

用神经网络方法对雷达资料 进行降水类型的分类

王 静 程明虎

(中国气象科学研究院,北京 100081)

提 要: 利用不依赖先验统计模型的多层前馈神经网络模型对合肥的新一代 S 波段 A 系列雷达 2001—2003 年的降水资料进行了三种降水类型的分类,并将训练完成后的网络应用于一次降水过程。利用单隐层的多层前馈神经网络模型,在取适当参数时,已经可以较好地对雷达资料进行对流云降水、层状云降水和混合云降水三种降水类型的分类。同时验证了:训练集样本的数量和顺序、隐层神经元的数目以及学习率的选择等都将影响分类的成功率。

关键词: 神经网络 雷达资料 降水类型

Precipitation Echo Classification of Radar Reflectivity with Artificial Neural Network

Wang Jing Cheng Minghu

(Chinese Academy of Meteorological Sciences, Beijing 100081)

Abstract: A Back-Propagation (BP) Model of Artificial Neural Network (ANN) is used for the partitioning of radar reflectivity into convective and stratiform-cloud precipitation classifications with the CINRAD-SA data from 2001 to 2003 in Hefei. The trained ANN is applied in a precipitation process. It is proved that the single hide-layer BP model of ANN can be used to classify the different precipitation echoes with a high success-rate. It is also validated that: the success-rate is influenced by following factors: the amount and the in-put-order of the training-database, the nerve cell number of the hided-layer and the choice of the learning rate.

Key Words: ANN CINRAD-SA radar data precipitation classification

引 言

我国是一个雨季集中、暴雨多发的国家,对于降水的研究早已超过了一个科学研究的范围。对降水时空分布的了解,有着深远的社会影响,并且可以产生明显的经济效益^[1]。

依据 $Z-I$ 关系式($Z=AI^b$)^[2],理论上可以由雷达接收的回波强度 Z 定量估测出目标降水区域或单点的降水强度 I 。合理的 $Z-I$ 关系主要取决于降水系统中滴谱的分布状况,然而滴谱分布在一般情况下是随时间和空间,特别是随不同的降水类型而变化的。对于某一区域,在某一季节, $Z-I$ 关系主要取决于降水类型。因此不同降水类型的区分,对于得到更合适的 $Z-I$ 关系,以便更精确的进行雷达定量测量降水有着非常重要的意义。

对于如何进行不同降水类型的分类,现有的方法总的来说,分为确定性方法和统计性方法两大类^[3-4]。但是这两类方法都基于两点:第一,基于一个假设,即地面雨量计的观测值是正确的;第二,依赖于判断者的经验。因此,在理论上和推广应用上都存在着不足。

随着计算机技术的发展,雷达图像的识别已经由目视判读向自动化发展。传统的计算机识别方法基于统计决策理论,当假设的统计概率模型不成立时,其识别性能会大大下降。神经网络的发展解决了这一问题。在国内外的研究中,神经网络分别被应用于对卫星资料进行云分类^[5-7]以及进行天气预测^[8-10],都取得了很好的结果。随着我国已可获取高质量、稳定的雷达资料,将神经网络方法应用于雷达资料降水类型的分类,已日益可行而重要。

1 人工神经网络模型及 BP (Back-Propagation) 算法

1.1 人工神经网络模型

神经网络模拟人脑神经系统的工作原

理,它通过学习获得知识,神经网络学习的过程就是各个权和神经元的自动调节过程。调节的准则是使神经网络的输出与实际的模式尽可能地接近,如果有误差,则自动调节权使误差减小直至满足要求。与传统的统计方法相比较,该方法不需要获取待识别模式的先验概率。

人工神经网络目前已有几十种网络模型。根据不同人工神经网络的特点和研究的实际问题,本试验选择多层前馈网络,其模型如图 1 所示。本文的传递函数(也称激活函数)选择连续可微的 Sigmoid 函数。

