

青海省降雨落区预报自动化系统

王繁强¹⁾ 徐文金²⁾ 陈杰伦³⁾ 王 莘¹⁾

(1)青海省气象台, 西宁 810001; 南京气象学院 2)气象学系; 3)计算机与信息工程系, 南京 210044)

摘 要 选用 08 时(北京时间)常规资料和数值预报产品作为系统输入资料, 将关键区要素场作切贝雪夫正交多项式展开, 取切贝雪夫系数作预报因子, 由 B-P 神经网络进行训练。分别建立了全省各片未来 24 h 有无大于(等于)5 mm、10 mm 和 25 mm 降水天气过程的人工神经元降水落区预报网络。从学习结果看, 历史概括率均达 96% 以上, 经 1994 年 6、7、8 三个月的业务试用, 效果较好。

关键词 B-P 算法, 降水落区预报, 自动化系统

分类号 P457.6

人工神经网络的基本处理单元是人工神经元, 它是对生物神经细胞功能的一种机器仿真形式, 通过对人脑神经元结构、特性和大脑认知功能的模仿构成新型信号、信息处理系统。它是由大量简单的非线性元件(类似于人脑中的神经元 Neuron)高度并联互联而成的非线性网络系统。由于气象问题中存在着大量的非线性关系, 采用传统的线性相关处理方法, 显然难以拟合大气运动的非线性特点。因此, 如果利用人工神经网络的非线性决策特性来处理天气预报问题, 将更符合大气运动的非线性特点, 具有明显的优越性, 王耀生^[1]、张承福^[2]等人在这方面已作了大量的应用研究工作。

我们用切贝雪夫正交多项式将气象要素场展开, 以切贝雪夫多项式系数场的特征为预报因子, 作为 B-P 网学习的输入资料, 取连续单调增 S 函数作为神经元的处理函数, 由 B-P 神经网络进行训练。分别建立了全省各片未来 24 h 有无大于(等于)5 mm、10 mm 和 25 mm 降水过程的人工神经元降水落区预报网络。

1 软 件

我们引进以 MS-WINDOWS 3.1 为操作系统、用 TPW(Turbo Pascal for Windows)语言编程的、在 386 或 486 系统微机上开发的计算机软件——智能暴雨预报系统(WAI85)[†]。在此基础上建立了青海省降雨落区预报自动化系统。它是以数值预报产品、动力诊断分析、用户所在省域或地区的实况降水资料数据库等为基础、以 B-P 算法为核心的降水预报系统。由数据库、物理量分析、神经网络学习与业务使用和用户参数自定义四个模块组成。其业务流程见图 1。

