

基于改进教学优化算法的 Hermite 正交基神经网络混沌时间序列预测

李瑞国 张宏立 范文慧 王雅

Hermite orthogonal basis neural network based on improved teaching-learning-based optimization algorithm for chaotic time series prediction

Li Rui-Guo Zhang Hong-Li Fan Wen-Hui Wang Ya

引用信息 Citation: [Acta Physica Sinica](#), 64, 200506 (2015) DOI: 10.7498/aps.64.200506

在线阅读 View online: <http://dx.doi.org/10.7498/aps.64.200506>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn/CN/Y2015/V64/I20>

您可能感兴趣的其他文章

Articles you may be interested in

基于时频差的正交容积卡尔曼滤波跟踪算法

[A tracking algorithm based on orthogonal cubature Kalman filter with TDOA and FDOA](#)

物理学报.2015, 64(15): 150502 <http://dx.doi.org/10.7498/aps.64.150502>

基于混沌理论和改进径向基函数神经网络的网络舆情预测方法

[Internet public opinion chaotic prediction based on chaos theory and the improved radial basis function in neural networks](#)

物理学报.2015, 64(11): 110503 <http://dx.doi.org/10.7498/aps.64.110503>

多元混沌时间序列的多核极端学习机建模预测

[Multivariate chaotic time series prediction using multiple kernel extreme learning machine](#)

物理学报.2015, 64(7): 070504 <http://dx.doi.org/10.7498/aps.64.070504>

短期风速时间序列混沌特性分析及预测

[Chaotic characteristics analysis and prediction for short-term wind speed time series](#)

物理学报.2015, 64(3): 030506 <http://dx.doi.org/10.7498/aps.64.030506>

交通流突变点的无标度特征分析

[Analysis of scale-free characteristic on sharp variation point of traffic flow](#)

物理学报.2014, 63(24): 240509 <http://dx.doi.org/10.7498/aps.63.240509>

基于改进教学优化算法的 Hermite 正交基神经网络混沌时间序列预测*

李瑞国^{1)†} 张宏立¹⁾ 范文慧²⁾ 王雅³⁾

1)(新疆大学电气工程学院, 乌鲁木齐 830047)

2)(清华大学自动化系, 北京 100084)

3)(新疆大学机械工程学院, 乌鲁木齐 830047)

(2015年5月8日收到; 2015年6月3日收到修改稿)

针对传统预测模型对混沌时间序列预测精度低、收敛速度慢及模型结构复杂的问题, 提出了基于改进教学优化算法的 Hermite 正交基神经网络预测模型. 首先, 将自相关法和 Cao 方法相结合对混沌时间序列进行相空间重构, 以获得重构延迟时间向量; 其次, 以 Hermite 正交基函数为激励函数构成 Hermite 正交基神经网络, 作为预测模型; 最后, 将模型参数优化问题转化为多维空间上的函数优化问题, 利用改进教学优化算法对预测模型进行参数优化, 以建立预测模型并进行预测分析. 分别以 Lorenz 系统和 Liu 系统为模型, 通过四阶 Runge-Kutta 法产生混沌时间序列作为仿真对象, 并进行单步及多步预测对比实验. 仿真结果表明, 与径向基函数神经网络、回声状态网络、最小二乘支持向量机及基于教学优化算法的 Hermite 正交基神经网络预测模型相比, 所提预测模型具有更高的预测精度、更快的收敛速度和更简单的模型结构, 验证了该模型的高效性, 便于推广和应用.

关键词: Hermite 正交基神经网络, 改进教学优化算法, 混沌时间序列, 预测

PACS: 05.45.Tp

DOI: 10.7498/aps.64.200506

1 引言

混沌现象是由非线性确定系统产生的一种不规则运动, 广泛存在于气象、水文、通信及经济等众多领域^[1]. 混沌具有内随机性、整体稳定局部不稳定、短期可预测而长期不可预测性等特征^[2]. 近年来, 随着混沌理论研究的不断深入及其在信号处理、自动控制、电力及金融短期预测等领域中的广泛应用^[3], 混沌时间序列的建模和预测已成为混沌领域的一个非常重要的研究方向^[4].

几乎所有的混沌时间序列预测模型都是建立在相空间重构的基础之上^[5], 其关键是最佳延迟时间 τ 和最小嵌入维数 m 的确定^[6]. 最佳延迟时间 τ 的确定方法主要有: 自相关法、平均位移法及互信

息法等. 最小嵌入维数 m 的确定方法主要有: 饱和关联维数法、伪最近邻域法、真实矢量场法及 Cao 方法等, 以及同时确定这两个参数的方法——C-C 法. 本文采用自相关法和 Cao 方法相结合进行相空间重构, 过程简单、程序易于实现.

除对混沌时间序列进行相空间重构外, 其预测模型的选定也是一关键环节. 传统的混沌时间序列预测模型主要包括: 自回归滑动平均 (auto regressive moving average, ARMA) 模型、差分自回归滑动平均 (auto regressive integrated moving average, ARIMA) 模型等线性模型^[7], 以及反馈传递神经网络、径向基函数 (RBF) 神经网络、小波神经网络、Elman 神经网络、回声状态网络 (echo state network, ESN) 等神经网络模型^[8], Bernstein 多项式模型、Volterra 自适应滤波器模型、支持向

* 国家自然科学基金 (批准号: 61463047) 和自治区研究生科研创新项目 (批准号: XJGR12015039) 资助的课题.

† 通信作者. E-mail: 1141567852@qq.com

量机 (support vector machine, SVM) 模型、加权极端学习机 (weighted extreme learning machine, WELM) 模型及最小二乘支持向量机 (least square support vector machine, LSSVM) 模型等非线性模型^[9]. 线性模型结构简单, 但预测精度低. 非线性模型中, 传统的神经网络在网络训练时样本及网络结构都需要选择^[10], 且收敛速度慢、易陷入局部最优^[11]. Elman 神经网络较传统神经网络增加了关联层, 使得结构更加复杂、建模时间增长^[12]; ESN 网络结构复杂、所需训练样本容量大, 建模时可能会由于某些矩阵规模过大而呈现病态特性, 从而导致输出权值幅值过大, 并降低网络的泛化性能和预测性能^[13]; Bernstein 多项式模型为提高运算速度, 删除了一些交叉项, 降低了对连续函数的逼近能力; Volterra 自适应滤波器模型会随着级数增大记忆单元增多, 滤波系数和计算幂次增加过快^[14], 导致参数过多、效率低下; SVM 模型及 LSSVM 模型均采用结构风险最小化准则, 不存在局部极值、隐含层节点数的选择问题, 但结构复杂、过程烦琐^[15]; WELM 模型不仅待优化参数多、结构复杂, 同时也存在样本及隐含层激励函数选择的问题, 使得模型参数训练复杂度增大、运行效率和实时性降低^[16].

本文结合相空间重构理论、Hermite 正交基函数性质、最佳平方逼近理论及神经网络理论^[17], 采用 Hermite 正交基函数作为神经网络隐含层的激励函数, 设计了多输入 Hermite 正交基神经网络预测模型, 并利用改进教学优化 (improved teaching-learning-based optimization, ITLBO) 算法对预测模型进行参数优化. 最后, 通过与其他预测模型的对比实验, 验证了所提预测模型的有效性和实用性.

2 混沌时间序列的相空间重构

在现实中, 对于一个复杂的高维混沌系统, 能采集到的通常是一组关于时间的一维数据——混沌时间序列, 但这样的一维标量信息无法透彻地反映系统在相空间中各维的变化情况. 为把一个形式上为一维的混沌时间序列变换到高维空间中以再现系统所蕴藏的演化信息, 相空间重构理论被提出. 相空间重构理论说明, 由一维混沌时间序列可重构一个与其原动力学拓扑等价的相空间, 以分析该序列的性质和规律. 因此, 相空间重构是混沌时间序列预测的基础. 本文利用自相关法和 Cao 方法

相结合分别确定最佳延迟时间 τ 和最小嵌入维数 m .

2.1 最佳延迟时间 τ 的确定

利用自相关法确定最佳延迟时间 τ . 对于混沌时间序列 $\mathbf{Y} = [y_1, y_2, \dots, y_n]^T$, 通过计算自相关函数 $Cl(\tau)$ 来判断 y_i 和 $y_{i+\tau}$ 的相关性, 进而确定最佳延迟时间 τ . 定义自相关函数为

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i,$$

$$Cl(\tau) = \frac{\frac{1}{n-\tau} \sum_{i=1}^{n-\tau} (y_{i+\tau} - \bar{y})(y_i - \bar{y})}{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2}. \quad (1)$$

由 (1) 式可知, 延迟时间过小, 重构吸引子相邻点的相关性太强, 吸引子的分析极容易被噪声干扰; 延迟时间过大, 原本较近的向量亦会被拉远, 导致不确定的系统状态. 研究表明, 选 $Cl(\tau)$ 下降到初始值的 $(1 - 1/e)$ 对应的 τ 为最佳延迟时间.

