

基于双隐层 BP 网络的沈阳地区 雷暴潜势预报模型^①

林中冠¹, 栾健², 王迪¹, 张春龙³, 栾澜⁴

1. 辽宁省气象服务中心, 沈阳 110166; 2. 中国气象局气象干部培训学院 辽宁分院, 沈阳 110166;
3. 黑龙江省气象灾害防御技术中心, 哈尔滨 151030; 4. 沈阳农业大学 农学院, 沈阳 110866

摘要: 利用沈阳市 2007—2011 年 6、7、8 月探空因子与闪电定位资料, 选取适合当地雷暴潜势预报的 9 个因子, 构建双隐层 BP 人工神经网络预报模型, 对当地雷暴的潜势发展进行预报. 结果表明: 基于双隐层 BP 人工神经网络的预报模型临界成功指数 CSI=57.46%、探测概率 POD=79.38%、虚假报警率 FAR=32.46% 和总指数 TS=74.89%, 预报效果较令人满意; 通过 12 h 和 6 h 潜势预报结果对比发现, 6 hBP 神经网络模型预报效果好, 更适用于沈阳地区雷暴潜势预报.

关键词: 双隐层; BP 网络; 雷暴; 潜势预报; 模型

中图分类号: P427.42

文献标志码: A

文章编号: 1673-9868(2017)02-0084-08

雷暴是一种对流旺盛的天气系统, 其发生常伴有强降水、大风、冰雹、龙卷等灾害性天气现象, 造成了大量的人身伤亡以及财产损失^[1-5], 因此雷暴活动的潜势预报是天气预报业务的重要组成部分之一, 如何提高雷暴潜势预报效率对防灾减灾具有重要意义^[6-8].

雷暴产生机制的复杂性决定了他的发生具有显著的非线性特征, 人工神经网络由于具有很强的非线性映射能力和并行性、适应性、容错性及自学习能力, 特别适用于解决因果关系复杂的非确定推理、判断、预测和分类等问题, 因此在雷暴潜势预报研究领域越来越受到重视, 在人工神经网络的研究应用中, BP (Back-Propagation, 反向传播算法) 人工神经网络被各领域广泛应用^[9-14]. 相对传统的数理统计方法而言, BP 人工神经网络模型可以求解非线性问题, 同样对样本大小的要求也可以相对少得多.

陈勇伟^[15]等利用南京地区 2008 年的闪电定位资料和探空资料并基于 BP 人工神经网络模型, 预报了南京地区 2009 年 6—8 月的雷暴活动潜势, 此模型的 POD(探测概率)为 80.9%, FAR(虚假报警率)为 9.5%, CSI(临界成功指数)为 74.5%. 杨仲江^[16]等利用闪电定位资料结合探空资料, 采用双隐层 BP 人工神经网络方法对太原地区的雷暴天气进行了潜势预报研究, 他发现在解决分类问题上双隐层 BP 人工神经网络比单隐层 BP 网络更具优势. 而与多元统计回归法相比, 双隐层 BP 网络可以获得更高的雷暴预报评分及更可靠的结果. 以上研究表明: 利用 BP 人工神经网络模型进行雷电潜势预报是可行的, 但是由于研究地域的不同, 选取的预报因子不同, 网络模型的构建方式不同, 预报结果判定方式也有所不同.

因此, 本文利用沈阳市 2007—2011 年 6、7、8 月探空资料与闪电定位资料, 分析该地区的雷暴潜势预报因子, 进一步建立 BP 人工神经网络模型, 对该地区的雷暴潜势预报进行研究, 对比分析 6 h 和 12 h 雷暴潜势预报效果, 以期获得更高的预报准确率.

① 收稿日期: 2015-11-09

基金项目: 辽宁气象局博士科研专项资助项目(D201603).

作者简介: 林中冠(1986-), 男, 辽宁沈阳人, 本科, 工程师, 主要从事气象科技服务相关研究.

通信作者: 栾健, 工程师.

1 资料来源与方法

1.1 资料来源

本研究利用沈阳市 2007—2011 年 6、7、8 月探空因子与闪电定位资料. 探空数据来自沈阳站(54342), 一日 2 次, 分别是北京时间 8:00 点和 20:00 点, 将每次探空得到的因子与探空后 12 h 内闪电定位系统探测到的雷暴发生与否组成样本对, 经统计样本数量为 910 个样本. 这里定义: 若探空后 12 h 内, 沈阳站附近 50 km 范围内发生一次及以上的雷暴即认为该样本是雷暴样本, 否则为非雷暴样本.

