

BP 神经网络模型在重庆伏旱 预测中的应用研究^①

李永华¹ 刘 德¹ 金 龙² 高阳华¹

(1. 重庆市气象局, 400039; 2. 广西气象局)

提 要

采用气象要素定义伏旱指数,利用小波分析等方法分析重庆地区伏旱变化特征,最后采用 BP 神经网络模型对伏旱进行预测试验,结果表明,重庆伏旱变化具有明显的阶段性特征,而基于 BP 神经网络模型的伏旱预测模型预测效果良好,可以应用于实际预测。

关键词: BP 神经网络 伏旱 预测

引 言

伏旱是重庆地区主要的气象灾害,出现的频率高,造成的损失大。据统计,仅 2001 年出现的伏旱给重庆市造成的直接经济损失约 30 亿元。因此,准确预测伏旱,从而为政府部门提供防灾减灾的科学决策依据具有重要的意义。伏旱同其它气候变化一样,具有十分显著的非线性演变特征,而传统的线性统计方法往往难以取得理想的预测效果。人工神经网络(ANN)是一个具有高度非线性的大规模的动力系统,其特色在于信息的分布式存储和并行协同处理。神经网络系统具有集体运算的能力和自适应的学习能力,它还具有很强的容错性。运用神经网络模型可实现函数逼近、数据聚类、模式分类、优化计算等功能,因此,神经网络广泛应用于人工智能、自动控制、机器人、统计学等领域的信息处理中,近年来在大气科学领域中也有较多的应用^[1-4]。从近些年来神经网络模型理论在各个领域应用的发展情况来看,基于人工神经网络的气候预测技术的应用前景是光明的。

人工神经元模型是对生物神经元模型的简化、抽象和模拟,它是神经网络的基本处理

单元。在这个简化抽象过程中,从不同的角度来考虑神经元及其网络所模拟的对象特点以及现有技术实现的可能性,就产生了形式各异的神经元模型。本文采用通用性好、较为成熟的 BP 神经网络模型对重庆地区伏旱进行预测试验,同时与采用其它传统统计方法的预测试验结果进行对比,从而考察其在伏旱预测中的应用前景。

1 伏旱定义及变化特征

对伏旱的定义,通常是采用时段雨量累计和的计算标准,即连续 20 天雨量小于 30mm 进入伏旱期,日雨量大于 25mm 或过程雨量大于 30mm 为伏旱结束,再根据累计雨量及伏旱天数定义伏旱等级,这样的定义方法仅仅考虑了降雨量,并且在确立预报对象标准时难以客观定量。为此,我们利用伏旱期(6 月下旬~8 月上旬)各站的高温日数(D ,指 $\geq 35^{\circ}\text{C}$ 的日数)、极端最高气温(T_M)和总降水量(R)的标准化量定义了伏旱指数(F)(如式(1)),该伏旱指数同时考虑了气温和降水对伏旱的贡献,经对比计算,其正常值为 1.8,伏旱指数(F)越大,表明伏旱越重,反之则轻,与历史实况比较,较为客观定量地刻画了伏旱特征。

^① 国家自然科学基金项目(40075021)及重庆市气象局科技计划项目(ZL200204)资助

$$F = \frac{1}{n} \sum_{i=1}^n \frac{D_i + T_{M_i}}{R_i} \quad (1)$$

式中 n 为测站数。

图 1a 是重庆地区 1959~2001 年伏旱指数 (F) 的变化曲线(实线)及多项式拟合曲线(虚线),从图 1a 可以看出,近 40 余年来,重庆地区 20 世纪 60 年代前期、70 年代末到 80 年代以及 90 年代末期伏旱相对较轻,其余时间均有较重的伏旱,其中 1959、1961、1966、1971、1975、1990、1992、2001 年伏旱严重,尤以 2001 年为最。

图 1b 是重庆地区 1959~2001 年伏旱指数 (F) 的墨西哥帽函数 (MHF) 小波变换结果(图中纵坐标为尺度,横坐标为时间(年),虚线为负等值线,实线表示小波变换系数的

