

一种混合优化的结构自适应极限过程神经网络及应用

刘志刚^{1†}, 许少华², 杜娟¹, 肖佃师³

(1. 东北石油大学 计算机与信息技术学院, 黑龙江 大庆 163318; 2. 山东科技大学 计算机科学与工程学院, 山东 青岛 266590; 3. 中国石油大学(华东) 非常规油气与新能源研究院, 山东 青岛 266580)

摘要: 为解决过程神经网络的隐层结构和训练速度问题,在极限学习机的基础上,提出一种混合优化的结构自适应极限过程神经网络. 首先,采用在隐层中逐次增加过程神经元节点直至满足输出误差的方式完成模型结构自适应;然后,为消除冗余节点,提出对新增临时节点输出实施Gram-Schmidt正交化完成相关性判别;最后,构建一种量子衍生布谷鸟算法,对新增节点输入权函数正交基展开系数实施寻优. 仿真实验以Mackey-Glass和页岩油TOC预测为例,通过对比分析验证所提出方法的有效性,仿真结果表明所得模型的逼近效率和训练速度有明显提高.

关键词: 过程神经网络; 极限学习; Gram-Schmidt正交化; 布谷鸟算法; 自适应结构

中图分类号: TP183

文献标志码: A

Structure-adapted extreme process neural network and application based on hybrid optimization

LIU Zhi-gang^{1†}, XU Shao-hua², DU Juan¹, XIAO Dian-shi³

(1. School of Computer and Information Technology, Northeast Petroleum University, Daqing 163318, China; 2. College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China; 3. Institute of Unconventional Oil & Gas and New Energy, China University of Petroleum, Qingdao 266580, China)

Abstract: For the hidden layer structure and learning speed of the process neural network, a structure-adapted extreme process neural network is proposed based on the extreme learning machine. Firstly, the adaptive model structure is achieved by incrementally adding the process neuron until output error meets accuracy requirements. Then, in order to eliminate the redundant nodes, a correlation discrimination method is proposed, in which Gram-Schmidt orthogonalization is executed to new node output. Finally, a quantum-inspired cuckoo search algorithm is proposed, which is used to optimize the parameters generated by input weight function orthogonal basis expansion. The effectiveness is validated by the comparative analysis based on Mackey-Glass and shale oil TOC prediction. The simulation result shows that the approximation efficiency and training speed of the proposed model are obvious improved.

Keywords: process neural network; extreme learning; Gram-Schmidt orthogonalization; Cuckoo algorithm; adaptive structure

0 引言

现实世界中,许多系统的输入信号都是与时间有关的函数过程. 为此,文献[1]提出了过程神经网络(PNN),改进了神经元的聚合机制,实现了连续输入信号的处理. 截至目前,在正交基展开PNN^[2]的基础上相继有并联^[3]、反馈^[4]、离散^[5]等模型被提出. 相对于普通神经网络,这些模型对非线性时变系统具有更好的适用性. Haykin在文献[6]中证明具有最小结

构的神经网络具有更好的逼近能力,隐层节点少易导致训练误差过大,节点多易导致模型泛化能力较差,因此自适应结构PNN具有重要的研究意义. 但是,因为PNN正交基展开后结构较为复杂,此方面的研究还很少. 文献[7]提出了使用粒子群算法优化PNN参数和模型结构,取得了一定的应用效果,但仍存在以下不足: 1) 每个粒子代表模型参数的一组可行解,训练过程等价于对多个PNN同时学习,训练非常耗

收稿日期: 2017-04-10; 修回日期: 2017-07-06.

基金项目: 国家自然科学基金项目(61170132, 41330313, 41602141); 黑龙江省自然科学基金项目(F2015021).

责任编辑: 王燕舞.

作者简介: 刘志刚(1979—),男,副教授,博士,从事极限学习机的研究; 许少华(1962—),男,教授,博士生导师,从事过程神经网络、智能计算等研究.

[†]通讯作者. E-mail: dqpilzg@163.com

时;2) PNN正交基展开后参数成倍增长,对所有参数寻优属于高维优化问题,算法极易陷入局部最小,影响网络逼近效率.

