

改进回溯搜索优化回声状态网络时间序列预测^①



胡 率, 肖治华, 饶 强, 廖荣涛

(国网湖北省电力有限公司 信息通信公司, 武汉 430077)

通讯作者: 胡 率, E-mail: hushuai.hs@hotmail.com

摘 要: 回声状态网络 (Echo State Network, ESN) 网络结构简单且耦合“时间参数”, 在时间序列预测研究中具有重要的理论和应用价值. 本文提出使用自适应回溯搜索算法 (Adaptive Backtracking Search optimization Algorithm, ABSA) 优化 ESN 输出连接权值矩阵, 克服标准线性回归方法造成的网络过拟合问题. ABSA 使用自适应变异因子策略替换标准 BSA 中随机给定变异因子的策略, 实现 BSA 在收敛精度和收敛速率之间的平衡. 实验表明, 采用 ABSA 优化的 ESN 能够比未优化的 ESN 和采用其他进化算法优化的 ESN 获得更好的预测精度.

关键词: 时间序列预测; 回声状态网络; 回溯搜索算法; 预测模型; 进化算法

引用格式: 胡率, 肖治华, 饶强, 廖荣涛. 改进回溯搜索优化回声状态网络时间序列预测. 计算机系统应用, 2020, 29(1): 236–243. <http://www.c-s-a.org.cn/1003-3254/7255.html>

Time Series Forecasting Based on Echo State Network Optimized by Improved Backtracking Search Optimization Algorithm

HU Shuai, XIAO Zhi-Hua, RAO Qiang, LIAO Rong-Tao

(Information & Communication Branch, State Grid Hubei Electric Power Co. Ltd., Wuhan 430077, China)

Abstract: Echo State Network (ESN) owns simple network structure and is coupled with a time parameter and thus it shows important theoretical and application values in time series forecasting. In this study, we propose to optimize the output weight matrix by Adaptive Backtracking Search optimization Algorithm (ABSA) to overcome overfitting problem caused by linear regression algorithm. ABSA adopts adaptive mutation factor strategy to replace the strategy of randomly given mutation factor in standard BSA to achieve the balance between convergence accuracy and convergence rate. Experimental results show that the ESN optimized by ABSA outperforms the basic ESN without optimization and the ESNs optimized by other EAs.

Key words: time series forecasting; Echo State Network (ESN); Adaptive Backtracking Search optimization Algorithm (ABSA); forecasting model; evolutionary algorithm

现实世界中存在各种各样的时间序列数据, 分析挖掘数据的隐藏信息, 基于合理的假设和推理, 建立能够拟合复杂系统行为的模型, 并依据该模型对序列未来的发展规律做出估计和判断符合科学规律^[1]. 作为时间序列分析的一个研究分支, 时间序列预测在实际中广泛应用. 但是, 进行时间序列预测依然存在巨大的挑

战, 主要原因是时间序列数据往往表现出复杂的特性, 比如时序性、高维性、非线性和非周期性等.

人工神经网络 (Artificial Neural Network, ANN) 是一种以数据为驱动的、自适应的人工智能模型, 它具有非常良好的非线性映射能力. 从理论角度来讲, ANN 可以以任意精度无限逼近任何动态系统的行为^[2]. ANN

① 基金项目: 国网湖北省电力有限公司科技项目 (52153317000B)

Foundation item: Scientific and Technological Project of State Grid Hubei Electric Power Co. Ltd. (52153317000B)

收稿时间: 2019-05-22; 修改时间: 2019-06-21, 2019-07-22; 采用时间: 2019-07-24; csa 在线出版时间: 2019-12-27

几乎能够模拟任何线性和非线性的时间序列数据生成过程,因而成为准确且应用广泛的预测模型^[3].

2004年,Jaeger等^[4]提出了一种全新的递归神经网络——回声状态网络(Echo State Network, ESN). ESN最显著特征是网络训练过程简单:只有储备池(隐藏层)到输出层的输出连接权值矩阵需要计算,其他各层之间的连接权值矩阵在网络初始化阶段随机生成并保持不变. ESN能够克服传统递归神经网络的网络结构难以确定、训练效率低下、收敛速度慢、容易陷入局部最优等缺点,因而应用广泛. 例如网络流量预测^[5]、变风量空调内模控制^[6]和非线性卫星信道均衡^[7]等. 标准的 ESN 在计算输出连接权值矩阵时会采用最小二乘估计或者岭回归等线性方法,这类方法在节约计算成本的同时会造成网络过拟合. 因此,对 ESN 的输出连接权值矩阵进行优化以提升其性能具有现实意义.

本质上,对 ESN 的输出连接权值矩阵进行优化是一个离散、高维和强非线性的复杂问题. 研究指出进化算法(Evolutionary Algorithm, EA)在解决此类问题时性能卓越. 宗宸生等^[8]运用改进粒子群算法优化 BP 神经网络的初始权重,建立粮食产量预测模型,该模型具有更高的预测精度和较大的适应度. 朱海龙等^[9]利用 GA 优化 BP 神经网络的初始权值和阈值进而建立胎儿体重预测模型,模型收敛速度和预测精度都得到提升. 针对 ESN 的优化,田中大等^[5]利用经典的 GA 对 ESN 储备池参数进行优化,进行网络流量预测时获得不错的精度. Chouikhi 等^[10]应用 PSO 对 ESN 的固定权值矩阵进行预训练,以降低连接权值矩阵在具体应用问题上的随机性. 也有学者使用进化算法对 ESN 的储备池到输出层之间的连接进行优化,例如 Wang 等^[11]提出使用 BPSO 对 ESN 的储备池到输出层之间的连接进行优化,优化模型在系统识别和时间序列预测问题上获得更好的实验结果. 本文将从提升 ESN 的预测精度出发,提出使用自适应回溯搜索算法(Adaptive Backtracking Search Optimization Algorithm, ABSA)对 ESN 的输出连接权值矩阵进行优化.

