

41-45

神经网络方法在煤层气测井资料解释中的应用

侯俊胜 王颖

(中国地质大学·北京·100083)

P 631.84

P 618.130.8

根据煤层气储层及其测井响应特征,提出应用BP神经网络(简称BP网络)方法进行煤层气测井资料的定量解释。通过对华北某地煤层气测井资料的处理,获得了满意的解释结果,证明该方法是有有效和实用的。并对应用过程中的若干问题进行了分析讨论。

关键词 煤层气储层 BP网络 测井 资料解释 问题讨论

煤层气储层其最大特点就是既是煤层气的储集层,又是它的生气源岩,并且煤储层包含了极为发育的自然裂隙。煤层气储层这一包括基质孔隙和裂缝孔隙的双重孔隙结构特点,使其测井资料受多种因素控制,采用单一测井信息或某一储层参数(例如孔隙度或其它参量)进行煤层气储层研究,除在一些特殊情况下可取得满意效果外,很难普遍推广应用。当前最佳选择是利用综合测井资料进行煤层气储层研究。由于事先难以确定煤层气储层参数与综合测井响应之间定量关系,因而只能提出若干假设条件,从而建立一个简化数学模型(例如,假设它们之间存在某种线性或非线性关系)。然而,这些假设条件的提出没有严格理论指导,可能仅是一种经验推测,给解释结果带来一定误差。应用神经网络进行煤层气测井资料定性和定量分析,其优点是不用考虑具体的数学模型,就能用“隐式”表达出煤层气储层参数与测井响应之间的复杂关系,从而实现煤层气识别和储层参数计算目的。目前,人工智能的进展已建立了近百种神经网络模型和算法,普遍认为BP神经网络(简称BP网络)是方法较完善、应用较广的网络模型。

1 BP神经网络方法

1.1 BP网络的基本特征

BP网络是一前向多层神经网络,该网络的基本拓扑结构如图1。它包括输入层、隐含层和输出层。

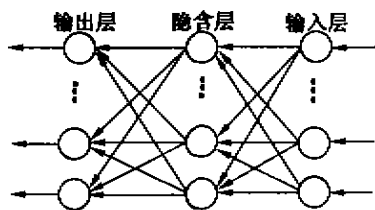


图1 BP网络基本拓扑结构

其中每一层都包括若干神经元,同一层中的神经元之间没有相互联系,各层神经元个数以及隐含层的层数视具体问题而定。

BP网络包括两个过程,即学习过程和工作过程。所谓学习过程,就是向网络提出一定数量的训练(学习)样本(训练样本就是提供输入层信息的同时,还应该提供输出层的期望值),将网络的实际响应和期望响应进行比较,通过修改各层神经元之间的连接权系数,以便使期望输出与实际输出之间误差达到最小,最后即可求得一组连接权系数。权系数确定以后,整个网络系统也就随之确定。所谓工作过程就是将待处理的测井信息送入输入层,对应输出层的输出即为网络的工作结果。例如,对于煤层气测井资料解释问题,假设输入层为自然伽玛、电阻率和声波时差等测井响应信息,其输出层可为煤层气的含气量等储层评价参数。

1.2 BP网络基本计算公式

根据神经网络理论,BP网络的神经元输入 $I_k^{(l)}$ 和输出 $O_k^{(l)}$,计算公式为

$$I_k^{(l)} = \sum_{j=0}^l w_{kj}^{(l)} \cdot O_j^{(l-1)} \quad (1)$$

$$O_k^{(l)} = \int_s [I_k^{(l)}] = \frac{1}{1 + e^{-I_k^{(l)}/T_0}} \quad (2)$$

其中: $P = 0, 1, 2, \dots, P_G$; P_G 为训练样本总数。 $l = 0, 1, 2, \dots, l_G$; l_G 为网络层数。 $i = 0, 1, 2, \dots, i_G$; i_G 为 l 层的神经元个数。 $w_{ij}^{(l)}$ 为网络连接权系数。 T_0 是温度调节系数。 L 为第 $(l-1)$ 层神经元个数。