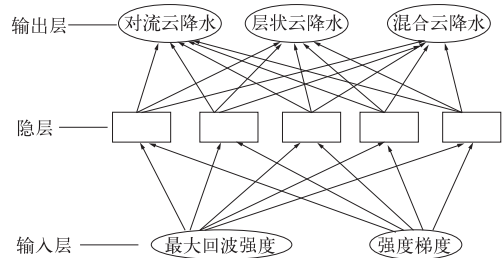


图 1 BP 网络模型

1.2 BP 算法

在多层前馈网络中,对每一个(输入、输出)模式对,其误差平方 E 为

$$E = \frac{1}{2} \sum_l (L_l - O_l)^2 \quad (1)$$

下标 i 代表第 i 个神经元, L 代表理想模式, O 代表多层前馈网络的实际输出。

利用梯度下降法,输出层权的调节量 δT_{ij} 为

$$\delta T_{ij} = AD_i O_j \quad (2)$$

$$\text{其中 } D_i = O_i(1 - O_i)(L_i - O_i) \quad (3)$$

式中 A 是学习的步长,称为学习率, O_j 是与输出层相连的第 j 隐层的各个神经元的输出, D_i 为输出层的误差项。

对于隐层,

$$D_k = O_k(1 - O_k) \sum_l D_l T_{lk} \quad (4)$$

式中 D_k 为 k 隐层的误差项, O_k 是比 k 隐层低一层的神经元的输出值, T_{lk} 为 l 层与 k 层之间的连接权重。

权 T 的调节是通过将误差项反向传播完成的。通过调节 T 使输出的误差 E 逐渐减小直到满足要求。该算法也叫 Back-Propagation 算法, 简称 BP 算法^[5-7]。

2 降水类型分类原理

根据不同类型降水回波在 PPI (Plane Position Indicator) 上的特征及神经网络方法的可行性, 本文仅选取了三种降水类型的回波强度的特征作为识别因子, 取最大中心回波强度 Z_{\max} 及回波梯度 G 两个特征量来进行判断。其中回波梯度为最大回波强度中心点与其周围 8 个格点的回波强度差的绝对值的平均值。

通过对所用资料的仔细分析, 因为论文所用的对流云降水和混合云降水的资料来自中国暴雨项目外场试验, 且观测资料的时间都是在 2001—2003 间每年的 6 月中旬到 7 月下旬, 观测的个例都是发生在江淮地区的暴雨过程, 因此对流云降水回波的 Z_{\max} 及 G 就有较大的值。而混合云降水回波中的则存在着密集的对流泡, 且存在多个强度很大的 Z_{\max} , 而 G 却不是很大。同时, 由于层状云降水资料的观测时间是冬季, 因此层状云的降水回波中, 只存在较小的 Z_{\max} , 而由于层状云降水比较均匀, 因此其 G 值也相应的较小。训练网络时对三种降水类型的定义分别为: 层状云连续性降水: $Z_{\max} < 30\text{dBz}$; $G < 5\text{dBz}$; 对流云降水: $Z_{\max} > 45\text{dBz}$; $G > 10\text{dBz}$; 混合云降水: $Z_{\max} > 40\text{dBz}$; $G < 5\text{dBz}$ 。

3 人工神经网络的训练过程

3.1 资料

本文中所用的资料来自位于合肥的新一

代 S 波段 A 系列雷达, 该雷达位于 $31^{\circ}52'1''\text{N}$ 、 $117^{\circ}15'28''\text{E}$; 天线的海拔高度为 165m, 最大遮挡仰角为 1.5° , 最佳观测仰角为 0.5° ; 其天线增益为 45.69dB, 峰值功率为 690kW, 波长为 10cm。

所用资料是利用该雷达的原始体扫资料经过补漏测方位处理、水平方向双线性插值、垂直方向线性插值后得到的直角坐标下的 1~15km 的不同层次的回波强度的 CAPPI (Constant Altitude PPI) 资料。该资料的水平格距为 1.5km, 垂直格距为 1km。其中, 对流云降水和混合云降水用的是 2001—2003 年中国暴雨项目外场试验的资料 (2001 年是 6 月 17 日—7 月 15 日; 2002 年是 6 月 10 日—7 月 15 日; 2003 年是 6 月 22 日—7 月 17 日); 层状云降水用的则是 2003 年 11 月 5 日—12 月 4 日雷达探测层状云降水所得的资料。