零或正等值线,等值线间隔为 0.2),从图 1b 中可以看到,从年代际尺度变化而言,近 40 余年来,20 世纪 70 年代后期到 80 年代末期变化相对稳定,伏旱相对较轻,而 20 世纪 60 年代到 70 年代中期和 90 年代前中期伏旱相对严重。从更小的年际尺度来看,变化相对复杂一些,但也表现出阶段性变化特征,如对 4 年时间尺度来说,伏旱大致有七个变化阶段,它们是 1959~1962 年的相对重旱期、1963~1969 年的相对轻旱期、1970~1978 年的相对重旱期、1979~1987 年的相对轻旱期、1988~1994 年的相对重旱期、1995~2000 年的相对轻旱期、2001 年的相对重旱期。

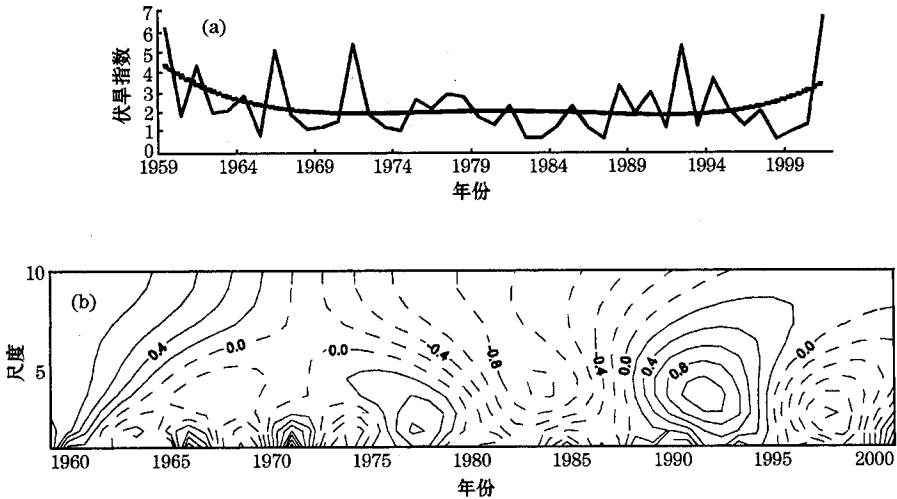


图 1 重庆地区 1959~2001 年伏旱指数(a)及其小波变换(b)

2 BP 神经网络的原理和方法

BP 神经网络即误差反向传播神经网络 (Back propagation Neural Network),它是一种单向传播的多层前向网络,也是一种最常用的神经网络。网络第一层称为输入层,最末一层为输出层,中间各层均称为隐含层。同层神经元节点间没有任何耦合。输入信息从输入层依次向输出层传递,每一层的输出只影响下一层的输入。

本文拟采用算法较成熟、应用较广泛的前向三层 BP 网络算法,对伏旱指数序列作拟合和预测试验,它包括一个输入层、一个隐

含层和一个输出层。该方法的计算公式和步骤见文献[5]。

3 BP 神经网络模型预测试验

3.1 建模方案

我们用重庆地区伏旱指数 (F) 序列建立 BP 神经网络预测模型。建立模型的资料样本为 1959~1995 年、1959~1996 年、1959~1997 年、1959~1998 年、1959~1999 年、1959~2000 年,并分别作 1996 年、1997 年、1998 年、1999 年、2000 年、2001 年的指数预测。

我们在建立 BP 网络预测模型时,将伏

旱指数序列作为学习矩阵的期望输出。在选择作为输入矩阵的预报因子时,考虑了三种不同的方案:①仅考虑影响伏旱指数的强信号(物理因子),其做法是用伏旱指数序列与待选物理因子序列普查相关,考虑到模型的实际应用,因子的选取定为上年及当年1~3月达到0.05相关显著性水平以上的前12个因子;②仅考虑伏旱指数本身不同时间尺度的周期变化,其做法是采用达到0.01相关显著性水平以上的前12个延拓均生函数^[6]序列;③既考虑影响伏旱指数的强信号(物理因子),又考虑伏旱指数本身不同时间尺度的周期变化,各选6个因子进行组合。为了满足BP网络节点函数的条件及有效提高网络训练速度,我们将学习矩阵的训练样本数据标准化在0.1~0.9之间:

$$\hat{x}_i = \frac{x_i - s_1}{s_2 - s_1} \quad (2)$$

其中 x_i 、 \hat{x}_i 分别是实况值和标准化值, $s_1 = (9x_{\min} - x_{\max})/8$, $s_2 = (9x_{\max} - x_{\min})/8$, x_{\max} , x_{\min} 分别为各样本序列中的最大值和最小值。