2.2 最小嵌入维数 m 的确定

利用 Cao 方法确定最小嵌入维数 m . 相空间重构后,

$$\mathbf{Z}_{m,i} = [y_i, y_{i+\tau}, \dots, y_{i+(m-1)\tau}]^T$$

$$(i = 1, 2, \dots, n - (m-1)\tau),$$

$$\mathbf{Z}_m = [\mathbf{Z}_{m,1}, \mathbf{Z}_{m,2}, \dots, \mathbf{Z}_{m,l}], \quad (2)$$

其中, $\mathbf{Z}_{m,i} (i = 1, 2, \dots, n - (m-1)\tau)$ 为重构延迟时间向量; \mathbf{Z}_m 为重构相空间. 为求取最小嵌入维数 m , 给出如下两个定义式.

定义

$$e_{m,i} = \frac{\min_k \{\|\mathbf{Z}_{m+1,i} - \mathbf{Z}_{m+1,k}\|_2\}}{\min_j \{\|\mathbf{Z}_{m,i} - \mathbf{Z}_{m,j}\|_2\}}$$

$$(i, j, k = 1, 2, \dots, n - m\tau; j, k \neq i), \quad (3)$$

其中, $\|\cdot\|_2$ 为二范数符号, $e_{m,i} (i = 1, 2, \dots, n - m\tau)$ 表示 $m + 1$ 维和 m 维重构相空间中的第 i 个重构延迟时间向量与其距离最近的重构延迟时间向量之间的二范数的比值.

定义

$$E_m = \frac{1}{n - m\tau} \sum_{i=1}^{n-m\tau} e_{m,i},$$

$$EE_m = |E_{m+1} - E_m|, \quad (4)$$

其中, $|\cdot|$ 为绝对值符号, EE_m 表示 E_{m+1} 与 E_m 之间变化的差值绝对值的大小.

如果序列 \mathbf{Y} 是由混沌吸引子产生的, 则对于任意给定的很小正数 ε , 当大于或等于某一 m 后, 总存在 $EE_m < \varepsilon$, 那么取此时的 m 为最小嵌入维数最为合宜, 否则会影响相空间重构的质量: 嵌入维数过低, 会出现吸引子的自交性; 嵌入维数过高, 会使点与点之间的距离太大.

3 Hermite 正交基神经网络预测模型

对一维的混沌时间序列进行相空间重构后, 将其映射到 m 维的相空间中, 对于预测分析只是第一步工作, 更关键的是预测模型的选定, 这会影响预测精度、建模时间等. 针对传统预测模型的不足, 本文根据相空间重构理论、Hermite 正交基函数的性质及最佳平方逼近理论, 结合神经网络拓扑, 提出了 Hermite 正交基神经网络预测模型.

3.1 理论基础

3.1.1 Hermite 正交基函数及其性质

在区间 $(-\infty, +\infty)$ 上权值为 e^{-x^2} 的 Hermite 正交基函数, 其表达式为

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n(e^{-x^2})}{dx^n} \quad (n = 0, 1, \dots), \quad (5)$$

其满足递推关系 (6) 式,

$$\begin{cases} H_0(x) = 1, \\ H_1(x) = 2x, \\ H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x) \end{cases} \quad (n = 1, 2, \dots), \quad (6)$$

其满足正交性 (7) 式,

$$\begin{aligned} & \int_{-\infty}^{+\infty} e^{-x^2} H_m(x) H_n(x) dx \\ &= \begin{cases} 0, & m \neq n, \\ 2^n n! \sqrt{\pi}, & m = n. \end{cases} \end{aligned} \quad (7)$$

3.1.2 最佳平方逼近理论

定理 对于 $f(x) \in C[a, b]$, 存在 $C[a, b]$ 上的一个正交函数簇 $\varphi(x) = \text{span}\{\varphi_0(x), \varphi_1(x), \dots, \varphi_p(x)\}$, 使得

$$\|S^*(x) - f(x)\|_2^2$$

$$\begin{aligned} &= \min_{S(x) \in \varphi(x)} \{\|S(x) - f(x)\|_2^2\} \\ &= \min_{S(x) \in \varphi(x)} \int_a^b \rho(x) [S(x) - f(x)]^2 dx, \end{aligned} \quad (8)$$

其中, $S^*(x)$ 和 $S(x)$ 分别为 $f(x)$ 的最佳逼近函数和逼近函数, $\rho(x)$ 为权值系数.

若 $f(x)$ 只存在一组离散点集 $\{(x_i, f_i) | f_i = f(x_i); i = 1, 2, \dots, l\}$, 则可根据已知的数据点对原曲线进行最佳拟合, 使得样本点上的误差平方和最小, 即

$$\begin{aligned} S(x) &= \sum_{j=1}^q a_j \varphi_{j-1}(x), \\ \delta &= \sum_{i=1}^l [S^*(x_i) - f_i]^2 \\ &= \min_{S(x) \in \varphi(x)} \left\{ \sum_{i=1}^l [S(x_i) - f_i]^2 \right\}, \end{aligned} \quad (9)$$

其中, $q = p + 1$ 为 $\varphi(x)$ 的维数, $a_j (j = 1, 2, \dots, q)$ 为权系数.

以上便是对曲线最佳平方逼近理论的描述. 结合上述理论基础, 本文提出了 Hermite 正交基神经网络预测模型以解决混沌系统复杂的动力学特性分析问题.

3.2 Hermite 正交基神经网络

理论上, 一个包含输入层-隐含层-输出层的神经网络可以逼近任意的非线性函数^[18], 但传统的神经网络结构复杂、执行效率低. Hermite 正交基函数递推关系简单, 且相对于 Chebyshev 正交基函数、Laguerre 正交基函数及 Legendre 正交基函数等对自变量无范围限定, 数据处理简单、计算量小. 因此, 根据 3.1 节理论基础, 并结合神经网络拓扑结构, 本文将 Hermite 正交基函数作为网络隐含层激励函数, 提出了 Hermite 正交基神经网络预测模型.

混沌系统动力学特性的研究几乎都是建立在相空间重构的基础之上. 对于混沌时间序列 $\mathbf{Y} = [y_1, y_2, \dots, y_n]^T$, Takens 定理证明了存在光滑曲线 $f(x)$, 使得

$$\begin{aligned} \tilde{y}_{i+(m-1)\tau+s} &= f(\mathbf{Z}_{m,i}) \\ (i &= 1, 2, \dots, n - (m-1)\tau - s), \end{aligned} \quad (10)$$

其中, s 为预测步数, m 为最小嵌入维数, τ 为最佳延迟时间, $\mathbf{Z}_{m,i} (i = 1, 2, \dots, n - (m-1)\tau - s)$ 为重构延迟时间向量, 如 (2) 式.

根据(10)式, 构建一 $m - q - 1$ 结构的 Hermite 正交基神经网络, 且各层神经元均采用恒等映射, 并使网络中所有神经元的阈值都为0, 网络如图1所示.

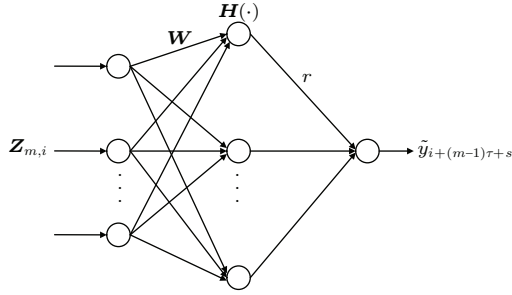


图1 Hermite 正交基神经网络

Fig. 1. Hermite orthogonal basis neural network.

图1中, $\mathbf{W} = [\mathbf{W}_1, \mathbf{W}_2, \dots, \mathbf{W}_q]$ 为输入层到隐含层的权值矩阵, 且 $\mathbf{W}_i = [w_1, w_2, \dots, w_m]^T$ ($i = 1, 2, \dots, q$) 为输入层到隐含层第 i 个神经元的权值向量; $\mathbf{H}(\cdot) = [H_0(\cdot), H_1(\cdot), \dots, H_P(\cdot)]^T$ 为隐含层激励函数向量, 且 $p = q - 1$; $\mathbf{r} = [r_1, r_2, \dots, r_q]^T$ 为隐含层到输出层的权值向量.

由于网络中各层神经元均采用恒等映射, 且所有神经元的阈值都为0, 所以, 隐含层第 i 个神经元输入为

$$net_i = \mathbf{Z}_{m,i}^T \mathbf{W}_i; \quad (11)$$

隐含层第 i 个神经元输出为

$$o_i = H_{i-1}(net_i); \quad (12)$$

输出层神经元输出为

$$\tilde{y}_{i+(m-1)\tau+s} = [o_1, o_2, \dots, o_q] \mathbf{r}. \quad (13)$$

网络训练时, 采用序列预测值与观测值的误差平方和 $E_p = \sum_{i=1}^n (\tilde{y}_i - y_i)^2$ 作为目标函数, 即 E_p 最小或达到预定指标时的权值向量 \mathbf{W}_i ($i = 1, 2, \dots, q$) 和 \mathbf{r} 为网络的参数. 其中, \tilde{y}_i ($i = 1, 2, \dots, n$) 为预测模型的预测值, y_i ($i = 1, 2, \dots, n$) 为实际的观测值. 传统的神经网络常利用梯度下降法对网络进行权值的训练, 而且存在网络隐含层节点数的选择及训练样本的选取等问题, 不仅算法过程烦琐、程序复杂、效率低, 同时其精度亦不是很高, 而且网络结构复杂、权值参数过多, 导致收敛速度慢且极易出现早熟等现象. 实验结果表明, Hermite 正交基神经网络隐含层一般选取3个节点便可达到精度及效率上的要求, 若节点增

加, 精度变化不大甚至会出现过拟合现象, 而且由于参数的增多导致网络结构更复杂, 同时延缓建模时间, 降低了时效性. 因此, 本文的 Hermite 正交基神经网络预测模型隐含层节点数均为 $q = 3$. 此外, 所提预测模型对训练样本亦无严格要求, 而且样本容量通常少于传统神经网络样本的数量.