极限学习机(ELM)是Huang在2004年提出的一种前馈神经网络快速学习算法^[8],理论分析和实验已证明,该算法可极大提高神经网络的学习速度(通常10倍以上)^[9-10],截至目前,以ELM为研究主题的被SCI收录的论文共2606篇,其中372篇为IEEE文献,ELM已成为国际研究热点^[11-13].根据ELM理论和连续PNN、离散PNN自身特性,文献[14-15]分别提出了FE-PNN、QE-DPNN,较大地提高了PNN学习速度.

本文在前期研究工作的基础上,为解决PNN隐层结构和训练速度问题,提出一种混合优化的结构自适应极限过程神经网络(OSE-PNN).首先,在模型隐层中,固定已有节点的权函数正交基展开系数,根据输出误差逐次增加新节点,实现网络结构的自适应增长;然后,提出使用Gram-Schmidt正交化完成隐层新增节点与现有节点的相关性分析;最后,提出一种量子衍生布谷鸟算法完成新增节点权值参数优化.仿真实验中,通过Mackey-Glass时间序列预测验证其有效性,并实际应用于页岩油总有机碳含量(TOC)的预测中.

1 过程神经网络PNN

过程神经元内部包括时空聚合两个算子,输入和隐层间的权值是与时间有关的函数,具体见文献[1-2].相对于隐层, $w_{ij}(t)$ 为输入权值, v_j 为输出权值,PNN的输入输出映射关系为

$$y = g\left(\sum_{j=1}^m v_j f\left(\sum_{i=1}^n \int_0^T w_{ij}(t)x_i(t)dt - \theta_j\right)\right). \quad (1)$$

为简化时域聚合计算,在输入函数空间中选取一组正交基 $b_1(t), b_2(t), \dots, b_l(t)$,对输入 $x_i(t)$ 和权函数 $w_{ij}(t)$ 在满足拟合精度条件下展开,即

$$x_i(t) = \sum_{l=1}^L a_i^{(l)} b_l(t), \quad w_{ij}(t) = \sum_{l=1}^L w_{ij}^{(l)} b_l(t). \quad (2)$$

根据正交基的正交特性,式(1)整理为

$$y = g\left(\sum_{j=1}^m v_j f\left(\sum_{i=1}^n \sum_{l=1}^L w_{ij}^{(l)} a_i^{(l)} - \theta_j\right)\right). \quad (3)$$

2 OSE-PNN模型和算法

2.1 隐层新增节点的相关性分析

PNN结构包括:1)输入到隐层、隐层到输出的线性加权;2)隐层非线性映射.最优网络结构的PNN隐层所有神经元对训练样本的输出应线性无关,输出

为隐层输出向量的线性组合^[6].记OSE-PNN的样本集为 $(x_{k1}(t), \dots, x_{kn}(t); d_k)_{k=1}^K$,隐层节点 m 个.由式(3)可知,第 j 个隐层节点对样本集的输出向量 h_j 为

$$h_j = (f(a_{i,l}^{(1)}, w_{ij,l}, \theta_j), \dots, f(a_{i,l}^{(K)}, w_{ij,l}, \theta_j)), \quad (4)$$

其中 $f(a_{i,l}^{(k)}, w_{ij,l}, \theta_m) = f\left(\sum_{i=1}^n \sum_{l=1}^L a_{i,l}^{(k)} w_{ij,l} - \theta_j\right)$.

由此可知,具有 m 个隐层节点OSE-PNN的隐层输出矩阵为

$$H_m = (h_1, h_2, \dots, h_m)^T. \quad (5)$$

隐层新增节点与相邻节点的相关性如下.

引理1^[16] 设 $\alpha_1, \alpha_2, \dots, \alpha_s$ 是一组线性无关的向量,则存在一组正交向量 $\beta_1, \beta_2, \dots, \beta_s$,使得 $\alpha_1, \alpha_2, \dots, \alpha_i$ 与 $\beta_1, \beta_2, \dots, \beta_i (i = 1, 2, \dots, s)$ 等价.

定理1 记OSE-PNN隐层新增节点为 $m+1$,该节点对训练样本输出向量为 h_{m+1} ,隐层各节点输出的Gram-Schmidt正交向量为 $\{s_1, s_2, \dots, s_{m+1}\}$,若 $\|s_{m+1}\|_2 = 0$,则节点 $m+1$ 与相邻节点 m 线性无关.