对某一给定的输入样本,以输出层实际输出 y_{pi} 与其期望输出 d_{pi} 的误差平方和

$$\epsilon = \sum (d_{pi} - y_{pi})^2 / 2 \quad (3)$$

作为目标函数,用最速下降法即可导出 $w_{ij}^{(l)}(t)$ 的迭代调整公式

$$W_y^{(l)}(t-1) = W_y^{(l)}(t) + \Delta_p W_y^{(l)}(t) \quad (4)$$

$$\Delta_p W_y^{(l)}(t) = \eta \cdot \delta_{P_i}^{(l)}(t) \cdot O_{P_i}^{(l+1)}(t) + \alpha \cdot \Delta_p W_y^{(l)}(t-1) \quad (5)$$

其中

(1)当 l 为输出层时

$$\delta_{P_i}^{(l)}(t) = 2(d_{P_i} - O_{P_i}^{(l)}) \cdot \int_s [I_{P_i}^{(l)}]$$

(2)当 l 为隐含层时

$$\delta_{P_i}^{(l)}(t) = \sum_{k=0}^{Q-1} [\delta_{P_k}^{(l+1)} \cdot W_{ik}^{(l+1)}] \cdot \int_s [I_{P_i}^{(l)}]$$

t 为学习次数, η 为学习步长(速率), α 为动量因子(惯性系数), Q 为第 $(l+1)$ 层的神经元个数, $\int_s [I_{P_i}^{(l)}]$ 为函数 $\int_s [I_{P_i}^{(l)}]$ 对 $I_{P_i}^{(l)}$ 的一阶偏导数。

1.3 BP 网络学习的计算步骤

根据前面讨论,不难总结出 BP 网络的学习步骤,即:

1)确定网络结构。即给定输入层与输出层的神经元数目、隐含层数及各层的神经元数目以及学习步长、动量因子和终止误差等网络参数;

2)网络状态初始化。即将所有权系数都设置成较小的随机数;

3)输入学习样本;

4)计算隐含层与输出层的输出;

5)计算输出层误差;

6)判断输出层误差 E_p 是否小于给定的误差极限 $\varepsilon_{\max}$, 满足要求,结束学习,否则转向 7)。

7)判断学习次数 IT 是否小于最大学习次数 $IT_{\max}$, 满足要求,转向 8), 否则结束学习;

8)调整连接权系数;

9)转向 3)。

2 方法实际应用及有关问题讨论

2.1 煤层气测井资料解释基本步骤

根据 BP 网络方法原理以及煤层气测井资料解释自身特点,可总结出如下步骤:

1)选择输入输出特征量(分别对应网络输入层和输出层神经元个数);

2)选择学习样本;

3)测井数据的预处理;

4)测井数据归一化。其计算公式为 $X'_K = (X_K - X_{\min}^{(k)}) / (X_{\max}^{(k)} - X_{\min}^{(k)})$, 其中 X'_K 为归一化值, X_K 为输入层矢量的第 K 个特征参数, $X_{\max}^{(k)}$ 和 $X_{\min}^{(k)}$ 分别为第 K 个特征参数最大与最小值;

5)将学习样本数据输入 BP 网络进行学习,求出一组连接权系数;

6)输入工作样本数据进行煤层气识别计算;

7)输出计算结果(表格或图示)。

2.2 应用实例

为检验 BP 网络在煤层气测井资料解释中应用效果,利用华北某煤田煤层气综合测井资料进行了煤层气含量等储层参数计算。选择深度 D (单位: m)、深侧向电阻率 RD (单位: $\Omega \cdot m$)、浅侧向电阻率 RS (单位: $\Omega \cdot m$)、声波时差 AC (单位: $\mu s/m$)、密度 DEN (单位: g/cm^3) 和自然伽玛 GR (单位: API) 共 6 个测井信息特征量作为输入层的特征参数,即 BP 网络输入层共有 6 个神经元;所要计算的煤层气储层参数分别为裂隙孔隙度 $PORF$ (单位用小数表示,简记 PF)、基质孔隙度 $PORB$ (单位用小数表示,简记 PB) 以及地层含碳量 V_c (单位用小数表示)、含灰量 V_a (单位用小数表示) 和含气量 V_g (单位: m^3/t)。因此,当仅计算 PF 、 PB 时,网络输出层共有 2 个神经元,当计算 V_c 、 V_a 、 V_g , 网络输出层共有 3 个神经元,当 5 个特征量同时计算时,输出层共有 5 个神经元。根据神经网络的映射定理,具有 3 层(一个输入层、一个隐含层、一个输出层)的网络结构,即可实现任意连续函数从输入空间向输出空间的映射,因此选择一个隐含层,根据经验和有关计算,选择隐含层有 6 个神经元,这就确定了 BP 网络的拓扑结构。