可见, Hermite 正交基神经网络不存在隐含层节点难以确定及样本选择等问题, 预测模型结构简单、待优化参数少, 只需找到维数较低的模型待优化参数向量 $\theta = [\mathbf{W}_i^T, \mathbf{r}^T]$. ITLBO 算法作为一种并行优化算法, 全局搜索能力强, 操作简单、收敛速度快, 因此被用来训练模型待优化参数向量 θ . 下面来证明利用 ITLBO 算法训练模型待优化参数向量 θ 的高效性.

证明 令 $n - (m - 1)\tau - s = n_0$, 则预测值与观测值的误差平方和 $E_p = \sum_{i=1}^{n_0} |\mathbf{H}^T(\mathbf{Z}_{m,i}^T \mathbf{W}_c) \mathbf{r} - y_{i+n-n_0}|^2$, 其中, $\mathbf{H}(\cdot) = [H_0(\cdot), H_1(\cdot), \dots, H_P(\cdot)]^T$ 为隐含层激励函数向量, 且 $p = q - 1$; $\mathbf{Z}_{m,i}$ ($i = 1, 2, \dots, n_0$) 为重构延迟时间向量, 如(2)式; $\mathbf{W}_c = \mathbf{W}_i = [w_1, w_2, \dots, w_m]^T$ ($i = 1, 2, \dots, q$) 为输入层到隐含层的权值向量; $\mathbf{r} = [r_1, r_2, \dots, r_q]^T$ 为隐含层到输出层的权值向量. 因此, 模型待优化参数向量 $\theta = [\mathbf{W}_c^T, \mathbf{r}^T]$, 且 E_p 两端分别对 w_j ($j = 1, 2, \dots, m$) 和 r_k ($k = 1, 2, \dots, q$) 求偏导, 有

$$\frac{\partial E_p}{\partial w_j} = 2 \left\{ \sum_{i=1}^{n_0} |\mathbf{H}^T(\mathbf{Z}_{m,i}^T \mathbf{W}_c) \mathbf{r} - y_{i+n-n_0}| \times \left\{ \frac{\partial \mathbf{H}^T(\mathbf{Z}_{m,i}^T \mathbf{W}_c)}{\partial \mathbf{W}_c} \times \frac{\partial \mathbf{W}_c}{\partial w_j} \times \mathbf{r} \right\} \right\}, \quad (14)$$

$$\frac{\partial E_p}{\partial r_k} = 2 \left\{ \sum_{i=1}^{n_0} |\mathbf{H}^T(\mathbf{Z}_{m,i}^T \mathbf{W}_c) \mathbf{r} - y_{i+n-n_0}| \times \left\{ \mathbf{H}^T(\mathbf{Z}_{m,i}^T \mathbf{W}_c) \times \frac{\partial \mathbf{r}}{\partial r_k} \right\} \right\}. \quad (15)$$

(14)和(15)式均是 θ 的多元函数, 结合以上两式并根据多元函数极值的必要条件有 $\sum_{i=1}^{n_0} |\mathbf{H}^T(\mathbf{Z}_{m,i}^T \mathbf{W}_c) \mathbf{r} - y_{i+n-n_0}| = 0$, 因此, $\mathbf{H}^T(\mathbf{Z}_{m,i}^T \mathbf{W}_c) \mathbf{r} - y_{i+n-n_0} = 0$ ($i = 1, 2, \dots, n_0$), 即

$$[\mathbf{H}(\mathbf{Z}_{m,1}^T \mathbf{W}_c), \mathbf{H}(\mathbf{Z}_{m,2}^T \mathbf{W}_c), \dots, \mathbf{H}(\mathbf{Z}_{m,n_0}^T \mathbf{W}_c)]^T \times \mathbf{r} = [y_{n-n_0+1}, y_{n-n_0+2}, \dots, y_n]^T. \quad (16)$$

令

$$[\mathbf{H}(\mathbf{Z}_{m,1}^T \mathbf{W}_c), \mathbf{H}(\mathbf{Z}_{m,2}^T \mathbf{W}_c), \dots, \mathbf{H}(\mathbf{Z}_{m,n_0}^T \mathbf{W}_c)]^T = \mathbf{A},$$

$$[y_{n-n_0+1}, y_{n-n_0+2}, \dots, y_n]^T = \mathbf{Y}_0,$$

则 $\mathbf{A} \in n_0 \times q (n_0 \gg q)$ 的矩阵, 为求解 (16) 式方程, 将其两端同乘以 \mathbf{A}^T , 有

$$\mathbf{A}^T \mathbf{A} \mathbf{r} = \mathbf{A}^T \mathbf{Y}_0. \quad (17)$$

因为 $\mathbf{A}^T \mathbf{A} \mathbf{x} = 0$ 与 $\mathbf{A} \mathbf{x} = 0$ 有相同的解, 因此, $\text{rank}(\mathbf{A}^T \mathbf{A}) = \text{rank}(\mathbf{A}) = q$, 即 $\mathbf{A}^T \mathbf{A}$ 和 \mathbf{A} 均满秩. 由克莱姆法则知, (17) 式存在一组惟一的解 $\boldsymbol{\theta} = [\mathbf{W}_c^T, \mathbf{r}^T]$; 根据问题的实际性知, 此时 E_p 取得极小值亦即最小值. 因此, 可将预测模型参数优化问题转化为函数优化问题. 通过 ITLBO 算法极强的寻优能力, 可以跳出局部最优以快速收敛到真值, 从而找到 $\boldsymbol{\theta}$, 即当教学次数 $g \geq g_0$ 后,

$$\boldsymbol{\theta} = \lim_{g \geq g_0} \arg \{ \min_{\mathbf{X}} [fitness(\mathbf{X})] \}, \quad (18)$$

其中, $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N]^T$, 且 $\mathbf{X}_i = [x_{i_1}, x_{i_2}, \dots, x_{i_D}]^T (i = 1, 2, \dots, N)$, 被比作算法中的学生个体; $fitness(\mathbf{X}_i) = \{E_p | \boldsymbol{\theta} = \mathbf{X}_i\} (i = 1, 2, \dots, N)$ 为目标函数, 作为一种评优准则.

可以看出, 利用 ITLBO 算法优化预测模型参数, 既可避免传统梯度下降法训练神经网络的烦琐性和低效性, 又能利用 ITLBO 算法的优点, 防止出现早熟及陷入局部最优等缺点, 理论上充分证明了基于 ITLBO 算法的 Hermite 正交基神经网络的高效性.

4 基于 ITLBO 算法的预测模型参数优化

4.1 预测模型参数优化原理

利用 Hermite 正交基神经网络做预测时, 关键是模型参数的优化, 其原理如图 2 所示.

图 2 中, $\mathbf{Y} = [y_1, y_2, \dots, y_n]^T$ 为原始混沌时间序列, $\tilde{\mathbf{Y}} = [\tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_n]^T$ 为模型预测序列. 从图 2 可以看出, 在待优化参数向量 $\boldsymbol{\theta} = [\mathbf{W}_c^T, \mathbf{r}^T]$ 的作用下, 系统稳态输出为

$$e_{ss} = \min_{\boldsymbol{\theta}} \{ (\tilde{\mathbf{Y}} - \mathbf{Y})^T (\tilde{\mathbf{Y}} - \mathbf{Y}) \}, \quad (19)$$

即通过不断调整 $\boldsymbol{\theta}$, 使系统期望输出与真实输出的误差平方和最小. 因此, 目标函数取为

$$F(\boldsymbol{\theta}) = (\tilde{\mathbf{Y}} - \mathbf{Y})^T (\tilde{\mathbf{Y}} - \mathbf{Y}). \quad (20)$$

利用 (20) 式, 将参数优化问题转化为多维空间上的函数优化问题, 即寻找 $F(\boldsymbol{\theta})$ 的最小值点,

$$\boldsymbol{\theta}_b = \arg \{ \min_{\boldsymbol{\theta}} \{ F(\boldsymbol{\theta}) \} \}. \quad (21)$$

由于目标函数为多变量函数, 因此, 可能存在多个局部最优解. 传统的梯度下降算法、遗传算法 [19] (genetic algorithm, GA)、粒子群优化 (particle swarm optimization, PSO) 算法 [20] 及量子粒子群优化 (quantum particle swarm optimization, QPSO) 算法等 [21] 在训练神经网络时极易陷入局部最优, 且操作复杂、收敛速度慢. 而 ITLBO 算法作为教学优化 (teaching-learning-based optimization, TLBO) 算法 [22] 的改进算法, 则很好地弥补了传统算法的不足, 是一种具有高效搜索性的智能优化算法. 因此, 本文利用 ITLBO 算法对预测模型进行参数优化.

