

模型参数对赤霉病人工神经网络 预报模型精度的影响^①

居为民

高 革 金 龙 陈 宁

(南京大学大气科学系, 南京 210093)

(江苏省气象科学研究所)

提 要

运用最优化相关普查方法,选取确定了对江苏省太湖地区小麦赤霉病发生程度有着显著指示意义的预报因子,将其作为输入变量经多层前馈型神经网络的BP算法进行学习训练,建立了赤霉病病穗率的人工神经网络预报模型。分析了结构参数对模型效果的影响情况,发现训练的总误差平方和对模型的效果影响最为显著,历史样本的拟合率随着总误差平方和的减小而稳定上升,但总误差平方和取值偏小时模型对独立样本的预报精度下降;当总误差平方和取适当值使模型稳定时,隐含层节点数、动量因子和学习因子对模型效果的影响可以忽略不计。

关键词: 小麦赤霉病 人工神经网络 模型参数

引 言

长江中、下游地区小麦生长发育后期最主要的病害之一是小麦赤霉病,其病情程度受到小麦抽穗扬花期的温、湿度条件控制,是典型的“气象型”病害,这就为从气象角度预报赤霉病的发生程度提供了可行性^②,前期及近期500hPa大气环流因子又控制了天气条件,且关系密切相关,故必然影响到赤霉病的发生发展。

作者曾利用500hPa大气环流特征量作为太湖地区小麦赤霉病流行程度长期预报因子,建立了回归模型^[1],但由于环流因子对赤霉病发生发展的作用是综合的,相互关系是非线性的,使赤霉病的历史拟合及精确预报变得非常困难和复杂。基于误差反向传播学习算法(BP)的多层前馈型人工神经网络模型引入了多层隐节点,解决了非线性样本问题,并具有高速度、强容错能力及稳健性等优良特性,已被广泛应用于信息处理、自动控制

和管理工程等各个领域,在中、长期天气预报中的应用也有不少成功的例子^[2],但在病虫害预报上的应用才刚起步^[3]。本文以太湖地区赤霉病为例,利用74项500hPa大气环流特征量作为因子,进行该方法在病虫害预报中应用的研究,讨论了模型参数对模型精度的影响作用,以便为该方法的实际应用提供有益的借鉴。

1 赤霉病病穗率预报模型的建立

1.1 网络预报模型的基本原理

典型的BP网络由输入层、隐含层和输出层组成,各层之间实行全连接,同一层单元之间不存在连接。已经证明含有一个隐含层的BP网络可以逼近任意有界连续非线性变量,通常在选择网络层数时,一般选包含一个隐含层的三层网络^[4]。所以在建立模型时采用了隐含层数为1的3层前馈型网络,其各层单元间的连接图及该模型的数学原理可参见文献^[4]。

^① 本文得到中国气象局青年基金和江苏省局课题资助

^② 江苏省植物保护站,江苏省农作物主要病虫害预测预报办法,内部资料

1.2 BP网络的实现步骤

BP网络在实际建模时按照以下步骤实现,即:

①首先,随机地给出一组从输入层到隐含层的连接权函数 v_{hi} 和一组从隐含层到输出层的连接权函数 w_{ij} ,以及隐含层单元的阈值 θ_i 和输出层单元的阈值 γ_j 。

②根据学习样本输入及连接权矩阵,计算隐含层的激活值:

$$b_i = f \sum_{h=1}^n (a_h v_{hi} + \theta_i) \quad (1)$$

式中: $i = 1, 2, \dots, p$, 为隐含层节点数; $h = 1, 2, \dots, n$, 为输入层节点数; f 为激励函数, 其表达式:

$$f(x) = 1/(1 + e^{-x}) \quad (2)$$

③计算输出层单元的激活值:

$$c_j = f \sum_{i=1}^p (b_i w_{ij} + \gamma_j) \quad (3)$$

式中 $j = 1, 2, \dots, q$, 为输出层节点数。

④计算输出单元的一般化误差 d_j :