在建立伏旱指数BP神经网络预测模型时,三个方案均选取了12个相关显著的因子作为输入矩阵,即输入层节点为12,输出节点为1(伏旱指数),隐层节点为10,学习因子为0.7,动量因子为0.9,总体误差为0.001。将学习矩阵加载到网络输入端,进行反复的学习训练,当误差函数趋于稳定并达到规定限值,训练结束,便得到相应的预测模型。

同样,采用三个方案所选择的12个相关显著的因子也可以建立相应的逐步回归预测模型。

3.2 结果分析

图2是重庆地区伏旱指数实况值及BP神经网络模型和逐步回归模型拟合值变化曲线,其建模样本为1959~2001年,因子的选择采用方案1。可以看出,逐步回归模型对

伏旱指数拟合较理想,但BP神经网络模型对伏旱指数的拟合更优于逐步回归模型,为了定量比较两种模型的拟合效果,特定义下列三个特征指标:

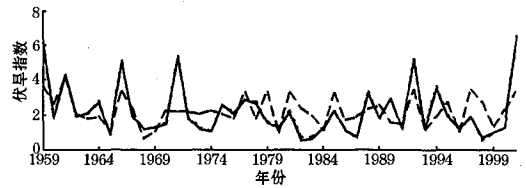


图2 重庆地区伏旱指数实况值与拟合值变化曲线

I. 平均绝对百分比误差:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{|y_i|} \quad (3)$$

II. 均方根误差:

$$MSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (4)$$

III. 平均绝对误差:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (5)$$

式(3)~(5)中 y_i 、 \hat{y}_i 分别代表实况值和拟合值。各指标统计结果如表1。从表1可以看出,但BP神经网络模型的拟合精度明显高于逐步回归模型。

表1 两种预测模型统计评价指标比较

特征指标	BP神经网络预测模型	逐步回归预测模型
MAPE	0.03	0.40
MSE	0.10	1.22
MAE	0.10	0.95

以上从模型的拟合情况进行了比较,BP神经网络预测模型的拟合精度高于逐步回归预测模型,但预测模型的拟合能力高并不能代表其实际预测能力也高,而预测模型的优劣还应体现在其预报能力的高低上。因此,我们采用上述三种方案(主要是选择因子的差异)分别采用BP神经网络模型和逐步回归模型作预测试验,以考察BP神经网络模型的预测能力,结果如表2。表2中“√”表示(实况值-1.8)与(预测值-1.8)同号,“×”表示(实况值-1.8)与(预测值-1.8)异

号。

从表 2 可以看出,总体上,以上三种方案中两种预测方法对伏旱指数的趋势预测比较理想。在作为预测检验的六个独立样本(1996~2001 年)中,1996、1998、1999、2000 年属于(偏)涝年,而 1997、2001 年属(偏)旱年,从三个方案来看,BP 神经网络模型的预测效果均好于相应的逐步回归模型,表现在平均绝对误差前者小于后者,且准确率总体上前者大于后者,说明采取同样的因子,前者的预测效果较优。另外,采取三个不同的方案(考虑的因子不同),其预测效果也是不同的,总体而言,对 BP 神经网络模型来说,三种方案预测效果均较好,但对逐步回归模型来说,方案 3 既考虑影响伏旱指数的强信号(物理因子),又考虑伏旱指数本身不同时间

尺度的周期变化时,预测效果较好,其次较好的是第 2 种方案,说明在作气候预测时,既要考虑较优的计算方法,还要考虑物理因子的稳定性及物理意义。从上述的分析中可以看出,BP 神经网络模型在伏旱指数预测中优于逐步回归模型,但选取不同的输入因子也在一定程度上会影响预测结果,因此,在使用该模型时,同样也应充分考虑输入因子的物理意义。从预测试验看,个别年份预测效果并不理想,甚至趋势预测结果与实况相反,如方案 1 对 1998 年伏旱的预测,三个方案对 2001 年的偏旱趋势虽基本正确,但伏旱指数绝对误差较大,异常偏旱未能准确预测,这说明,数理统计模型的拟合效果好而实际预报不一定好,同样,神经网络模型学习能力较强,推广能力虽然也明显好于逐步回归,但仍明显低于其对历史样本的拟合效果。因此,无论是传统的数理统计模型或是神经网络模型,都要用一定量的独立样本作为预测集对模型进行检验,效果满意才能用于实际预报。如果我们很好地应用神经网络方法,能比传统的数理统计方法取得更好的效果,尤其是它具有很好的处理非线性问题的能力,而大量的短期气候预测问题都具有非线性演变特征,因此这种方法具有广阔的应用前景。