证明 OSE-PNN隐层中新增节点输出的Gram-Schmidt正交向量^[16]为

$$s_{m+1} = h_{m+1} - \frac{[s_1, h_{m+1}]}{[s_1, s_1]} s_1 - \dots - \frac{[s_m, h_{m+1}]}{[s_m, s_m]} s_m. \quad (6)$$

为表述方便,记 $\mathfrak{S}(\cdot)$ 表示新增节点与相邻节点的输出向量线性相关, $\lambda(\cdot)$ 表示线性无关.

1) 由引理1可知,当 $\lambda(h_m, h_{m+1})$ 成立时, $\lambda(s_m, s_{m+1})$ 成立,即 $\text{span}\{h_1, h_2, \dots, h_n\} = \text{span}\{s_1, s_2, \dots, s_n\}$,其中 $\text{span}(s_1, s_2, \dots, s_n)$ 表示由 (s_1, s_2, \dots, s_n) 张成的向量空间;

2) 当 $\mathfrak{S}(h_m, h_{m+1})$ 成立时,记 $h_{m+1} = \alpha h_m$,则有

$$\begin{aligned} s_{m+1} &= h_{m+1} - \sum_{i=1}^m \frac{[s_i, h_{m+1}]}{[s_i, s_i]} s_i = \\ &= \alpha h_m - \sum_{i=1}^m \frac{[s_i, \alpha h_m]}{[s_i, s_i]} s_i = \\ &= \alpha \left(s_m - \frac{[s_m, h_m]}{[s_m, s_m]} s_m \right) = (0, 0, \dots, 0)^T. \end{aligned}$$

因此 $\|s_{m+1}\|_2 = 0$. \square

2.2 隐层新增节点的输出权值计算

OSE-PNN逐次添加新节点后,判断新增节点与其他节点的相关性.若线性相关,则丢弃,并重新生成新节点加入到网络中,重复直到达到最小输出误差.由于隐层所有节点的输出均线性无关,消除线性相关的冗余节点,提高模型结构紧凑性.此外,

根据增量极限学习机 (EI-ELM)^[13] 理论, 固定 OSE-PNN 现有隐层节点的输入权函数的基展开系数和输出权值, 在新增节点 $w_{i,m+1}(t)$ 基展开系数 ($w_{i,m+1}^{(1)}, \dots, w_{i,m+1}^{(L)}$) 确定时, 其输出权值计算如下:

$$v_{m+1} = \frac{[e_m, h_{m+1}]}{\|h_{m+1}\|_2}. \quad (7)$$

其中: e_m 为具有 m 个隐层节点的 PNN 的输出误差, 计算方式如下:

$$e_m = e_{m-1} - v_m h_m. \quad (8)$$

2.3 隐层新增节点的输入权值优化

隐层增加单个节点后, EI-ELM 权值参数仅增加 n 个, 而 PNN 受正交基展开机制影响, 参数增加数量为 $L \times n$, 因此 EI-ELM 新增节点权值随机赋值的方式不适用于 PNN, 很难在较少迭代次数内获取较优参数. 标准布谷鸟搜索算法 (SCS) 是 2009 年由 Yang 等提出的一种新型元启发算法^[17], 已证明其性能接近或略高于粒子群 (PSO)、差分进化 (DE) 和蜂群 (ABC) 算法. 本文为加快 OSE-PNN 输出误差减小速度, 引入 Bloch 球面坐标的量子编码机制^[18], 提出一种新的量子衍生布谷鸟算法 (QICS), 用于 OSE-PNN 输入权值优化. QICS 算法策略如下:

1) 个体编码: 种群个体的量子比特采用 Bloch 球面坐标, 其三链编码结构如下:

$$p_i = \begin{bmatrix} \cos \phi_{i1} \sin \theta_{i1} & \cos \phi_{i2} \sin \theta_{i2} & \cdots & \cos \phi_{iD} \sin \theta_{iD} \\ \sin \phi_{i1} \sin \theta_{i1} & \sin \phi_{i2} \sin \theta_{i2} & \cdots & \sin \phi_{iD} \sin \theta_{iD} \\ \cos \theta_{i1} & \cos \theta_{i2} & \cdots & \cos \theta_{iD} \end{bmatrix}. \quad (9)$$

其中: $\theta_{ij} = 2\pi \times r$, $\phi_{ij} = \pi \times r$, $r \in (0, 1)$ 是随机数; D 是问题空间维数; $i = 1, 2, \dots, N$, N 是种群规模.

