

BP 神经网络在沉积微相自动识别中的应用

刘俊玲, 吴学东, 王东寨, 贺正刚

(中国石化 西北油田分公司 塔河采油三厂, 新疆 轮台 841600)

摘要:提出了基于测井数据深度挖掘和前馈式(BP)人工神经网络算法的沉积微相识别方法。在测井数据较少、井多的条件下深入挖掘有限的测井数据,获取蕴含沉积学意义的参数,提高了测井数据的利用率。通过一系列实验,研究了BP人工神经网络拓扑结构的优选准则,并提出了成长型网络训练方法。最后利用建立的样本集和自然样本进行网络训练和微相识别,准确率达到了83%以上。在测井数据不足、微相特征复杂的条件下实现了高效率、高准确度的沉积微相识别。

关键词:BP神经网络;测井数据深入挖掘;成长型神经网络;沉积微相自动识别

中图分类号:TE2

文献标识码:A

Application of BP neural network to sedimentary micro-facies identification

Liu Junling, Wu Xuedong, Wang Dongzhai, He Zhenggang

(No.3 Tahe Oil Production Plant, SINOPEC Northwest Company, Luntai, Xinjiang 841600, China)

Abstract: A method of sedimentary micro-facies identification based on logging data and BP neural network was proposed in this paper. Through deeply exploring limited logging data, sedimentological sample indexes were gained and the utilization rate of logging data was improved. A series of experiments were conducted in order to find the optimization criterion of the BP artificial neural network and a growing network training method was put forward. Finally, a actual case of net training and micro-facies identification by using sample set and natural samples was analyzed, which showed an accuracy ratio for 83% and realized both high efficiency and precision of micro-facies identification.

Key words: BP neural network; further study of logging data; growing artificial neural network; identification of microfacies

沉积微相的分析与识别是地质研究中一项基础而又复杂的工作,目前采用的常规手段是通过单井解释模版进行判别,工作量大,重复性高,且人为误差较大。近年来,有学者采用自动化方法进行沉积微相判别:(1)靳松等利用不同沉积微相对应测井曲线段之间实测数据的变差函数(变程、主方向、次方向)特征差异进行微相判别^[1],这种方法将地质统计学算法中经典的变差函数^[2]引入到沉积微相判别中。该方法本质上是提取不同沉积微相的曲线特征差异,并不能直接识别沉积微相;(2)唐为清等利用前馈式神经网络算法进行沉积微相判别^[3-4]。通过不同微相的样本进行网络训练,训练完成之后将待分析样本投入到网络中去计算,即可得出判别结果,该方法具有判别速度快、正确率较高的特点^[5],是一种潜力较大的算法。

目前利用BP神经网络判别沉积微相还存在以下问题:(1)判别指标缺乏地质意义,前人进行沉积微相判别主要基于多条测井曲线的幅度

值^[6],但测井曲线的幅度值并不是真正意义上的沉积相标志;(2)测井数据失真,许少华等人对测井数据进行样本化处理的时候采用了二值点阵法^[4],在处理过程中为了提高运行速度,必须舍去部分测点数据,造成数据失真;(3)针对沉积微相识别的BP神经网络的规模和拓扑结构的选定缺乏指导原则,神经网络学习过程耗时长,亟待优化^[7-9]。本文针对以上问题,提出了测井曲线深度挖掘方法,建立了基于测井曲线的沉积相识别指标,在此基础上通过大量实验提出了BP神经网络结构定义及高效训练方法。

1 训练样本准备与测井数据深度挖掘

1.1 研究区状况

研究区X油田A组发育三角洲相,主要亚相为三角洲前缘,发育的沉积微相有:水下分流河道、河口坝、河道间(也作泥岩)、前缘席状砂、远岸滩坝、近岸滩砂等沉积微相,不同类型的沉积微相特

表 1 测井曲线对应的样本指标
Table 1 Parameters of samples
calculated from well log curve

| 指标 | GR | SP | ML1 | ML2 |
|---------|----|----|-----|-----|
| 幅度均值 | √ | √ | | √ |
| 含油性 | | | √ | √ |
| 变异系数 | √ | √ | | |
| 变差函数主方向 | | √ | | |
| 韵律系数 | | √ | | |
| 厚度系数 | | | √ | |