1 相关知识

1.1 回声状态网络

回声状态网络因为其相对简单的网络结构、训练算法和性能而广受关注. ESN 是储备池计算方法之一,采用“储备池”作为信息处理媒介,储备池将网络的输

入信号映射到高维的复杂动态状态空间. 储备池的神经元之间的连接在网络初始化阶段随机生成并在整个训练过程中保持不变^[12]. 正是由于储备池计算特征, ESN 相对于传统的 RNN 具有明显的优势. 标准结构的 ESN 如图 1 所示,可以分为 3 个组成部分:一个输入层(K 个神经元),一个储备池(N 个神经元)和一个输出层(L 个神经元).

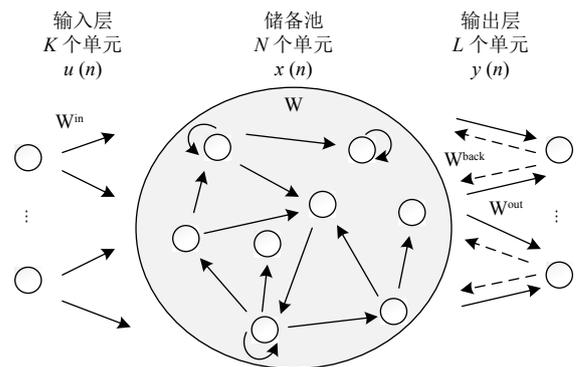


图 1 回声状态网络的标准结构

ESN 在时刻 n 的输入,储备池神经元的状态和输出层神经元输出分别表示如下:

$$u(n) = (u_1(n), u_2(n), \dots, u_K(n))^T \quad (1)$$

$$x(n) = (x_1(n), x_2(n), \dots, x_N(n))^T \quad (2)$$

$$y(n) = (y_1(n), y_2(n), \dots, y_L(n))^T \quad (3)$$

如图 1 所示,在标准结构的 ESN 中,当输入信号 $u(n+1)$ 输入到网络时,储备池内部神经元在时刻 $n+1$ 的状态和网络的输出分别按照方程 (4) 和 (5) 进行更新:

$$x(n+1) = f(W^{\text{in}}u(n+1) + Wx(n)) \quad (4)$$

$$y(n+1) = f^{\text{out}}(W^{\text{out}}u(n+1)) \quad (5)$$

其中, W^{in} ($W^{\text{in}} \in \mathbb{R}^{N \times K}$)、 W ($W \in \mathbb{R}^{N \times N}$) 和 W^{out} ($W^{\text{out}} \in \mathbb{R}^{L \times N}$) 分别代表网络的输入层到储备池、储备池内部神经元和储备池到输出层之间的连接权值矩阵. $f = [f_1, f_2, \dots, f_N]$ 表示储备池中神经元的输出函数. 一般来说, f_i ($i = 1, 2, \dots, N$) 采用“S-型”函数,比如双曲正切函数 \tanh . $f^{\text{out}} = [f_1^{\text{out}}, f_2^{\text{out}}, \dots, f_L^{\text{out}}]$ 表示网络输出层神经元的输出函数. 通常情况下, f_i^{out} ($i = 1, 2, \dots, L$) 会取恒等函数.

1.2 回溯搜索算法及其改进

回溯搜索算法(Backtracking Search optimization Algorithm, BSA)是由 Civicioglu 在 2013 年提出的一

种属于进化算法范畴的新智能算法^[13]. BSA 自从被提出就吸引了众多学者进行研究并应用到不同的领域, 比如数值优化^[13,14]、自动发电控制^[15]和功率流^[16]等. 标准的 BSA 主要包括 5 个基本算子: 初始化 (Initialization)、选择 I (Selection-I)、变异 (Mutation)、交叉 (Crossover) 和选择 II (Selection-II), 其算法流程如图 2 所示.

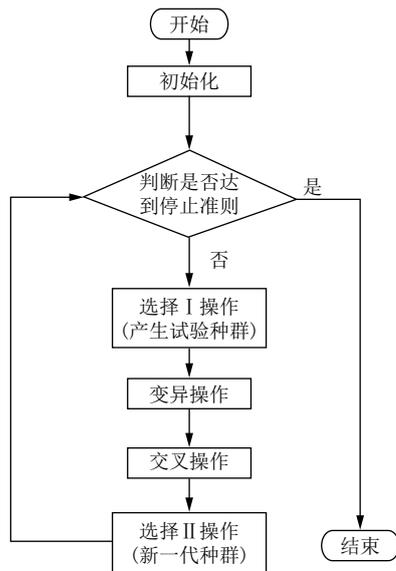


图 2 标准 BSA 算法流程

在标准 BSA 中, 变异因子 F 在初始化过程中随机给定并在整个过程中保持不变, 这种策略不能平衡 BSA 算法的搜索能力和收敛性之间的关系. 如果变异因子 F 取值大, BSA 的全局搜索能力会很强而获得的全局最优解可能展现出低精度. 如果变异因子 F 取值小, BSA 在迭代过程中会加速收敛而其全局搜索能力会减弱. 可以说, 变异因子 F 是平衡 BSA 的全局搜索能力和收敛性的关键参数. 为了充分考虑变异因子 F 对 BSA 性能的影响, 本文提出自适应回溯搜索算法 (ABSA). ABSA 使用自适应变异因子策略去替换标准的随机给定策略, 其策略方程表达如式 (6) 所示:

$$F = F_{\min} + (F_{\max} - F_{\min}) \times e^{1 - \frac{GenM}{GenM - G + 1}} \quad (6)$$

其中, F_{\max} 和 F_{\min} 分别代表变异因子 F 的最大值和最小值. $GenM$ 是 BSA 的最大迭代次数, G 是当前迭代次数. e 是自然常数. 从式 (6) 可以得出: 变异因子的取值是随算法迭代过程的进行而不断改变的. 在开始迭代阶段, F 取值较大用于扩大搜索范围, 保证搜索时种群的多样性; 在随后的迭代过程中, 变异因子 F 取值逐渐

变小, 种群的搜索范围将缩小且种群中已经获得的优秀个体将会以较大概率被保留.

2 预测模型

2.1 模型介绍

为了验证本文拟采用的自适应回溯搜索算法优化回声状态网络 (ABSA_ESN) 的性能, 将设计其他的四种预测模型用于与 ABSA_ESN 进行比较. 具体而言, 本文将设计以下 5 种预测模型:

- (1) ESN. 标准的 ESN 模型.
- (2) GA_ESN. 使用遗传算法 (Genetic Algorithm, GA) 优化 ESN 输出连接权值矩阵.
- (3) DE_ESN. 使用差分进化算法 (Differential Evolution, DE) 优化 ESN 输出连接权值矩阵.
- (4) BSA_ESN. 使用标准回溯搜索算法 (BSA) 优化 ESN 输出连接权值矩阵.
- (5) ABSA_ESN. 使用自适应回溯搜索算法 (ABSA) 优化 ESN 输出连接权值矩阵.

以上 5 种预测模型中, GA_ESN、DE_ESN、BSA_ESN 和 ABSA_ESN 是四个优化的 ESN 模型. 其中, GA、DE、BSA 和 ABSA 都属于进化算法 (Evolutionary Algorithm, EA) 的范畴.

2.2 ESN 的优化流程

利用 ABSA (GA, DE 或者 BSA) 优化 ENS 的输出连接权值矩阵 W^{out} 的流程图如图 3 所示.

ESN 的 W^{out} 优化流程分为以下几个步骤:

Step 1. 设置 ESN 参数并收集网络输入产生的状态矩阵.

Step 2. 设置 ABSA (GA, DE 或 BSA) 的参数并生成初始化种群. 设置 ABSA 的初始迭代代数 $G = 0$.

Step 3. 判断 ABSA (GA, DE 或 BSA) 的迭代是否达到最大迭代次数. 如果当前迭代次数 G 等于最大迭代次数 $GenM$, 则 ABSA (GA, DE 或 BSA) 的迭代过程结束, 获得最优种群和最优个体; 否则, 流程继续执行.

Step 4. 依次执行 ABSA (GA, DE 或 BSA) 的选择 I (或没有)、自适应变异 (或变异)、交叉和选择 II (或选择) 等几个操作, 获得当前迭代过程 G 的新种群.

Step 5. 计算新种群中每个个体的适应值, 按照贪婪选择原则更新当前迭代过程的种群. 比较当前迭代过程获得的最优个体与当前获得的全局最优个体的适应值, 适应值较小的将作为新的全局最优适应值且相

应的个体更新为全局最优个体。

Step 6. 更新 $G = G + 1$. 返回 Step 3.

Step 7. 将进化算法获得的最优个体设置为 ESN 的最优输出连接权值矩阵 W^{out} , ESN 网络训练完成。

Step 8. 将测试集中样本按顺序输入到训练好的 ESN 中, 进行预测。

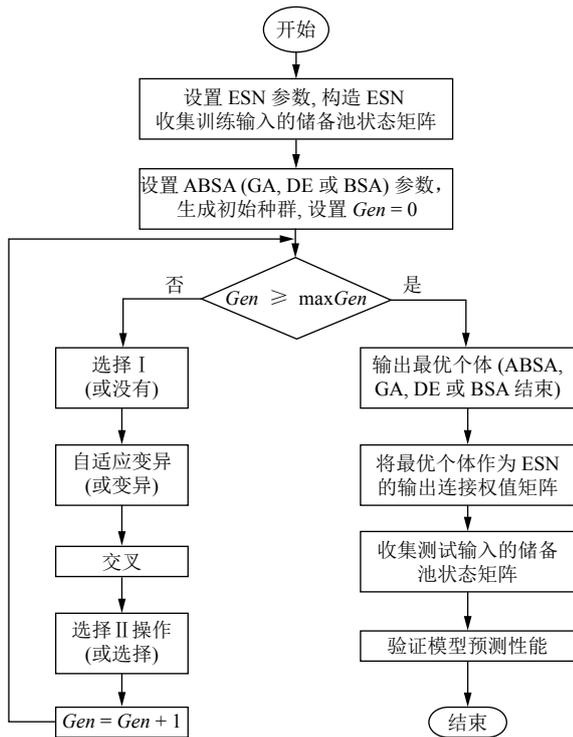


图3 利用 ABSA (GA, DE 或 BSA) 优化 ESN 的 W^{out} 的流程图

2.3 模型合理性分析及模型比较

本文所采用的 5 种模型的建模方法在上述 2.1 和 2.2 小节已经介绍. 本小节将对 5 种模型进行比较并讨论选择它们的合理性.