多个因子中有些与雷暴发生与否相关性不大, 应予以甄选. 由于对流参数物理量是连续型因子, 雷暴发生与否为 0, 1 变量, 所以不可直接求出其相关系数, 而是求出其双点序列相关系数^[16].

78 个因子双点序列相关系数(r)大于或等于 0.25 的有 12 个, 序列号分别是 1,2,3,4,7,9,13,17,18,57,72,73, 如表 1 所示. 其中序列号 57 因为样本缺失过多, 只有 95 个可用, 所以不予考虑. 1,2 号因子分别是总指数 T 和修正总指数 mTT, 3,4 号因子分别是 K 指数和修正 K 指数, 为了避免信息重复, 所以保留相关系数略大的修正总指数和修正 K 指数. 因此, 参与预报的因子共计 9 个, 分别是: 修正总指数、修正 K 指数、沙氏指数、条件稳定度指数、条件对流稳定度指数、Teffer 指数、Charba 总指数、瑞士雷暴指数、抬升指数这 9 个因子, 样本 168、645 号有缺失, 至此, 共有样本 908 个, 其中雷暴样本 402 个, 非雷暴样本 506 个.

表 1 选取的 9 个因子名称及双点序列相关系数

因子英文缩写	因子名称	与雷暴发生与 否的相关系数	因子英文缩写	因子名称	与雷暴发生与 否的相关系数
mTT	修正总指数	0.27	Teffer	Teffer 指数	0.25
mK	修正 K 指数	0.28	ChTT	Charba 总指数	0.26
SI	沙氏指数	0.28	SWISS	瑞士雷暴指数	0.31
IL	条件稳定度指数	0.27	LI	抬升指数	0.27
CIL	条件对流稳定度指数	0.27			

908 个样本中, 3/4 的样本用来训练网络, 1/4 的样本用于检验. 为避免样本之间的相关性影响预报结果, 故这里采取随机选取的方式来进行样本划分. 因此样本被分为 681 个训练样本和 227 个独立检验样本. 908 个样本中, 雷暴样本有 402 个, 占 44.27%; 681 个训练样本中, 雷暴样本有 305 个, 占 44.79%; 227 个独立检验样本中, 雷暴样本有 97 个, 占 42.73%. 网络输出采用两级分类, 将有雷暴样本设计成 $[0.9, 0.1]$, 无雷暴样本设计成 $[0.1, 0.9]$, 并将训练样本对进行归一化, 数值范围归一到 $[0.1, 0.9]$, 独立检验样本按照与训练样本同样的方法进行归一化处理.

1.2 研究方法

BP 神经网络可有效地用于复杂的非线性函数的逼近, 一个 3 层的前馈网络能够实现任意精度的连续函数映射^[18], 两个隐层的神经网络则可以解决各种分类问题^[19]. 因此本文尝试构建两个隐藏层的 BP 网络, 下文称为双隐层 BP 网络, 双隐层 BP 网络模型示意图如图 1 所示.

1.2.1 输入输出数据的处理

训练样本集是影响 BP 网络性能优劣的关键, 神经网络是以样本在事件中的统计几率来进行训练和预测的. 归一化是归纳统一样本的统计分布性, 可以简化计算, 缩小量值, 加快网络的收敛. 归一化在 $[0, 1]$ 之间是统计的概率分布, 归一化在 $[-1, +1]$ 之间是统计的坐标分布. 兼顾 sigmoid 函数的定义, 这里采取式(1)对输入输出样本进行 $[0.1, 0.9]$ 之间的归一化.

$$\mathbf{X} = 0.1 + \frac{\mathbf{P} - \mathbf{P}_{\min}}{\mathbf{P}_{\max} - \mathbf{P}_{\min}} \times 0.8 \quad (1)$$

这里, \mathbf{P} 表示归一化前的输入数据; \mathbf{P}_{\min} 表示矩阵 \mathbf{P} 的最小值; \mathbf{P}_{\max} 表示矩阵 \mathbf{P} 的最大值; \mathbf{X} 表示归一化之后的输入矩阵.

1.2.2 训练算法

选择算法对网络进行训练时, 在网络参数很多, 需要考虑存储容量问题时, 选择共轭梯度法. 经试验, 这里选择 Scaled 共轭梯度法 trainscg.