征复杂,而测井曲线类型较少(主要有自然电位、自然伽马、微梯度、微电位 4 种曲线),研究区内共有井 576 口,部分井存在曲线不全、不同时期的测井曲线存在较大差异,给沉积微相识别带来一定难度。为了准确识别沉积微相,需要对现有测井曲线进行归一化^[10],进行沉积相标志挖掘,并采用容错性能较强的 BP 神经网络方法。

1.2 测井数据深度挖掘

由于研究区测井曲线资料较少,曲线本身不具有相标志意义。因此,本文从沉积学和统计学的角度出发,设计了曲线幅度均值、含油性、韵律指数、厚度系数等多个指标(表 1):

- ①幅度均值反映不同沉积微相的岩性—电性特征;
- ②含油性反映不同含油饱和度对岩电关系的影响,用于排除沉积微相识别中饱和度差异的影响;
- ③变异系数反映沉积微相单元内隔夹层出现频率,指示沉积微相类型;
- ④变差函数主方向指示曲线形态特征^[1],采用混合编程方法^[11],利用 GSLIB 程序库中的 gam.f^[12]程序提取参数;
- ⑤韵律指数反映沉积微相的垂向韵律特征,选用 SP 较为平滑的曲线;
- ⑥厚度系数通过沉积微相的厚度差异区分不同微相。

编写了基于 MFC 的沉积相识别指标提取软件,在测井曲线的基础上一致性、批量提取各井不同层内不同岩性段的沉积相识别指标。

2 神经网络拓扑结构的选定与网络训练

2.1 神经网络拓扑结构的选定

选定神经网络的拓扑结构是整个神经网络模型构建中最基本的一环,主要包括层次结构和单层神经元数量关系。所选样本的数量和复杂度在一定程度上决定了所需神经网络的拓扑结构。前人在对神经网络的拓扑结构与其收敛速率上进行过

初步研究,认为 3 层(即只有一个隐含层)的神经网络即可适应沉积微相判别^[13-20]。但是在实践中,特别是测井数据不足的情况下,这种简易、局限的网络设置方法并不能适应网络训练的需求。

针对神经网络拓扑结构与其收敛性关系,选取研究区不同部分已经完成沉积微相识别的样本作为训练样本,在不同神经网络拓扑结构下进行训练实验。为了保证数据的兼容性,每个样本包含 42 个指标,误差精度设为 0.000 001。在训练过程中观测训练所需时间作为评价网络收敛效率的指标。整个测试分为单隐含层、多隐含层 2 个阶段。

2.1.1 单隐含层

定义只有一个隐含层的 BP 神经网络结构,研究隐含层神经元数目对收敛效率的影响。采用不同的隐含层节点数量进行 20 次实验,并对结果进行回归分析(图 1)。

在单隐含层的条件下,将隐含层神经元个数设置为输入层的 0.8~1.5 倍,神经网络能够较快收敛,隐含层神经元数目过少将导致网络无法收敛,而隐含层神经元过多则会导致整个网络复杂,无法在有限时间内达到收敛。

2.1.2 多隐含层

在多个隐含层下对神经网络的收敛特性进行研究,重点对 4~6 层的神经网络进行对比实验。考虑单隐含层中的实验结论,设置第一隐含层节点个数为 50,对第二隐含层设定不同值进行 18 次实验,并对目标函数收敛速度进行回归分析(图 2)。