(1) ESN 是最先进的时间序列预测模型之一. ESN 因为其网络结构简单、训练效率高且耦合“时间参数”, 因而能够有效地应用于时间序列预测问题研究. 然而, 标准的 ESN 计算地输出连接权值矩阵容易致使网络陷入过拟合状态, 影响 ESN 性能的发挥.

(2) GA_ESN, DE_ESN 和 BSA_ESN 是优化的 ESN 预测模型. GA、DE 和 BSA 3 种智能算法的算法流程、算子和参数个数不尽相同. 使用 3 种智能算法对 ESN 进行优化的目的是克服标准 ESN 容易陷入网络过拟合状态的缺陷, 增加优化方法的多样性. 相对于

GA 和 DE, BSA 在全局寻优时只有一个参数需要设定, 在很大程度上能够简化算法参数设置过程, 进而节约计算成本.

(3) ABSA_ESN 是采用改进 BSA 优化的 ESN 预测模型. 标准 BSA 的参数变异因子在算法初始化过程中随机给定并在整个过程中保持不变, 这种策略难以平衡 BSA 算法的搜索能力和收敛速率. ABSA 是采用自适应变异因子策略改进的 BSA, 其能够有效地克服上述缺陷. 具体而言, ABSA 在迭代初期变异因子取值较大, 能够扩大搜索范围从而保证种群的多样性. 此时, ABSA 的收敛速度较慢且收敛精度较低. 随着迭代过程的进行, 变异因子的取值将逐渐变小, 种群的搜索范围将围绕已经获得的优秀个体展开以保证优秀个体将会以较大概率被保留. 此时, ABSA 的收敛速度和收敛精度逐步加快和提高. 总体而言, ABSA_ESN 在算法的参数设置、收敛速度和收敛精度上较 GA_ESN, DE_ESN 和 BSA_ESN 等均有优势.

3 数值实验和结果分析

3.1 实验设置

为了验证本文所提出的 ABSA_ESN 和其他四个用于对比的模型预测性能, 两个混沌时间序列将被采用作为实验数据集. 完成所有实验的个人电脑配置如下: 操作系统是 64 位 Windows 10 专业版、处理器是 Intel(R) Core(TM) i7-7700 CPU @3.6 GHz 和内存为 16 GB. 运行实验的软件环境是 Python 3.6 和 Matlab 2016b.

3.2 数据集及性能评估标准

本文将采用两个混沌时间序列数据集用于验证所提出的五种预测模型的性能. 第一个混沌时间序列是非线性自回归滑动模型 (Nonlinear Autoregressive Moving Average, NARMA), 这是一种有输入的非线性自回归移动平均数数列, 该模型已经成功应用于非线性系统^[17]. 第二个混沌时间序列是 Mackey-Glass 时间序列^[18]. 以上两个时间序列模型经常被使用作为测试时间序列预测算法的性能.

3.2.1 NARMA 模型

NARMA 模型方程式表示如下:

$$y(t) = 0.7e(t - \delta) + 0.1 + (1 - y(t - 1))y(t - 1) \quad (7)$$

其中, $e(t)$ 为独立分布在 $(0, 1)$ 之间的时刻 t 的模型输入, $y(t)$ 是时刻 t 的模型输出, δ 为时滞参数. 通常取 $\delta=3$.

在本实验中,我们设定 NARMA 模型的样本大小为 1000. 训练集、验证集和测试集样本的产生过程如下: 首先, 随机产生 1000 个取值范围在(0, 1)的样本作为系统输入 $e(t)$. 根据式 (7) 计算出 NARMA 方程式相应的期望输出 $y(t)$. 第二, 分别取系统输入样本和期望输出样本的前 80%(样本编号 1-800) 样本作为训练集样本. 第三, 紧接着选取 10%(样本编号 801-900) 样本作为确定回声状态网络模型结构的验证样本. 最后, 选取剩余 10%(样本编号 901-1000) 样本作为预测集样本. 本文设计验证集的主要目的是为了在训练阶段选择合适的 ESN 模型结构. 含有 1000 个样本的系统输入如图 4 及其对应的系统输出如图 5 所示.