首先,从井 A 选择 64 个样本数据,它包括 6 个输入特征量(D 、 RD 、 RS 、 DEN 、 AC 、 GR) 和 5 个输出特征量(PF 、 PB 、 V_c 、 V_a 、 V_g)。再从 64 个样本中每隔 2、3、4 个样本进行抽样,得到样本数分别为 32、22、16 3 种学习样本,选择动量因子 $\alpha = 0.9$, 学习步长 $\eta = 0.7$, 温度因子 $T = 1.0$, 误差 $E = 0.01$, 迭代次数 $I = 10000$, 利用 BP 网络分别对上述 3 种学习样本进行学习,使计算值和期望值之间的误差达到精度要求,从而计算出网络的连接权系数,即建立了所用测井信息与储层参数之间的关系。再用上述 64 个工作样本(图 2)对 3 种不同的学习样本进行工作,得出 3 种不同的计算结果;其次,把输出层 5 个特征量分为 PF 、 PB 和 V_c 、 V_a 、 V_g 两组分别计算,同时还采用 32、22、16 个学习样本分别学习,得出了 3 组不同输出层特征量计算结果与期望值的平均误差(表 1),图 3 是利用 32 个学习样本学习,同时计算 5 个输出特征量($PF1$ 、 $PB1$ 、 V_{c1} 、 V_{a1} 、 V_{g1}) 与期望值(PF 、 PB 、 V_c 、 V_a 、 V_g) 的对比曲线。

表1 平均误差计算表

输出量 个数	学习 样本 个数	平均误差 ($\times 10^{-3}$)				
		裂隙孔 隙度 (PF)	基质 孔隙度 (PB)	碳质 含量 (Vc)	灰质 含量 (Va)	煤层气 含量 (Vg)
5	16	6.545	5.968	5.969	28.29	
	22	5.548	6.772	3.998	31.78	
	32	3.993	4.692	4.185	21.91	
2	16	18.43	12.77			
	22	12.75	14.50			
	32	22.03	18.33			
3	16			5.309	23.67	84.52
	22			2.761	25.22	65.64
	32			2.702	18.53	62.26

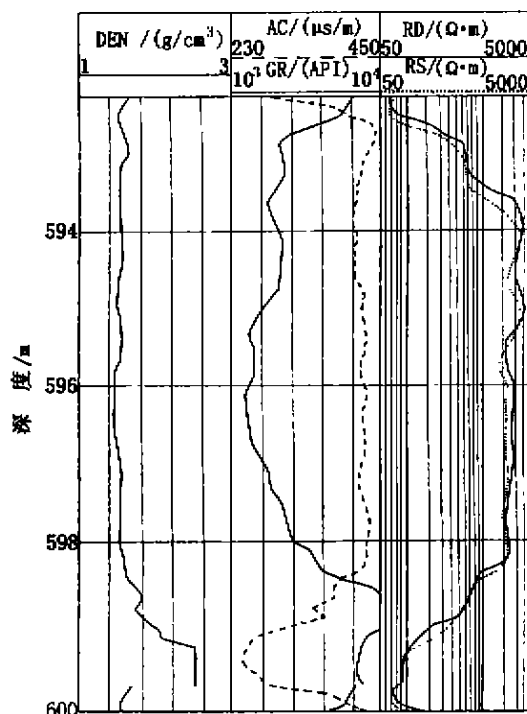
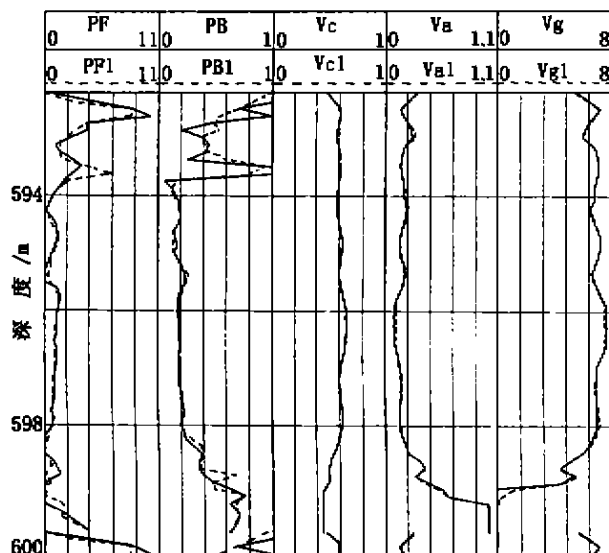