2) Bloch 球面纬线搜索: 由 Bloch 球面几何关系可知, 沿球面纬线搜索时, $\Delta\theta$ 确定个体的进化过程, $\Delta\phi$ 丰富种群多样性, 则每个量子位向最优个体 p_{best} 逼近时的量子旋转门幅角分别为

$$\begin{cases} \Delta\theta_{ij}(t+1) = \alpha_0 \frac{\eta \times \mu}{|v|^{1/\beta}} (\theta_{ij}(t) - \theta_{\text{best},j}(t)), \\ \Delta\phi_{ij}(t+1) = \text{rand}(0, 1) \Delta\theta_{ij}(t+1). \end{cases} \quad (10)$$

3) 个体贪婪择优: 个体 $p_i(t)$ 纬线搜索更新后记为 $\hat{p}_i(t+1)$, 按贪婪择优策略选择 $p_i(t)$ 、 $\hat{p}_i(t+1)$ 中较优的个体作为 $p_i(t+1)$ 进入下一代种群, 其中每个个体的适应度由三链中的适应度最优链表示, 即

$$p_i(t+1) = \max(f(p_i(t)), f(\hat{p}_i(t+1))). \quad (11)$$

4) 个体变异策略: 记发现个体 $p_i(t)$, 随机选择个体 $p_m(t)$ 和 $p_n(t)$, 在 Biased 随机走动交叉变异过程中, 变异后量子个体为 $\hat{p}_i(t)$. 对 $p_i(t)$ 和 $\hat{p}_i(t)$ 贪婪择优, 其中 $\hat{p}_i(t)$ 的 Bloch 球面坐标的角度如下:

$$\begin{cases} \hat{\theta}_{ij}(t) = \theta_{ij}(t) + r_1(\theta_{mj}(t) - \theta_{nj}(t)), \\ \hat{\phi}_{ij}(t) = \phi_{ij}(t) + r_2(\phi_{mj}(t) - \phi_{nj}(t)). \end{cases} \quad (12)$$

其中: r_1 和 r_2 是缩放因子, 是 $(0, 1)$ 之间的随机数.

定理 2 QICS 沿 Bloch 球面实施纬线搜索更新过程中只需调整 θ , 且对全局最优解中每一维分量, 球面上有无数个点坐标表示的量子比特与之对应.

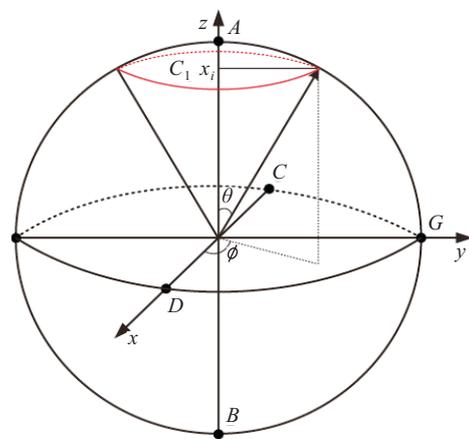


图 1 全局最优解与 Bloch 球面坐标的对应关系

证明 1) 记全局最优解 $X = (X_1, X_2, \dots, X_D)$ 从问题空间 Ω 映射到 $I^D = [-1, 1]^D$ 后为 $p = (x_1, x_2, \dots, x_D)$. 设当前最优解中的某一维 x_i 对应的 Bloch 球的转角分别为 θ_i 和 ϕ_i , 以 $\sin \theta_i$ 为半径, 在 Bloch 球面上做垂直于 z 轴的纬线圆周 C_1 , 其上任意一点坐标为 $(\cos \phi \sin \theta_i, \sin \phi \sin \theta_i, \cos \theta_i)$, 并且 C_1 所有点的 z 坐标都为 x_i , 此时圆周 C_1 上有无数个坐标对应的量子比特与 x_i 对应. 2) 由图 1 给出的 Bloch 球面几何关系可知, 个体沿球面做纬线搜索即可完成最优解寻优, 并且最优解 x_i 仅与 θ_i 大小有关, 与 ϕ 无关, ϕ 可增加量子比特多样性. \square

由定理 2 可知, QICS 沿 Bloch 球面纬线搜索, 问题最优解对应的球面个体在数量上得到极大扩充, 大大提高了获得问题最优解的概率, 加快了算法收敛速度.

