射的状态未知，无法直接用线性 PCA 求取主元变化矩阵，必须通过求 Mercer 核来完成特征提取过程. 运算复杂度高达 $O(N^3)$(N 是样本个数)，而储备池中的节点状态能够显性求出，可直接应用 PCA 技术进行储备池状态空间重构和特征提取，运算复杂度仅为 $O(d^3)$(d 为储备池维数，通常有 $d<N$). 最后，在 KPCA 中，测试样本的主元矩阵同样需通过核函数计算得到，过程繁琐且运算量大. 而在 RPCA 中可直接通过变换矩阵计算得到. RPCA 的模型如图 1 示，具体运算步骤如下.

Step1：对输入值进行整理，有

$$U=[\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_N]^{\mathrm{T}},\ U\in R^{N\times l}.$$

Step2：根据问题确定合适的储备池回声属性，

设定参数 W_x,W_{in} 和 $\rho(W_x)$,将输入代入式(5),得到储备池状态矩阵

$$X=[\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_N]^{\mathrm{T}},\ X\in R^{N\times d}.$$

Step3:求取 X 的相关矩阵 P[12],$P\in R^{d\times d}$. 不妨设 P 的 r 个特征值为 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_r>0$,对应的单位特征向量依次为 $\boldsymbol{q}_1,\boldsymbol{q}_2,\cdots,\boldsymbol{q}_r$.

Step4:选取 $0.8<\delta_0<1$,求取累积贡献率为

$$\delta_h=\sum_{j=1}^{h}\lambda_j\Big/\sum_{i=1}^{r}\lambda_i,$$
$$j=1,2,\cdots,h,\ i=1,2,\cdots,r. \tag{6}$$

当有 $\delta_h>\delta_0$ 时,取变换矩阵 $Q=[\boldsymbol{q}_1,\boldsymbol{q}_2,\cdots,\boldsymbol{q}_h]$.

Step5:计算得到储备池状态主元矩阵为

$$\hat{X}=X\times Q. \tag{7}$$

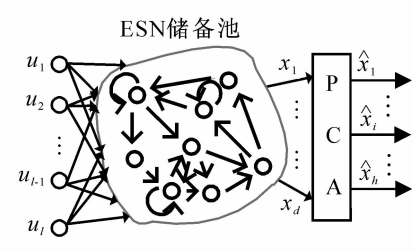

图 1 RPCA 模型图

由于储备池的扩维作用,经 RPCA 处理后的数据维数不一定小于其原始维数,即不保证 $h<l$.

2.2 基于 RPCA 的多变量时间序列预测模型

将 RPCA 方法用于多变量嵌入延迟向量重构中,并利用线性回归算法建立预测模型. 首先,构造如式(2)所示的多元嵌入延迟向量,并按照 Step2 ~ Step5 得到式(7)所示的主元矩阵.

运用RPCA方法对多元变量时间序列数据处理具有以下优点:

1) 对于复杂的多元变量,输入间存在较强的相关性,会影响建模质量. 与 PCA 可以解决线性输入间的相关性问题类似,用 RPCA 能够有效解决非线性输入间的相关性问题.

2) 储备池的回声状态属性,使得 RPCA 更适用于动态系统建模.

3) 由于多元变量的非线性部分在储备池内部被处理,RPCA 的输出向量与预测点间呈线性映射,简化了预测模型,可采用线性方法建模.

正是由于RPCA独特的动态特性和非线性处理特性,使得其适用于复杂时间序列处理. 对于由 RPCA 方法得到的主元矩阵,可采用线性回归方法逼近其与预测向量间的函数映射关系. 这里选用 Tikhonv 正则化方法构建直接预测模型[13],得到模型的线性系数和预测输出分别为

$$\boldsymbol{w}^*=(\hat{X}^{\mathrm{T}}\hat{X}+C^{-1}I)^{-1}\hat{X}^{\mathrm{T}}\boldsymbol{y}, \tag{8}$$
$$y(t+\eta)=\boldsymbol{w}^{*\mathrm{T}}\hat{\boldsymbol{x}}(t). \tag{9}$$

其中:$\hat{x}$ 为训练样本输入储备池主元矩阵,$\boldsymbol{y}$ 为相应的 $t+\eta$ 时刻的期望值. 式(8)中的正则项系数 C 具有平衡模型复杂度和预测精度的功能,本文用网格法对不同数据的 C 值进行筛选.

