

神经网络方法在静止卫星多通道资料 估算降水中的应用^①

熊秋芬^②

胡江林

(南京大学大气科学系, 南京 210093)

(武汉暴雨研究所)

夏 军

(中国科学院环境与水资源研究所)

提 要

使用人工神经网络(ANN)模型探讨了利用静止卫星多通道资料估算地面降水的一种新方法,对淮河和长江典型区域(24~36°N,108°E以东)一次暴雨过程的卫星遥感数据和地面水文站逐时降水资料的应用分析表明,该方法提供的客观定量的降水量估算的平均相关系数为0.57,较现行业务使用的方法效果更佳。

关键词: 人工神经网络 静止卫星多通道资料 定量估算降水

引 言

我国汛期发生的洪水大部分是长时间大范围的强降水造成的。根据大气环流的演变虽然对降水的分布、强度及持续时间等可作一定程度的预测,但现阶段对降水的定点定量监测和预报仍存在较大的困难,预报精度难以满足洪水预报的需求,因而对降水天气的监视和预测、对实况降水量及时全面的了解是防汛防洪工作的一项重要任务,可为洪水预报和防洪调度决策提供有参考价值的信息。然而,由于降水的强度和时空变率很大,现有的常规测站密度远不能完整、有效地观测强降水系统带来的充沛雨量,尤其是高原、山区等常规测站稀少的地区更是如此,这严重阻碍了洪水预报的进一步发展,难以满足对洪水监测和预报的需求。

自1960年气象卫星问世以来,利用气象卫星探测的高时空分辨率资料追踪和分析各

种尺度的天气系统和估算降水工作一直受到十分重视,各种利用卫星产品实时估算降水量的方法和系统不断提出和发展改进^[1]。美国NOAA的NESDIS发展了一个利用GOES红外资料估算降水量的系统并于1997年投入业务运用^[2],我国水利部信息中心使用云分类方法对数字云图估算面雨量^[3],国家卫星气象中心使用卫星云图和常规气象资料设计了一种自动估算降水方法^[4]。但由于降水和降水量与许多因子有关,目前国内业务应用中利用的卫星观测资料仍主要是波长为10.7 μm 的红外单通道资料,使用的方法也多为阈值法或线性回归等简单方法,因而估算的降水与实测降水之间仍有较大的差距。利用卫星的多通道资料和进一步采用更复杂的非线性方法来改进实况降水量的估算应仍有很大的潜力。

近年来,神经网络模型在模式识别、

^① 本文由湖北省气象局“利用卫星和地面气象资料估算面雨量研究”课题资助
^② 工作单位:湖北省武汉中心气象台

图像处理、信号分析和自动控制等应用研究领域取得了重要进展。人工神经网络技术为利用 GMS 卫星的多通道资料估算降水量提供了一条新的途径。由于神经网络是模拟生物神经网络的人工智能技术,具有并行处理、非线性、容错性和自适应、自学习的特点,因而广泛应用于模式识别、信号处理和自动控制以及遥感图象的分类。本文探讨了利用 GMS 卫星 4 个通道(远红外、近红外、水汽通道和可见光通道)的资料,采用非线性的人工神经网络方法来估算单站降水量,以提高降水的单站和面雨量的估算精度并用于洪水预报,进而提高暴雨和洪水的实时监测和短时降水的预报精度,延长其有效预见期。

1 人工神经网络方法及特点

1.1 人工神经网络方法简介

人工神经网络是一个可自动提取一组预报变量和另一组自变量之间非线性关系的数据处理系统,该网络的建立过程称为人工神经网络的训练过程,是用自适应算法递归迭代求解因变量与自变量之间的非线性关系。训练后的神经网络模型可用来估算或预报预测变量。

到目前为止,已经出现许多神经网络模型及相应的学习方法。其中反向传播神经网络模型(简称 BP 模型)是应用较广泛的一种,这种模型除输入层和输出层之间插入若干个隐含层,相邻层次之间的神经网络元之间用连接权系数作相互连接,而各层内的神经元之间没有连接,本文使用的神经网络模型只含 1 个隐含层, $f(x)$ 为网络激活函数采用 S 型函数:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (1)$$