3.2 训练集和检验集的建立

从已有的研究成果中得知, 神经网络训练集的样本质量和数量将直接影响网络的推广能力。因此所建立的训练集应当尽可能的包含所要分类三种降水类型中的各种情况; 同时, 为了网络有较快的学习速度, 训练集一定要全面且大小合适。

考虑到所研究对流系统的水平尺度及层状云、对流云系统的垂直特征, 从上述 1~15km 的各层次资料中选择 5~10km 间的层次的数据, 并且从各层资料中选取了具有代表性的 $30\text{km} \times 30\text{km}$ 范围的典型的层状云降水、混合云降水和对流云降水的回波强度矩阵各 100 个, 且每个矩阵中只有一个对流中心最大值存在, 来建立数据库。各类型的 100 个回波强度矩阵中取 20 个建立训练库, 剩下的 80 个作为检验库。

3.3 网络结构设计

基于 BP 算法的多层前馈网络的输入层、

输出层的设计完全根据问题的要求而定^[5-7]。本文所用的 BP 网络结构如图 1 所示。输入层神经元数目等于样本特征数据的维数,因此为 2。输出层根据所要识别的类型设计,因此设为 3 个神经元。

关于隐层的设计,目前还没有理论上的指导。基于单隐层的网络已经能够完成所需分类,因此本论文选了单隐层结构。隐层神经元的数目靠实验来确定。实验结果如表 1 所示。从表中可见,当单隐层取 6 个神经元时,网络的分类效果最好。

表 1 BP 网络隐层神经元数目实验结果

网络结构	系统误差	识别率	
		训练集	测试集
2—2—3	0.054295	38.33%	22.50%
2—4—3	0.005872	83.33%	79.58%
2—6—3	0.000384	98.33%	96.25%
2—8—3	0.001562	91.67%	87.08%
2—10—3	0.083741	78.33%	49.58%

3.4 参数的选择

学习率 A 对网络的收敛效果影响很大。借鉴诸多已有的研究结果^[5-10],本文中,将 A 设为 $[0, 1]$ 区间内的一个常数。通过多次实验得知,当 $A=0.7$ 时,网络的收敛效果较好。

因为处理不同的问题,甚至处理同一问题而训练集不同时, A 都将取不同的最优值。所以,在以后的工作中,作者将尝试 A 随学习过程变化的自动优化算法。

3.5 样本的输入顺序对网络性能的影响

在以前的研究中,曾得出结论:样本的输入顺序将会对网络的学习效果产生很大的影响。网络能很好的记忆后输入的样本,先输入类别信息被后输入的类别信息所覆盖,出现先学后忘的现象^[5-7]。本文的研究结果,验证了这一结论。

4 训练完成的 BP 网络在一次降水过程中的应用

4.1 过程介绍

此次降水过程取自中国暴雨项目外场试验个例库,其从 2001 年 6 月 16 日 15 时开始至 2001 年 6 月 19 日 08 时,是长江中下游一次中低层切变线与中尺度对流系统的相互作用造成的过程降水。该过程资料用的是 2001 年中国暴雨试验项目的雷达回波资料,观测起始时间为北京时间 2001 年 6 月 16 日 21 时 01 分,观测结束时间为世界时 2001 年 6 月 19 日 05 时 53 分,每隔 5 分钟左右观测一次。

4.2 BP 网络对对流性降水回波的识别

北京时间 2001 年 6 月 17 日 09 时 30 分,合肥雷达探测此降水过程的回波强度和径向回波速度图(图 2,见彩页)。从图上可以很明显的看到在该雷达的正东方向有两个强的对流性系统,其系统的中心 Z_{\max} 的回波都在 50dBz 以上,且 G 值也较大。

