第15卷 第1期 空间科学学报 Vol. 15, No. 1
1995 年 1 月 CHINESE JOURNAL OF SPACE SCIENCE Jan., 1995

人工神经网络在磁情预报中的应用*

顾左文 高玉芬 范国华
(国家地震局地球物理研究所，北京 100081)

摘 要

介绍了人工神经元网络在磁情预报中的应用研究. 并对具体应用于 C_9 指数预报的 MLP 的拓扑结构、预报效率及在预报时所表现出的特点和问题进行了讨论.

关键词 人工神经元网络，磁情预报，C_9 指数

一、概 述

人工神经元网络的提出由来已久. 但自八十年代以来，其理论研究及在实际物理问题中的应用才得到了迅速的发展. 人工神经元网络的种类很多. 其中应用较广的是MLP (Multi-layer Perceptron，多层感知器).

MLP 是一种前馈型、有指导训练的人工神经元网络. 前馈型，即指其信息流程从输入层经中间隐含层向输出层单向传输. 有指导训练，即指其各神经元的输入权系数及阈值是通过对学习事件(包括输入事件及标准输出——即指导) 的学习过程而确定的. MLP 所采用的学习方法通常为 BP 法 (Error Back propagation).

MLP 能精确地实现实数域中的任意连续函数 $\varphi: E^N \to R^M$. 它对应于这样一种特殊的数学方法：对所研究的物理过程的认识并非通过对其运动规律的理性、抽象的描述来实现，而是通过对该过程不同阶段的运动状态及相关物理过程的运动状态之具体描述的对应来实现的；并非通过具有特定意义的函数形式而是通过规模巨大而形式简单的计算过程来建立所需的数学映射的. 相对于其它数学方法，MLP 是一种“模糊”的方法，具有很浓的形象思维色彩.

由于 MLP 本身的特点及其在模式识别、市场观测等应用中所取得的较好效果，因此我们尝试着将其引入到磁情预报的应用研究中.

二、MLP 在 C_9 指数预测中的应用

C_9 指数的记录已有上百年的历史，章公亮等人曾对其进行了较深入的研究，得出了

* 本课题为国家自然基金支持项目. 国家地震局地球物理研究所论著登记号 95A0006.
本文于 1994 年 2 月 14 日收到

某些特点和规律[1]. 因此,作为一种新方法的实验,我们选择 C_9 指数作为预测对象.

若将预测过程视为一个输入、输出间的传输系统,则预测过程的研究便涉及如下内容: 输入量及输入形式的确定;输出量及输出形式的确定;传输形式的确定及传输过程的建立等.

1. 输入量及输入形式

地磁扰动受多方面因素的制约,最主要的因素是太阳活动的影响. 太阳活动引起物质量、能量、动量等向行星际空间传输,并在到达地球轨道后与磁层相互作用从而引起地磁场的扰动.但到目前为止,人们对上述整个物理过程的认识及行星际扰动观测资料的积累等方面仍是非常欠缺的. 因此,本文将注重于 C_9 指数本身所包蕴的规律而暂不将行星际扰动作为预测系统的输入量. 同时为了消除瞬变扰动的影响,突出共转扰动特征,故在本文中对 C_9 指数经过了时窗为五天的加权滑动平均(这与[1]中处理方法相同),即

$$C_9'(t) = \mathrm{Int}[0.15 \times [C_9(t+2) + C_9(t-2)] + 0.2 \times [C_9(t+1) + C_9(t-1)] + 0.3 \times C_9(t)]. \tag{1}$$

上式中的 Int 表示取整数函数. 我们将一个完整 Bartels 周的 C_9' 指数作为预测系统的输入量,记为 $\vec{X}i:\{X_1^i, X_2^i, \cdots, X_N^i\}$, 其中 $N=27$, 上标 i 表示不同的事件.

2. 输出量和输出形式

预测系统的输出量即为 C_9' 指数. 将与输入事件相衔接 14 天的 C_9' 指数作为输出量. 记为 $\vec{y}^i:\{y_1^i, y_2^i, \cdots, y_L^i\}$. 其中 $L=14$, 上标 i 表示不同的事件.