† 此系统为四川省气象科学研究所何险峰同志的软件产品

收稿日期: 1995-05-31; 改回日期: 1995-12-01

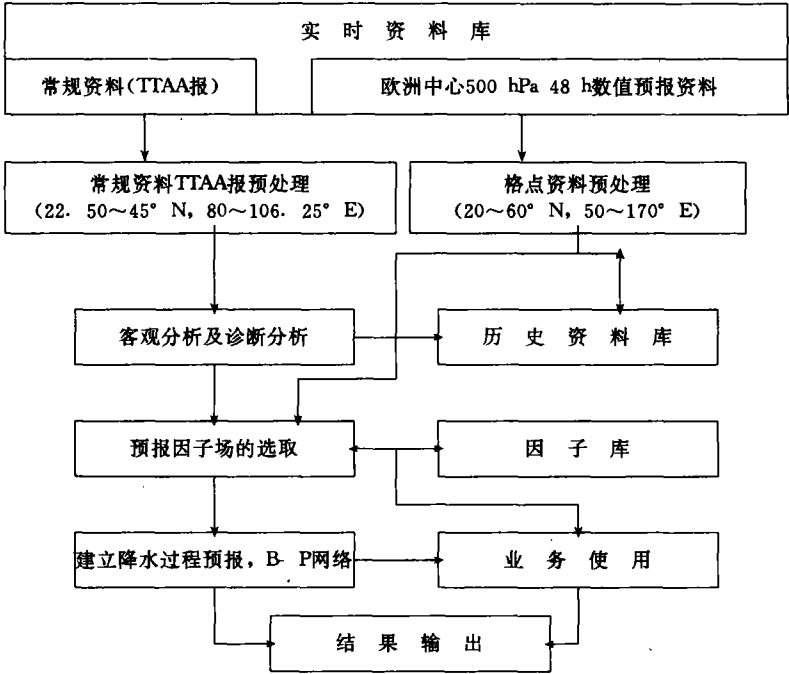


图 1 系统业务应用流程图
Fig. 1 Flow chart of the operational system

2 B-P 算法

B-P 算法是目前广泛应用的神经网络学习算法之一。一般网络是由一个输入层、一个输出层及一个或多个隐节点层组成。神经元由输入、处理、输出三部分构成,它们在网络图中用圆圈表示。神经元可以接受其他神经元输出的信号或系统外部输入的信号,作为它的输入内容,将这些信号经过一定的处理后得到一个结果,再以信号形式输出到其他神经元或系统外部(图 2)。它可以接受多个输入信号,输出一个信号,同时可以引向多个去处。信号可以是连续量、离散量或模糊量^[3],目前在气象预报业务中的应用大多是将各种神经元节点的输出函数定义为

$$Y_i = f \left(\sum_{j=1}^m W_{ij} X_j + \theta_i \right) \tag{1}$$

式中, Y_i 为第 i 个神经元的输出; W_{ij} 为第 j 节点对后一层 i 的节点的权值; m 为节点数; X_j 为输入信号的第 j 个分量; θ_i 为门限值。

为便于求误差函数时求导,一般取 $f(x)$ 为连续单调增 S 形函数。
即

$$f(x) = \frac{1}{1 + e^{-x}} \tag{2}$$

网络学习中,信息的处理由正向传播和反向传播两部分组成。正向传播过程是:从输入层输入模式,经隐单元层逐层处理并传向输出层,每一层神经元的状态只影响下一层神经元的状态。如果输出层不能得到期望的输出,则转入反向传播,即将误差信号沿原来的连通路返回,通过修改各层神经元的权值,使得误差信号最小,即使得网络对于所给定的输入模式 A 所产生的输出模式 Y 与希望的模式 D 十分接近或完全相同。

定义误差函数 r 为期望输出与实际输出之差的平方和

$$r = \frac{1}{2} \sum_{j=1}^n (V_{mj} - d_j)^2 \quad (3)$$

式中, d_j 为期望输出, 作为训练网络的教师信号; V_{mj} 是实际的输出, 是输入模式和数值向量 w 的函数。按照最优化理论的梯度下降法, 使网络权值沿误差函数的负梯度方向改变, 网络权值的更新量 Δw_{ij} 可由下式表示

$$\Delta w_{ij} \propto -\frac{\partial r}{\partial w_{ij}} \quad \text{或} \quad \Delta w_{ij} = -\alpha \frac{\partial r}{\partial w_{ij}}$$

式中, α 为学习步长, 取正参数。

现以一个有 n 个节点的输入层、 m 个节点的隐层和 L 个节点的输出层的两层半线性神经网络为例, 简要说明 B-P 网络的学习过程。

第一步 构造如图 3 所示的神经网络结构。

第二步 设置网络的初始权值为随机小量 ($-0.4 \sim 0.4$)。

第三步 确定学习精度 (0.1)。

第四步 输入样本, 计算输出 (信号向前传输)。

输入样本 $(x_1, x_2, x_3, \dots, x_n)$, 并指明对第 i 类期望输出 $(d_1, d_2, d_3, \dots, d_L)$, 网络的实际输出为

$$Y_i = 1 / [1 + \exp(-\sum_{j=1}^m U_j w_{ij} + \theta_i)]$$

$$U_j = 1 / [1 + \exp(-\sum_{k=1}^n x_k w_{jk} + \theta_j)]$$

其中, U_j 为第 j 个隐节点的输出。

第五步 计算误差, 向后传输, 修正权值。

对于输出层, $\delta_i = y_i(1 - y_i)(y_i - d_i)$, 对于隐节点层, $\delta_i = u_j(1 - u_j) \sum_{k=1}^L w_{kj} \delta_k$ 。从输出层到输入层依次反向修正权值时, 从输入层到隐节点层的权值改变为 $\Delta w_{ij}(t+1) = -\alpha \delta_i x_j + \gamma \Delta w_{ij}(t)$, 从隐节点层到输出层的权值改变为 $\Delta w_{ij}(t+1) = -\alpha \delta_i u_j + \gamma \Delta w_{ij}(t)$ 。式中, γ 为调整变量的参数。一般当学习步长取 $\alpha = 0.4 \sim 0.9$ 时, $\gamma = 0.7 \sim 0.9$ 。