2.4 OSE-PNN 的算法步骤

Step 1: 记 OSE-PNN 当前隐层节点数 $m = 1$, 训练误差为 η , 最大隐层节点数为 M , QICS 种群规模为 N , 最大进化次数为 G .

Step 2: 按式 (6) 对当前隐层节点输出向量实施

Gram-Schmidt正交化,得到 s_m .

Step 3: 增加隐层节点,设节点输入权值为 $w_{i,m,k}^{(\delta)}$ 、隐层阈值为 $\theta_{m,k}^{(\delta)}$, $k = 1, 2, \dots, N, l = 1, 2, \dots, L$,并按式(9)初始化种群,记当前优化次数 $\delta = 0$.

Step 4: 利用QICS算法对新增节点输入权值寻优.

1) 个体解空间变换后,按式(4)计算每条量子个体对应临时新增节点的输出向量 $h_{m,k}^{(\delta)}$.

2) 对 $h_{m,k}^{(\delta)}$ 实施Gram-Schmidt正交化后为 $s_{m,k}^{(\delta)}$.若 $\|s_{m,k}^{(\delta)}\| < \varepsilon$ 且 $\text{rank}(H_{m+1}) < m + 1$,则转3),否则转4),其中 ε 是一个很小的正数.

3) 丢弃第 k 条个体,并产生一条新个体.

4) 按式(7)计算每条量子个体对应的临时节点的输出权值,取式(8)倒数为个体适应度,按式(10)~(12)实施进化操作.

5) $\delta \leftarrow \delta + 1$,若 $\delta < G$,则转1),否则QICS进化过程结束,转Step 5.

Step 5: 记最优个体为 γ ,解空间变换后的数值为新增节点的权值参数和阈值,将该节点添加到OSE-PNN中,并记 $e_m = e_m^{(\gamma)}$.

Step 6: 若 $e_m < \eta$ 且 $m \leq M$,则转Step 3,否则结束.

3 Mackey-Glass时间序列预测

3.1 仿真实验

本文利用经典的Mackey-Glass时间序列预测验证所提方法的有效性,对比模型包括文献[2]正交基展开过程神经网络(OBPNN)、文献[3]双并联过程神经网络(DPPNN)、文献[7]粒子群优化结构的PSO-PNN. Mackey-Glass的时滞差分方程如下:

$$\frac{dx}{dt} = \frac{\alpha x(t-\tau)}{1+x(t-\tau)^{10}} - \beta x(t).$$

样本集构造过程请参照文献[15],正交基函数均为勒让德基函数,基展开项数 $L = 12$.其中OSE-PNN的初始隐层过程神经元节点 $m = 1$,训练误差精度 $\eta = 0.005$,权系数优化的QICS种群规模 $G = 20$,个体编码长度为12位,最大优化次数 $K = 50$.为方便对比,参照文献[3],OBPNN、DPPNN的学习速度参数均设置为 $\alpha = \beta = \gamma = 0.5$,惯性系数 $\eta = 0.1$.

对OSE-PNN的训练和预测能力进行验证,对比指标包括训练耗时Time、训练收敛率TCR、预测的相关系数 R 、平均绝对误差MAE、平均相对误差MRE、均方根误差RMSE.其中: R 越接近1、均方误差越接近0,预测效果越好.针对4种模型和100组训练样本,

各进行20次仿真实验,训练预测情况对比如表1所示.

表1 4种过程神经网络的实验结果对比

Model	Time/s	TCR	R	MAE	MRE	RMSE
OBPNN	73.46	0.60	0.969 3	0.046 5	0.133 5	0.058 2
DPPNN	134.05	0.65	0.975 6	0.126 2	0.107 3	0.050 7
PSO-PNN	263.36	0.85	0.977 2	0.040 2	0.058 7	0.038 9
OSE-PNN	28.59	0.90	0.997 8	0.011 7	0.045 0	0.014 2

从实验结果中可明显看出,OSE-PNN和PSO-PNN的预测能力优于其他两种PNN模型,其中OSE-PNN效果最为理想,且训练时间相对于PSO-PNN有明显缩短(9倍左右).此外,图2给出了OSE-PNN和效果次优的PSO-PNN的预测情况对比.