在存在 2 个隐含层的条件下,第二隐含层的节点数 1.5~3 倍于第一个隐含层的节点数,能够保证神经网络的快速收敛。

随着隐含层增加,网络拓扑结构越来越复杂,网络训练过程中的正向传播和误差反馈过程计算量和调节力度都受到了限制,导致收敛时间逐步增

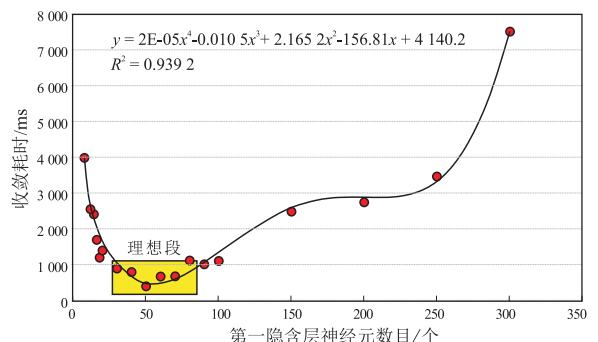


图 1 三层神经网络拓扑结构与训练耗时对比

Fig.1 3-layered neural network topological structure and time consumption contrast

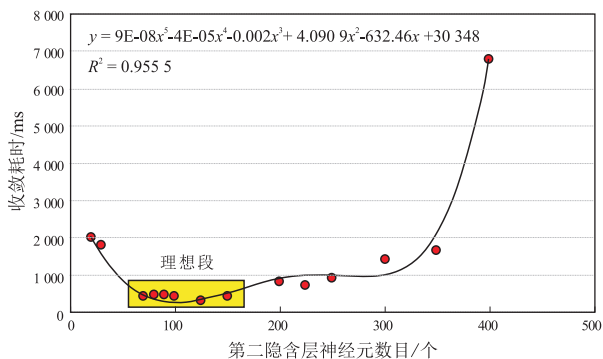


图 2 四层神经网络拓扑结构与训练耗时对比

Fig.2 4-layered neural network topological structure and time consumption contrast

表 2 多层神经网络拓扑结构与训练耗时

Table 2 Multi-layered neural network topological structure and time consumption contrast

| 拓扑结构 | 训练时长/ms | 是否收敛 | 备注 |
|--------------------------|------------|------|--|
| 42×50×125×10 | 3 360 | 是 | |
| 42×50×125×25×10 | 91 344 | 是 | |
| 42×50×125×50×10 | 26 781 | 是 | |
| 42×50×125×100×10 | 29 719 | 是 | 1. 训练过程采用 Intel Pentium T3200 处理器单线程运算。计量训练过程消耗时间; |
| 42×50×125×125×10 | 48 828 | 是 | |
| 42×50×125×150×10 | 18 437 | 是 | |
| 42×50×125×200×10 | 41 360 | 是 | |
| 42×50×125×250×10 | 48 922 | 是 | |
| 42×50×125×300×10 | 119 093 | 是 | 2. 训练精度:0.000 01; |
| 42×50×125×400×10 | >1 200 000 | 否 | 3. 冲量因子:0.5。 |
| 42×50×125×500×10 | >600 000 | 否 | |
| 42×50×125×180×10 | 78 469 | 是 | |
| 42×50×125×180×250×10 | 91 218 | 是 | |
| 42×50×125×180×250×200×10 | >1 200 000 | 否 | |

加。隐含层数过多将导致训练时间过长或无法收敛等问题,使得其失去实用意义(表 2)。

2.2 网络训练

2.2.1 网络训练的阶段性和冲量调整策略

进行网络训练的时候,观察整个训练过程中发现,在前期,网络误差在 90% 以上的阶段往往需要消耗整个调整过程的 90% 以上的训练时间,在达到一定的精度之后网络收敛速度大大提高。图 3 是 2 个三层神经网络的收敛过程,按照一定的步长进行训练程度统计,得到其训练过程。

在网络训练初期,进行权值调整具有一定的混乱性,而在后期调整中,出现了一定的调节失误。这是由于在随机权值的基础上进行调节有一定的盲目性和试探性,如果冲量因子过小,会限制权重调节的区间;而在后期调节中,网络权值的调节幅度随着网络误差的减小而减小,如果冲量因子较

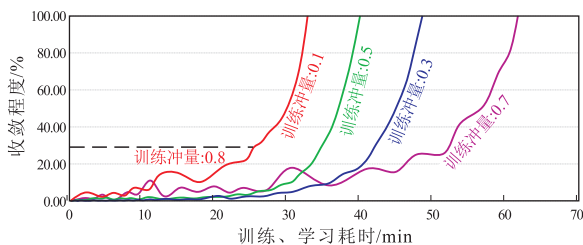


图 3 4 种冲量设置三层神经网络的训练过程

Fig.3 Training process of 3-layered neural networks (four impulses)