3 仿真实例

为验证本文方法的有效性,将其应用于 2 个二元变量时间序列仿真中,它们分别来自 Lorenz 混沌方程生成的混沌时间序列与年太阳黑子和黄河年径流实际观测序列.

引入两个性能评价指标(均方根误差 E_{RMSE} 和预测精度 E_{PA})来定量说明网络预测性能的好坏[14],即

$$E_{\text{RMSE}}=\Big(\frac{1}{S-1}\sum_{t=1}^{S}[P_i-O_i]^2\Big)^{1/2}, \tag{10}$$

$$E_{\text{PA}}=\frac{\sum_{t=1}^{S}[(P_i-P_m)(O_i-O_m)]}{(S-1)\sigma_P\sigma_O}. \tag{11}$$

其中:S 是测试样本个数,O_i 是某个变量的实际观测值,P_i 是该变量的预测输出,O_m 是实际观测值的平均,P_m 是预测值的平均,σ_O 和 σ_P 分别是观测值和预测值的标准差. 均方根误差 E_{RMSE} 反映了预测值对观测值的平均偏离程度,取值大于或等于零,预测无误差时等于零;预测精度 E_{PA} 反映了预测值和观测值在其均值附近的偏差之间的相关性,取值在 +1 ~ −1 间,预测无误差时为 1.

3.1 Lorenz 混沌方程 $x(t)$ 时间序列预测

Lorenz 混沌方程如下所示:

$$\begin{cases}\mathrm{d}x/\mathrm{d}t=a(y-x),\\ \mathrm{d}y/\mathrm{d}t=(c-z)x-y,\\ \mathrm{d}z/\mathrm{d}t=xy-bz,\end{cases} \tag{12}$$

其中 a,b,c 为常数,取不同值时,方程表现出不同的性质. 当取 $a=10,b=8/3,c=28,x(0)=y(0)=z(0)=1.0$ 时,系统产生混沌. 利用四阶 Runge-Kutta 方法迭代产生混沌时间序列,使用 $x(t)$ 和 $y(t)$ 时间序列共同预测 $x(t+\eta)$.

由 C-C 方法计算得到 $\tau_1=19,\tau_2=13,m_1=3,m_2=5$. 初始输入样本向量为 $\boldsymbol{x}(t)=\{x(t),\cdots,x(t-2\tau_1),y(t),\cdots,y(t-4\tau_2)\}=\{\boldsymbol{x}_1^{\mathrm{T}},\boldsymbol{x}_2^{\mathrm{T}}\}^{\mathrm{T}}$,预测输出为 $x(t+\eta)$. 当取 $\eta=1$ 时,经相空间重构可产生 2448 组数据,取前 1000 组为训练样本,后 1448 组为检测样本. 为了测试本方法在有噪声的情况下仍然可以较精确地预测. 训练样本中加入了噪声水平 10% 的噪声. 考虑到网络暂态作用,舍弃前 20 个样本. 令储备池节点数为 200,$\eta_0=0.999$. 经 RPCA 处理后原输入向量变为 9 维,其中前 5 位主元的累计方差贡献率已经超过 0.99,具体见表 1. $x(t+1)$ 预

测结果见图 2.

表 1　前 5 位主元特征值、方差贡献率及累积方差贡献率

序号 i	1	2	3	4	5
特征值	112.86	40.164	26.272	12.604	3.8672
贡献率	0.5718	0.2035	0.1331	0.0639	0.0196
累积贡献率	0.5718	0.7753	0.9084	0.9723	0.9919

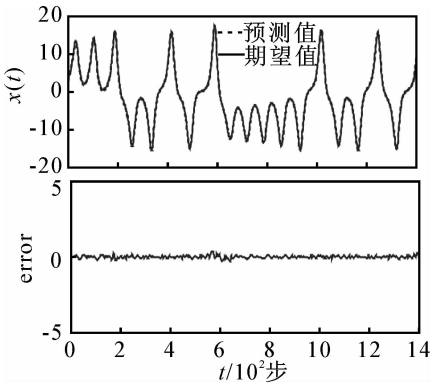

图 2　Lorenz 时间序列一步预测结果

分别采用单变量预测模型和神经网络系统辨识工具箱(NNSYSID)对 Lorenz 数据 $x(t)$ 序列进行预测仿真，并与本文方法进行性能比较. 其中单变量预测仍采用本文方法，仅将预测输入向量替换为 $\boldsymbol{x}_1$. 最终预测结果见表 2. 可以看出，由于 RPCA 能够有效提取多元变量综合信息，具有较好的预测性能. 单变量时间序列在预测时域为 $\eta=1$ 时预测性能较好，随着预测时域扩大，预测性能下降得较为严重. 这也说明单变量不能提供完整的数据信息.