$$d_j = C_j(1 - C_j)(C_j^k - C_j) \quad (4)$$

其中 $k = 1, 2, \dots, m$, 为样本数; C_j^k 为输出层单元 j 的期望输出。

⑤计算隐含层单元相对于每个 d_j 的误差 e_i 。

⑥调整隐含层单元到输出层单元的连接权 Δw_{ij}

⑦调整输出层单元的阈值:

$$\Delta \gamma_j = \alpha d_j \quad (5)$$

α 为人为给定的动量因子值。

⑧调整输入层到隐含层的连接权 Δv_{hi} 。

⑨调整隐含层单元的阈值:

$$\Delta \theta_i = \beta e_i \quad (6)$$

β 为人为给定的学习因子值。

⑩重复②到⑨的步骤,直到 $j = 1, 2, \dots, q$ 和 $k = 1, 2, \dots, m$, 实际输出与期望输出的误差 d_j 变得足够小为止,学习结束。通过学习训练后,网络从实际样本中获得了知识

信息,对于学习过的样本具有记忆和联想能力。利用训练好的网络对未参加学习的样本进行预报。

1.3 赤霉病预报模型的建立

建立赤霉病病穗率的人工神经网络预报模型时,以1952~1989年的资料数据做为训练样本用于建模,1990~1999年的资料作为独立样本用于模型效果检验。首先用滑动最优化相关普查方法在前一年1月到当年3月的74项500hPa大气环流因子中选择与当年赤霉病病穗率相关显著的因子;对初选上的因子进行独立性检验,剔除存在复共线性的因子,一共选择了7个因子输入到网络的输入层进行学习训练。这些因子为2月西太平洋副高强度指数、4月北美大西洋副高脊线、2月太平洋副高脊线、10~12月西太平洋副高北界的平均值、4月南海副高北界、10月亚洲区极涡面积指数、3月太平洋区极涡强度指数。

如何合理给定网络模型的参数,目前尚无客观定量化的指标,常用的方法是给定不同的参数值,在比较模式性能指标的基础上,确定计算量和模型精度都能满足要求的参数值。经反复对比试验,动量因子 $\alpha = 0.7$, 学习因子 $\beta = 0.5$, 隐含层节点数取6, 训练的全局误差平方和 E (全局误差平方和是所有样本输出值与实际值之差的平方和) 控制为0.005时,模型对独立样本的试报精度最高,对历史样本的拟合率也通过信度为0.05的可靠性检验。图1给出了每一年的拟合和试报结果,可以发现模型不仅对历史样本具有较高的拟合率,特别是对独立样本表现出了较好的预报能力。若预报的赤霉病等级与实况值完全相同认为预报正确,误差在一个等级以内的认为预报基本正确,则模型对1990~1999年的10年赤霉病试报的正确率为60%,基本正确率为90%,特别是对1990年和1998年两个病穗率在40%以上的赤霉病