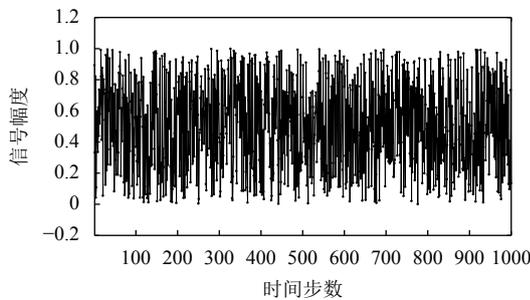


图4 NARMA 模型输入序列

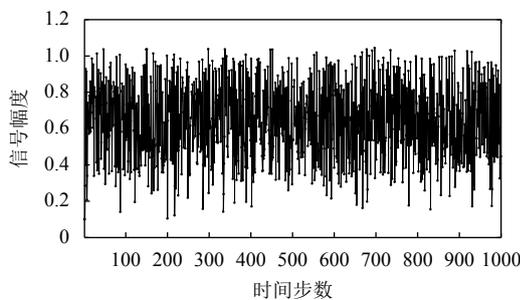


图5 NARMA 模型输出序列

3.2.2 Mackey-Glass 混沌系统

Mackey-Glass 时间序列是从时延差分系统导出, 其方程如式 (8) 所示:

$$\frac{dx(t)}{dt} = -cx(t) + \frac{ax(t-\tau)}{1+x^b(t-\tau)} \quad (8)$$

其中, $x(t)$ 是时间序列在时刻 t 时的取值. a 、 b 和 c 为参数, 通常取 $a = 0.2$ 、 $b = 10$ 和 $c = 0.1$. Farmer 对 Mackey-Glass 方程的行为特性做过深入研究, 研究结果表明当时滞参数 $\tau > 16.8$ 时该方程呈现混沌性, 并且 τ 值越大, 混沌程度越高^[19]. 本文依照一般情况设置时滞参数取

值为 $\tau = 17$.

在本实验中, 我们设定 Mackey-Glass 时间序列样本大小为 1000. 训练集、验证集和测试集样本的划分同 NARMA 模型一样, 即编号 1-800 的样本作为训练集, 编号 801-900 的样本作为验证集和编号 901-1000 的样本作为测试集. 使用二阶 Runge-Kutta 方法以步长为 0.1 构造的样本长度为 1000 的时间序列如图 6 所示.

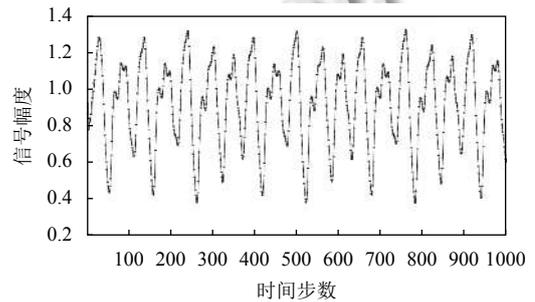


图6 Mackey-Glass 时间序列

3.2.3 性能评估标准

用于测量时间序列预测模型性能的指标多种多样, 本文将采用均方根误差 (Root Mean Square Error, RMSE) 作为评估预测精度指标, 其定义如式 (9) 所示:

$$RMSE = \sqrt{\frac{1}{k} \sum_{t=1}^k (\hat{y}_t - y_t)^2} \quad (9)$$

其中, \hat{y}_t 和 y_t 分别表示在时刻 t 时的预测值和实际值, k 表示预测值的个数.

3.3 参数设置与分析

3.3.1 参数设置

本文所提出的预测模型的性能受 ESN 结构选择和相应参数设置的影响. 结合本文所使用的两个时间序列数据集的特点, 拟采用 ESN 的结构参数如下: (1) 输入层单元的数目为 2(在每个时刻 t , ESN 有一个常数输入 0.1); (2) 储备池的大小从集合 {20, 30, 50} 中选择; (3) 输出层单元数目为 1. ESN 的其它的参数, 比如储备池内部连接权值矩阵谱半径 SR 和储备池稀疏性 SP 根据已有文献的推荐进行设置, 即 $SR = 0.8$ 和 $SP = 5\%$ ^[12]. 输入单元尺度 IC 及位移尺度 IS 分别设置为 $IC = 0.3$ 和 $IS = -0.2$.

关于 BSA 和 ABSA 的参数设置主要是基于一列实验和已有参考文献的推荐. BSA 和 ABSA 的具体参

数设置如下: (1) 种群大小均设置为 100; (2) 最大迭代次数均设置为 100; (3) 交叉概率 *mixrate* 均设置为 1.0^[13]; (4) BSA 的变异因子 *F* 设置为 $F=3 * randn$ ($randn \in N(0,1)$, $N(0,1)$ 是标准正太分布)^[13], 而 ABSA 的最大和最小的变异因子分别设置为 0.9 和 0.1. 另外, GA 和 DE 的参数分别设置如下: (1) 种群大小设置为 100; (2) 最大迭代次数设置为 100; (3) 交叉概率为 0.6; (4) 变异概率为 0.01.

3.3.2 自适应变异因子 *F* 取值分析

ABSA 的变异因子 *F* 取值如图 7 所示. 从图中可以看出, 在算法迭代初期变异因子 *F* 取值较大 (第一个迭代阶段, *F* 取值为最大设定值 0.9). 此时, ABSA 的种群中的个体以较大概率进行变异, ABSA 主要是在种群个体空间中进行全局寻优, 形成的新种群个体随机分布在个体空间之中且个体之间的关联性较小. 可以说, ABSA 在迭代初期的全局寻优能力非常强而局部寻优能力非常弱, 种群个体的相似性很小. 随着迭代过程的更迭, 变异因子 *F* 取值逐步减小, ABSA 种群中的优秀个体逐渐沉淀并且这些优秀个体以逐渐减小的概率进行变异. 此时, ABSA 逐渐由全局寻优转变为局部寻优. 新种群的形成主要是围绕已经获得的优秀个体进行局部寻优. 即随着迭代的进行, ABSA 的局部寻优能力逐渐加强而全局寻优能力逐渐减弱, 种群个体的相似性逐渐加强. 直到迭代进行到后期 (图 7 中的迭代次数达到 90 次左右), 变异因子 *F* 取值趋于稳定 (最小设定值 0.1), ABSA 已经由全局寻优过渡到局部寻优. 此时, ABSA 全局寻优能力非常弱而局部寻优能力非常强. 所以, ABSA 的自适应变异因子取值策略能够平衡算法的全局寻优能力和局部寻优能力.