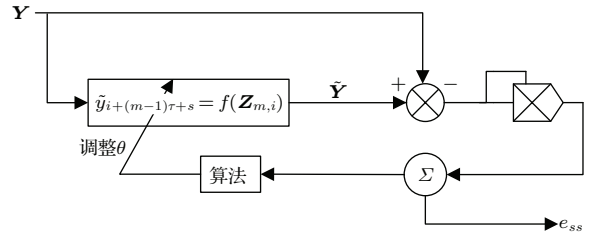


图 2 参数优化原理图

Fig. 2. Principle of parameter optimization.

4.2 ITLBO 算法

TLBO 算法是一种基于班级教学过程的优化算法, 包括教授阶段和学习阶段两个过程, 但其收敛慢、精度不高. 在 TLBO 算法的基础上, 本文提出了 ITLBO 算法, 通过在学习阶段之后加入反馈阶段, 增加较差学生与教师交流的机会, 便于保持学生学习的多样性, 加快较差学生向好学生靠拢的速度, 以提高算法的全局搜索能力及搜索速度.

4.2.1 教授阶段

设一个班级由教师 \mathbf{T} 和学生 \mathbf{S} 组成, 将目标函数 $f(\mathbf{S})$ 作为准则对每个学生进行评优. 其中,

$$\mathbf{S}_i = [s_{i_1}, s_{i_2}, \dots, s_{i_D}]^T \quad (i = 1, 2, \dots, N),$$

$$\mathbf{T} = \arg \left\{ \min_{i=1,2,\dots,N} f(\mathbf{S}_i) \right\},$$

$$\mathbf{S} = [\mathbf{S}_1, \mathbf{S}_2, \dots, \mathbf{S}_N]^T. \quad (22)$$

在教授过程中, 教师 \mathbf{T} 尽力使学生的平均水平 \mathbf{S}_M 接近自己的水平, 其中,

$$\mathbf{S}_M = \frac{1}{N} \sum_{i=1}^N \mathbf{S}_i$$

$$= \frac{1}{N} \left(\sum_{i=1}^N s_{i_1}, \sum_{i=1}^N s_{i_2}, \dots, \sum_{i=1}^N s_{i_D} \right), \quad (23)$$

即教师根据 \mathbf{T} 和 \mathbf{S}_M 的差异教授学生学习, 使学生进步更快. 定义教师水平和学生平均水平的差异为

$$\mathbf{D}_{M_i} = r_{i_1}(\mathbf{T} - T_{F_i} \mathbf{S}_M), \quad (24)$$

其中, r_{i_1} 为 $[0, 1]$ 上的随机数, 即学生的学习因子; T_{F_i} 为教师的教学因子, 且

$$\begin{aligned} T_{F_i} &= \text{round}(1 + r_{i_2}) \\ &= \begin{cases} 1, & 0 \leq r_{i_2} < 0.5, \\ 2, & 0.5 \leq r_{i_2} \leq 1. \end{cases} \end{aligned} \quad (25)$$

根据 (22)—(25) 式, 在教师教授后, 学生达到一个新的水平,

$$\mathbf{S}_{\text{new},i} = \mathbf{S}_i + \mathbf{D}_{M_i}. \quad (26)$$

再根据评优准则, 判断学生是否进步, 从而更新每个学生,

$$\mathbf{S}_i = \begin{cases} \mathbf{S}_{\text{new},i}, & f(\mathbf{S}_{\text{new},i}) < f(\mathbf{S}_i), \\ \mathbf{S}_i, & \text{其他.} \end{cases} \quad (27)$$

4.2.2 学习阶段

学习阶段为学生之间相互交流的过程, 学生能从比自己水平高的同学那里获取新的知识, 以提升自己的水平. 学生学习阶段可描述为: 在班级中, 随机选择两个学生 \mathbf{S}_i 和 \mathbf{S}_j , 且 $i \neq j$, 表明学生不和自己比较. 根据评优准则, 决定学生该向哪些同学看齐使其进步最快. 若学生 \mathbf{S}_i 优于 \mathbf{S}_j , 则 $\mathbf{S}_{\text{new},i}$ 向 \mathbf{S}_i 靠近; 反之, 向 \mathbf{S}_j 靠近. 学习阶段具体实现为

$$\begin{cases} \mathbf{S}_{\text{new},i} = \mathbf{S}_i + r_{i_3}(\mathbf{S}_i - \mathbf{S}_j), & f(\mathbf{S}_i) < f(\mathbf{S}_j), \\ \mathbf{S}_{\text{new},i} = \mathbf{S}_i + r_{i_3}(\mathbf{S}_j - \mathbf{S}_i), & f(\mathbf{S}_i) > f(\mathbf{S}_j), \end{cases} \quad (28)$$

其中, r_{i_3} 为 $[0, 1]$ 上的随机数, 即较自己水平高的同学对自己的影响因子. 向好学生学习之后, 再利用 (27) 式更新每个学生.

4.2.3 反馈阶段

经历过教授阶段、学习阶段之后, 由于在教师的教授和学生间的学习下, 学生水平有了一定的提高, 导致之前的教师水平不能满足学生的要求, 需要请来新的水平更高的教师. 利用 (22) 式对教师进行更新. 为了进一步提高学生的水平, 除了教师的教授和学生间的相互学习外, 学生还需与教师进行沟通交流, 具体实现为: 在班级中, 随机选择两个不

同的学生 \mathbf{S}_i 和 \mathbf{S}_j , 其中 $i \neq j$. 根据评优准则, 决定哪些学生应与教师进行反馈交流. 若学生 \mathbf{S}_i 优于 \mathbf{S}_j , 则选择差生 \mathbf{S}_j 与教师进行反馈学习; 反之, 选择 \mathbf{S}_i 与教师进行交流学习. 反馈阶段具体实现为

$$\begin{cases} \mathbf{S}_{\text{new},i} = \mathbf{S}_i + r_{i_4}(\mathbf{T} - \mathbf{S}_j), & f(\mathbf{S}_i) < f(\mathbf{S}_j), \\ \mathbf{S}_{\text{new},i} = \mathbf{S}_i + r_{i_4}(\mathbf{T} - \mathbf{S}_i), & f(\mathbf{S}_i) > f(\mathbf{S}_j), \end{cases} \quad (29)$$

其中, r_{i_4} 为 $[0, 1]$ 上的随机数, 即较差学生向教师反馈信息的能力. 反馈过程结束后, 利用 (27) 式更新每个学生. 反馈阶段不仅保留了班级多样性, 又使差生快速向好学生靠拢.

4.3 算法流程实现

在利用 ITLBO 算法训练预测模型时, 将待优化参数向量 $\boldsymbol{\theta} = [\mathbf{W}_c^T, \mathbf{r}^T]$ 比作算法中的班级成员, 在目标函数 (20) 式的评价下, 其流程如下.

步骤 1 初始化班级成员. 根据待优化参数向量的维数 $D = \dim(\boldsymbol{\theta})$, 随机生成一个由 N 个学生组成的班级 $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N]^T$, 且 $\mathbf{X}_i = [x_{i_1}, x_{i_2}, \dots, x_{i_D}]^T (i = 1, 2, \dots, N)$; 根据 (22) 式, 选出最优成员作为教师, 并教学 G 次. 同时, 通过目标函数计算每个学生的适应度, 如 (30) 式.

$$\begin{aligned} \tilde{y}_i &= y_i (i = 1, 2, \dots, n - n_0), \\ \tilde{y}_i &= \{H_0[\mathbf{Z}_{m,i-n+n_0}^T \mathbf{X}_j(1:m,:)], \\ &H_1[\mathbf{Z}_{m,i-n+n_0}^T \mathbf{X}_j(1:m,:)], \dots, \\ &H_p[\mathbf{Z}_{m,i-n+n_0}^T \mathbf{X}_j(1:m,:)]\} \mathbf{X}_j(m+1:D,:), \\ &(i = n - n_0 + 1, n - n_0 + 2, \dots, n; \\ &j = 1, 2, \dots, N), \\ F_j &= \sum_{i=1}^n (\tilde{y}_i - y_i)^2, \end{aligned} \quad (30)$$

其中, $\tilde{y}_i (i = 1, 2, \dots, n)$ 是对 $y_i (i = 1, 2, \dots, n)$ 的预测值; $\mathbf{Z}_{m,i-n+n_0}^T (i = n - n_0 + 1, n - n_0 + 2, \dots, n)$ 为重构延迟时间向量如 (2) 式, 且 $n_0 = n - (m - 1)\tau - s$.

步骤 2 教授阶段. 根据 (22)—(26) 式, 在教师的教授和指导下, 努力提高学生的个人水平, 并通过目标函数 (30) 和 (31) 式的新旧适应度来判断每个学生是否已进步, 从而根据 (27) 式更新班级成员.

$$\tilde{y}_{\text{new},i} = y_i (i = 1, 2, \dots, n - n_0),$$

$$\begin{aligned} \tilde{y}_{\text{new},i} = & \{[H_0[\mathbf{Z}_{m,i-n+n_0}^T \mathbf{X}_{\text{new},j}(1:m,:)], \\ & H_1[\mathbf{Z}_{m,i-n+n_0}^T \mathbf{X}_{\text{new},j}(1:m,:)], \dots, \\ & H_p[\mathbf{Z}_{m,i-n+n_0}^T \mathbf{X}_{\text{new},j}(1:m,:)]\} \\ & \times \mathbf{X}_{\text{new},j}(m+1:D,:), \\ & (i = n - n_0 + 1, n - n_0 + 2, \dots, n; \\ & j = 1, 2, \dots, N), \\ F_{\text{new},j} = & \sum_{i=1}^n (\tilde{y}_{\text{new},i} - y_i)^2, \end{aligned} \quad (31)$$

其中, $\mathbf{X}_{\text{new},j} = [x_{\text{new},j1}, x_{\text{new},j2}, \dots, x_{\text{new},jD}]^T (j = 1, 2, \dots, N)$ 为学生 $\mathbf{X}_j (j = 1, 2, \dots, N)$ 经历教师教授后的新值.