第六步 得新的权值为

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}(t+1)$$

第七步 递归。

显然, 当输出与期望不一致 ($d_i - y_i \neq 0$) 时, 则各节点权值得到调整, 重复第四步至第七步, 直至达到精度要求或产生振荡。

网络学习结束后, 权值保持不变, 工作过程仅相当于学习过程的第四步。每输入一个样本后, 输出层输出值最大的节点就代表了模式的类别。于是网络便自动完成了分类任务。

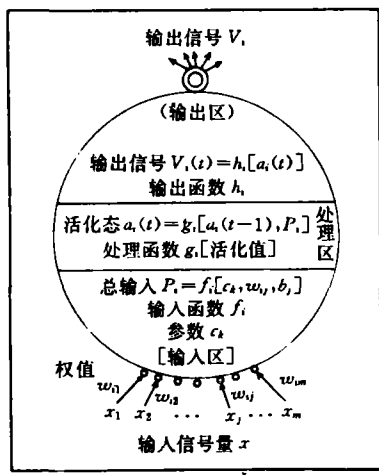


图 2 神经元功能模型

Fig. 2 Model of functions of the neural element

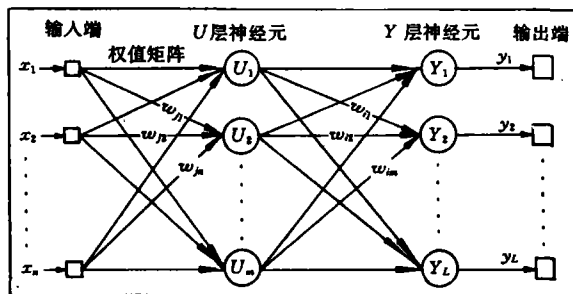


图 3 两层 B-P 网示意图

Fig. 3 Schematic of 2-level B-P neural network

3 资 料

选用欧洲气象中心 500 hPa 形势场的 48 h 数值预报产品作为资料 1, 缺资料时用 72 h 或 96 h 预报场代替, 范围为(20~60°N, 50~170°E)。资料 2 是省台“STYS”系统中 700~250 hPa (5 层)国内 125 个探空站点的陆地测站压、温、湿和风的报告。依赖于资料 2 进行客观分析, 从而计算散度、涡度、垂直速度、位温、假相当位温等动力诊断场^[5], 范围为(22.5~45°N, 80~106.25°E)。

4 预报因子

切贝雪夫正交多项式作为一个数学工具被应用到气象工作中来已有 40 多年的历史, 我国气象工作者已对其作了大量的研究^[6], 并用于单点降水预报、形势场预报等。我们将其用于降水落区预报的意义在于: 试图从切贝雪夫系数的几何意义上反映降水落区与形势场的变化关系, 即降水落区是由气象场的变化决定的。因子的选取有两类, 一类是不同关键区内欧洲气象中心 500 hPa 48 h 数值预报格点场和各种诊断场的二阶或三阶切贝雪夫正交多项式系数, 另一类是切氏调整指数 $C_{k_1}(t)$ ($C_{k_1}(t) = A_{k_1}(t-1)A_{k_1}(t)$, $A_{k_1}(t-1)$ 、 $A_{k_1}(t)$ 分别表示预报前一天和当天因子场的切贝雪夫系数), 用来表征二维图形在连续两个不同时刻是否发生跳跃变化。切氏系数的正负是决定多项式图形方向的决定因素, 如对二阶二维切氏系数 A_{20} 项, 当 $A_{20}(t-1) > 0$ 时, 若 $A_{20} > 0$, 则 $C_{20}(t) > 0$, 表示原始图形在时间演变中未发生质变; 而若 $A_{20}(t) < 0$, 则 $C_{20}(t) < 0$, 表示图形发生明显调整, 调整量由 $|C_{k_1}(t)|$ 反映。即在一定的环流背景下经过一段相对稳定的时间后, 在较短的时间内发生了质变。