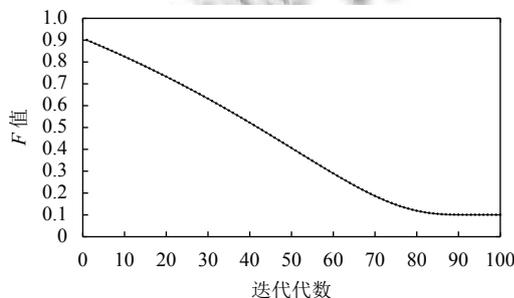


图 7 ABSA 的变异因子 *F* 取值曲线

3.4 参数设置与分析

对于本文所使用的 5 个预测模型中的每个预测模

型, 在训练阶段由于 ESN 储备池大小不同而会设计 3 个候选模型结构, 并且根据验证集的结果选择一个最优模型结构作为该预测模型的最优结构. 表 1 给出所有 5 种预测模型的最优模型结构. 5 种预测模型在两个时间序列数据集上的性能分别如图 8 (NARMA 模型) 和图 9 (Mackey-Glass 混沌系统) 所示, 其中“Actual”代表时间序列数据集的期望值. 在以 RMSE 为误差度量评价标准的前提下, 5 种预测模型在两个时间序列数据集上的误差如表 2 所示 (最优的预测结果用黑体特别标注).

表 1 5 种预测模型的最优模型结构

| 模型 | NARMA | Mackey-Glass |
|----------|--------|--------------|
| ESN | 2×50×1 | 2×50×1 |
| GA_ESN | 2×20×1 | 2×20×1 |
| DE_ESN | 2×50×1 | 2×20×1 |
| BSA_ESN | 2×50×1 | 2×30×1 |
| ABSA_ESN | 2×50×1 | 2×50×1 |

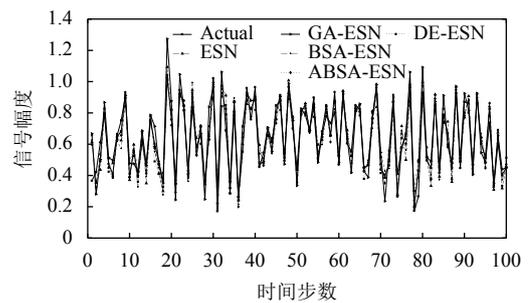


图 8 NARMA 上 5 种模型的预测值和期望值

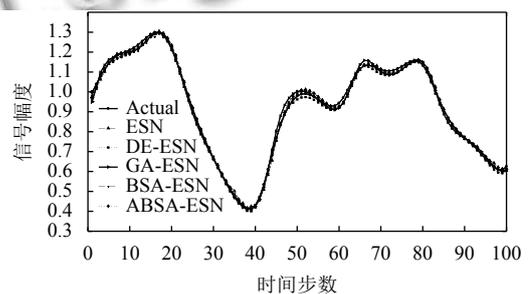


图 9 Mackey-Glass 上 5 种模型的预测值和期望值

表 2 5 种预测模型的最优结果 (以 RMSE 为度量方式)

| 模型 | NARMA | Mackey-Glass |
|----------|---------------|---------------|
| ESN | 0.0210 | 0.0198 |
| GA_ESN | 0.0282 | 0.0202 |
| DE_ESN | 0.0194 | 0.0175 |
| BSA_ESN | 0.0192 | 0.0162 |
| ABSA_ESN | 0.0176 | 0.0126 |

为了评估两个相比较模型之间的性能差异, 相对误差将被采用作为度量方法. 相对误差是一种流行的用于比较两个方法之间的性能的度量方法^[20]. 对于两种方法 A (基本方法) 和 B (对比方法), 其误差分别记为 δ_A 和 δ_B , 则相对误差为 $\frac{\delta_B - \delta_A}{\delta_A}$. 相对误差小于 0, 说明对比方法 B 比基本方法 A 性能更优. 另外, 相对误差的平均指两种比较方法在两个或两个以上数据集上的相对误差的平均值. 相对误差的均值小于 0 说明对比方法 B 相对于基本方法 A 在所有数据集上预测性能有改进.

3.4.1 优化的 ESN 与标准的 ESN

在本文中, 为了验证所提出的使用进化算法优化的 ESN (GA_ESN、DE_ESN、BSA_ESN 和 ABSA_ESN) 与标准的 ESN 之间的预测性能差异, 将使用标准的 ESN 作为基本方法, 使用 GA_ESN、DE_ESN、BSA_ESN 和 ABSA_ESN 作为对比方法进行预测性能比较. 表 3 给出了两两比较模型之间的相对误差.