由于 MLP 为二值输出,同时考虑到预测系统的容错性并使系统的拓扑结构尽量简单,我们采用分组求和制描述所预测的 C_9' 指数, 即

$$y_j^i = \sum_{k=1}^{K}\left[(M+1)^{k-1} \cdot \sum_{m=1}^{M} V_{(j-1)\cdot M\cdot K+(k-1)\cdot M+m}\right]. \tag{2}$$

为了简化 MLP 的拓扑结构,我们做如下定义:

$$C_9' = \begin{cases} C_9, & C_9 \leqslant 8 \\ 8, & C_9 = 9 \end{cases} \tag{3}$$

于是式 (2) 中的 $K=M=2$. 即每个预测的 C_9' 指数对应于 MLP 的 4 个二值输出 $V_{(j-1)\cdot 4+(k-1)\cdot 2+m}$.

3. 传输过程的确定

我们采用 MLP 作为观测过程的传输系统. 所要确定的是 MLP 的拓扑结构及神经元的状态转移函数和学习训练方法. 由于 MLP 本身的研究仍是不完备的,故其拓扑结构与所研究问题之间的对应关系没有明确而严谨的结论. 因此,在研究过程中我们先后采用了十三种拓扑结构不同的 MLP 进行尝试.

我们采用了双极性压缩 Sigmoid 函数作为 MLP 各隐层及输出层单元的状态转移函数. 并采用了串行的、带附加冲量项的 BP 法作为 MLP 的学习方法. 这两项设定的目的在于加快学习时的收敛速度,并避免可能导致学习失败的"麻痹"现象出现. 实践证明上述两项设定是有效的.

三、MLP 对 C_9' 指数的预测

为了探求适用的 MLP 的拓扑结构，我们采用少量的学习事件和预测检验事件对十三种 MLP 进行训练并对比检验其预测效果。学习事件和预测检验事件均由从 1974 年和 1976 年中各选取的五个不同的一个半 Bartels 周的 C_9' 指数所组成。其中一个完整 Barterl 周的 C_9' 指数作为学习或预测时的输入事件，相继的 14 天 C_9' 指数作为学习时的指导输出或预测检验时的标准输出。通过计算我们发现：(1) 这十三种 MLP 在学习时均很快收敛，即它们均能精确地认知学习事件。(2) 其预测效果均较理想且差异不大。说明它们均较好地把握了这十个学习事件和十个预测事件的某种内在规律并将其表现出来。