步骤3 学习阶段. 根据(28)式, 通过学生间的相互学习和监督, 不断完善自己, 并通过目标函数(30)和(31)式的新旧适应度来断定每个学生是否有所提高, 从而根据(27)式更新班级成员. 其中, $\mathbf{X}_{\text{new},j} (j = 1, 2, \dots, N)$ 为学生 $\mathbf{X}_j (j = 1, 2, \dots, N)$ 经历学习后的新值.

步骤4 反馈阶段. 经历教授阶段、学习阶段后, 根据(22)式, 推选出新的教师 \mathbf{T} ; 根据(29)式, 通过教师与差生的反馈交流和学习, 使差生快速向好学生转变, 并通过目标函数(30)和(31)式的新旧适应度来决断每个学生是否已向好的方向转化, 从而根据(27)式更新班级成员. 其中, $\mathbf{X}_{\text{new},j} (j = 1, 2, \dots, N)$ 为学生 $\mathbf{X}_j (j = 1, 2, \dots, N)$ 经历与教师交流后的新值.

步骤5 若当前教学次数 $g = G$, 则结束, 并根据(22)式选出当前教师作为待优化参数向量 θ 的解; 否则, 转至步骤2, 继续.

5 仿真实验及性能分析

为验证本文所提预测模型的高效性, 分别以连续多输入多输出 (multiple input and multiple output, MIMO) 的 Lorenz 系统和 Liu 系统为模型进行仿真实验. 为定量比较所提预测模型和对比预测模型的差异, 引入4个性能评价指标: 平均建模时间 (mean modeling time, MMT)、平均绝对值误差 (mean absolute error, MAE)、决定系数 (decision coefficient, DC) 及均方根误差 (root mean square error, RMSE). 其中,

$$\text{MMT} = \frac{1}{N_a} T_f, \quad (32)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |\tilde{y}_i - y_i|, \quad (33)$$

$$\text{DC} = \frac{\sum_{i=1}^n (\tilde{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (34)$$

$$\text{RMSE} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (\tilde{y}_i - y_i)^2}, \quad (35)$$

其中, N_a 为预测模型总的建模次数, T_f 为预测模型总的建模时间, $\tilde{y}_i (i = 1, 2, \dots, n)$ 是对 $y_i (i = 1, 2, \dots, n)$ 的预测值, $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ 为观测值的均值.

5.1 Lorenz 系统

1961年, 世界著名的气象物理学家 Lorenz 在研究对流不稳定时, 采用高截谱方法由 Saltzman 给出的描述对流的控制方程得到了被命名为 Lorenz 系统的3维非线性常微分方程, 如(36)式, 并首创性地发现了气象系统存在奇特的“蝴蝶效应”, 揭示了任何具有非周期行为的物理系统均将存在不可预期的现象^[23].

$$\begin{aligned} \dot{x} &= \sigma(y - x), \\ \dot{y} &= \gamma x - xz - y, \\ \dot{z} &= xy - bz, \end{aligned} \quad (36)$$

当 $\sigma = 10, \gamma = 28, b = 8/3$ 时, 系统呈现混沌现象, 其吸引子如图3所示.

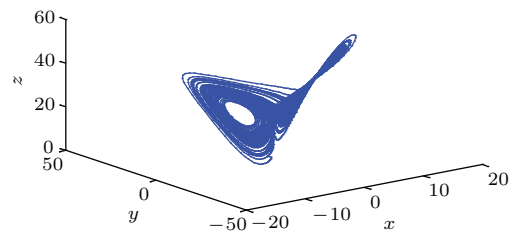


图3 Lorenz 吸引子

Fig. 3. Lorenz attractor.

仿真过程中, 初值 x_0, y_0 及 z_0 分别取为 $-1, 0$ 和 1 , 步长 h 取为 0.01 . 通过四阶 Runge-Kutta 法计算包含 1200 个数据点的混沌时间序列, 其中, 序列 x 的前 900 个点用来建立预测模型, 后 300 个点用来进行预测分析.

首先, 对序列 x 进行相空间重构, 取 $\varepsilon = 0.01$, 其两个重要参数的获得如图4所示.

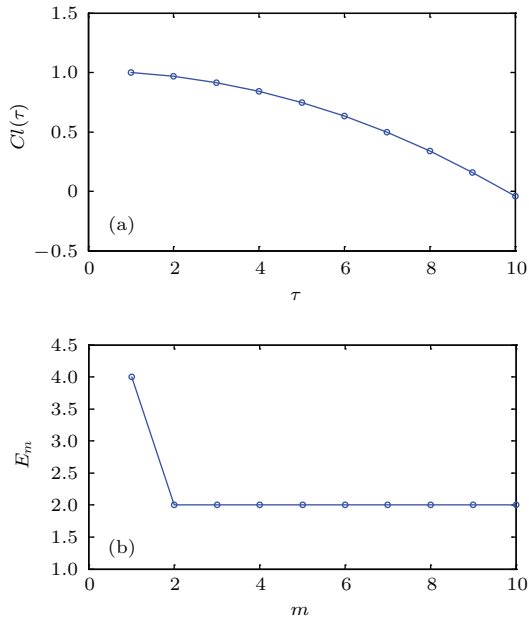


图4 相空间重构图 (a) $\tau-Cl(\tau)$ 关系曲线; (b) $m-E_m$ 关系曲线

Fig. 4. Phase space reconstruction: (a) Relationship curve of $\tau-Cl(\tau)$; (b) relationship curve of $m-E_m$.

结合自相关法, 从图4(a)可以看出, 最佳延迟时间 τ 为6; 结合Cao方法, 从图4(b)可以看出, 最小嵌入维数 m 为2. 因此, 根据(2)式可确定重构延迟时间向量.

其次, 在利用重构延迟时间向量和基于ITLBO算法的Hermite正交基神经网络建立预测模型时, Hermite正交基神经网络采用 $m-q-1$ 结构, 其中, $m=2, q=3$; ITLBO算法的初始化参数设置为: 优化参数向量维数 $D=5$, 班级成员数量 $N=40$, 教学次数 $G=30$. 为获取具有统计意义的结果, ITLBO算法独立运行20次, 根据每次得到最优解的平均值建立预测模型及进行预测分析, 其仿真曲线如图5所示(为方便, 将图、表中的基于ITLBO算法的Hermite正交基神经网络简记为ITLBO-HOB).

为定量验证本文所提预测模型的有效性和实用性, 将RBF神经网络、ESN网络、LSSVM模型及基于TLBO算法的Hermite正交基神经网络与之进行比较. 其中, 所有神经网络均采用 $m-q-1$ 结构且 $m=2$, 在权衡预测精度和建模时间的基础上, RBF神经网络隐含层节点数取 $q=5$, 且其隐含层激励函数采用Gauss核函数 $g(\mathbf{x}) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{c}_i\|_2^2}{\sigma_i^2}\right)$ ($i=1, 2, \dots, q$), \mathbf{c}_i ($i=1, 2, \dots, q$) 为第 i 个节点的

心向量, σ_i ($i=1, 2, \dots, q$) 为第 i 个节点的基宽, 并通过无监督学习算法确定 \mathbf{c}_i ($i=1, 2, \dots, q$) 和 σ_i ($i=1, 2, \dots, q$), 通过有监督学习算法确定隐含层到输出层的权值向量 $\mathbf{r} = [r_1, r_2, \dots, r_q]^T$; ESN网络隐含层激励函数采用双极性Sigmoid函数 $f(\mathbf{x}) = \frac{1 - e^{-\mathbf{x}}}{1 + e^{-\mathbf{x}}}$, 储备池规模取为40, 储备池内部连接权值谱半径取为0.8、输入单元尺度取为0.2、稀疏程度取为0.03, 并通过储备池学习算法建立预测模型; LSSVM模型采用组合核函数 $K = \rho K_{\text{poly}} + (1 - \rho) K_{\text{RBF}}$, K_{poly} 取为多项式函数 $[(\mathbf{x}, \mathbf{x}_i) + 1]^\alpha$ (\mathbf{x}_i 为输入向量), K_{RBF} 取为Gauss核函数 $g(\mathbf{x}) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{c}\|_2^2}{\sigma^2}\right)$, 且组合核函数比重 $\rho \in [0, 1]$, 多项式幂次 $\alpha \in (0, 1)$, 正则化变量 $\gamma \in [0.01, 1000]$, 核函数变量 $\sigma^2 \in [0.01, 1000]$, 并通过遗传算法确定 $\rho, \alpha, \gamma, \mathbf{c}$ 和 σ^2 , 其中, GA算法初始化参数设置为: 种群数量为40, 粒子维数为6, 进化代数为30, 交叉概率为0.8, 变异概率为0.2, 采用二进制编码, 长度为10; 基于TLBO算法的Hermite正交基神经网络与ITLBO算法的Hermite正交基神经网络建立预测模型时, 参数初始化设置相同. 根据上述各预测模型设置, 利用(32)–(35)式性能指标进行评价参数列于表1.