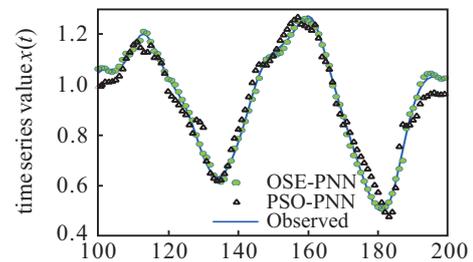


图2 Mackey-Glass预测结果与实际结果对比

3.2 实验结果分析

1) 预测能力对比分析.4种模型从高到低排列依次为:OSE-PNN、PSO-PNN、DPPNN、OBPNN.原因在于:OSE-PNN和PSO-PNN的网络结构均具有自适应机制,利用结构紧凑性提高模型的泛化能力,这一点是其他两种模型所不具备的.OSE-PNN优于PSO-PNN的原因在于:i) OSE-PNN通过正交化快速判别新增节点与相邻节点的相关性,由定理1可知,网络结构紧凑性相对于PSO-PNN更好;ii)由定理2可知,QICS的寻优能力较强,进一步提高了获得新增节点较优权值参数的可能性;iii) OSE-PNN将输出权值转化为最小二乘问题,误差向量与隐层节点输出向量正交,即 $e_m \perp \text{span}(h_1, h_2, \dots, h_m)$,此时输出权值 v 为最小范数二乘解,减小了模型的映射误差.

2) 训练时间对比分析.4种模型从快到慢依次为:OSE-PNN、DPPNN、OBPNN、PSO-PNN.原因在于:PSO-PNN对模型所有参数编码,每次进化需要计算多组可行解误差精度,所以训练时间最长.而OSE-PNN仅计算新增节点输出权值,且进化过程仅为消除隐层输出矩阵的病态性,种群评估次数和个体维数均明显减少,因此OSE-PNN训练时间最短.两者训练时间主要集中在权值参数优化过程,时间复杂度分别为

$$T_{\text{OSE-PNN}} = G_{\text{OSE-PNN}} \times N \times O(3 \times n \times L),$$

$$T_{\text{PSO-PNN}} = G_{\text{PSO-PNN}} \times N \times O(n \times m \times L).$$

其中: $G_{\text{OSE-PNN}} \ll G_{\text{PSO-PNN}}, m > 3$.

4 页岩TOC预测

近年页岩油气成为世界能源焦点,TOC预测是其重点研究内容^[19].由于测井曲线随深度、时间不断变化,其连续性特征能有效反映TOC变化,该类问题是非线性时变系统的预测问题.本文以A区B井为例,测井资料包括自然伽马GR、自然电位SP、深测向电阻率RLLD、密度DEN、声波时差AC等.为提高TOC预测精度,选取所提OSE-PNN作为预测模型,同时与其他PNN对比,此外还包括文献[19]的RBF神经网络和传统 $\Delta \log R$ 法.

1) 相关性分析:利用神经网络预测TOC,若将所有测井曲线作为特征,则会提高模型复杂性,降低网络泛化能力.本文根据相关系数优选对TOC较为敏

感的测井曲线,其中DEN、CNL和AC与TOC的相关系数为0.5652、0.5706、0.4439,优于SP、GR和RLLD,因此选DEN、CNL和AC作为模型特征输入.

2) 实验对比:针对B井的训练样本共计672组,其中过程神经网络模型中每个样本均为3个长度为10的离散序列.每种神经网络模型分别独立运行20次,训练结束后立即对45组测试样本进行预测,其中预测情况对比见表2. OSE-PNN的TOC预测结果中,各种指标均优于其他模型.此外,OSE-PNN对B井TOC预测纵向分布情况如图3所示.