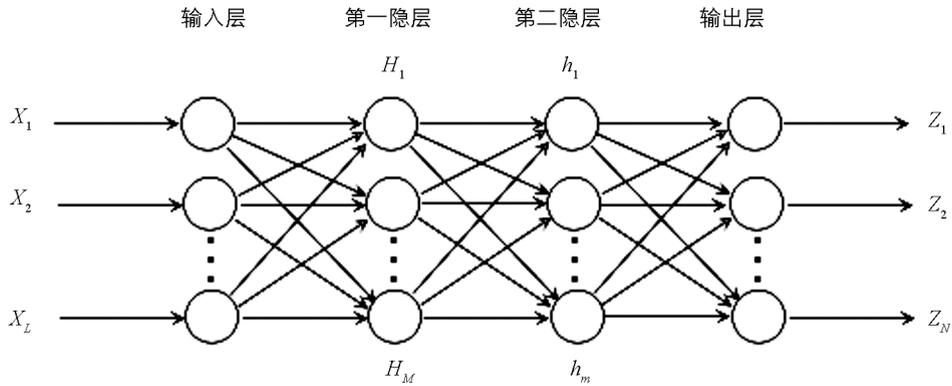


图 1 双隐层 BP 网络模型

1.2.3 传递函数

各层神经元的传递函数根据需要, 不同层内采取不同的传递函数组合, 并且一般常采用 S 形函数和线性函数组合, S 形函数能将跨度很大的数值压缩到一个很小的固定范围内. 输出层常采用线性函数, 这样整个网络的输出可以取实数域内任何值. 双隐层的网络结构, 需要 3 层传递函数, 选取 tansig 和 purelin 的结合, 在试验中, 最终确定 tansig-tansig-purelin 组合. tansig 的表达式如下:

$$\mathbf{A} = f(\mathbf{W} \cdot \mathbf{P} + \mathbf{b}) = \frac{1}{1 + e^{-(\mathbf{W} \cdot \mathbf{P} + \mathbf{b})}} \quad (2)$$

$$\mathbf{A} = f(\mathbf{W} \cdot \mathbf{P} + \mathbf{b}) = \mathbf{W} \cdot \mathbf{P} + \mathbf{b} \quad (3)$$

其中, \mathbf{P} 表示输入矩阵, \mathbf{W} 表示权值矩阵, \mathbf{b} 表示偏置矩阵, \mathbf{A} 表示输出.

1.2.4 过拟合问题的处理

怎样寻求泛化能力最大的网络是网络构建的一个难题, 网络的泛化性能是指神经网络对训练样本以外的新样本的适应能力. 在网络的训练过程中经常会出现过拟合现象, 即在网络训练过程中, 从某一次训练开始, 随着网络的训练能力提高, 仿真的能力反而下降, 解决这个问题的一种办法是: 在数据输入中, 给训练的数据分类, 分为正常训练数据、变量数据、测试数据, 3 种数据占用全部样本的 70%, 15%, 15%. 设定参数 maxfail 为 6 (最大失败次数), 在网络训练的过程中, 如果从某一步开始, 变量数据的误差不降反升, 则 Validation Checks 开始计数, 当计数到 6, 则认为网络陷入过拟合, 停止训练.

1.2.5 隐节点数的选取

目前隐节点的选取还没有一个统一的标准, 前人也进行了许多研究, 得到几种隐节点的计算公式^[19-22], 由于这些公式只是大致给出了隐节点的取值范围, 而不同的研究问题之间有很大的差别, 所以仅根据上述公式来确定隐节点数不合适, 这里将其作为参考, 尝试多种隐节点数组合, 根据网络训练的 MSE, 最终确定隐节点数.

2 结果分析

2.1 预报因子分析

将选取的 9 个探空因子分别统计雷暴样本、非雷暴样本的数值分布, 制成方框一端须图, 纵坐标为因子数值. 一般来说, 对于大量的统计测量数据, 均可视作是呈现正态分布, 在此将这 9 个因子视为呈现正态分布.

如图 2 所示, 在每个长方形上下底分别表示正态分布两端占该类样本总数 (有雷暴样本或无雷暴样本) 25% 个例的因子数值, 方框内的横线表示占该类样本总数 50% 个例的因子数值, 从长方形上下边延伸出的线段端点表示占该类样本 5% 个例的因子数值, “×”表示的是占该类样本总数 1% 个例的因子数值, 最上端和最下端的“+”表示该类样本的最大和最小值, 方框中的“□”表示该类样本均值.

1) 修正总指数

图 2(a) 给出了修正总指数的分布, 可以看出雷暴样本的数值偏大, 雷暴样本的中位数是 52.4 °C, 说明雷暴样本有半数 > 52.4 °C, 而在这一区间仅有约 25% 的非雷暴样本修正总指数的定义为

$$\text{修正总指数}(mTT) = \bar{T} + \bar{T}_d - 2T_{500} \quad (4)$$

下标 500 表示气压, \bar{T} 与 T_d 分别表示地面至 850 hPa 的平均温度与平均露点温度. 由定义可以看出 mTT 越大, 越容易发生对流天气.