取该时刻的 CAPPI 资料,从 5~10km 层资料中的 5km、7km 和 9km 三个高度层分别以 Z_{\max} 为中心取 30km×30km 的回波强度矩阵,其中 9km 处的降水回波已很弱,故舍弃(图 3 和图 4,见彩页)。

将上述矩阵输入训练好的 BP 神经网络进行识别,分类结果均为对流性降水。

4.3 BP 网络对混合性降水回波的识别

北京时间 2001 年 6 月 18 日 03 时 04 分,合肥雷达探测此降水过程的回波强度和径向回波速度图分别如(图 5,见彩页)所示。从图上可以很明显的看到在该雷达的西北方向有大片的混合云降水系统,其系统的中心 Z_{\max} 都在 50dBz 以上,但其 G 值并不大。

取该时刻的 CAPPI 资料,从 5~10km 层资料中的 5km、7km 和 9km 三个高度层分别以 Z_{\max} 为中心取 $30\text{km}\times 30\text{km}$ 的回波强度矩阵(图 6 和图 7,见彩页)。

将上述矩阵输入训练好的 BP 神经网络进行识别,分类结果均为混合性降水。

5 结 语

(1) 神经网络在对雷达资料进行三种降水类型的分类时可以达到较高的成功分辨率。当取适当参数时,对训练集和测试集都可取得 95% 以上的成功分辨率。

(2) 神经网络进行识别时不需要知道降水类型的先验统计模型,且所用特征维数仅为 2。因此,其相对于传统的统计方法更利于推广使用。

(3) 验证了以往研究中所得的一些结论:神经网络的性能受其结构和参数的影响较大;训练集的样本质量和数量是一个关键问题;训练网络时,样本的输入顺序将会很大的影响网络的学习效果。

参考文献

- [1] 王小兰,程明虎,崔哲虎. 对流性降水云辐射特性研究[J]. 气象, 2005, 31(9): 3-7.
- [2] 张培昌,杜秉玉,戴铁丕,等. 雷达气象学[M]. 北京: 气象出版社, 2001: 511.
- [3] Demott, C. A., R. Cifelli, and S. A. Rutledge, An improved method for partitioning radar data into convective and stratiform components. Preprints, 27th Conf. on Radar Meteorology[J], Vail, CO, Amer. Meteor. Soc., 1995: 233-236.
- [4] E. N. Anagnostou, C. Kummerow, Stratiform and convective classification of rainfall using SSM/I 85-GHz brightness temperature observation[J]. Amer. Meteor. Soc. 1997, 14: 570-575.
- [5] 师春香,吴蓉彰,项续康. 多阈值和神经网络卫星云图云系自动分割试验[J]. 应用气象学报, 2001, Vol(12): 70-78.
- [6] 师春香,瞿建华. 用神经网络方法对 NOAA2AVHRR 资料进行云客观分类[J]. 气象学报, 2002, Vol(60): 250-256.
- [7] 白慧卿,方宗义,吴蓉璋,等. 基于人工神经网络的 GMS 云图四类云系的识别[J]. 应用气象学报, 1998, Vol(9): 401-409.
- [8] 张承福. 人工神经网络在天气预报中的应用研究[J]. 气象, 1995, Vol(20): 43-47.
- [9] M. Hessami, F. Ancil and Alain A., Selection of an artificial neural network model for the post-calibration of weather radar rainfall estimation[J]. J. Data. Sci., 2004, 107-124.
- [10] R. J. Kuligowski and A. P. Barros, Experiments in short-term precipitation forecasting using artificial neural networks[J]. Amer. Meteor. Soc., 1997, 470-482.