大,就会出现过调节行为(图 3)。为此,可以在调整初期增大权重调节的力度,适当增大冲量因子;而在网络误差达到 80% 以下之后适当减小冲量因子,以免出现较大的调整失误(图 3)。

2.2.2 成长型神经网络

所谓成长型神经网络,就是在已经训练完成的神经网络的基础上通过强化训练和拓展训练不断拓展网络的映射能力。在样本总体发生改变,或者样本内部出现矛盾的情况下,之前的神经网络权值分布已经不能适应工作的需要,如果重新进行样本准备和网络训练将十分耗时,且有可能使网络目标函数难以收敛。考虑到之前的神经网络能够适应大部分样本,设计了成长型神经网络训练方法。主要针对以下 3 种情形:

(1) 对预测不准确,或者出现冲突的样本进行分析,除去样本集中的不良样本,将最新确认的样本群加入到样本集中,形成新的样本集,在原有网络权值的基础上进行网络迁移训练,纠正之前不良样本带来的影响。

(2) 对新增的样本类型(样本指标类型不变),为新增样本制定新的导师信号之后投入到之前的网络中进行拓展训练,强化网络的映射能力。

(3) 如果随着应用环境的变化,某些类型的样本退出了预测,则在进行网络判别时应当屏蔽退出样本对应的导师信号,退出模式识别过程。网络本身不需要进行演化训练。

成长型神经网络需要对既有神经网络数据进行时时保存和恢复。本文在程序设计过程中充分考虑了这一特点。采用成长型神经网络,完成首次训练之后,在后续增加或者修正样本之后,在低冲量值进行训练,神经网络很快就能完成训练(图 4),成长型神经网络训练方法是一种高效、可靠的方法。

3 实例研究

对 X 油田 A 组的各种沉积微相的测井曲线特

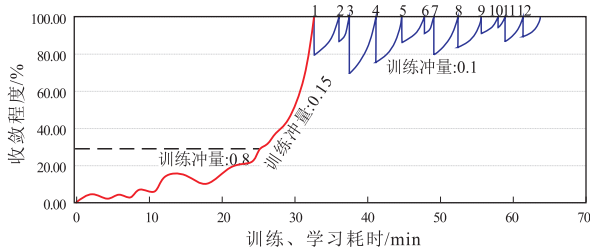


图 4 成长型神经网络训练过程

Fig.4 Training process of growing neural network

征进行分析,按照上文方法提取各个沉积微相各项指标统计数据,具体如表 3。

根据样本种类和样本指标,选定 3 种神经网络拓扑结构:A:9×15×7,B:9×15×20×7,C:9×15×20×30×7进行神经网络训练和识别。检验识别的标准有 2 个:对训练样本的识别准确率(C2)和对其他自然样本识别的准确率(C1)。由实际检验可得,对训练样本,判断的准确率为 100%;对其他自然样本训练的综合准确率约为 83%。在测井数据较少的条件下得到了良好的预测精度。此外,对于部分缺少 GR 数据的井,沉积微相识别的准确率约为 78%,表明神经网络具有良好的容差能力和稳定性(表 4)。

表 4 样本集构成与识别结果

Table 4 Sample organization and recognition result

| 项目 | | 分流河道 | 河口坝 | 泥岩 | 溢岸 | 远岸滩坝 | 席状砂 | 坝后 | |
|-------------|---|------------------------------|-----|-----|-----|------|-----|-----|-----|
| 基础样本数/个 | | 3 | 3 | 3 | 1 | 1 | 1 | 1 | |
| 自然样本数/个 | | 12 | 9 | 8 | 5 | 7 | 9 | 6 | |
| 待识别样本/个 | | 135 | 126 | 359 | 78 | 56 | 73 | 68 | |
| 样本总量 | | 964 个(69 个训练样本,895 个待识别自然样本) | | | | | | | |
| 识别 正确率/% | A | C1 | 76 | 82 | 87 | 83 | 96 | 81 | 69 |
| | | C2 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| | B | C1 | 78 | 76 | 90 | 85 | 85 | 85 | 78 |
| | | C2 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| | C | C1 | 82 | 83 | 86 | 81 | 89 | 79 | 73 |
| | | C2 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