2) 修正 K 指数

图 2(b) 给出了修正 K 指数的分布, 可以看出雷暴样本数值偏大, 中位数接近 40 °C, 非雷暴样本的中位数位于 30 °C 左右. 79.35% 的雷暴样本 mK 大于 30 °C, 59.95% 的雷暴样本 mK 大于 40 °C, 相应的非雷暴样本对应的比例为 54.35%, 41.90%. mK 的定义中考虑了温度之间率、底层水汽条件、中层饱和度以及地面温度状况, 因此 K 指数可以反映大气的层结稳定情况, mK 值越大表示气团低层越暖湿, 稳定度越小, 层结越不稳定, 因而越有利于对流产生.

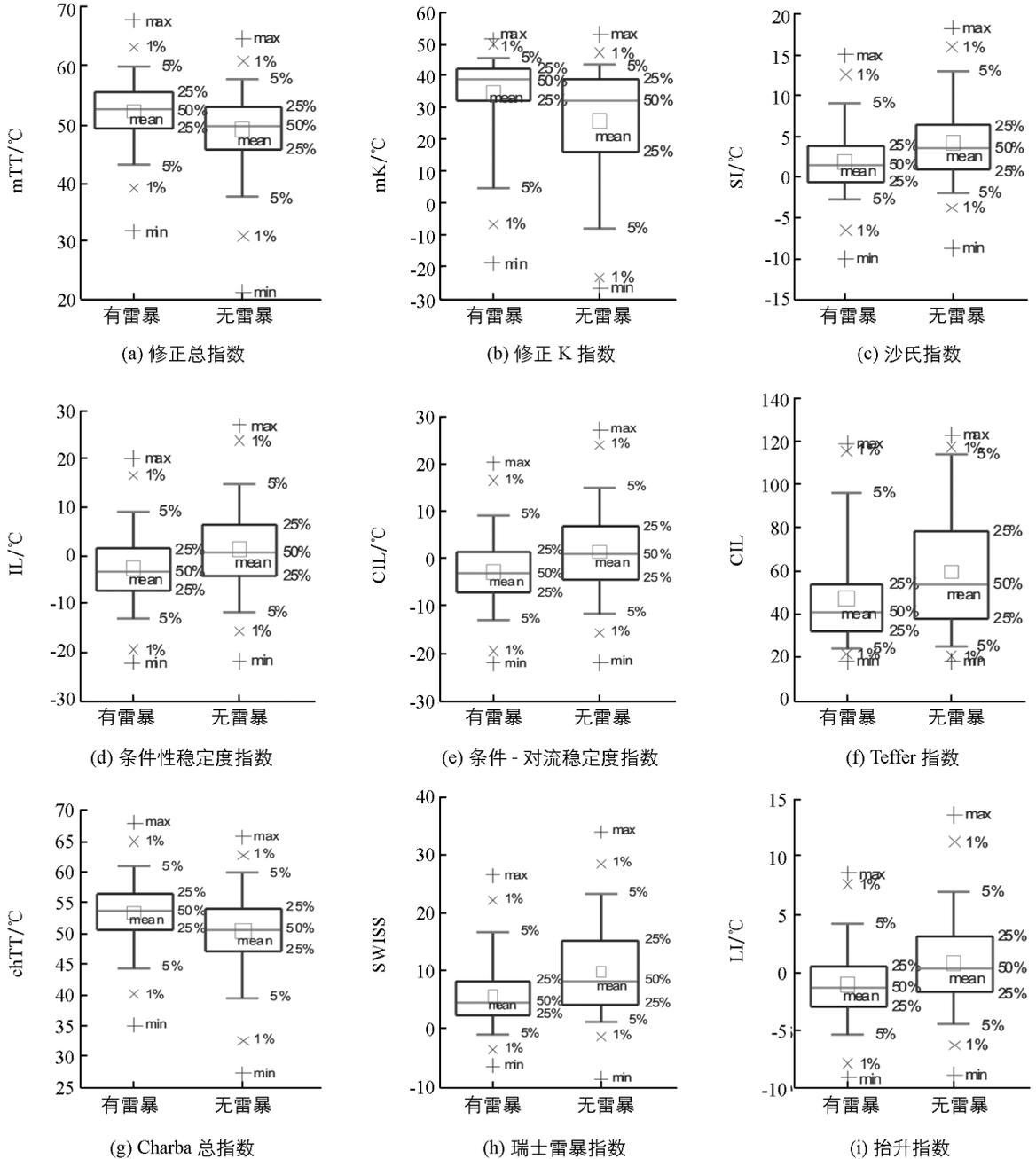


图 2 9 个探空因子的雷暴样本和非雷暴样本的数值分布

3) 沙氏指数 SI

图 2(c) 给出了沙氏指数的分布, 可以看出雷暴样本的沙氏指数偏低, 71.89% 雷暴样本的 SI 小于 3 °C,

而非雷暴样本在这一区间的比例有 45.65%。在 SI 小于 0 °C 区间, 雷暴样本比例达到 32.84%, 非雷暴样本仅有 15.81%。沙氏指数作为一个能够表征大气稳定状况的指数, 它定义: 850 hPa 等压面上的湿空气团沿干绝热线上升, 到达凝结高度后再沿湿绝热线上升至 500 hPa 时所具有的气团温度 T_{s850} 与 500 hPa 等压面上的环境温度 T_{500} 的差值。当 SI 小于 0 时, 大气层结不稳定, 且不稳定程度与负值成正比。反之, 当 SI 大于 0 时, 表示气层是稳定的。SI 虽然也是表示条件性稳定度的, 但只有当起始高度为 850 hPa, 上层为 500 hPa 时, 它才与 IL 类似。

4) 条件性稳定度指数

当 IL 小于 0, 为条件性不稳定; IL 为 0, 为中性; IL 大于 0, 为条件性稳定。图 2(d) 可以清晰地看出雷暴样本的中位数位于 0 以下, 而非雷暴样本的中位数位于 0 之上, 67.66% 的雷暴样本处于不稳定状态, 仅 46.44% 的非雷暴样本处于这一区间。

5) 条件对流稳定度指数