表1 性能指标评价表

Table 1. Evaluation of performance indicator.

预测步数	预测模型	性能指标			
		MMT/s	MAE	DC	RMSE
1	ITLBO-HOB	9.6660	0.0802	0.9895	0.1169
	TLBO-HOB	17.2996	0.2638	1.0774	0.3161
	LSSVM	27.6229	0.5889	0.9860	0.8458
	ESN	101.5057	1.5285	0.3499	2.3110
	RBF	58.7455	1.5119	0.6562	2.5218
6	ITLBO-HOB	9.8364	0.1505	1.0132	0.2181
	TLBO-HOB	16.9952	0.2548	1.0806	0.3213
	LSSVM	27.5891	1.1303	1.1785	1.7071
	ESN	102.0852	1.5297	0.7960	2.0804
	RBF	59.1397	0.8827	0.4977	2.2769
12	ITLBO-HOB	9.7803	0.2167	0.9917	0.3476
	TLBO-HOB	16.9959	0.3895	0.9005	0.6318
	LSSVM	27.7131	0.8721	0.9491	1.2525
	ESN	101.5318	1.5186	0.9364	2.0509
	RBF	59.1820	0.9994	0.2999	2.4113
18	ITLBO-HOB	9.7226	0.3044	0.9999	0.5619
	TLBO-HOB	17.0390	0.4167	1.0680	0.5634
	LSSVM	28.0880	0.9821	0.9129	1.3576
	ESN	102.5531	1.0990	0.4393	1.9358
	RBF	59.8733	1.4381	0.2114	2.6178

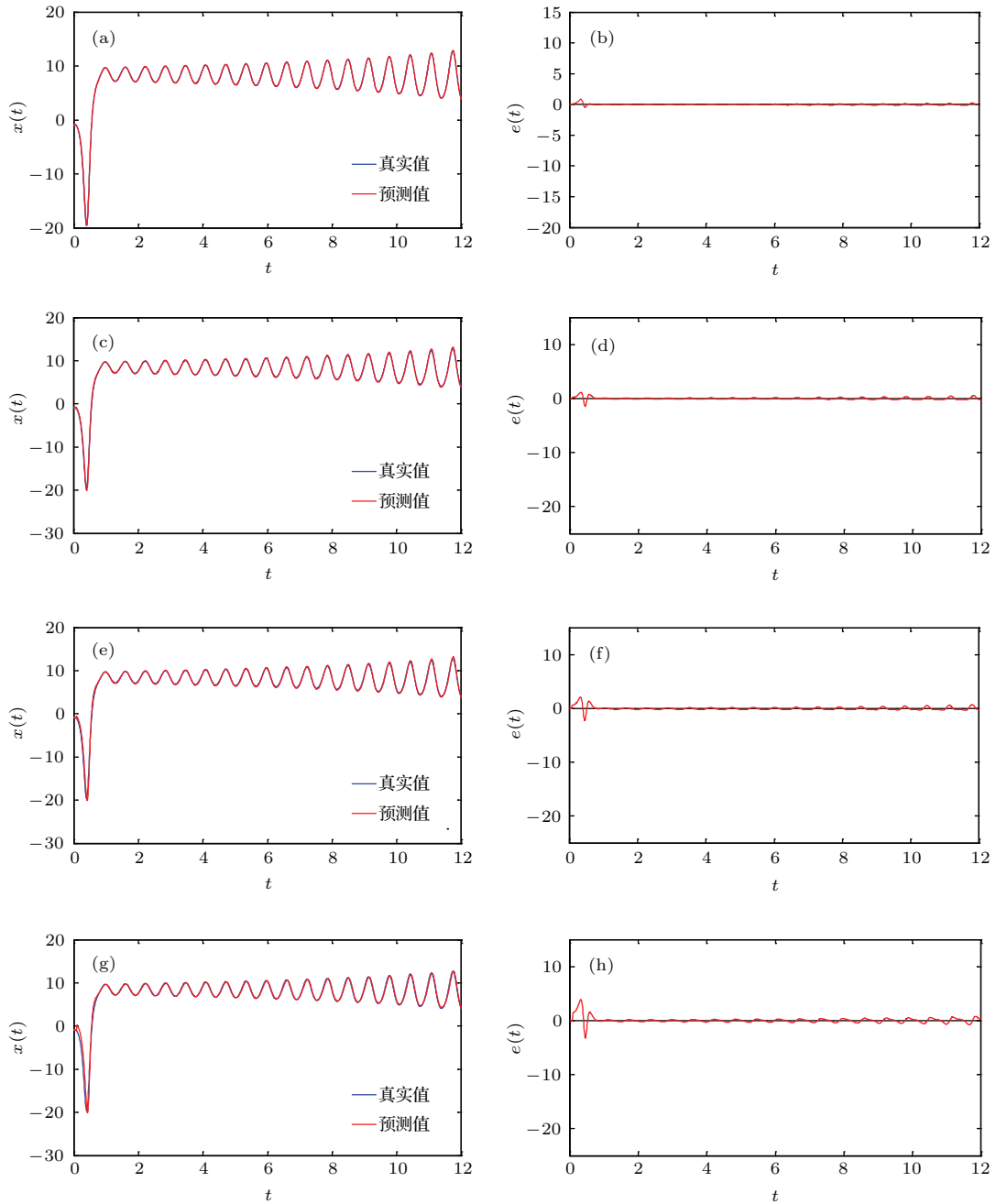


图5 (网刊彩色) ITLBO-HOB 仿真曲线 (a), (b) 单步预测及误差曲线; (c), (d) 6步预测及误差曲线; (e), (f) 12步预测及误差曲线; (g), (h) 18步预测及误差曲线

Fig. 5. (color online) Simulation curve of ITLBO-HOB: (a), (b) single-step prediction and its error curve; (c), (d) 6-step prediction and its error curve; (e), (f) 12-step prediction and its error curve; (g), (h) 18-step prediction and its error curve.

5.2 Liu系统

最近, 刘崇新等在电路实验中提出了一种新的混沌系统——Liu系统, 如(37)式. 在适当的参数控制下, 系统的吸引子呈“双翅膀形”, 与Lorenz系统有着相似的吸引轨道, 但与其拓扑不等价^[24]. 作为一种新型混沌系统, Liu系统在保密通信、同步与

反同步控制中可发挥重要的作用^[25].

$$\begin{aligned} \dot{x} &= a(y - x), \\ \dot{y} &= bx - kxz, \\ \dot{z} &= -cz + hx^2. \end{aligned} \quad (37)$$

当 $a = 10, b = 40, k = 1, c = 2.5, h = 4$ 时, 系统呈现混沌现象, 其吸引子如图6所示.

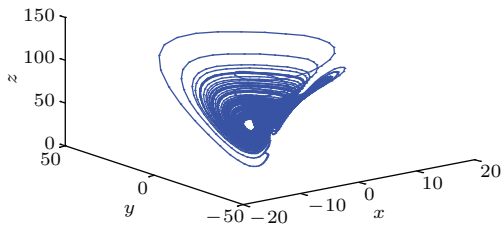


图6 Liu吸引子
Fig. 6. Liu attractor.

仿真过程中, 初值 x_0, y_0 及 z_0 分别取为 $-1, 0$ 和 1 , 步长 h 取为 0.01 . 通过四阶 Runge-Kutta 法计算包含 1200 个数据点的混沌时间序列, 其中, 序列 x 的前 1000 个点用来建立预测模型, 后 200 个点用来进行预测分析.

首先, 对序列 x 进行相空间重构, 取 $\varepsilon = 0.01$, 其两个重要参数获得如图 7 所示.

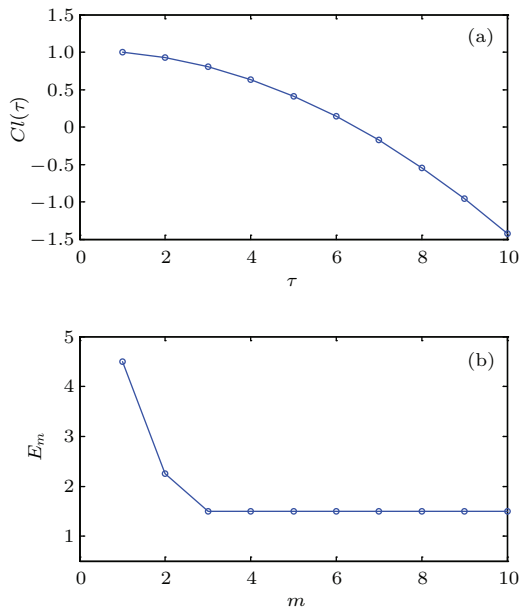


图7 相空间重构图 (a) $\tau-Cl(\tau)$ 关系曲线; (b) $m-E_m$ 关系曲线
Fig. 7. Phase space reconstruction: (a) Relationship curve of $\tau-Cl(\tau)$; (b) relationship curve of $m-E_m$.

结合自相关法, 从图 7(a) 可以看出, 最佳延迟时间 τ 为 4; 结合 Cao 方法, 从图 7(b) 可以看出, 最小嵌入维数 m 为 3. 因此, 根据 (3) 式可确定重构延迟时间向量.