具体的算法可参阅文献[5]。

1.2 BP 神经网络的特点

与线性模型比较,神经网络模型主要具有以下优点:

(1)具有强大的学习功能:神经网络的连接权和连接结构都可通过网络训练得到,并不需人们预先掌握两组变量之间的关系,因而建立的网络模型完全是客观定量的。

(2)非线性:虽然线性回归模型中可通过因子之间的非线性变换来发现非线性关系,但非线性变换需人为地预先确定。神经网络的非线性是通过网络激活函数 $f(x)$ 导入的,可自动解决非线性问题。

(3)神经网络的容错性和冗余性:容错性是指如果训练的样本足够多,则即使部分样本数据出错,神经网络可自动识别这种错误并更正这些错误。冗余性则是如果训练样本足够多,即使网络的输入部分有相关不大的因子,也不会严重影响网络模型的预报效果。

2 实例

2.1 资料和处理方法

1996 年 6 月 28 日长江和淮河流域发生了一次大暴雨过程,1 小时最大雨量达 64.2mm。这里选取 1996 年 6 月 28 日 08 时至 18 时逐时的 GMS 4 通道数字化资料和地面 1111 个水文测站的逐时雨量资料来建立神经网络模型,该模型的输入是 4 个通道的数字化资料和地面测站的经度、纬度(6 个节点),输出是地面降水量(1 个节点),隐含层节点为 6。计算框图如图 1 所示。

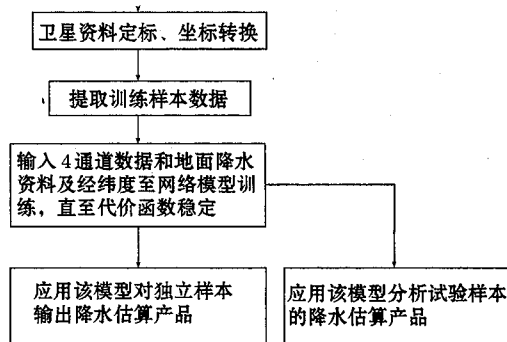


图 1 数据计算流程图

输入神经网络进行训练的所有样本的输入输出因子 X_i 均作如下 0、1 归一化处理:

$$a_i = \frac{X_i - X_{\min}}{d} \quad (2)$$

其中 $d = 1.05(X_{\max} - X_{\min})$, 这里 X_i 是输入或输出因子, X_{\max} 、 X_{\min} 为因子 X_i 的最大值和最小值, a_i 即为标准化因子。

2.2 模型训练效果分析

将 08~18 时共 11 个时次的 12221 个样本用神经网络模型进行训练, 为检验神经网络模型的训练效果, 我们将神经网络模型的估计降水量与实测降水量和文献[3]的估计降水量作对比分析, 文献[3]是水利部信息中心使用云分类方法对数字云图估算面雨量的业务方法, 该方法的要点是将卫星云图上的云分门别类, 对每类云根据气候概率确定其降水量。

表 1 给出了 08~18 时共 11 个时次的神经

表 1 各时次估计的降水与实测降水之间的相关系数

时次	神经网络	文献[3]
08	0.33	0.27
09	0.54	0.43
10	0.66	0.28
11	0.75	0.27
12	0.71	0.24
13	0.69	0.22
14	0.63	0.24
15	0.44	0.07
16	0.50	0.26
17	0.46	0.20
18	0.52	0.27
平均	0.57	0.25

神经网络模型与文献[3]估计的全部 1111 个测站与实测降水量之间的相关系数。从表 1 中可见, 所有相关系数都通过信度为 99.9% 的检验, 且神经网络模型所估算的结果在各个时次的相关系数都大于文献[3]分类方法所得的相关系数, 说明神经网络模型的估计效果要好于文献[3]的分类方法的效果。特别是 10~15 时这 6 个时次, 前者的相关系数都比后者高 1 倍以上, 有时达到高 2 倍以上, 文献[3]分类方法除第 9 时的分类方法较好达 0.4 外, 其余的均在 0.3 以下, 而神经网络模型除第 8 时的系数在 0.4 以下外, 其余有 10 个时次都在 0.4 以上, 且最高达 0.75。11 个时次的平均相关系数更显示神经网络模型方法比文献[3]有质的提高, 即相关系数从 0.25 提高到 0.57。