表 3 两种比较模型之间相对误差 (以 ESN 为基本方法)

| 对比模型 | 数据集 | | MEAN |
|--------------|---------|--------------|---------|
| | NARMA | Mackey-Glass | |
| GA_ESN&ESN | 0.3429 | 0.0202 | 0.1815 |
| DE_ESN&ESN | -0.0762 | -0.1162 | -0.0962 |
| BSA_ESN&ESN | -0.0857 | -0.1338 | -0.1338 |
| ABSA_ESN&ESN | -0.1619 | -0.2628 | -0.2628 |

从表 3 可以得出: 标准的 ESN 预测模型无论是在单个数据集还是在两个数据集作为整体的性能评估中, 其性能都只比 GA_ESN 优越而比 DE_ESN、BSA_ESN 和 ABSA_ESN 等预测模型差. 具体而言, GA_ESN 相对于 ESN 的性能单独在 NARMA 和 Mackey-Glass 上分别下降 34.29% 和 2.02%, 在两个数据集的整体性能上平均下降 18.15% (如表 3 最后一列“MEAN”所示). 而 DE_ESN、BSA_ESN 和 ABSA_ESN 相对于 ESN 单独在两个数据集上的性能分别提升 7.62%/11.62%、8.57%/18.18% 和 16.19%/36.36%, 在两个数据集的整体性能上平均提升 9.62%、13.38% 和 26.28% (如表 3 最后一列“MEAN”所示). 总体来说, 使用进化算法对 ESN 的输出连接权值矩阵 W^{out} 进行优化能够提升 ESN 的预测性能.

3.4.2 ABSA_ESN 与其他 4 种模型

由 3.4.1 小节可知, 使用进化算法对 ESN 的输出连接权值矩阵 W^{out} 进行优化相对于标准的 ESN 具有一定的性能优势. 本节将具体讨论本文所提出的 ABSA-

ESN 相对于标准的 ESN 和其他 3 种优化的 ESN 的优势体现. 具体而言, 将 ABSA-ESN 作为对比方法, 而 ESN、GA-ESN、DE-ESN 和 BSA-ESN 等分别作为基本方法, 然后比较它们两两之间的模型性能. 表 4 列出了两两相比较模型之间的相对误差.

表 4 两种比较模型之间相对误差
(以 ABSA_ESN 为对比方法)

| 对比模型 | 数据集 | | MEAN |
|------------------|---------|--------------|---------|
| | NARMA | Mackey-Glass | |
| ABSA_ESN&ESN | -0.1619 | -0.3636 | 0.1815 |
| ABSA_ESN&GA_ESN | -0.3759 | -0.3762 | -0.0962 |
| ABSA_ESN&DE_ESN | -0.0923 | -0.2800 | -0.1338 |
| ABSA_ESN&BSA_ESN | -0.0833 | -0.2222 | -0.2628 |

从表 4 可以得出: ABSA_ESN 预测模型比所提出的其他 4 种预测模型 ESN、GA_ESN、DE_ESN 和 BSA_ESN 的预测精度大幅提升, 具体体现在以下两个方面:

(1) 当单独关注 NARMA 或者 Mackey-Glass 数据集时, ABSA_ESN 能够比 ESN、GA_ESN、DE_ESN 和 BSA_ESN 等模型在两个数据集上都能获得更好的预测结果, 主要表现在 ABSA_ESN 在两个数据集上的预测精度提升比例分别为 16.19%/36.36%、37.59%/37.62%、9.23%/28% 和 8.33%/22.22%.

(2) 当关注所有 5 种预测模型在两个数据集上的总体性能时, ABSA_ESN 模型比其他 4 个模型在预测精度上也有大幅提升. ABSA_ESN 模型相对于 ESN、GA_ESN、DE_ESN 和 BSA_ESN 等模型的平均相对误差提升比例分别为 26.28%、37.61%、18.64% 和 15.28%. 这个信息如表 4 中最后一列“MEAN”所示.

由此可见, 本文提出的使用自适应回溯搜索算法优化的回声状态网络模型 (ABSA_ESN) 能够比未优化的回声状态网络模型 (ESN) 和采用其他进化算法优化的回声状态网络模型 (GA_ESN、DE_ESN 和 BSA_ESN) 在两个混沌时间序列数据集上获得更好的预测精度.

4 结束语

本文提出使用自适应回溯搜索算法优化回声状态网络的时间序列预测模型 (ABSA_ESN), 以解决标准回声状态网络中使用的线性方法求解输出连接权值矩阵容易陷入过拟合的问题, 从而提升回声状态网络的预测性能. 为了验证本文所提出 ABSA_ESN 在时间序

列预测问题上的可行性和有效性, 本文还设计了4种预测模型用于比较, 它们分别是标准的回声状态网络(ESN)、遗传算法的优化回声状态网络(GA_ESN)、差分进化算法优化的回声状态网络(DE_ESN)和标准回溯搜索算法优化的回声状态网络(BSA_ESN). 通过对实验结果进行分析可以得出如下结论: (1) 使用进化算法优化的ESN(GA_ESN、DE_ESN、BSA_ESN和ABSA_ESN)相对于标准的ESN在预测性能上具有一定的优势; (2) ABSA_ESN相对于4种对比模型能够获得更好的预测精度. 值得注意的是本文使用的是基本结构的回声状态网络, 其自身性能可能会受到一定的限制. 在未来, 可以研究使用其他更加有效的进化算法优化具有复杂结构的回声状态网络的输出连接权值矩阵, 进一步提升回声状态网络的性能.