条件性稳定度指数考虑的是一小块空气上升得到的, 而对流性稳定度考虑的是整层空气抬升得到的, 常把 IL 与 IC 相加称为条件-对流稳定度指数, 也有称为位势稳定度指数, 因此当 CIL 小于 0, 为对流性不稳定; CIL 为 0, 为中性; CIL 大于 0, 为对流性稳定。由图 2(e) 可以看出近 70% 的雷暴样本处于不稳定状态下。

6) Teffer 指数

图 2(f) 给出了 Teffer 指数的分布, 可以看出雷暴样本该指数数值偏小, 雷暴样本有 75% 位于 53.59 以下, 而位于此区间的非雷暴样本仅有 50%。

7) Charba 总指数

由图 2(g) 可以看出, 雷暴样本的值较大, 雷暴样本有 50% 位于 54 °C 以上, 而这个区间非雷暴样本仅有 25%, 该指数用于表征对流稳定度, 数值越大, 越不稳定, 越容易发生强对流天气。

8) 瑞士雷暴指数

瑞士雷暴指数用于确定是否有雷暴发生, 它类似强天气威胁指数, 这是一个无量纲的指数, 一般来讲, 当 SWISS 小于 5.1 时预报有雷暴。图 2(h) 给出了瑞士雷暴指数的分布, 54.73% 的雷暴样本该指数小于 5.1, 而非雷暴样本只有 29.64% 小于 5.1。

9) 抬升指数 LI

图 2(i) 给出了抬升指数的分布, 可以看出雷暴样本的沙氏指数偏低, 67.66% 雷暴样本的 LI 小于 0 °C, 而非雷暴样本在这一区间的比例有 46.44%; 雷暴样本的均值及中位数均在 0 °C 以下, 而非雷暴样本均在 0 °C 以上。抬升指数表征的是气块从低层 900 m 高度沿干绝热线上升, 到达凝结高度后, 再沿湿绝热线上升至 500 hPa 时所具有的温度 T_s 与 500 hPa 等压面上的环境温度 T_{500} 的差值。当 LI 小于 0 时, 且不稳定程度与负值成正比。反之, 当 LI 大于 0 时, 表示气层是稳定的。

总的来说, 这 9 个因子表征着大气层结稳定度、对流稳定度、底层水汽等条件, 而这些条件与对流的产生、发展有紧密的联系, 因此一定程度上包含了雷暴的信息, 并且这些因子与辽宁地区雷暴发生与否相关性较高, 因此利用这 9 个因子作为辽宁地区雷暴潜势预报的因子是可行的。