所选因子为以下关键区内各要素场的切氏多项式展开系数 A_{k_1} 和切氏调整指数 C_{k_1} 及 500 hPa $T-T_d$ 的平均值。

(1) 欧洲中心 500 hPa 48 h 数值预报场(EH50)。关键区为(22.5~45°N, 85~105°E)。

(2) 700、500、400、300、250 hPa 形势分析场(HIGH)。500~250 hPa 关键区 1 为(32.5~35.0°N, 80~100°E), 关键区 2 为(25~35°N, 80~100°E)。700 hPa 关键区为(32.5~45°N, 80~100°E)。

(3) 700、500 hPa 水汽通量(QELX)。700 hPa 关键区为(32.5~42.5°N, 82.5~105°E), 500 hPa 关键区为(25~47.5°N, 82.5~105°E)。

(4) 700、500、400、300、250 hPa 假相当位温平流(QSE0)。700 hPa 关键区为(32.5~45°N, 80~105°E), 500~250 hPa 关键区为(25~45°N, 80~105°E)。

(5) 700、500、400、300、250 hPa 温度平流(TRDE)。700 hPa 关键区为(35~42.5°N, 82.5~100°E), 500~250 hPa 关键区为(30~42.5°N, 82.5~100°E)。

(6) 700、500、400、300、250 hPa 风场 u 分量(UWND)。700 hPa 关键区为(35~45°N, 80~106.25°E), 500~250 hPa 关键区为(22.5~45°N, 80~106.25°E)。

(7) 700、500、400、300、250 hPa 风场 v 分量(VWND)。关键区同(6)。

(8) 700、500、400、300、250 hPa 涡度场(VORT)。700 hPa 关键区为(35~43.75°N, 81.25~100°E), 500~250 hPa 关键区为(30~43.75°N, 81.25~100°E)。

(9) 700、500、400、300、250 hPa 涡度平流(VRDE)。700 hPa 关键区为(35~42.5°N, 82.5~100°E), 500~250 hPa 关键区为(30~42.5°N, 82.5~100°E)。

(10) 500 hPa 温、露差($T - T_d$)。关键区 1 为(35~42.5°N, 85~100°E), 关键区 2 为

(26.25~35°N, 85~100°E), 关键区 3 为(26.25~42.5°N, 95~106.25°E)。

5 预报结果及检验

根据青海省气候特点及历史降水个例的统计分析, 将全省降水落区分成如图 4 所示的三个区, 全省为一区, 共四个区。第一区为全省; 第二区为东北部(NE), 包括海东、海北、黄南北部和海南的部分地区; 第三区为东南部(SE), 包括玉树、果洛、海南南部、黄南南部; 第四区为西部(West)。

按 24 h 降水量等级划分情况及省内降水特点, 将降水过程分为 5 mm、10 mm、25 mm 三个等级, 分别建立了未来 24 h 有无 ≥ 5 mm、 ≥ 10 mm、 ≥ 25 mm 降水过程的人工神经元降水落区预报网络。网络学习分单输出和多输出两种情况。

单输出: 第一区 50 个站中有 10 站以上出现降雨, 为一次大范围降雨天气过程。

多输出: 全省 4 个区中, 有 3 站/区以上出现降雨, 该区为一次降雨天气过程。

将上述资料(1981~1990 年 5~9 月)输入系统, 由 B-P 网分别对 ≥ 5 mm、 ≥ 10 mm、 ≥ 25 mm 的降水过程进行独立训练, 结果如下:

(1) 大于(等于)5 mm 降水过程

对于单输出: 大约在 B-P 网络自动训练到 70 次后, 由网络输出判定的天气类别与实际出现天气类别的概括率达 988/996=99%。

对于多输出: 大约在 B-P 网络自动训练到 90 次后, 2501 次的 ≥ 5 mm 降水过程, 网络可以预报出 2438 次, 概括率可达 97%。

(2) 大于(等于)10 mm 降水过程

对于单(多)输出: 约在 B-P 网络自动训练到 65(80)次后, 概括率达到 545/548=99%(1172/1216=96%)。

(3) 大于(等于)25 mm 降水过程

对于单输出和多输出: 大约在 B-P 网络自动训练到 19 次后, 概括率均达到 100%。

从以上各级降水过程的 B-P 网学习结果看, 其概括率均达 96%以上, 并通过了学习结果历史检验, 完全能满足业务预报需要。实际使用时, 通过人机对话形式由机器自动输入当日 08 时 TTAA 报资料和当日接收到的欧洲中心 500 hPa 48 h 数值预报资料, 进行降雨落区预报。