为了简便起见，我们仅给出其中有代表性的五个 MLP 的拓扑结构及其预测检验结果。

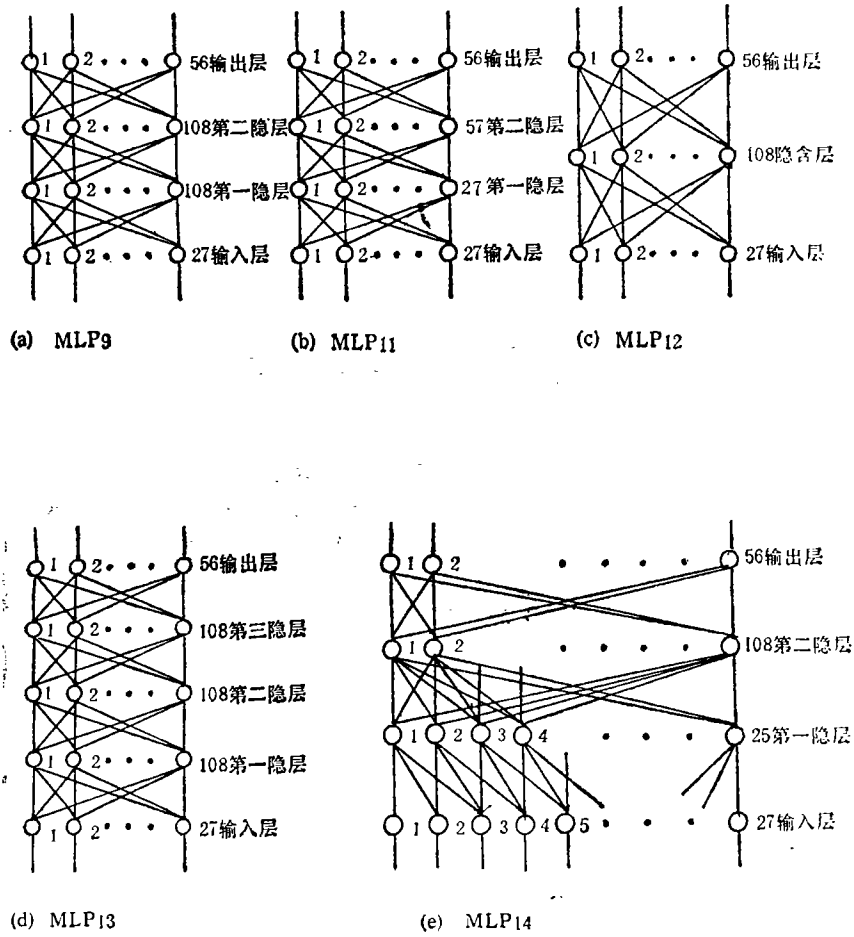


图 1　五个 MLP 的拓扑结构

Fig. 1　The Topology of 5 MLP

为了更好地对比不同 MLP 的预测效果，我们引用周晓燕论文*中的预测效率的定义：

$$Eff(i) = 1.0 - \frac{(\bar{y}^{i*} - \bar{y}^{i})^T \cdot (\bar{y}^{i*} - \bar{y}^{i})}{\bar{y}^{i*T} \cdot \bar{y}^{i*}} \tag{4}$$

及

$$Effo = \frac{1}{NP} \sum_{i=1}^{NP} Eff(i). \tag{5}$$

其中 $\bar{y}^i$ 为不同预测事件，$\bar{y}^{i*}$ 为相应的标准输出事件．上标 T 表示转置．NP 为预测事件数．$Eff(i)$ 为每个预测事件的预测效率．$Effo$ 为所有预测事件的平均预测效率，它们越趋于 1 说明预测效率越好．

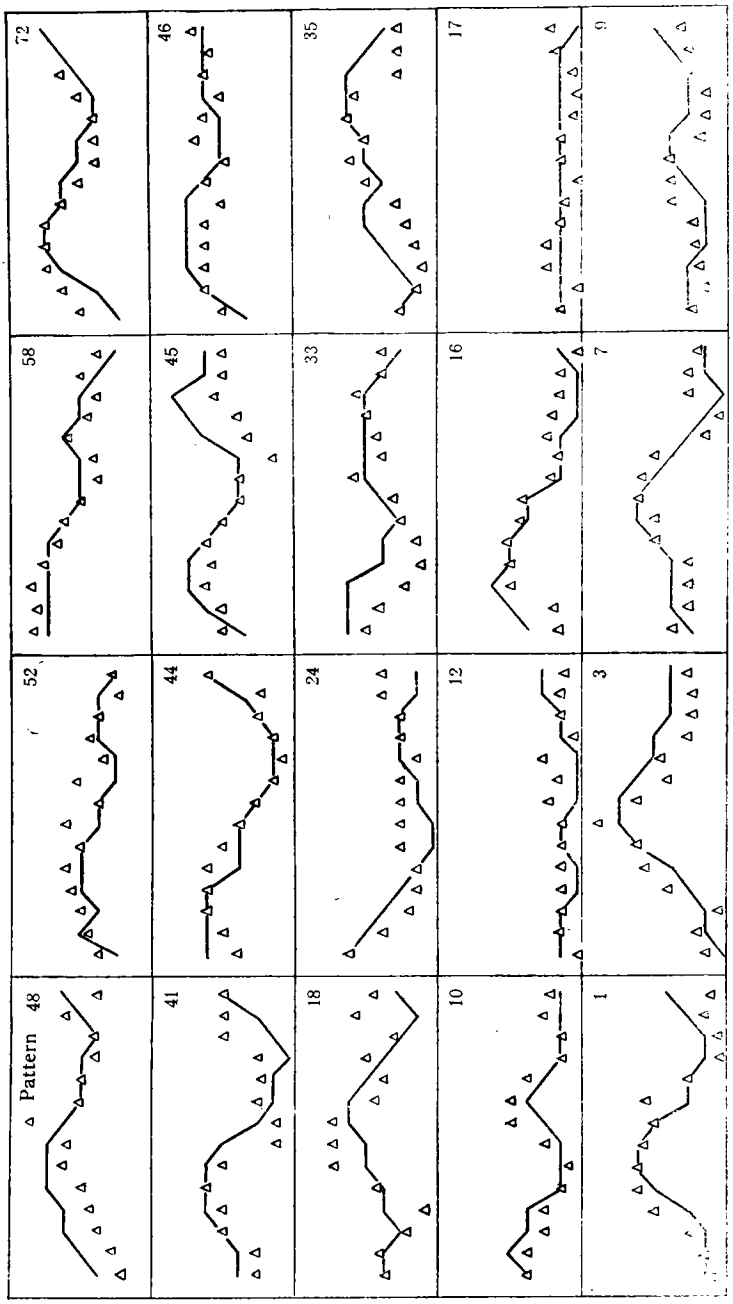
从表 1 可以看出，尽管五个 MLP 的拓扑结构有很大的差异，但其预测效率的差异却不显著．它说明如 MLP12 那样简单的拓扑结构已经基本满足了本问题的需要．在此基础上再增加网络层数及人工神经元的数量，其预测效率也不会有显著的提高．而考虑到在 MLP12 的基础上再增加的隐含层数和神经元个数会大大增加计算时的空间及时间的开销，这便存在一个预测效率与计算效率相配合的最佳选择问题．即在确定本问题所适用的 MLP 的最佳拓扑结构时，不仅要考虑其预测效率的好坏，亦要考虑到其计算效率．这个问题对我们的研究来讲并不很重要，但当把本研究的结果推进到实用预报时便是个必须要考虑的问题．