2.2 双隐层 BP 网络预报结果

为了验证预报方法的优劣, 这里采用独立样本来检验预报结果, 并用 4 个指数作为评估标准, 分别是: 临界成功指数 CSI^[22]、探测概率 POD^[23]、虚假报警率 FAR^[24] 和总指数 TS^[25]。经多次试验, 隐节点数设为 7, 20, 采取两级分类输出, 即: 当预报该样本为雷暴样本的概率大于预报其为非雷暴样本的概率时, 将其预报为雷暴样本, 反之为非雷暴样本, 预报结果见表 2。

表 2 双隐层 BP 神经网络预报结果

	实况有	实况无	预 报 结 果	
预报有雷暴	$a = \text{雷暴}$	$b = \text{空报}$	$a = 77$	$b = 37$
预报无雷暴	$c = \text{漏报}$	$d = \text{无雷暴}$	$c = 20$	$d = 93$

$$C_{\text{CSI}} = \frac{a}{a+b+c} = 57.46\% \quad (5)$$

$$T_{\text{TS}} = \frac{a+d}{a+b+c+d} = 74.89\% \quad (6)$$

$$F_{\text{FAR}} = \frac{b}{a+b} = 32.46\% \quad (7)$$

$$P_{\text{POD}} = \frac{a}{a+c} = 79.38\% \quad (8)$$

$C_{\text{CSI}}, T_{\text{TS}}, F_{\text{FAR}}, P_{\text{POD}}$ 分别为成功指数、探测概率、虚假报警率和总指数的值。

2.3 6 h 实效 BP 神经网络模型雷暴潜势预报试验

雷暴是局地强对流天气,其发生发展往往比较迅速,12 h 预报模型常常由于实效过长导致预报结果达不到理想效果.因此,为寻求最佳的预报模型,调整实效,尝试建立 6 h 实效的雷暴潜势预报模型。

因此,908 个样本被分为 3 类,6 h 内有雷暴样本 276 个,6 h 以上 12 h 以内有雷暴样本 126 个,12 h 以内无雷暴样本 506 个.针对 6 h 实效,雷暴样本是 276 个,无雷暴样本是 632 个,雷暴样本比例 30.4%,需要删除一些非雷暴样本来提高雷暴样本的比例,这么做的目的是为了网络能够掌握每一类样本的特征.处理方法是,6 h 以上 12 h 以内有雷暴样本随机删除一半,12 h 以内无雷暴样本随机删除一半,至此,剩余的样本中,雷暴样本 276 个,非雷暴样本 316 个,雷暴样本比例达到了 46.62%,将这 592 个样本随机抽取其中 3/4 用来训练,1/4 用来独立检验.这一模型中,训练样本共有 444 个,检验样本有 148 个.将样本采取与 2.2 同样的处理方法进行预处理并建模,得到的双隐层 BP 网络两层隐节点数分别为 16,13,预报结果如表 3。

表 3 6 h 双隐层 BP 神经网络预报结果

	实况有	实况无	预报结果	
预报有雷暴	$a = \text{雷暴}$	$b = \text{空报}$	$a = 55$	$b = 18$
预报无雷暴	$c = \text{漏报}$	$d = \text{无雷暴}$	$c = 17$	$d = 58$