其次, 在利用重构延迟时间向量和基于 ITLBO 算法的 Hermite 正交基神经网络建立预测模型时, Hermite 正交基神经网络采用 $m-q-1$ 结构, 其中, $m=3, q=3$. ITLBO 算法的初始化参数设置为: 优化参数向量维数 $D=6$, 班级成员数量 $N=40$, 教学次数 $G=30$. 为获取具有统计意

义的结果, ITLBO 算法独立运行 20 次, 根据每次得到最优解的平均值建立预测模型及进行预测分析, 其仿真曲线如图 8 所示.

表2 性能指标评价表
Table 2. Evaluation of performance indicator.

预测步数	预测模型	性能指标			
		MMT/s	MAE	DC	RMSE
1	ITLBO-HOB	11.6979	0.1418	0.9929	0.2162
	TLBO-HOB	23.2578	0.6647	1.0165	0.8082
	LSSVM	31.1595	0.7329	0.9034	1.0376
	ESN	138.5305	2.7877	0.7208	3.4274
	RBF	81.7044	3.5285	0.3036	4.4355
8	ITLBO-HOB	11.7143	0.1928	0.9903	0.3110
	TLBO-HOB	23.2318	0.4936	0.9244	0.6116
	LSSVM	30.9557	0.8853	0.8631	1.2971
	ESN	138.6610	2.6871	0.5386	3.6071
	RBF	80.9393	3.6161	0.2325	4.5734
16	ITLBO-HOB	11.6613	0.3582	0.9802	0.5639
	TLBO-HOB	23.2532	0.5711	1.0461	0.8148
	LSSVM	31.0952	1.3311	0.9471	1.8375
	ESN	138.5403	2.6657	0.6090	3.5016
	RBF	80.8359	2.8948	0.3284	3.8798
24	ITLBO-HOB	11.5040	0.5721	0.9525	0.8800
	TLBO-HOB	23.1910	0.8363	0.8035	1.1641
	LSSVM	30.0054	1.7164	0.5556	2.3360
	ESN	138.3648	2.1114	1.3054	3.0436
	RBF	80.7794	3.4671	0.4717	4.3432

根据 (32)—(35) 式, 利用 RBF 神经网络、ESN 网络、LSSVM 模型及基于 TLBO 算法的 Hermite 正交基神经网络与基于 ITLBO 算法的 Hermite 正交基神经网络做对比. 其中, 所有神经网络均采用 $m-q-1$ 结构且 $m=3$, 在权衡预测精度和建模时间的基础上, RBF 神经网络隐含层节点数取 $q=6$, 且其隐含层激励函数采用 Gauss 核函数 $g(\mathbf{x}) = \exp\left(-\frac{\|\mathbf{x}-\mathbf{c}_i\|_2^2}{\sigma_i^2}\right)$ ($i=1, 2, \dots, q$), \mathbf{c}_i ($i=1, 2, \dots, q$) 为第 i 个节点的中心向量, σ_i ($i=1, 2, \dots, q$) 为第 i 个节点的基宽, 并通过无监督学习算法确定 \mathbf{c}_i ($i=1, 2, \dots, q$) 和 σ_i ($i=1, 2, \dots, q$), 通过有监督学习算法确定隐含层到输出层的权值向量 $\mathbf{r} = [r_1, r_2, \dots, r_q]^T$; ESN 网络隐含层激励函数采用双极性 Sigmoid 函数 $f(\mathbf{x}) = \frac{1-e^{-\mathbf{x}}}{1+e^{-\mathbf{x}}}$, 储备池规模取为 60, 储备池内部连接权值谱半径取为 0.9, 储备池输入单元尺度取为 0.4, 储备池稀疏程度取为 0.05, 并通过储备池学习算法建立预测模型; LSSVM 模型采用组合核函数 $K = \rho K_{\text{poly}} + (1-\rho) K_{\text{RBF}}$, K_{poly} 取为多项式函

数 $[(\mathbf{x}, \mathbf{x}_i) + 1]^\alpha$ (\mathbf{x}_i 为输入向量), K_{RBF} 取为 Gauss 核函数 $g(\mathbf{x}) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{c}\|_2^2}{\sigma^2}\right)$, 且组合核函数比重 $\rho \in [0, 1]$, 多项式幂次 $\alpha \in (0, 1)$, 正则化变量 $\gamma \in [0.01, 1000]$, 核函数变量 $\sigma^2 \in [0.01, 1000]$, 并通过 GA 算法确定 $\rho, \alpha, \gamma, \mathbf{c}$ 和 σ^2 , 其中, 遗传算法初始化参数设置为: 种群数量为 40, 粒子维数为 7,

进化代数为 30, 交叉概率为 0.8, 变异概率为 0.2, 采用二进制编码, 长度为 10. 基于 TLBO 算法的 Hermite 正交基神经网络与 ITLBO 算法的 Hermite 正交基神经网络建立预测模型时, 参数初始化设置相同. 按照上述各预测模型设置, 其性能指标评价参数列于表 2.

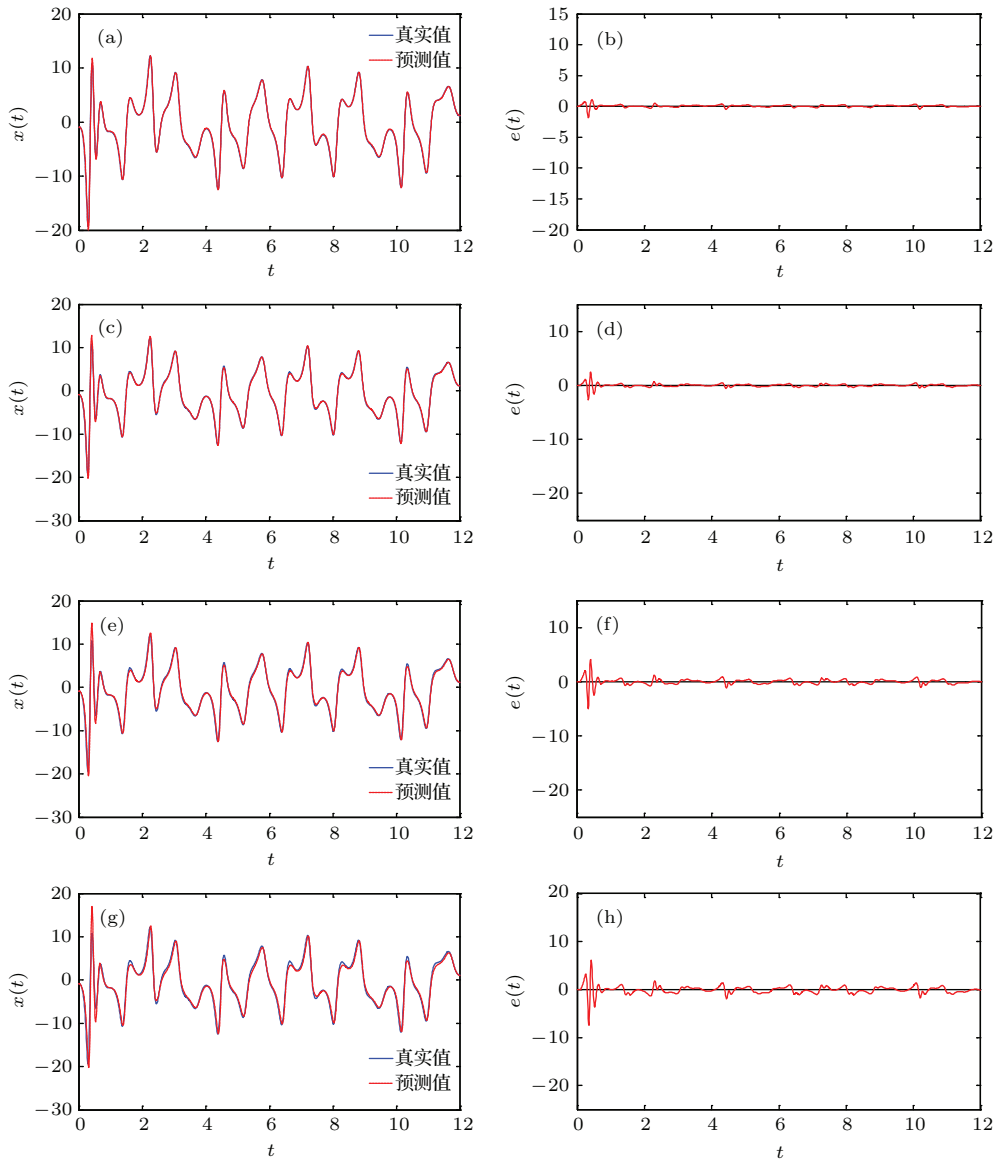


图 8 (网刊彩色) ITLBO-HOB 仿真曲线 (a), (b) 单步预测及误差曲线; (c), (d) 8 步预测及误差曲线; (e), (f) 16 步预测及误差曲线; (g), (h) 24 步预测及误差曲线

Fig. 8. (color online) Simulation curve of ITLBO-HOB: (a), (b) single-step prediction and its error curve; (c), (d) 8-step prediction and its error curve; (e), (f) 16-step prediction and its error curve; (g), (h) 24-step prediction and its error curve.