表 2　3 种方法 $x(t)$ 预测性能比较

预测时域	本文方法		单变量预测		NNSYSID	
	E_{RMSE}	E_{PA}	E_{RMSE}	E_{PA}	E_{RMSE}	E_{PA}
$\eta=1$	0.0543	0.9999	0.0785	0.9998	0.0735	0.9999
$\eta=20$	0.7547	0.9955	0.8573	0.9901	0.8263	0.9922

3.2　黄河年径流时间序列预测建模

太阳黑子作为表示太阳活动强弱的一项重要指标，对全球气候、日地环境均具有重要作用，黄河径流也受到太阳黑子的影响. 因此在对黄河年径流的预测中，除考虑年径流本身的历史状态和演变特性，还应考虑太阳黑子变化的作用. 本文选用河南省三门峡水文站 1700 ～ 1997 年观测到的年太阳黑子数和黄河年径流构成二变量时间序列，如图 3 所示.

仍采用前文方法选择 $\tau_1=\tau_2=1$，$m_1=m_2=6$，即初始输入样本向量为

$$\boldsymbol{x}(t)=\{x(t),\cdots,x(t-6),y(t),\cdots,y(t-6)\}=\{\boldsymbol{x}_1^{\mathrm{T}},\boldsymbol{x}_2^{\mathrm{T}}\}^{\mathrm{T}}.$$

其中：$x(t)$ 为黄河年径流时间序列，$y(t)$ 为太阳黑子时间序列，预测输出为 $t+1$ 年的黄河年径流量 $x(t+\eta)$. 当取 $\eta=1$ 时，可形成 292 组样本，取前 272 个样本为训练样本，对后 20 年径流量进行预测. 用 RPCA 对输入进行处理，仍取 $\eta_0=0.999$，储备池维数为 200，经 RPCA 处理后降为 11 维. 取前 5 个主元成分对应特征值、方差贡献率及累积方差贡献率，如表 3 所示.

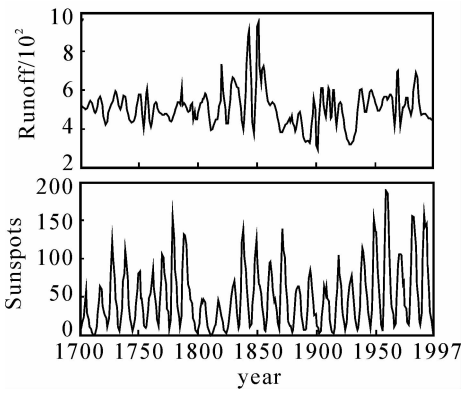

图 3　黄河年径流和年太阳黑子二变量时间序列

表 3　前 5 位主元特征值、方差贡献率及累积方差贡献率

序号 i	1	2	3	4	5
特征值	98.187	54.911	28.022	8.7075	3.316
贡献率	0.4909	0.2802	0.1401	0.0435	0.0166
累积贡献率	0.4909	0.7711	0.9112	0.9547	0.9713

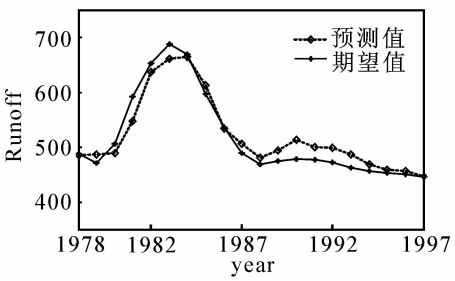

图 4　黄河年径流预测值和观测值比较

将 RPCA 输出的主元矩阵用于 Tikhonv 正则化建模，得到 1978 ～ 1997 年黄河年径流预测曲线如图 4 所示. 与单变量预测模型及 NNSYSID 方法比较，具体结果如表 4 所示.

表 4　3 种方法黄河径流预测性能比较

预测时域	本文方法		单变量预测		NNSYSID	
	E_{RMSE}	E_{PA}	E_{RMSE}	E_{PA}	E_{RMSE}	E_{PA}
$\eta=1$	20.414	0.9741	22.196	0.9297	31.821	0.9142
$\eta=2$	45.250	0.8645	51.332	0.8229	67.472	75.320

从表 4 可以看出，基于 RPCA 的多变量预测模型能够综合提取各个变量的有效信息，与其他两种方法相比，可实现对黄河年径流较高精度的预测.