注:A、B、C 表示 3 种神经网络拓扑结构;C1 为对其他自然样本的识别,C2 为对训练样本的识别。

(3)采用实例分析了神经网络的训练过程,明确了早期采用大冲量,中后期减小冲量的神经网络高效训练方式。

(4)编写了基于 BP 神经网络的沉积微相智能识别软件,设计实现了“成长型”神经网络,提高了沉积微相识别工作中计算的效率和成果的继承性与拓展性。

参考文献:

[1] 靳松,朱筱敏,钟大康.变差函数在沉积微相自动识别中的应用[J].石油学报,2006,27(3):57-60.
 [2] Deutsch C V,Journal,A G. GSLIB:geostatistical software library and

表 3 3 种沉积微相曲线特征及指标统计

Table 3 Curve features and parameters statistics of 3 kinds of microfacies

| 微相 | | 分流河道 | 河口坝 | 泥岩 |
|------------|-----|-----------|----------|---------|
| 曲线幅度 均值 | SP | 0.239 8 | 0.469 8 | 0.163 4 |
| | GR | 0.440 5 | 0.386 6 | 0.522 5 |
| | ML1 | 0.273 5 | 0.277 5 | 0.171 7 |
| | ML2 | 0.284 5 | 0.312 4 | 0.189 1 |
| 含油性 | | 0.046 6 | 0.047 8 | 0.019 1 |
| 变异系数 | | 0.235 5 | 0.549 6 | 0.234 7 |
| 变差函数主方向 | | 135°/315° | 38°/218° | 0°/180° |
| 韵律系数 | | 1.112 3 | 0.921 3 | 1.000 0 |
| 厚度/m | | 2.703 8 | 10.488 4 | 4.978 4 |

4 结论与认识

(1)通过测井数据的深入挖掘,获得了多种具有沉积相标志意义的参数,提高了沉积微相自动识别的地质合理性和测井数据利用率。

(2)采用大量神经网络训练实验与分析,得到了沉积微相识别所需的神经网络结构,一般 3~4 层神经网络能够高效、准确地完成沉积微相识别,隐含层节点数应当为上一层节点数的 0.5 倍至数倍。

user's guide [M].New York: Oxford University Press,1992:340.
 [3] 唐为清,郭荣坤,王忠东,等.沉积微相测井资料神经网络判别方法研究[J].沉积学报,2001,19(4):581-585.
 [4] 许少华,刘扬,梁久祯,等.基于遗传—BP 算法和图像处理的沉积微相识别[J].石油学报,2002,23(3):48-51.
 [5] 许少华,陈可为,梁久祯,等.基于遗传—BP 算法和图像处理的沉积微相识别[J].大庆石油学院学报,2001,25(3):51-54.
 [6] E M Iloghalu. "Application of neural networks technique in lithofacies classifications used for 3-D reservoir geological modelling and exploration studies;a novel computer-based methodology for depositional environment interpretation" [C]// AAPG Annual Convention.Salt Lake City,Utah,May 2003:11-14.

的液体由上引液管吸入,并经由在线含水传感器后由回管返回罐内;按计算仪的“暂停”按钮,循环泵依程序指令停止,待 5 s 后计算仪的 PLC 自动记录并储存在线含水传感器发出的含水信号 C_A 。此时关闭上引液管的上引液管截止阀,打开中引液管的中引液管截止阀,再按计算仪的“测试”按钮,循环泵依程序指令动作,将立式储罐内的液体由中引液管吸入,并经由在线含水传感器后由回管返回罐内;按计算仪的“暂停”按钮,循环泵依程序指令停止,待 5 s 后计算仪的 PLC 自动记录并储存在线含水传感器发出的含水信号 C_B 。测得上引液管端口、中引液管端口两组含水值后,计算仪的 PLC 记录储存立式储罐自配的液位传感器的液位信号,并依计算仪中 S7-200PLC 的 PLC 控制软件所内嵌的计算公式(公式 1),计算出立式储罐内凝析油的平均含水值。

$$\text{平均含水} = \frac{h}{H} \times \frac{C_B}{\left(\frac{C_B}{C_A}\right)^{\frac{H-E}{h}}} \times \frac{\left[\left(\frac{C_B}{C_A}\right)^{\frac{H}{h}} - 1\right]}{\ln\left(\frac{C_B}{C_A}\right)} \quad (1)$$