$$C_{\text{CSI}} = \frac{a}{a+b+c} = 61.11\% \quad (9)$$

$$T_{\text{TS}} = \frac{a+d}{a+b+c+d} = 76.35\% \quad (10)$$

$$F_{\text{FAR}} = \frac{b}{a+b} = 24.66\% \quad (11)$$

$$P_{\text{POD}} = \frac{a}{a+c} = 76.39\% \quad (12)$$

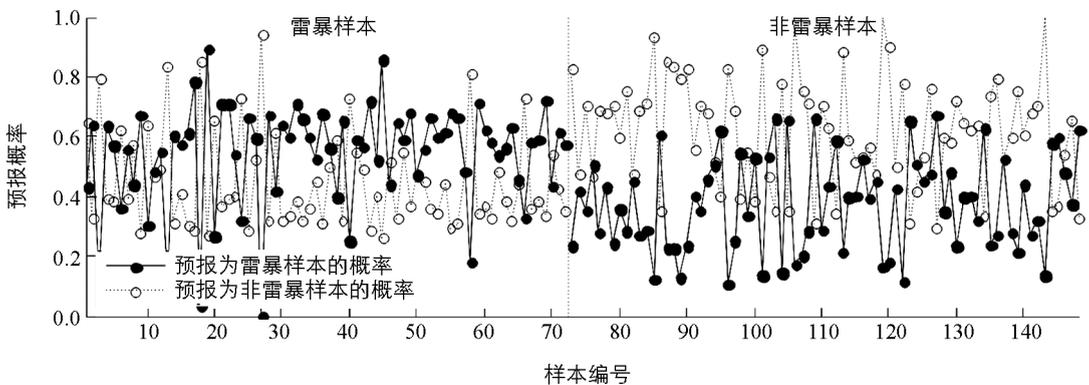


图 3 6 h 实效双隐层 BP 网络模型雷暴潜势预报结果

与 12 h 实效的预报结果对比发现,6 h 实效的预报结果更好,CSI 评分提高了 3.65%,总指数提高了

1.46%, 虚假报警率降低 7.8%, 探测率略有降低, 降低 2.99%。

2.4 预报误差分析

通过统计 12 h 实效双隐层两级输出 BP 网络预报错误的样本发现: 空报的 37 个样本中探空前 1 h 发生雷暴的样本有 3 个, 探空前 3 h 发生雷暴的样本有 9 个, 探空前 6 h 发生雷暴的样本有 13 个, 这些雷暴的发生对大气电场以及探空因子产生的影响可能还没有消失, 容易导致空报。漏报的 20 个样本中, 探空前 12 h 未发生雷暴, 探空后仅发生 1 次闪电的样本有 7 个, 平均发生在探空后 6.4 h。这 7 次雷暴发生的地点有 6 个距离探空站在 37 km 之外, 平均距离 36.97 km。由于雷暴发生距离探空时间较长, 且距离探空站点距离较远, 探空时这些因子可能还没有受雷暴影响, 容易导致漏报。

同样, 统计 6 h 实效 BP 网络模型预报错误的样本发现: 空报的 18 个样本中 6 h 实效前后 1 h 发生雷暴的样本有 5 个, 6 h 实效前后 3 h 发生雷暴的样本有 7 个。漏报的 17 个样本中, 探空前 6 h 未发生雷暴, 探空后仅发生 1 次闪电的样本有 5 个, 平均发生在探空后 3.2 h。这 5 次雷暴发生的地点有 4 个距离探空站在 29 km 之外, 平均距离 26.73 km。

3 结论与讨论

1) 利用和雷暴相关的 9 个因子: 修正总指数、修正 K 指数、沙氏指数、对流稳定度指数、条件对流稳定度指数、Teff 指数、Charba 总指数、瑞士雷暴指数、抬升指数作为辽宁地区雷暴潜势预报的因子。

2) 利用 BP 人工神经网络模型对雷暴天气进行潜势预报研究, 临界成功指数为 57.46%、探测概率为 79.38%、虚假报警率为 32.46% 和总指数为 74.89%, 预报效果较令人满意。

3) 建立 12 h 和 6 h 潜势预报 BP 神经网络模型, 通过预报结果对比发现, 6h BP 神经网络模型更适用于沈阳地区雷暴潜势预报。对预报误差进行初步分析, 发现雷暴发生与探空的时间差、距离、雷暴强度、陡度、密集程度都对探空因子产生影响, 从而影响预报精度。雷暴临近预报还是需要结合卫星云图、雷达回波等数据进行综合预报来提高预报精度。

参考文献:

- [1] 栾健, 李家启, 肖稳安, 等. 基于信息扩散理论的雷电天气关注度研究 [J]. 西南师范大学学报(自然科学版), 2013, 38(9): 143-149.
- [2] 栾健, 李家启, 肖稳安, 等. 闪电定位资料的误差订正 [C] //浙江省气象学会, 上海市气象学会. 江苏镇江: 第九届长三角气象科技论坛论文集, 2012.
- [3] 李家启, 栾健, 王鹏, 等. 闪电定位资料质量控制中小幅值地闪范围的研究 [J]. 高电压技术, 2014, 40(3): 727-731.
- [4] 王清川, 寿绍文, 田晓飞, 等. 廊坊市雷电灾害易损性分析、评估及易损度区划 [J]. 干旱气象, 2009, 27(4): 402-403.
- [5] 金龙, 陈宁, 林镇山. 基于人工神经网络的集成预报方法研究和比较 [J]. 气象学报, 1999, 57(2): 198-207.
- [6] 李家启, 江孟蜀, 王代新, 等. 基于库仑定律的经验保护角法形成机理分析 [J]. 西南师范大学学报(自然科学版), 2012, 37(11): 28-31.
- [7] 李家启, 王代新, 江孟蜀, 等. 基于 IEC62305 避雷针最大保护角法防雷效果分析 [J]. 西南大学学报(自然科学版), 2012, 34(11): 43-46.
- [8] 李家启. 基于 LLS 的雷电流参数随海拔变化特征分析 [J]. 西南大学学报(自然科学版), 2013, 35(5): 77-84.
- [9] 栾健, 李家启, 肖稳安, 等. 基于信息扩散理论的雷电天气关注度研究 [J]. 西南师范大学学报(自然科学版), 2013, 38(9): 143-149.
- [10] 李志鹏, 邵宪友, 张东兴, 等. 基于 BP 神经网络的电控发动机故障诊断研究 [J]. 重庆理工大学学报(自然科学版), 2015, 29(8): 6-11.
- [11] 栾健, 李家启, 肖稳安, 等. 闪电定位系统监测小幅值闪电的特征分析 [C] //中国气象局第 29 届年会论文集. 沈阳: 中国气象学会, 2012: 1-11.
- [12] 李家启, 栾健, 王鹏, 等. 闪电定位资料质量控制中小幅值地闪范围的研究 [J]. 高电压技术, 2014, 40(3): 727-731.

