

定性和定量长期预报模型的综合分析*

金 龙

(江苏省气象科学研究所, 南京, 210008)

苗春生

(南京气象学院, 南京, 210044)

陈 宁

(江苏省气象科学研究所, 南京, 210008)

罗 莹

(江苏省气候应用所, 南京, 210009)

摘 要

根据相同的 500 hPa 和海温场预报因子, 利用神经网络灵活可变的拓扑结构, 分别构造了定性和定量的降水长期预报模型。并在同等条件下, 建立了逐步回归预报方程。通过对比分析表明, 这种定性和定量相结合的神经网络综合预报分析方法, 是增强预报结果可靠性和稳定性的一种有效途径。该预报建模方法具有比较合理的分析依据, 值得进一步探索、应用。

关键词: 综合分析, 神经网络, 定性, 定量。

1 引 言

在月、季降水量长期预报的业务工作中, 统计预报方法仍占有重要位置^[1]。从根本上讲, 统计预报模型对历史样本的拟合效果和实际预报能力除了与预报建模的方法有关以外, 预报因子的好坏具有最重要的影响。然而对于一个相同的预报问题, 即在预报量和预报因子相同的条件下, 通过减少预报问题“解域”中解的个数, 建立简单的定性预报模型(只预报正负距平趋势类别), 使定性预报比相应的定量预报更“容易”, 预报结果更准确。然后以定性预报结果为背景, 进一步建立定量预报, 以适应旱涝长期预报的实际工作需要。下面是试验用神经网络方法来进行这种从定性到定量的预报建模方法的尝试。

2 基本方法和预报因子

目前, 中国国内在长期预报的业务工作中采用的统计预报方法, 其大部分均属于线性的顺序处理技术。而当前国内外迅速发展的人工神经网络方法在众多的学科领域中得到了应用^[2~4]。由于神经网络方法所具有的自适应学习和非线性映射能力, 使得在大气科学的预报建模研究中也显示了很好的应用前景^[5,6]。该方法的主要特点是比较适用于处理因果关系不甚明确, 推理规则不太确定的复杂非线性问题, 并且有灵活可变的网络拓扑结构, 从而为我们对某一预报系统建立单输出的定量预报模型和多输出的定性预报模型提供了可能。

本文在实际建立定性和定量的长期预报模型时, 均采用了广泛使用的前馈网络模型, 只是在建立定性预报模型时, 是将预报量输出设计为多输出的网络模型结构, 而定量预报则采用单输出的前馈网络模型。前馈网络模型的基本算法见文献[5]。并且, 实际的预报建

* 初稿时间: 1998 年 9 月 28 日; 修改稿时间: 1999 年 3 月 24 日。

资助课题: 中国气象局“九五”课题。

模均以江苏省南京、苏州、南通、淮阴 4 个站 1952 ~ 1994 年历年 6 ~ 7 月平均降水量作为预报对象,并考虑在实际预报中,预报因子资料取得的时效性,分别进行了往年 6 ~ 12 月,及当年 1 ~ 3 月的 500 hPa 月平均高度场和月平均海温场的相关场普查计算。取大于等于 4 个相连的成片相关格点作为一个预报因子,共得到前期 500 hPa 13 个预报因子和 14 个前期海温场预报因子(见表 1)。

表 1 前期 500 hPa 和海温场预报因子相关系数

序号	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}
500 hPa	0.52	-0.41	0.51	0.45	0.39	0.38	-0.36	-0.41	0.49	-0.34	-0.33	-0.36	0.35	
SST	0.45	0.46	-0.33	-0.40	0.39	-0.39	0.44	-0.41	-0.45	0.41	0.55	-0.39	-0.35	-0.34

由于本文拟建立定性和定量预报模型时,是采用前馈网络方法,该方法是将预报系统的预报因子与预报量处理为输入与输出之间的一种非线性映射关系,并且实现这种关系并不需要知道预报系统内部的线性或非线性结构,而是通过实际学习样本来模拟得到预报因子与预报量的内在关系。但是所建立的预报模型对历史样本的拟合效果和预报模型的预报推广能力,除了与网络模型的拓朴结构及各参数有关以外,其本质上还取决于预报系统的预报因子与预报量的内在关系。然而,前馈网络方法本身并不提供如何从众多预报因子中选择预报因子。为此,作者首先采用逐步回归方法^[5],从这 27 个预报因子中自动筛选出 5 个预报因子,其中海温场 2 个预报因子分别为 x_7 和 x_{11} , 500 hPa 预报因子为 3 个,分别是 x_5 , x_8 和 x_9 (见表 1)。将这 5 个预报因子分别用于建立前馈网络的定性和定量预报模型。

3 定性预报模型

由于长期天气过程变化的物理机制,目前尚不能完全了解,本文中所选用的 5 个相关因子只是表明其在统计上与预报量存在一定的统计相关关系,因此先从定性的角度建立预报模型,以减少预报模型“解域”中解的个数,从而减小预报“难度”,以提高预报模型的预报准确性,这是本文重点讨论研究的问题。为此,在建立定性预报模型时,首先将预报量按正负距平分成两类,这在前馈网络学习矩阵的期望输出中,将预报量历史样本的正距平设计成 0 1,而负距平设计成 10。由于前馈网络的节点函数是选用了 Sigmoid 函数,其值域为 $[0, 1]$,因此,利用文献[5]的方法,将 5 个预报因子作标准化处理。这样就构造了一个输入节点为 5,输出节点为 2 的前馈网络学习矩阵。通过对学习矩阵作学习训练,当网络模型达到设定的误差收敛标准时,便可以通过网络的连接权系数和相应的阈值计算出预报值。