图2 井A 64个测井样本数据

图3 井A输出特征量与期望值对比(32个学习样本)曲线
从上述讨论可看出,计算结果与期望值基本吻

合,计算误差可以接受。在此基础上,计算了同区域内井B的测井资料。以上述的64个样本作为学习样本,工作样本数据见图4,工作样本数为321,隐含层神经元数为3,其它网络参数同上。把学习样本和工作样本进行规一化,利用测井资料的储层参数计算结果见图5。已知该区煤层的地球物理特征一般为:高电阻率、低密度、高时差、低自然伽玛。由图5可见,PF, PB, Vc, Va, Vg的计算结果与该区煤层的地球物理特征基本符合。

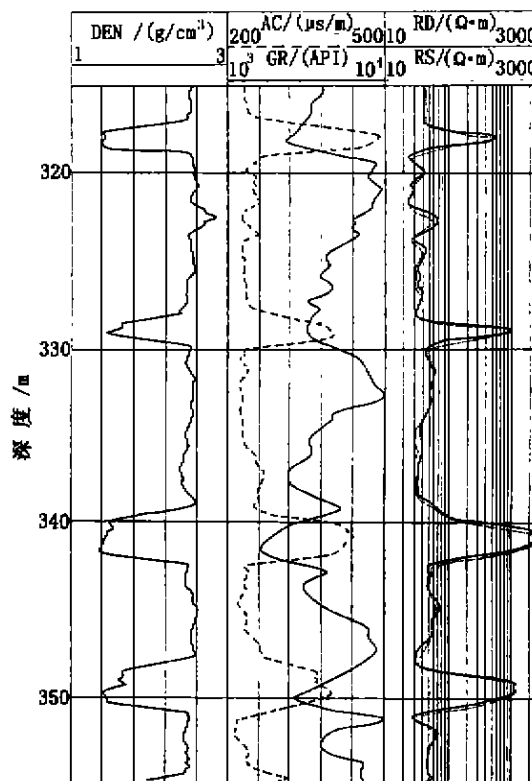


图4 井B测井数据

2.3 问题讨论

在实际计算过程中,参数选取、学习样本数选择及一些计算技巧值得很好讨论。

2.3.1 隐含层数的选择

在实例中,选择了一个隐含层,计算结果达到一定精度要求,且计算时间较为节省。当取隐含层数较大时,发现计算精度并没有很大提高,计算时间反而明显增加。同时,隐含层太大,将使网络记忆输入模式而不是归纳输入特性,使学习了的网络降低了处理非样本输入信息能力。所以取一个隐含层完全可以达到测井资料解释的要求,增加隐含层数没有实际意义。