进一步对比每个时次逐站两种方法的降水量与实测降水量之间差别, 从中可发现文献[3]方法所估计的降水测站数偏多, 降水量偏小, 而神经网络模型所估计的降水测站数比文献[3]方法大大减少, 且降水量与实测降水量较接近。例如图 2 给出了 10 时的部分测站的实测降水量、神经网络方法估计量和文献[3]方法所估计的降水量三者之间的对比。从图 2 中可见对应实测降水大的测站, 神经网络方法估计的降水也相应较大; 对应没有实测降水的测站, 神经网络方法估计的降水也相应为 0。而文献[3]方法所估计的

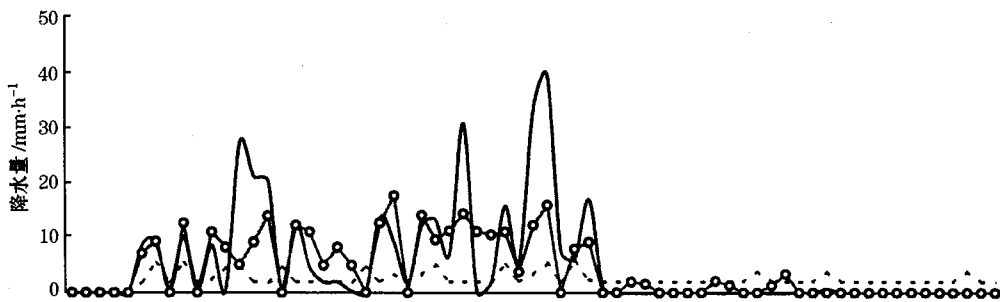


图 2 部分测站的实测降水和两种方法估计的降水量系列

峰值最大的为实测降水序列, ○○线为使用神经网络方法估计的降水序列, 虚线为使用文献[3]方法估计的降水序列

降水在降水较大的测站只有很小的峰值,而在实测没有降水的地方该方法却估计了很多测站的弱降水。其它时次和其它测站与图2类似。图2的情况基本反映了两种方法所估计的降水与实测降水之间的关系。但神经网络方法在估计每小时20mm以上的强降水时也存在一定的困难。

下面再分析一下所估计全部测站的平均降水量的情况,表2显示两种方法都对平均降水量有一定程度的高估,神经网络方法约高估35%左右。文献[3]方法高估约55%,显然神经网络模型效果较文献[3]方法好些。进一步分析还可发现前8个时次估计的平均降水量神经网络方法都好于文献[3]方法,而后3时次相反,说明其中的关系较复杂,可能与降水的发生、发展、成熟、减弱等暴雨生命周期有关。

表2 各时次估计和实测的所有测站平均降水量

时次	实测	神经网络	文献[3]
08	0.70	1.08	1.25
09	0.91	1.04	1.27
10	0.92	1.32	1.48
11	0.78	1.26	1.60
12	0.70	1.67	1.71
13	0.85	1.51	1.79
14	1.09	1.13	1.87
15	1.27	1.36	1.74
16	1.47	1.96	1.90
17	1.67	2.94	1.59
18	1.44	2.58	1.74

2.3 独立检验效果分析

要确定模型的预测效果,必须进行独立样本统计检验。

将08~17时共10个时次的样本分成两部分,即用9个时次的样本用神经网络模型进行训练得到网络模型参数,用另外1个时次的样本作独立检验。例如,将09~17时样本训练得到网络模型参数,作出08时的降水估计,再以此类推,直至做出08~17时的降水估计。计算结果如表3和表4所示,每个

时次的独立样本的相关系数都通过信度为99.9%的检验。比较表3和表1可知,使用神经网络方法预测降水的独立样本的相关系数与模型回归训练的相关系数基本相当,即该方案是可行的。独立样本的平均降水量也基本反映了降水的强度,即没有明显的高估或低估平均降水量(可理解为面雨量)。