大发生年作出了准确的预报;误差最大的是 1991 年,预报结论是中等偏重,而实况是偏

轻发生。从模型 10 年的试报效果来看,模型是可靠的。

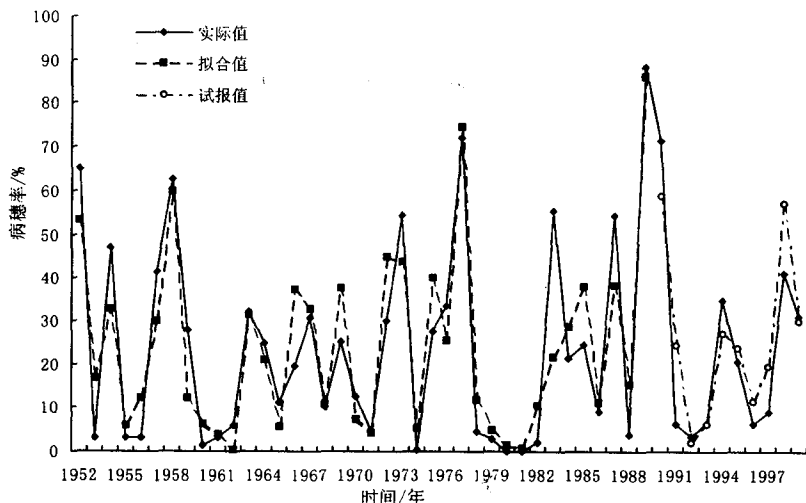


图 1 神经网络模型对赤霉病病穗率历史样本的拟合和独立样本的预报结果

2 网络参数对模型精度的影响

为了分析动量因子 (α)、学习因子 (β)、隐含层节点数 (N_h)、训练的全局误差平方和 (E) 等参数对模型性能的影响,定义均方误差做为评价的定量化指标。即:

$$MSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (7)$$

对历史样本用 MSE_1 表示,对独立样本用 MSE_2 表示。表 1 给出了 E, N_h, α, β 取不同数值时,计算得到的 MSE_1 和 MSE_2 的结果,可以发现: MSE 随着 N_h 的变化而变化。因此可获得以下几个特点:

表 1 模型参数对模型效果的影响分析结果

E	N_h	0.001		0.002		0.003		0.004		0.005		0.006	
		MSE_1	MSE_2	MSE_1	MSE_2	MSE_1	MSE_2	MSE_1	MSE_2	MSE_1	MSE_2	MSE_1	MSE_2
$\alpha=0.9$ $\beta=0.7$	4	4.87	27.86	6.99	23.43	8.37	21.35	9.12	11.95	10.57	10.08	12.72	11.88
	5	5.29	31.66	6.91	18.24	8.01	12.32	9.54	12.40	10.80	10.02	13.47	11.40
	6	5.20	31.54	7.13	24.46	8.49	20.28	9.35	13.48	10.71	9.88	13.19	11.79
	7	5.37	32.61	6.58	24.95	7.99	16.31	9.32	11.56	10.71	10.21	13.15	11.88
$\alpha=0.7$ $\beta=0.5$	8	5.15	25.01	6.73	25.65	8.02	14.96	9.38	12.87	10.70	10.60	13.00	11.52
	4	5.69	30.81	6.74	23.37	8.11	22.64	9.20	19.43	10.37	9.71	11.65	12.57
	5	4.67	24.88	6.47	16.44	9.50	12.42	9.40	12.36	10.44	9.58	11.91	12.27
	6	5.23	27.33	6.74	21.59	8.18	17.79	9.32	13.48	10.40	9.90	11.87	12.39
$\alpha=0.5$ $\beta=0.5$	7	4.77	24.42	6.46	16.94	8.02	14.23	9.33	12.28	10.39	9.52	11.89	12.28
	8	4.68	22.54	6.52	17.63	8.06	13.60	9.36	13.14	10.38	10.00	11.89	12.06
	4	5.11	28.35	6.73	22.29	8.14	15.70	9.31	12.82	10.57	9.91	11.59	12.77
	5	4.58	26.54	6.55	16.26	8.15	13.04	9.46	12.65	10.48	9.81	11.69	12.31
$\alpha=0.5$ $\beta=0.5$	6	4.84	30.62	6.75	24.41	8.19	20.28	9.39	13.49	10.44	9.89	11.68	12.48
	7	4.71	23.85	6.55	17.01	8.11	13.76	9.42	12.55	10.43	9.92	11.69	12.56
	8	4.67	25.28	6.56	19.27	8.14	16.35	9.43	13.56	10.41	10.00	11.70	12.25