2.3.2 隐含层单元数的选择

实例中,选择隐含层单元数为6,即有6个神经元,在用22个样本计算Vc、Va、Vg三个特征量时取

3 个神经元进行计算并做误差对比(表 2)。

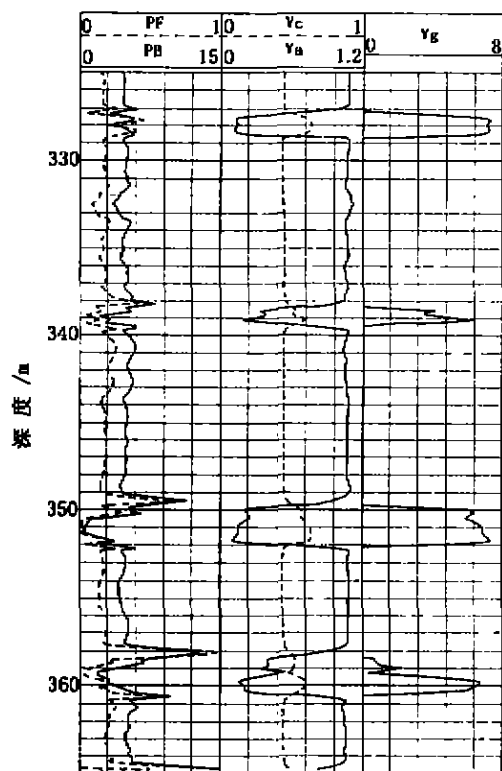


图 5 井 B 储层参数的计算结果

由表 2 可见,虽存在一定精度差异,但并不很大。较少的神经元数可以减少计算时间,所以在不同情况下,可根据精度要求自由选择隐含层单元数,但不宜太大。

表 2 隐含层不同神经元数计算误差对比表

神经元数	误差		
	碳质含量(V_c)	灰质含量(V_a)	煤层气含量(V_g)
6	2.671×10^{-3}	2.522×10^{-2}	6.564×10^{-2}
3	2.093×10^{-3}	2.647×10^{-2}	7.671×10^{-2}

2.3.3 网络其它参数的选择

1) 动量因子 α 的选择:动量项是动量因子 α 乘以新旧权向量之差,动量项的变化率应保持在某个量级上,才能经迭代步骤逼近于解,非无限小的 α 将能抑制振荡的产生,但可能使学习过程变慢。

2) η 的选择:参数 η 必须设置成一个适当小的值,避免使网络在最速下降过程中产生振荡。但较小的值会使网络学习过程需要更多的迭代次数, η 值大可有快速的学习过程,但结果可能产生振荡。

2.3.4 学习样本的选择

学习样本个数的选择对结果有较大的影响(表 1)。学习样本越多,计算精度越高;但在实际工作中,出于工作时间、经济效益等方面考虑,不宜取很多样本进行学习。在达到精度要求情况下,以选取

尽量少的学习样本为宜。

2.3.5 输出层特征参量数的选择

由表 1 可见,在单独计算 3 个特征量(V_c 、 V_a 、 V_g)时,比同时计算 5 个特征量(PF 、 PB 、 V_c 、 V_a 、 V_g)的误差明显要小。这是因为 BP 网络使用了最优化中最普遍的一种沿梯度下降算法,这种算法从某一点 X_i 出发,沿梯度方向求目标函数 $\int (X_1, X_2, \dots, X_n)$ 的极小点 X_{i+1} ,即求出: $\int (X_1, X_2, \dots, X_n) = \min \int (X_1, X_2, \dots, X_n)$ 。参数越多(n 值大),找到的 X_{i+1} 就可能不是我们所需的极小值,而是局部极小值。这样计算误差就会增大,所以输出层特征量不宜太多。

2.3.6 学习样本数据和工作样本数据归一化

当学习样本不是从工作样本中抽样得到的时候,分别对两个样本进行归一化,不同的极值使归一化结果不统一,在反归一时就造成了不可避免的错误。我们对学习样本和工作样本进行统一归一化,人为设定极值,在反归一时确保计算结果的正确性。

3 结论

BP 网络可以在煤层气测井资料解释中发挥重要作用,但它自身还存在许多问题:

1) 从数学观点看,它是一个非线性优化问题,这就不可避免地遇到最常见的局部极小,致使问题得不到很好解决。

2) 学习过程中神经网络的所有权系数都要调整,而且为了防止振荡现象出现,学习步长必须取为较小的数值,尤其是在接近最优时步长必须相当小,这就会降低学习速度。

3) 网络中隐含层节点个数(神经元数)的选取尚无理论上的指导,而是根据经验选取。

参考文献

- 1 杨建业,等.煤层气藏的储集特征及储层评价.西安地质学报,1995,17(3):77~82
- 2 地矿部华北石油局.煤层气论文集.郑州:河南科技出版社,1990
- 3 潘和平,等.利用模糊模式识别煤成气层.地球科学,1993,18(1):84~94
- 4 潘和平,等.测井资料解释煤成气方法研究.现代地质,1994,8(1):119~125
- 5 李纪森,等.应用密度测井资料计算煤岩组分.地球物理测井专刊.石油物探,1994,64~71
- 6 柳孟文,等.煤层气识别方法.中国地球物理学会年刊,1995,323
- 7 李庆谋,等.煤层气地球物理测井评价研究.中国地球物理学会年刊,1996,133

INTERPRETATION OF WELL LOGGING DATA FOR COALBED METHANE USING BP NEURAL NETWORK

Hou Junsheng, Wang Ying

Based on the signatures of coalbed methane reservoir and its well logging data, this paper proposes to use BP neural network on the interpretation of well logging data of coalbed methane reservoir. According to the computation for the well logging data for coalbed methane reservoir in some area of North China, the computing results show BP neural network method is reliable and effective, and some applied problems are discussed.

Key words coalbed methane reservoir, BP neural network, well logging, discussion



第一作者简介:

侯俊胜 男, 1963年生。1984年毕业于武汉地质学院物探系金属物探专业, 1988年在中国地质大学(北京)物探系获硕士学位, 1993年在东北大学地质系获博士学位, 1995年在中国地质大学(北京)完成第一站博士后科研工作。现任中国地质大学(北京)物探系副教授, 主要从事应用地球物理教学和科研工作。

通讯地址: 北京市海淀区学院路29号 中国地质大学(北京)物探系 邮政编码: 100083

(上接第35页)

2 武耀诚, 等. 地质数据处理及微机应用. 北京: 地质出版社, 1992

3 吴锡生. 化探数据处理方法. 北京: 地质出版社, 1993

WEIJIAGOU GOLD DEPOSIT IN SHANDONG: APPLICATION OF TECTONIC GEOCHEMISTRY TO PROSPECTING

Wen Hanjie, Xiao Huayuan, Yu Guangjun

Based on the study of some gold deposits in Zhaoyuan, Shandong, the principle and method of tectonic geochemistry are summarized. The Weijiagou gold deposit is exemplified to detail the procedure and application of tectonic geochemical method.

Key words prospecting by tectonic geochemistry, prediction of mineralization, gold deposit



第一作者简介:

温汉捷 男, 1971年生。1993年毕业于昆明理工大学国土开发与城乡建设系矿产勘察专业, 1996年在昆明理工大学国土开发与城乡建设系获硕士学位。现在中国科学院地球化学研究所攻读博士学位, 主要从事矿床地球化学和找矿预测研究。

通讯地址: 贵州省贵阳市观水路73号 中国科学院地球化学研究所矿床开放室 邮政编码: 550002

(上接第37页)

2 张庆海. 盐矿钻探技术与工艺. 北京: 地质出版社, 1992

3 杜德科 Л. М. 盐类地下水溶开采. 北京: 地质出版社, 1985

THE ENVIRONMENT OF GYPSUM AND SALT MINES AND MEASURES TAKEN TO CURB POLLUTION IN THE YINGCHENG AREA

Xiang Wu, Bao Zhengyu

The environment of the gypsum and salt mines in the Yingcheng area, Hebei, is discussed. The studies indicate that the affects of gypsum and salt on environment result from inorganic salt pollution. In addition, other dangerous incidents such as ground depression may occur. Finally, some suggestion are put forward on measures to take to curb pollution.

Key words saline environment, pollution, gypsum mine, salt mine



第一作者简介:

向武 男, 1968年生。1990年毕业于中国地质大学地球化学系地球化学专业, 1998年在中国地质大学地球科学学院获硕士学位。现任中国地质大学地球化学所工程师, 主要从事环境地球化学的教学与科研工作。

通讯地址: 湖北省武汉市 中国地质大学地球化学研究所 邮政编码: 430074