- [13] 李家启, 申双和, 刘俊, 等. 重庆地区地闪活动空间分布及其趋势分析 [J]. 西南大学学报(自然科学版), 2012, 34(5): 77—84.
- [14] 蒋勇军, 况明生, 李林立, 等. 重庆市自然灾害的综合区划及评价 [J]. 西南师范大学学报(自然科学版), 2003, 28(4): 627—632.
- [15] 陈勇伟, 郑涛, 王汉堃, 等. 基于 BP 神经网络模型的雷电潜势预报 [J]. 干旱气象, 2013, 31(3): 595—601.
- [16] 杨仲江, 蔡波, 刘旸. 利用双隐层 BP 网络进行雷暴潜势预报试验——以太原为例 [J]. 气象, 2013, 39(3): 377—382.
- [17] 刘宸钊, 卓伟, 裴军林. 基于对流参数的雷暴预报方法研究 [J]. 高原山地气象研究, 2010, 2(5): 0022—04.
- [18] 丛爽. MATLAB 工具箱的神经网络理论与应用 [M]. 合肥: 中国科学技术大学出版社, 2003.
- [19] 袁曾任. 神经网络及其应用 [M]. 北京: 清华大学出版社, 1999.
- [20] MIRCHANDANI GAGAN, GAO WEI. On Hidden Nodes for Neural Nets [J]. IEEE Transactions on Circuits and Systems, 1989, 36(5): 661—664.
- [21] 高大启. 有教师的线性基本函数前向神经网络结构研究 [J]. 计算机学报, 1998, 21(4): 80—86.
- [22] 蒋乐怡, 黄炎, 姚祖庆, 等. 地面形势识别在上海强对流预报中的应用 [J]. 气象, 2000, 26(3): 25—32.
- [23] 孔燕燕, 彭治班, 赵秀英, 等. 寻找降雹预报指标集的一种方法: 逐步消空法 [J]. 气象, 2000, 26(10): 10—15.
- [24] 寿绍文, 励申申, 姚秀萍. 中尺度气象学 [M]. 北京: 气象出版社, 2003: 349—358.
- [25] MADDOX R A, DOSWELL C A. An Examination of Jet Stream Configurations, 500 mb Vorticity Advection and Low-Level Thermal Advection Patterns During Extended Periods of Intense Convection [J]. Mon Wea Rev, 1985, 110: 185—197.

A Thunderstorm Potential Forecast Model of Shenyang Based on the Double Hidden Layer BP Artificial Neural Network

LIN Zhong-guan¹, LUAN Jian², WANG Di¹,
ZHANG Chun-long³, LUAN Lan⁴

1. Liaoning Meteorological Service Center, Shenyang 110166, China;

2. China Meteorological Administration Training Center Liaoning Branch, Shenyang 110166, China;

3. Meteorological Disaster Prevention Technology Center in Heilongjiang Province, Harbin 151030, China;

4. Shenyang Agricultural University Agronomy Courtyard, Shenyang 110866, China

Abstract: In an experiment reported in this paper, nine factors suitable for thunderstorm potential forecast of Shenyang were selected based on the sounding factors and the lightning location data from 2007 to 2011 in June, July and August in this city, and a double hidden layer BP artificial neural network prediction model was built to forecast the local thunderstorm potential development. The experiment showed that the double hidden layer BP artificial neural network prediction model gave satisfactory results, with a critical index of success of CSI=57.46%, a detection probability of POD=79.38%, a false alarm rate of FAR=32.46% and a total index of TS=74.89%. A comparison of the potential forecast results between 12 h and 6 h demonstrated that the forecast effect of the 6h BP neural network model had better effect and was, therefore, more suitable for the thunderstorm potential forecast in Shenyang.

Key words: double hidden layer; BP neural network; thunderstorm; potential forecast; model