4 结 论

结合 RPCA 方法，提出一种新的多元时间序列预测模型. 其中，RPCA 是一种建立在 ESN 储备池和 KPCA 理论基础之上，适用于动态系统的新型非线性主成分分析方法. 它将相点映射到网络储备池空间上进行特征提取. 对于复杂多变量时间序列，利用 RPCA 对多元嵌入延迟向量进行再处理可有效解决输入间的相关性问题. 同时，由于 RPCA 的输出与预测向量间呈动态线性映射关系，降低了建模复杂度，可应用 Tikhonv 正则化方法建立预测模型. 将预测模型分别用于 Lorenz 人工数据 $x(t)$时间序列和太阳黑子-黄河径流实测数据中，并与单变量预测模型和 NNSYSID 进行预测性能比较. 仿真结果表明，该方法对多元时间序列具有较好的预测性能，进而也反映出 RPCA 具有适应时变特性的能力，能够有效地提取出复杂变量的动态信息.

参考文献(References)

[1] Karunasinghea D S K, Liong S Y. Chaotic time series prediction with a global model: Artificial neural network [J]. J of Hydrology, 2006, 323(1-4): 92-105.

[2] Elman J L. Finding structure in time[J]. Cognitive Science, 1990, 14(2): 179-211.

[3] Takens F. Daynamical systems and turbulence, lecture notes in mathematics[M]. Berlin: Sringer, 1981.

[4] Peng D Z, Zhang Y. Dynamics of generalized PCA and MCA learning algorithms[J]. IEEE Trans on Neural Networks, 2007, 18(6): 1777-1784.

[5] Hartmann H, Becker S, King Lorenz. Predicting summer rainfall in the Yangtze river basin with neural networks[J]. Int J of Climatology, 2008, 28(7): 925-936.

[6] Mika S, Schölkopf B, Smola A. Kernel PCA and de-noising in feature space [C]. Advances in Neural Information Processing Systems. Cambridge: MTI Press, 1999: 524-536.

[7] Alzate C, Suykens J A K. Kernel component analysis using an epsilon-insensitive robust loss function [J]. IEEE Trans on Neural Networks, 2008, 19(9): 1583-1598.

[8] 赵峰，张军英. 一种 KPCA 的快速算法[J]. 控制与决策，2007，22(9)：1044-1049.
(Zhao F, Zhang J Y. Fast algorithm about KPCA[J]. Control and Decision, 2007, 22(9): 1044-1049.)

[9] Jaeger H, Haas H. Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication[J]. Science, 2004, 304(5667): 78-80.

[10] Shi Z W, Han M. Support vector echo-state machine for chaotic time-series prediction[J]. IEEE Trans on Neural Networks, 2007, 18(2): 359-372.

[11] Kim H S, Eykholt R, Salas J D. Nonlinear dynamics, delay times, and embedding windows[J]. Physica D, 1999, 127(1/2): 48-60.

[12] 成世学，严颖，张诒兰. 概率统计[M]. 北京：中国人民大学出版社，1994.
(Cheng S X, Yan Y, Zhang Y L. Probability statistics [M]. Beijing: China Renmin University Press, 1994.)

[13] Atiya A F, El-Shoura S M, Shaheen S I, et al. Comparison between neural-network forecasting techniques-case study: River flow forecasting [J]. IEEE Trans on Neural Networks, 1999, 10(2): 402-409.

[14] Chen J L, Islam S, Biswas P. Nonlinear dynamics of hourly ozone concentrations: Nonparameteric short term prediction[J]. Atmospheric Environment, 1998, 32(11): 1839-1848.

(上接第 1525 页)

[10] Gong D W, Guo G S. Interactive genetic algorithms with interval fitness of evolutionary individuals [J]. Dynamics of Continuous, Discrete and Impulsive Systems, Series B, 2007, 14(S2): 446-450.

[11] Gong D W, Guo G S, Lu L, et al. Adaptive interactive genetic algorithms with individual interval fitness[J]. Progress in Natural Science, 2008, 18(3): 359-365.

[12] Hornik K J. On the approximate realization of continuous mapping by neural network [J]. Neural Networks, 1989, 2(3): 183-192.

[13] 阎平凡. 对多层前向神经网络研究的几点看法[J]. 自动化学报，1997，23(1)：129-139.
(Yan P F. Some views on the research of multilayer feedforward neural network [J]. Acta Automatica Sinica, 1997, 23(1): 129-139.)