王 静等：用神经网络方法对雷达资料进行降水类型的分类

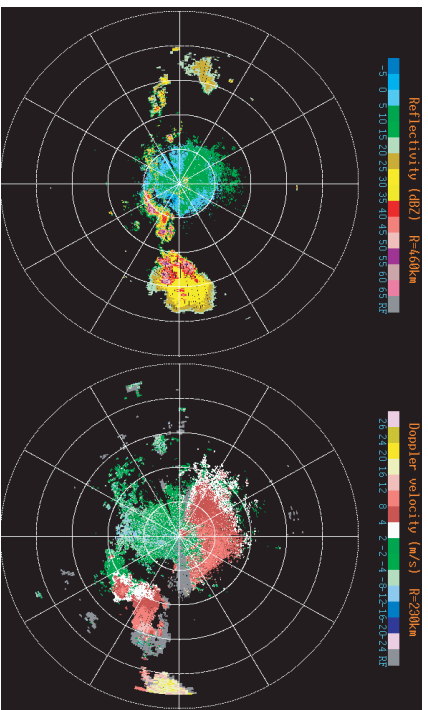


图 2 2001年6月17日09时30分合肥多普勒雷达回波强度和径向回波速度图(仰角: 0.5°)

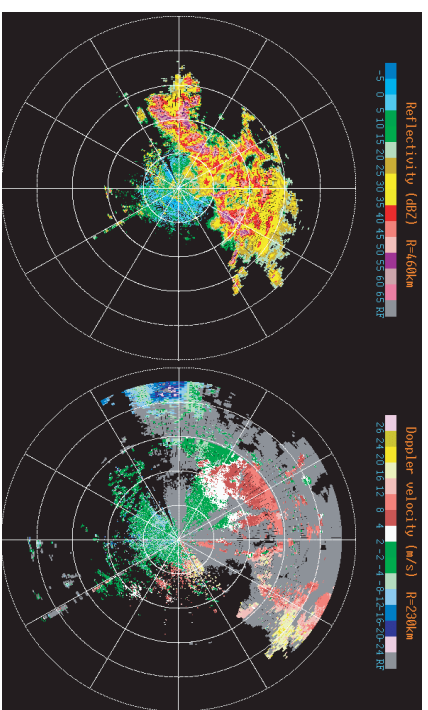


图 5 2001年6月18日03时04分合肥多普勒雷达回波强度和径向回波速度图(仰角: 0.5°)

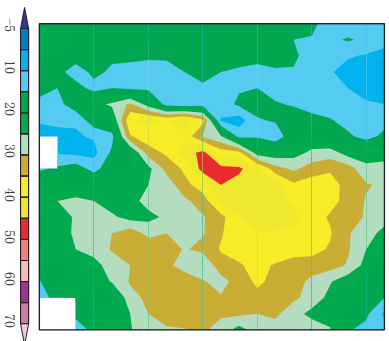


图 3 2001年6月17日09时30分7km高度回波强度图

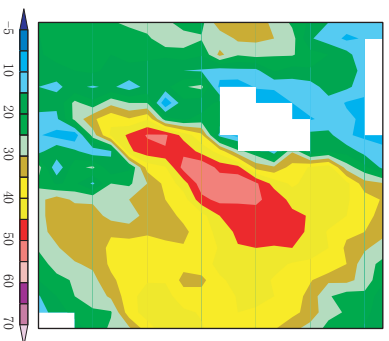


图 4 2001年6月17日09时30分5km高度回波强度图

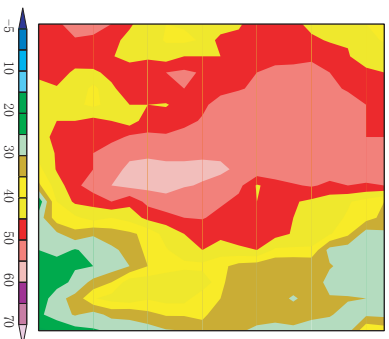


图 6 2001年6月18日03时04分7km高度回波强度图

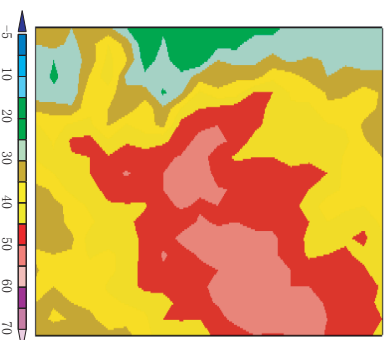


图 7 2001年6月18日03时04分7km高度回波强度图