参考文献

- 1 Lee Giles C, Lawrence S, Tsoi AC. Noisy time series prediction using recurrent neural networks and grammatical inference. *Machine Learning*, 2001, 44(1-2): 161-183.
- 2 Zhang GQ, Patuwo BE, Hu MY. Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, 1998, 14(1): 35-62. [doi: [10.1016/S0169-2070\(97\)00044-7](https://doi.org/10.1016/S0169-2070(97)00044-7)]
- 3 Wang L, Wang ZG, Qu H, *et al.* Optimal forecast combination based on neural networks for time series forecasting. *Applied Soft Computing*, 2018, 66: 1-17. [doi: [10.1016/j.asoc.2018.02.004](https://doi.org/10.1016/j.asoc.2018.02.004)]
- 4 Jaeger H, Haas H. Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication. *Science*, 2004, 304(5667): 78-80. [doi: [10.1126/science.1091277](https://doi.org/10.1126/science.1091277)]
- 5 田中大, 高宪文, 李树江, 等. 遗传算法优化回声状态网络的网络流量预测. *计算机研究与发展*, 2015, 52(5): 1137-1145. [doi: [10.7544/issn1000-1239.2015.20131757](https://doi.org/10.7544/issn1000-1239.2015.20131757)]
- 6 王华秋, 王斌, 龙建武. 回声状态网络在变风量空调内模控制中的应用. *重庆理工大学学报(自然科学)*, 2017, 31(6): 120-126, 153.
- 7 杜晟磊. 基于回声状态网络的非线性卫星信道盲均衡. *科技创新与应用*, 2017, (30): 19-21.
- 8 宗宸生, 郑焕霞, 王林山. 改进粒子群优化BP神经网络粮食产量预测模型. *计算机系统应用*, 2018, 27(12): 204-209. [doi: [10.15888/j.cnki.csa.006651](https://doi.org/10.15888/j.cnki.csa.006651)]
- 9 朱海龙, 陶晶, 俞凯, 等. 基于GA-BP神经网络的胎儿体重预测分析. *计算机系统应用*, 2018, 27(3): 162-167. [doi: [10.15888/j.cnki.csa.006252](https://doi.org/10.15888/j.cnki.csa.006252)]
- 10 Chouikhi N, Ammar B, Rokbani N, *et al.* PSO-based analysis of Echo State Network parameters for time series forecasting. *Applied Soft Computing*, 2017, 55: 211-225. [doi: [10.1016/j.asoc.2017.01.049](https://doi.org/10.1016/j.asoc.2017.01.049)]
- 11 Wang HS, Yan XF. Optimizing the echo state network with a binary particle swarm optimization algorithm. *Knowledge-Based Systems*, 2015, 86: 182-193. [doi: [10.1016/j.knsys.2015.06.003](https://doi.org/10.1016/j.knsys.2015.06.003)]
- 12 Jaeger H. The "echo state" approach to analysing and training recurrent neural networks-with an erratum note. Bonn: German National Research Center for Information Technology GMD Technical Report, 2001.
- 13 Civicioglu P. Backtracking search optimization algorithm for numerical optimization problems. *Applied Mathematics and Computation*, 2013, 219(15): 8121-8144. [doi: [10.1016/j.amc.2013.02.017](https://doi.org/10.1016/j.amc.2013.02.017)]
- 14 Wang S, Da XY, Li MD, *et al.* Adaptive backtracking search optimization algorithm with pattern search for numerical optimization. *Journal of Systems Engineering and Electronics*, 2016, 27(2): 395-406. [doi: [10.1109/JSEE.2016.00041](https://doi.org/10.1109/JSEE.2016.00041)]
- 15 Madasu SD, Kumar MLSS, Singh AK. Comparable investigation of backtracking search algorithm in automatic generation control for two area reheat interconnected thermal power system. *Applied Soft Computing*, 2017, 55: 197-210. [doi: [10.1016/j.asoc.2017.01.018](https://doi.org/10.1016/j.asoc.2017.01.018)]
- 16 Chaib AE, Boucekara HREH, Mehasni R, *et al.* Optimal power flow with emission and non-smooth cost functions using backtracking search optimization algorithm. *International Journal of Electrical Power & Energy Systems*, 2016, 81: 64-77.
- 17 庞世伟, 于开平, 邹经湘. 基于时变NARMA模型的非线性时变系统辨识. *工程力学*, 2006, 23(12): 25-29. [doi: [10.3969/j.issn.1000-4750.2006.12.005](https://doi.org/10.3969/j.issn.1000-4750.2006.12.005)]
- 18 MacKey MC, Glass L. Oscillation and chaos in physiological control systems. *Science*, 1977, 197(4300): 287-289. [doi: [10.1126/science.267326](https://doi.org/10.1126/science.267326)]
- 19 Farmer JD. Chaotic attractors of an infinite-dimensional dynamical system. *Physics D: Nonlinear Phenomena*, 1982, 4(3): 366-393. [doi: [10.1016/0167-2789\(82\)90042-2](https://doi.org/10.1016/0167-2789(82)90042-2)]
- 20 Bauer E, Kohavi R. An empirical comparison of voting classification algorithms: Bagging, boosting, and variants. *Machine Learning*, 1999, 36(1-2): 105-139.