通过预测效率的对比，发现 MLP9 的预测效率最好．于是我们又从 1963 年至 1987 年共 25 年的期间内选取了 91 个连续的一个半 Barterl 周的 C_9' 指数做为学习事件对 MLP9 进行了学习训练．在学习了 1036 遍后收敛，即 MLP9 无误差地认知了这 91 个学

表 1 五个 MLP 的预报效率

		MLP9	MLP13	MLP14	MLP11	MLP12
	平均效率	0.900	0.886	0.841	0.882	0.865
事件序号	事件起始时间	预 报 效 率				
3	740210	0.884	0.927	0.934	0.886	0.915
4	740405	0.972	0.972	0.875	0.973	0.952
5	740722	0.933	0.961	0.896	0.931	0.944
6	740914	0.922	0.925	0.850	0.883	0.866
7	741107	0.891	0.809	0.843	0.873	0.886
8	760209	0.961	0.966	0.935	0.964	0.976
9	760403	0.898	0.898	0.887	0.893	0.632
10	760623	0.755	0.666	0.680	0.677	0.805
11	760816	0.911	0.908	0.902	0.908	0.868
12	761105	0.869	0.832	0.606	0.835	0.808

* 周晓燕，地磁扰动的系统辨识及其预报方法的研究，理学博士学位论文，1991.



——检验事件　△预测事件

图 2　MLP 对 1964—1987 年部分 C_9' 指数预测结果

Fig. 2　Part of C_9' Index prediction made by MLP9 from 1963 to 1987

表 3 MLP9 对 96 个事件的预报效率 平均效率 0.684

No	Eff	No	Eff	No	Eff	No	Eff	No	Eff
1	0.815	2	1.000	3	0.902	4	1·000	5	0.651
6	0.852	7	0.897	8	1.000	9	0.788	10	0.831
11	0.369	12	0.036	13	1.000	14	0.660	15	0.134
16	0.825	17	0.376	18	0.769	19	0.571	20	1.000
21	0.741	22	0.617	23	1.000	24	0.672	25	0.638
26	0.244	27	0.361	28	1.000	29	0.273	30	0.099
31	0.640	32	1.000	33	0.856	34	1.000	35	0.847
36	1.000	37	0.894	38	0.741	39	0.048	40	0.425
41	0.775	42	0.850	43	0.793	44	0·934	45	0.905
46	0.963	47	1.000	48	0.908	49	0.768	50	0.877
51	0.669	52	0.883	53	1.000	54	0.786	55	0.795
56	0.748	57	0.900	58	0.977	59	1.000	60	0.695
61	0.714	62	0.879	63	1.000	64	0.592	65	0.735
66	0.631	67	1.000	68	0.767	69	0.723	70	0.663
71	1.000	72	0.933	73	0.762	74	0.771	75	0.763
76	0.845	77	1.000	78	0.794	79	0.698	80	1.000
81	0.716	82	0.853	83	0.681	84	0.583	85	0.808
86	0.659	87	1.000	88	0.896	89	0.843	90	0.454
91	0.667	92	1.000	93	0.530	94	0.558	95	0.688
96	0.667								