(1)在所有的参数中全局误差平方和 E 对模型效果的影响最为显著。对历史样本而言,随着 E 的减小,模型的训练次数增加, MSE_1 稳定地下降,标志着模型的拟合率上升;而对于独立样本,当 E 取值偏小时 ($E < 0.005$), MSE_2 反而上升,预报准确率下降。这是因为过度训练所引起的模型剧烈振荡,使得网络的推广性变差;但若 E 取值过大 ($E = 0.006$), 训练次数太少,识别率太低, MSE_2 也会上升,这与文献[5]和[6]的结论是一致的。所以,在实际应用时,应特别注意 E 值的选取或训练次数的控制,可以通过多样本试报试验确定。根据多次计算的经验、误差的收敛情况和拟合率,本文 E 的值取 0.005。

(2)总体而言,隐含层节点数 (N_h) 对模型精度的影响较小,但是对独立样本试报精度的影响大于对历史样本拟合率的影响,特别是在 E 取值偏小的情况下更是如此;当 E 取值偏小时,模型的精度随着 N_h 振荡;如果 E 取值适当, N_h 对模型精度的影响变得非常小。

(3)在 E 取值偏小时,动量因子 α 以及学习因子 β 同样会引起模型预报精度的振荡;随着 α 、 β 取值的减小,模型总体上对历史样本的拟合精度上升,对独立样本的试报精度略有下降。

3 结论

(1)人工神经网络模型对非线性样本具

有比一般回归模型更好、更强的拟合和预报能力,用于受多种气候要素综合影响的小麦赤霉病的预报是可行的。

(2)神经网络模型的精度受到模型参数的影响,特别是模型学习的全局误差平方和或训练次数的合理选择是模型取得理想预报精度的关键,应避免片面追求对历史样本的高拟合率,而形成“过度训练”,使得模型的预报精度降低。

(3)隐含层因子数 N_h 、动量因子 α 和学习因子 β 对模型预报精度的影响较小,但在 E 取值偏小时,会引起模型的预报精度剧烈振荡,失去稳定性。

参考文献

- 1 高苹,居为民.冬小麦赤霉病流行程度长期预报模型.气象,1998,24(6):55~57.
- 2 Jing Long, Luo Ying, Guo Guang and Lin Zhengshan. Study on mixed model of neural network for farmland flood/drought prediction. Acta Meteorologica Sinica, 1997,11(3):362~374.
- 3 文新辉,陈开周,牛明浩.一种新的昆虫神经预测预报方法.系统科学与数学,1995,15(1):64~74.
- 4 王永骥,涂健.神经网络控制.北京:机械工业出版社,1999.
- 5 张立明.人工神经网络的模型及其应用.上海:复旦大学出版社,1993:22~37.
- 6 冯国章,李佩成.人工神经网络结构对径流预报精度的影响分析,自然资源学报,1998,13(2):169~174.
- 7 王成刚,吴宝俊,朱官忠. BP 网络在鲁西南地区西南涡降水量级预报中的应用试验.气象科学,1999,19(1):158~164.

The Influence of Parameters on Prediction Accuracy of ANN Model for Winter Wheat Scab

Ju Weimin Gao Ping Jin Long Chen Ning

(Meteorological Institute of Jiangsu Province, Nanjing 210008)

Abstract

By dint of optimum correlation method the predictors, which are good indicators of winter wheat scab occurrence in Taihu area in Jiangsu province, were input into the feed-forward multi-layer artificial neural network (ANN) based on back-propagation study. ANN prediction model for occurrence of winter wheat scab was developed. The influence of model parameters on the fitting and prediction accuracy of model was studied. It can be concluded that total error square sum (E) has the most outstanding impact on model function than other parameters. The smaller is the value of E , the higher is fitting accuracy for historical samples. Extreme small E will result in lower prediction accuracy for independent samples. The action of number of neurons in the hidden layer (N_h), study factor (α) and momentum factor (β) may be ignored when suitable E was given to make model stable.

Key Words: winter wheat scab artificial neural network model parameter