表2 TOC预测的实验对比

Model	R	MAE	MRE	RMSE
RBF	0.6496	1.2214	0.1764	1.3589
OBPNN	0.6930	1.0317	0.1573	1.2473
DPPNN	0.7429	1.1684	0.1648	1.0891
PSO-PNN	0.8205	0.9135	0.1485	0.9443
OSE-PNN	0.8677	0.7507	0.1212	0.8979
$\Delta \log R$	0.5137	1.6889	0.2776	2.1208

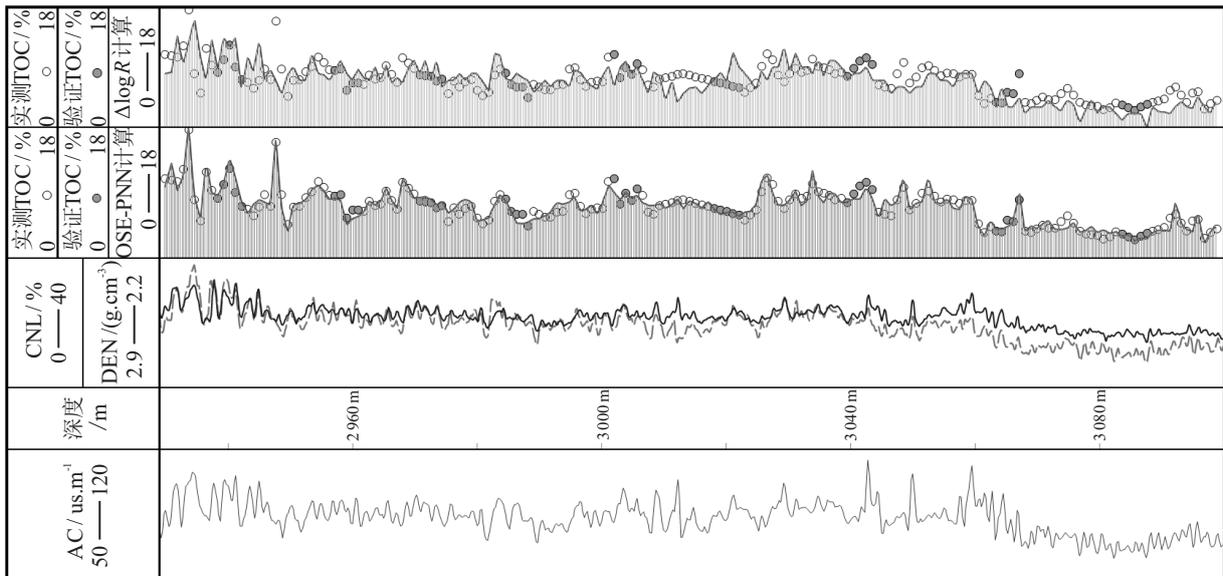


图3 B井的TOC预测垂向分布

该结果进一步验证了OSE-PNN在页岩TOC预测问题中的适用性,较好地完成了A区TOC垂向分布预测.

$\Delta \log R$ 法的前提是假定声波和孔隙的差异变化仅由有机质含量变化引起,本质上是假定TOC与电阻率、声波之间存在线性映射关系.但A区电阻率与TOC相关性较差,相关系数仅为0.203,其变化对TOC影响较小,使得 $\Delta \log R$ 法的测井响应线性假定与实际不符.其次,由于过程神经网络内部具有时空聚合算子,减小了时变信号映射时的细节丢失,而普通神经网络受同步瞬时输入限制,无法有效体现测井连续

信号对TOC的时间累积效应.因此,4种过程神经网络的逼近效率和预测精度高于RBF. OSE-PNN具备较好的模型结构紧凑机制和模型参数求解能力,使其TOC预测精度高于其他3种PNN模型.

5 结论

过程神经网络模型通过内部的时空聚合算子可以更加有效地完成时变过程信号的映射逼近问题,但由于其自身具有复杂性,在模型隐层规模上一直缺少自适应结构方面的研究,从而影响了模型的非线性映射逼近效率,成为走向实际工程应用的瓶颈问题.本文通过引入极限学习理论,在固定已有节点权值的情

况下,对于新增过程神经元节点分别从节点相关性分析、节点的输入权函数的正交基展开系数、节点输出权值3个方面,建立与自适应机制相应的模型学习和参数求解方法.其中相关性分析消除模型结构中的冗余节点;节点输出权值根据极限学习理论计算,满足最小二乘理论;节点输入权值参数利用量子布谷鸟算法完成寻优,加速了模型的训练收敛过程,从而为过程神经网络结构自适应研究提供了新思路.