青海省降水预报的人工神经网络训练完成后, 于 1994 年 6、7、8 月份进行了业务试用。表 1 是 25 mm、10 mm、5 mm 降水预报 B-P 网络的使用结果。表中第一列为降水级别、落区名和预报合计; 其他各列是: $Fu1$ 为预报有降水的正确次数, $Fore1$ 为预报有降水的次数; $All1$ 为实况有降水的次数; $Accu = Fu1/Fore1$ 为降水预报准确率; $Prob = Fu1/All1$ 为降水预报概括率; $Csi = Fu1/[Fu1 + (All1 - Fu1) + (Fore1 - Fu1)]$ 为降水预报成功指数。 ≥ 25 mm、 ≥ 10 mm、 ≥ 5 mm 情况下业务使用次数分别为 67、67、65 次。

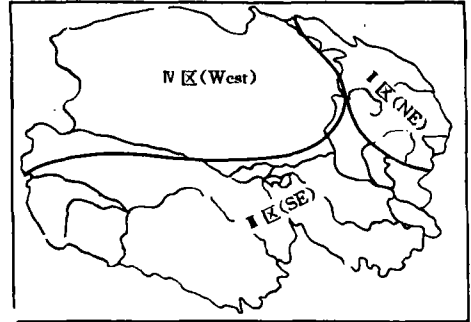


图 4 降水落区示意图

Fig. 4 Predicted fallout regions

表 1 降水预报结果

Table 1 Predicted 25 mm rainfall for 24 hr to come

	≥25 mm						≥10 mm						≥5 mm					
	Fit1	Fore1	All1	Accu	Prob	Csi	Fit1	Fore1	All1	Accu	Prob	Csi	Fit1	Fore1	All1	Accu	Prob	Csi
≥10	0	0	0	0.00	0.00	0.00	9	15	10	0.60	0.90	0.56	23	36	23	0.64	1.00	0.64
NE	0	0	1	0.00	0.00	0.00	6	8	11	0.75	0.55	0.46	10	16	15	0.62	0.67	0.48
SE	0	0	0	0.00	0.00	0.00	8	12	11	0.67	0.73	0.53	19	24	22	0.79	0.86	0.70
West	0	0	0	0.0	0.0	0.0	0	0	0	0.00	0.00	0.00	0	0	0	0.00	0.00	0.00
≥3	1	1	1	1.00	1.00	1.00	13	16	16	0.81	0.81	0.68	26	38	26	0.68	1.00	0.68
合计	1	1	2	1.00	0.50	0.50	36	51	48	0.71	0.75	0.57	78	114	86	0.68	0.91	0.64

由表可见,对于≥25 mm 降水过程,在预报时段内,只有东北部出现一次,网络未能报出,故有 $Accu = 0, Prob = 0, Csi = 0$ 的结果,但对于≥3 站的全省片,网络正确报出,有 $Accu = 1, Prob = 1, Csi = 1$ 的结果,从而得以补救。

对于≥10 mm 降水过程,西部未出现规定的降水过程(预报无,实况也无),全省片≥10 站、≥3 站和 SE 片(≥3 站)的降水过程预报准确率分别为 60%、81%和 67%;降水预报概括率分别为 90%、81%和 73%;降水预报成功指数分别为 56%、68%和 53%,效果较好。东北片降水预报准确率为 75%;降水预报概括率为 55%;降水预报成功指数为 46%,效果较差。各片合计:预报准确率为 71%;降水预报概括率为 75%;降水预报成功指数 57%,总的预报结果还是令人满意的。

对于≥5 mm 降水过程,西部同样也未出现规定的降水过程(预报无,实况也无),全省片≥10 站、≥3 站和 SE 片(≥3 站)的降水过程,预报准确率分别为 64%、68%和 79%;降水预报概括率分别为 100%、100%和 86%;降水预报成功指数分别为 64%、68%和 70%,效果较好。东北片预报准确率为 62%;降水预报概括率为 67%;降水预报成功指数为 48%,效果仍较差。各片合计:预报准确率为 68%;降水预报概括率为 91%;降水预报成功指数 64%,总的预报结果也还是令人满意的。