分降水强度统计的预测和观测的测站数显示,预测和观测具有相同量级或只差一级的测站数最多。模型预测的降水强度与实际较接近。但表中也显示人工神经网络模型仍不能预测每小时大于25mm的降水(这时模型一般预测成10~25mm或2~10mm)。模型不能预测降水峰值的原因可能是训练样本中大降水的样本占总样本的比例偏少,因而模型预测强降水极值的能力较差,这也是各种统计模型都有的弱点。

表3 神经网络模型交差检验的相关系数和雨量预测平均值

时次	相关系数	平均偏差
08	0.31	0.2
09	0.51	0.6
10	0.61	0.7
11	0.70	0.1
12	0.68	-0.4
13	0.71	-0.5
14	0.58	0.0
15	0.29	-0.1
16	0.43	-0.6
17	0.40	-0.5
平均	0.52	-0.0

表4 神经网络模型预测和观测的雨量测站数对比(独立样本统计)

观测	预测				
	无降水	0~2mm	2~10mm	10~25mm	≥25mm
无降水	5706	1544	815	48	0
0~2mm	488	577	634	87	0
2~10mm	103	177	378	204	0
10~25mm	30	39	102	129	0
≥25mm	2	6	19	22	0

3 讨论

本文的工作是利用神经网络模型运用于

洪水暴雨期间估算降水量的一个实例,结果令人鼓舞,利用人工神经网络的学习与记忆能力,探讨多通道的卫星资料信息与地面降水量之间复杂的非线性关系,为利用多通道的卫星资料估算降水量提供了一条新的途径。本文的工作表明:使用多通道的卫星资料 and 人工神经网络方法可以用来估算降水量,与现有使用单通道红外云图资料和简单阈值和函数方法相比有相当的改进。其主要原因一是多通道卫星资料比单通道卫星资料提供了更多的信息;其二是神经网络方法比传统的阈值法或非线性法有更大的优越性,能更好地分析与揭示卫星资料与降水之间的复杂非线性关系。该方法客观自动,实用性较强,估算结果可直接应用于洪水业务预报中。

需要指出的是,由于资料限制,本项工作所做的工作仍是初步的,对不同的季节降水估计的模型参数有可能发生变化。同时,由于降水的复杂性,现阶段仅利用 GMS 4 通道

资料要精确估计每个象素点的降水量也是困难的,进一步加入气象要素和物理量场资料有可能改进估计精度。但无论如何,该方法建立了利用 4 个通道 GMS 资料估算降水的复杂非线性模型,估算的精度较现有业务使用的方法有很大改进,是一种从理论和实践上都大有发展前途的实时估算区域降水量方法。

参考文献

- 1 Adder, F.R., and A.J. Negri. A satellite infrared technique to estimate tropical convective and stratiform rainfall. *J. Appl. Meteor.* 1988, 27:30—51.
- 2 Gilberto A. Vicente Roderick A. Scofield, and W. Paul Menzel. The operational GOES infrared rainfall estimation technique. *Bull. A. M. S.*, 1998, 79:1883—1898.
- 3 杨扬,戚建国. 数字卫星云图估算面雨量的应用试验. *气象*, 1995, 21:35—39.
- 4 江吉喜等. 一种利用卫星资料的实时降水量自动计算方法. *气象科技*, 1984:39—45.
- 5 胡江林. 神经网络模型用于湖北省月降水量的探讨. *暴雨·灾害*, 北京:气象出版社, 1999(1):36—41.

A Rainfall Estimation Technique Based on the Stationary Satellite Mutli-channel Data Using Artificial Neural Network Models

Xiong Qiufen

(Department of Atmospheric Sciences of Nanjing University, 210093)

Hu Jianglin

(Wuhan Heavy Rain Institute)

Xia Jun

(Hydrology and Water Resources Institute, Chinese Academy of Sciences)

Abstract

The rainfall estimation technique based on the artificial neural network model is developed. Using the hourly GMS four channel's data and rainfall records from hydrographical station, the study of a heavy rain case which occurred in Huai River and Yangtze River basins (east of 108°E, from 24° to 36°N) shows that the average correlative coefficient between the quantitative precipitation estimation and the observation rainfall is 0.57, which is far higher than that in operation.

Key Words: artificial neural network stationary satellite multi-channel data quantitative precipitation estimation