表 2 两种预测模型的独立样本预测效果比较

	年份	实况	BP 网络 预测值	绝对 误差	逐步回归 预测值	绝对 误差	
方 案 1	1996	1.4	1.6(√)	0.2	1.4(√)	0	
	1997	2.1	2.7(√)	0.6	4.3(√)	2.2	
	1998	0.7	4.2(×)	3.5	3.5(×)	2.8	
	1999	1.1	1.0(√)	0.1	2.1(×)	1.0	
	2000	1.4	1.4(√)	0	3.2(×)	1.8	
	2001	6.6	5.3(√)	1.3	2.5(√)	4.1	
	预报准确率 平均误差			5/6(83%)	3/6(50%)		0.95
方 案 2	1996	1.4	1.2(√)	0.2	1.6(√)	0.2	
	1997	2.1	2.8(√)	0.7	3.2(√)	1.1	
	1998	0.7	1.2(√)	0.5	0.9(√)	0.2	
	1999	1.1	2.5(×)	1.4	3.5(×)	2.4	
	2000	1.4	1.6(√)	0.2	1.9(×)	0.5	
	2001	6.6	3.1(√)	3.5	2.5(√)	4.1	
	预报准确率 平均误差			5/6(83%)	4/6(67%)		1.08
方 案 3	1996	1.4	1.3(√)	0.1	1.4(√)	0	
	1997	2.1	2.8(√)	0.7	3.2(√)	1.1	
	1998	0.7	2.5(×)	1.8	1.4(√)	0.7	
	1999	1.1	1.6(√)	0.5	3.1(×)	2.0	
	2000	1.4	1.7(√)	0.3	1.8(√)	0.4	
	2001	6.6	2.1(√)	4.5	2.8(√)	3.8	
	预报准确率 平均误差			5/6(83%)	5/6(83%)		1.32

4 结 语

本文采用 BP 神经网络模型对重庆地区伏旱指数进行了拟合和独立样本预测试验,结果均表现出比逐步回归模型优越的性能,表明该预测模型在气候预测中有较好的应用前景,但 BP 神经网络模型也表现出预测结果不够稳定、可调参数较多(人为干预较多)、误差函数还会出现振荡、不收敛等模型本身的固有问题,需要进一步深入研究。同时,在利用该模型作预测时,还要充分考虑因子的物理意义,这样才能使预测效果更加稳定,更加准确。

参考文献

- 1 Jin Long(金龙), Luo Ying(罗莹), Lin Zhenshan(林振山). Study on mixed model of neural network for farmland flood/drought prediction. Acta Meteorological Sinica, 1997, 11(3): 364—373.
- 2 金龙, 陈宁, 林振山. 基于人工神经网络的集成预报方法研究和比较. 气象学报, 1999, 57(2): 198~207.
- 3 金龙, 罗莹, 袁成松等. 农业气象干旱防灾减灾的重要途径——一种新的土壤湿度预报方法. 南京大学学报, 1996, 32(特辑): 100~107.
- 4 李永华, 刘德, 金龙. 基于BP神经网络的汛期降水预测模型研究. 气象科学, 2002, 22(4): 461~467.
- 5 省地气象台短期预报岗位培训教材. 北京: 气象出版社, 1998: 248~252.
- 6 魏凤英, 曹鸿兴. 长期预测的数学模型及其应用. 北京: 气象出版社, 1990: 42~46.

Study on Late-summer Drought Prediction Model in Chongqing with BP Neural Network

Li Yonghua¹ Liu De¹ Jin Long² Gao Yanghua¹

(1. Chongqing Meteorological Bureau, 400039; 2. Guangxi Autonomous Region Meteorological Bureau)

Abstract

The late-summer drought indexes are defined by using meteorological elements. And the variations of late-summer drought in Chongqing is made with analysis of wavelet. The late-summer drought prediction model based on BP neural network is established, too. The results show that the variations of late-summer drought in Chongqing contain obvious periodic features. The model is a superior and useful one for the actual operational forecasting.

Key Words: BP neural network late-summer drought prediction