从整个试报结果看:

(1)该网络的预报准确率和概括率均较高,从合计值看, Csi 最高达 64%,最低为 50%,说明该网络判断有无降水天气过程的预报能力较强,就全省片≥3 站标准的 10 mm 降水过程与对应标准的大降水专家系统相比,均高于专家系统。可见 B-P 网具有较高使用价值。

(2)除东北片预报效果较差外,其他各片效果均较好。我们认为造成东北片预报效果不好的主要原因之一是因子的选择。因青海省东北部降水的天气过程主要受起源于东欧或西亚、中亚地区的西方槽的影响,当冷空气由新疆南下时,由于阿尔金山、祁连山的影响,低层冷空气很难侵入青海腹地,往往沿河西走廊南下,之后沿祁连山东部湟水河、黄河河谷侵入青海,造成东北部地区的降水。而在因子场的选取中并未考虑地形的作用。

(3)就各片降水预报成功指数看:5 mm 降水预报成功指数明显高于 10 mm,而 10 mm 降水预报成功指数又明显高于 25 mm。造成这一特点的主要原因是随着降水等级的提高,其降水天气过程的气候概率也就越低。

通过对降雨预报结果与实况的对比分析,得知网络对天气趋势的预报能力较高,即网络预报降雨的周期变化与降雨实况的周期变化相吻合。

6 结 语

由于青海省地处高原,气象要素的日变化大,动力、热力因素复杂,给降水预报带来了许多困难,引入人工神经网络作降水等级的有无预报,从某种意义上看,是模仿了预报员用特征相似做预报的思维方式,方法简便、客观,历史拟合好,预报准确率较高。初步实现了降水预报高速、客观、自动化的要求。

在青海省引进人工神经网络作降水天气过程的分类预报,仅是初步尝试,还有不当或不完善之处,有待进一步吸取专家的有益经验,研究改进,选取效果最好的预报因子,对网络进行更新,从而获得历史拟合率高、实际预报效果好的 B-P 网络。进而逐步加入其他有效手段使系统更加完善,最终形成具有综合决策预报能力的高速、客观、定量、自动化的降雨预报系统。

参 考 文 献

- 1 王耀生. 人工智能、模式识别在气象领域应用的现状与展望. 气象, 1994, 20(6): 9~14
- 2 张承福. 人工神经网络在天气预报中的应用研究. 气象, 1994, 20(6): 43~47
- 3 Rogers S K, Kabrisky M. An introduction to biological and artificial neural networks for pattern recognition, SPIE Press, 1991
- 4 陈景耀. 青海省夏季中一大雨预报专家系统知识库设计. 青海气象, 1987, (5): 7~9
- 5 董晓敏, 田盛培. 天气诊断分析方法简介. 北京: 气象出版社, 1986
- 6 屠其璜, 王俊德, 丁裕国, 等. 气象应用概率统计学. 北京: 气象出版社, 1984. 427~445
- 7 青 天. 青藏高原应用气候与天气. 青海: 青海人民出版社, 1987. 81~138
- 8 戴加洗主编. 青藏高原气候. 北京: 气象出版社, 1990. 63~97, 171~196

AUTOMATIC SYSTEM FOR PREDICTING RAIN FALLOUT REGIONS OVER QINGHAI PROVINCE

Wang Fanqiang¹⁾ Xu Wenjin²⁾ Chen Jilun³⁾ Wang Shen¹⁾

(1) Qinghai Meteorological Bureau, Xining 810001, 2) Department of Meteorology,

3) Department of Computer Science/Information Engineering, NIM, Nanjing 210044)

Abstract With 0800 h convenient data and numerical products as input, the element field of a key region is expanded as the Chebyshev orthogonal polynomial, whose coefficients are taken as predictors to be trained with the aid of the B-P neural network, whereupon is established an artificial rainfall fallout region forecasting network for categories 5, 10 and 25 mm for 24 h in advance. Tests show that >95% of the historical events are predicted with successful operational application in summer (JJA) 1994.

Keywords B-P algorithm, fallout region forecasting, automatic system