5.3 性能分析

从图 5 和图 8 可以看出, 不论是单步预测还是多步预测, 基于 ITLBO 算法的 Hermite 正交基神

经网络的预测序列几乎与原序列重合, 误差线与零水平线十分接近, 验证了所提预测模型的高精度性. 由表 1 和表 2 可以看出, 不论是单步预测还是多步预测, 基于 ITLBO 算法的 Hermite 正交基神

经网络的MMT明显缩短,约是其他对比预测模型的8%—50%;所提预测模型的MAE明显减小,约是其他对比预测模型的5%—20%;相对于其他预测模型,所提预测模型的DC更接近于1;所提预测模型的RMSE明显降低,约是其他对比预测模型的4%—25%.表1和表2为图5和图8做了补充和完善,进一步验证了基于ITLBO算法的Hermite正交神经网络收敛速度快、预测精度高且稳定的特性.同时,通过性能指标对比,也说明ITLBO算法较TLBO算法明显增强了全局搜索能力、提高了收敛速度,很好地改进了TLBO算法.

此外,通过与其他预测模型对比知,Hermite正交神经网络隐含层节点数少、结构简单,而且待优化参数向量维数低、未知参数少,过程简单、易于实现优化操作.同时,ITLBO算法作为一种高效性的智能优化算法,避免了传统其他优化算法的缺点,而且操作简单、寻优能力强、收敛速度快.因此,结合两者的优点,利用ITLBO算法优化Hermite正交神经网络参数,极大地提高了运行效率:1)所提预测模型结构简单、待优化参数少,使得操作简单、易于实现,缩短建模时间;2)结合ITLBO算法的优点,操作简单、搜索能力强、收敛速度快,既能提高搜索精度,进而提高预测精度,又能提高搜索速度,从而减少建模时间.

通过以上分析知,基于ITLBO算法的Hermite正交神经网络预测模型结构简单、预测精度高、建模时间短,很好地说明了该预测模型的有效性和实用性,可进一步推广和应用.

6 结 论

混沌现象在人类社会及自然界中普遍存在,使人们对混沌的认识和研究不断加强,尤其是有关于混沌时间序列预测的问题,成为目前研究的焦点和热点问题.但由于混沌系统的复杂性,使得其产生的混沌时间序列具有很强的非线性,给预测和应用带来了很大的麻烦.因此,结合相空间重构理论、Hermite正交基函数的性质、最佳平方逼近理论及神经网络理论,本文提出了基于ITLBO算法的Hermite正交神经网络预测模型.通过与对比预测模型分别进行单步和多步预测实验相比较,仿真结果证实了所提预测模型具有预测精度高、收敛速度快及模型结构简单的优良特性,便于在生产、生活中应用,具有一定的理论意义和实用价值,尤其

在交通预测、股市分析、天气预报、水文监测、故障诊断、保密通信及同步与异步控制等领域可发挥极其重要的作用.

参考文献

- [1] Xing H Y, Zhang Q, Xu W 2015 *Acta Phys. Sin.* **64** 040506 (in Chinese) [行鸿彦, 张强, 徐伟 2015 物理学报 **64** 040506]
- [2] Zhao Y P, Wang K K 2013 *Acta Phys. Sin.* **62** 240509 (in Chinese) [赵永平, 王康康 2013 物理学报 **62** 240509]
- [3] Lü S X, Wang Z S, Hu Z H, Feng J C 2014 *Chin. Phys. B* **23** 010506
- [4] Wang L, Zou F, Hei X, Yang D, Chen D, Jiang Q, Cao Z 2014 *Neural Comput. Appl.* **25** 1407
- [5] You R Y, Huang X J 2011 *Chin. Phys. B* **20** 020505
- [6] Lee S H, Chung K Y, Lim J S 2014 *Pers. Ubiquit. Comput.* **18** 1315
- [7] Zhao N, Yu F R, Sun H, Yin H, Wang G 2015 *Wirel. Netw.* **21** 1227
- [8] Li S, Zhang C, Shi M 2015 *J. Shanghai Jiaotong Univ. (Sci.)* **20** 224
- [9] Zhao L, Shui P, Jiang F, Qiu H, Ren S, Li Y, Zhang Y 2014 *Earth Sci. Inform.* **7** 59
- [10] Bansal A, Chen T, Zhong S 2011 *Neural Comput. Appl.* **20** 143
- [11] Yu X, Wang B, Batbayar B, Wang L, Man Z 2011 *J. Global Optim.* **51** 271
- [12] Li P H, Chai Y, Xiong Q Y 2013 *Acta Autom. Sin.* **39** 1511 (in Chinese) [李鹏华, 柴毅, 熊庆宇 2013 自动化学报 **39** 1511]
- [13] Han M, Xu M L, Wang X Y 2014 *Chin. J. Comput.* **37** 2268 (in Chinese) [韩敏, 许美玲, 王新迎 2014 计算机学报 **37** 2268]
- [14] Kohli A K, Rai A 2013 *Circ. Syst. Signal Pr.* **32** 223
- [15] Aiyer B G, Kim D, Karingattikkal N, Samui P, Rao P R 2014 *KSCE J. Civ. Eng.* **18** 1753
- [16] Han M, Wang X Y 2013 *Control Theory Appl.* **30** 1467 (in Chinese) [韩敏, 王新迎 2013 控制理论与应用 **30** 1467]
- [17] Lazzús J A 2011 *Chin. Phys. Lett.* **28** 110504
- [18] Pandey A, Thapa K B, Prasad R, Singh K P 2012 *J. Indian Soc. Remote* **40** 709
- [19] Hosseini E S, Esmaelzadeh V, Eslami M 2015 *Wireless Pers. Commun.* **80** 1579
- [20] Sardashti A, Daniali H M, Varedi S M 2013 *Meccanica* **48** 1681
- [21] Zhang S, Zhang Y, Zhu J 2015 *J. Mech. Sci. Technol.* **29** 605
- [22] Pawar P J, Rao R V 2013 *Int. J. Adv. Manuf. Tech.* **67** 1955
- [23] Liao X X, Luo Q 2010 *Sci. Sin. Technol.* **40** 1086 (in Chinese) [廖晓昕, 罗琦 2010 中国科学: 信息科学 **40** 1086]
- [24] Huang Y 2014 *Acta Phys. Sin.* **63** 080505 (in Chinese) [黄云 2014 物理学报 **63** 080505]
- [25] Wang Z L 2010 *Nonlinear Dynam.* **59** 455

Hermite orthogonal basis neural network based on improved teaching-learning-based optimization algorithm for chaotic time series prediction*

Li Rui-Guo^{1)†} Zhang Hong-Li¹⁾ Fan Wen-Hui²⁾ Wang Ya³⁾

1) (School of Electrical Engineering, Xinjiang University, Urumqi 830047, China)

2) (Department of Automation, Tsinghua University, Beijing 100084, China)

3) (School of Mechanical Engineering, Xinjiang University, Urumqi 830047, China)

(Received 8 May 2015; revised manuscript received 3 June 2015)

Abstract

Chaos phenomenon which exists widely in nature and society affects people's production and life. It has great important significance to find out the regularity of chaotic time series from a chaotic system. Since chaotic system has extremely complex dynamic characteristics and unpredictability, and chaotic time series prediction through traditional methods has low prediction precision, slow convergence speed and complex model structure, a prediction model about Hermite orthogonal basis neural network based on improved teaching-learning-based optimization algorithm is proposed. Firstly, according to the chaotic time series, autocorrelation method and Cao method are used to determine the best delay time and the minimum embedding dimension respectively, then a phase space is reconstructed to obtain the refactoring delay time vector. Secondly, on the basis of phase space reconstruction and best square approximation theory, combined with the neural network topology, a prediction model about Hermite orthogonal basis neural network with excitation functions based on the Hermite orthogonal basis functions is put forward. Thirdly, in order to optimize the parameters of the prediction model, an improved teaching-learning-based optimization algorithm is proposed, where a feedback stage is introduced at the end of the learning stage based on the teaching-learning-based optimization algorithm. Finally, the parameter optimization problem is transformed into a function optimization problem in the multidimensional space, then the improved teaching-learning-based optimization algorithm is used for parameter optimization of the prediction model so as to establish it and analyze it. Lorenz and Liu chaotic systems are taken as models respectively, then the chaotic time series which will be used as simulation object is produced by the fourth order Runge-Kutta method. The comparison experiments with other prediction models are conducted on single-step and multi-step prediction for the chaotic time series. The simulation results and numerical analysis show that compared with radial basis function neural network, echo state network, least square support vector machine prediction model and Hermite orthogonal basis neural network based on teaching-learning-based optimization algorithm, the proposed prediction model has the mean absolute error and root mean square error reduced significantly, has a decision coefficient close to 1, meanwhile, has a mean modeling time shortened greatly. So the proposed prediction model can improve the prediction precision, accelerate the convergence speed and simplify the model structure, thus the prediction model is effective and feasible, which makes it promoted and applied easily.

Keywords: Hermite orthogonal basis neural network, improved teaching-learning-based optimization algorithm, chaotic time series, prediction

PACS: 05.45.Tp

DOI: 10.7498/aps.64.200506

* Project supported by the National Natural Science Foundation of China (Grant No. 61463047) and the Graduate Student Scientific Research Innovation Program of the Autonomous Region, China (Grant No. XJGRI2015039).

† Corresponding author. E-mail: 1141567852@qq.com