习事件．然后，我们又从该期间内选取了 96 个预测事件对 MLP9 进行预测效率的检验．在挑选学习及预测检验事件时避开了太阳活动的上升期及高峰期．

由于 MLP 在学习时精确地认知所有的学习事件，因此表 2 中所列的 20 个与学习事件相同的预测事件的预测效率均为 1.0．除去这部分事件后，预测效率在 0.6 以上的预测事件有 60 个，占 78.9%；预测效率在 0.7 之上的预测事件有 42 个，占 55.2%；平均预测效率为 0.684． 显然，这一平均效率不如 MLP9 对 1974 年和 1976 年 C_1' 指数的预测效率．其直接解释为 1974 年和 1976 年地磁扰动的 27 天重现性显著．

四、讨 论

MLP 是一种有效的数学工具，具有很强的认知能力，如 MLP9 能精确地认知和重建其所学的选自 25 年 91 个学习事件，这是其它诸如拟合、回归等方法所不能比拟的．但与其认知能力相比，MLP 所表现的判断能力要逊色一些．如表 2 中所列出的几个预测，预测效率相当低．这是由于在我们的研究中所设定的 MLP 的拓扑结构及输入量和输入形式比较适合于对 C_1' 指数 27 天重现性的描述，而对于 27 天重现性不好的事件来说，其所适用的 MLP 的拓扑结构及输入量、输入形式的设定仍有待于进一步的研究，这也涉及到对 C_1' 指数本身的研究．

在我们的研究中仅利用 MLP 建立了输入事件与输出事件在数据空间中的映射，即经过学习，MLP 仅掌握了数据空间所表现出的能较容易被认识的规律． 前面曾指出，

MLP 不是一种抽象、理性的方法.它虽能精确地建立起数据空间中的映射,但却不能很好地归纳、提取出数据空间中所包蕴的更深刻、更普遍的规律. 为了弥补 MLP 的这一弱点，必须采用其它方法提取出这种更深刻、更普遍的规律，此即是数据的预处理. 它是 MLP 在实际问题应用中的一个重要环节. 具体地说就是采用某种数学方法将输入、输出事件从数据空间转换到某种特征空间中去. 然后再利用 MLP 所特有的长处来建立特征空间中的映射. 我们认为这是今后的研究方向之一.

最后,在我们设计 MLP 的输入量时舍去了行星际物理场参量. 显然,这样做从物理上讲是不完备的. 其结果是 MLP 对于 C_9 指数所描述的大量瞬变扰动无法认知和预报. 因此,以何种行星扰动参量或太阳活动参量的何种形式的表述作为 MLP 的输入量,使其能精确认知地磁场的瞬变强烈扰动并对其进行报预是我们今后将要深入研究的问题.

参　考　文　献

[1] 章公亮,高玉芬,陆晨和冯忠孝,中国科学A辑,第8期,第883页,1986.
[2] 钟义信,潘新安和杨义先,«智能理论与技术——人工智能与神经网络»,北京,人民邮电出版社,1992.
[3] 杨行峻和郑君里,«人工神经网络»,北京,高等教育出版社,1992.
[4] 周继成等,«人工神经网络——第六代计算机的实现»,北京,科学普及出版社,1993

THE APPLICATION OF ARTIFICIAL NEURAL NETWORKS TO THE PREDICTION OF GEOMAGNETIC INDEX

Gu Zuo-wen　Gao Yu-fen　Fan Guo-hua
(Institute of Geophysics, State Seismological Bureau, Beijing 100081)

Abstract

The application of Aritficial Neural Networks to the prediction of Geomagnetic Index is introduced in this paper. The topology of one of such Networks-Multi-Layer perceptron, which is applied to the prediction of C_9 Index, the efficiency of prediction and the features and problems in the practice of prediction are discussed.

Key words Aritficial Neural Networks, Prediction of Geomagnetic Index, C_9 Index