式中: h 表示上引液管中心线与中引液管中心线间的距离; E 表示中引液管中心线与罐底间的距离; H 表示液位传感器探测到的储罐内的液体高度;

C_A 表示在线含水传感器测出的上引液管引出的液体含水; C_B 表示在线含水传感器测出的中引液管引出的液体含水。

3 结论

采用二点式立式储罐均匀液体平均含水测试装置进行凝析油含水测试,一是减少了岗位人员上罐取样次数,降低了岗位人员的劳动强度;二是消除了人员上罐取样发生跌滑或吸入过量有害气体的安全隐患;三是减少了手工取样和蒸馏法进行离线含水检测过程中的人为干扰因素,提高了含水检测的准确性;四是在线含水检测节约了凝析油含水分析时间。笔者研发的二点式立式储罐均匀液体平均含水测试装置,已获国家知识产权局授权的实用新型专利。该项技术的运用,为其他油田类似的问题提供了可借鉴的经验和实例,具有较强的推广价值和应用前景。

参考文献:

- [1] 金德玺.GB/T4756-1998 石油液体手工取样法[S].1998.
- [2] 薄艳红,李季成.GB/T8929-2006 原油水含量的测定(蒸馏法)[S].2006.
- [3] 惠学先,解更存,张永红.油气田天然气凝析油含水测定研究[J].工业计量,2001(51):188.

(编辑 徐文明)

(上接第 55 页)

- [7] 杨斌,匡立春,孙中春,等.神经网络及其在石油测井中的应用[M].北京:石油工业出版社,2005.
- [8] 张昭昭,乔俊飞,杨刚.自适应前馈神经网络结构优化设计[J].智能系统学报,2011,6(4):312-317.
- [9] 王金荣,刘洪涛.测井沉积微相识别方法及应用[J].大庆石油学院学报,2004,28(4):18-20.
- [10] 王楠.测井曲线模式识别及其在地层对比中的应用[D].黑龙江大学硕士论文,2008.
- [11] 李君,李少华,毛平,等.VC++结合 Fortran 升级地质统计学算法[J].物探与化探,2009,33(6):715-717.
- [12] Carr J R, Miranda F P. The semivariogram in comparison to the co-occurrence matrix for classification of image texture[J].IEEE Transactions on Geoscience and Remote Sensing,1998,36(6):1945-1952.
- [13] Funabashi K. On the approximate realization of continues mapping by neural networks[J].Neural Networks,1989,2:183-192.
- [14] Lippmann R P. Review of neural networks for speech recognition[J].Neural Computation,1989,1:1-38.
- [15] Cyberno G. Approximation by superpositions of a sigmoidal function[J].Mathematics of Control, Signals and Systems,1989,2(4):303-314.
- [16] 孙鲁平,首皓,赵晓龙,等.基于微电阻率扫描成像测井的沉积微相识别[J].测井技术,2009,33(4):379-383.
- [17] 周金应,桂碧雯,李茂,等.基于岩控的人工神经网络在渗透率预测中的应用[J].石油学报,2010,31(6):985-988.
- [18] K Aminian, S Ameri. Application of artificial neural networks for reservoir characterization with limited data[J].Journal of Petroleum Science and Engineering,2005,49:212-222.
- [19] 宋延杰,杨艳,杨青山,等.过程神经网络在厚层细分水淹解释中的应用[J].测井技术,2009,33(4):340-344.
- [20] Fischetti A, Andrade A. Porosity images from well logs[J].Journal of Petroleum Science and Engineering,2002,36:149-158.

(编辑 叶德燎)