参考文献(References)

- [1] He X G, Liang J Z. Some theoretical issues on procedure neural networks[J]. *Engineering Science*, 2000, 2(12): 40-44.
- [2] 许少华, 何新贵. 基于函数正交基展开的过程神经网络学习算法[J]. *计算机学报*, 2004, 27(5): 645-650. (Xu S H, He X G. Learning algorithms of process neural networks based on orthogonal function basis expansion[J]. *Chinese J of Computers*, 2004, 27(5): 645-650.)
- [3] Ding G, Zhong S. Time series prediction by parallel feedforward process neural network with time-varied input and output functions[J]. *Neural Network World*, 2005, 15(2): 137-147.
- [4] Luan S G, Zhong S S, Li Y. Hybrid recurrent process neural network for aero engine condition monitoring[J]. *Neural Network World J*, 2008, 2(18): 133-145.
- [5] Li P C, Shi G Y. Numerical integration-based discrete process neural networks algorithm and applications[J]. *System Engineering — Theory & Practice*, 2013, 33(12): 3216-3222.
- [6] Haykin S. *Neural Networks: A comprehensive foundation*[M]. Englewood Cliffs: Prentice Hall, 1998: 76-78.
- [7] 刘坤, 谭莹, 何新贵. 基于粒子群优化的过程神经网络学习算法[J]. *北京大学学报: 自然科学版*, 2011, 47(2): 238-244. (Liu K, Tan Y, He X G. Particle swarm optimization based learning algorithm for process neural networks[J]. *Acta Scientiarum Naturalium Universitatis Pekinensis*, 2011, 47(2): 238-244.)
- [8] Huang G B, Zhu Q Y, Siew C K. Extreme learning machine: A new learning scheme of feedforward neural networks[C]. *Proc Int Joint Conf on Neural Networks*. Budapest: IEEE, 2004: 985-990.
- [9] Huang G B. Extreme learning machine for regression and multiclass classification[J]. *IEEE Trans on Systems*, 2012, 42(2): 513-516.
- [10] Kasun L L C, Zhou H M, Huang G B. Representational learning with extreme learning machine for big data[J]. *IEEE Intelligent System*, 2013, 28(6): 31-34.
- [11] Zhang J, Feng L, Wu B. Local extreme learning machine: Local classification model for shape feature ex-traction[J]. *Neural Computing & Applications*, 2016, 27(7): 2095-2105.
- [12] Nayak P K, Mishra S, Dash P K. Comparison of modified teaching-learning-based optimization and extreme learning machine for classification of multiple power signal disturbances[J]. *Neural Computing & Applications*, 2016, 27(7): 2107-2122.
- [13] Huang G B, Chen L. Enhanced random search based incremental extreme learning machine[J]. *Neurocomputing*, 2008, 71(16/17/18): 3460-3468.
- [14] 刘志刚, 李盼池. 基于ELM-PNN算法的第24周太阳黑子预测预报[J]. *控制与决策*, 2017, 32(4): 642-646. (Liu Z G, Li P C. Prediction and forecast of sunspot numbers in 24 cycle based on ELM-PNN algorithms[J]. *Control and Decision*, 2017, 32(4): 642-646.)
- [15] 刘志刚, 许少华, 李盼池. 基于QPSO和极限学习的离散过程神经网络及学习算法[J]. *控制与决策*, 2016, 31(12): 2241-2247. (Liu Z G, Xu S H, Li P C. Discrete process neural networks and learning algorithms based on QPSO and extreme learning machine[J]. *Control and Decision*, 2016, 31(12): 2241-2247.)
- [16] Shores T S. *Applied linear algebra and matrix analysis*[M]. Berlin: Springer, 2007: 116-119.
- [17] Yang X S, Deb S. Cuckoo search via levy flights[C]. *Proc of the World Congress on Nature & Biologically Inspired Computing*. Coimbatore: IEEE Publications, 2009: 210-214.
- [18] Li P C, Li S Y. Quantum-inspired evolutionary algorithm for continuous spaces optimization based on bloch coordinates of qubits[J]. *Neurocomputing*, 2008, 72(1/2/3): 581-591.
- [19] Tan M J, Liu Q. A dynamic adaptive radial basis function approach for total organic carbon content prediction in organic shale[J]. *Geophysics*, 2013, 78(6): 445-459.

(责任编辑: 闫妍)