另外,为了能客观地比较分析这种定性的前馈网络预报模型的实用性能,本文在前面的 500 hPa 和海温场的相关因子普查计算时,预报量只用了 43 个样本,即 1952 ~ 1994 年资料序列,而取 1995 ~ 1997 年资料作为独立样本,1998 年作实际预报,进行预报模型检验。由该预报模型的计算结果(表略)发现,由于定性预报模型的设计对预报只有很“低”的定性趋势要求,只需要预报模型能分辨降水量的正距平和负距平。因此,预报模型的拟合效果是 100% 的正确。并且利用该预报模型对 3 a 独立样本进行实际预报检验,结果表明,3 a 的趋势预报全部正确(表略)。这主要是因为,由于这种定性预报模型,只要求能够作出未来降水量是偏多(正距平)和偏少(负距平)的简单趋势预报,所以模型的预报和拟合效

果是令人满意的。从该模型的独立样本预报及历史样本拟合效果的综合分析, 不难看出, 该模型具有较好的识别推广能力和稳定性。但是, 这种降低了要求的定性预报模型预报结果, 与实际长期预报业务工作希望能有比较定量的预报结论存在一定差距。能否以这样的定性预报模型预报结果作为预报背景, 在此基础上, 再作进一步的定量预报, 以实现先定性后定量的综合预报分析方法。

4 定量预报模型

为了便于进行同等条件下的比较, 在以上定性预报模型的背景条件下, 进一步建立定量预报模型时, 所选用的 5 个预报因子不变, 且同样是用 1952~1994 年 43 个样本建立预报模型, 1995~1997 年 3 a 样本作为独立预报样本检验。这样在前馈网络的定量预报模型中, 输入节点同样为 5, 而输出节点为 1, 隐节点为 7。同样先根据文献[5]的方法, 对网络学习矩阵作标准化处理。然后对学习矩阵作学习训练, 最后得到实际的预报和拟合结果。图 1 给出了前馈网络定量预报模型对历史样本的拟合效果。从图 1 中可以看出, 该预报模型对历史样本也有较好的拟合效果, 尤其对极端多雨和少雨年, 拟合值与实况值也比较吻合。但是, 进一步分析发现, 如果将该预报模型 43 a 的历史样本拟合值按相应实况值正负距平分类对照发现, 43 个样本中有 4 a 拟合值与实况值趋势不符。进一步利用该预报模型对 1995~1997 年 3 a 独立样本进行预报试验(见表 2), 3 a 独立样本的预报平均相对误差为 27.3%。

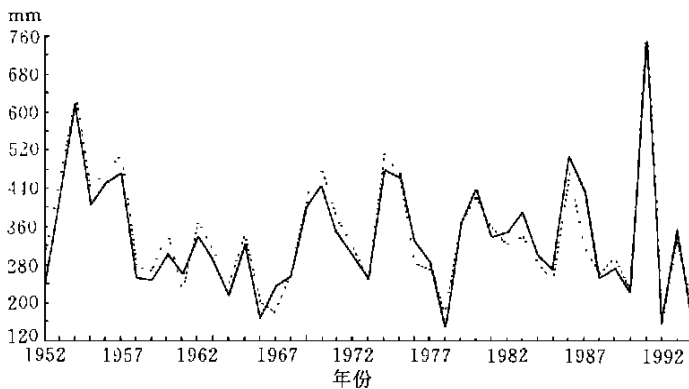


图 1 神经网络定量预报模型的历史样本拟合

表 2 神经网络和回归定量预报模型的预报检验

年 份	实 况	神经网络预报		逐步回归预报	
		预 报	相对误差 (%)	预 报	相对误差 (%)
1995	344.5	454.2	31.8	562.0	63.0
1996	513.0	273.7	- 46.6	452.7	- 11.7
1997	259.4	268.7	3.5	354.4	36.6
绝对值平均			27.3		36.5

5 两种预报模型的综合分析和实际预报

在进行两种预报模型的综合对应分析时,首先对比一下定性和定量预报模型对历史样本的拟合情况(表略),结果发现,定性预报模型对正负距平的两类趋势拟合准确率是100%。而用同样的5个预报因子建立的定量预报模型,同样将拟合值按正负距平分类,其趋势拟合准确率为 $39/43=90.6\%$ 。同时由图1已经看到,定量预报模型的历史样本拟合值与实况已经比较吻合。但是按正负距平分类的趋势拟合准确率还是明显比定性预报模型差,其主要原因可能是,一开始在建立定性预报模型时,只要求预报模型对正负距平两类趋势进行识别,这使得预报因子容易“胜任”。从一般的分析知道,目前还无法用完全定量的数学方法描述长期天气的变化过程,而根据预报量与预报因子统计相关建立预报模型时,降低数值定量的预报要求,可能更符合一般统计相关预报因子的实际预报能力,从而使这些统计预报因子能在要求较低的定性预报模型中,发挥其实际具有的预报能力,减少对这些预报因子要求“过高”而增加预报出错的可能性。但是在建立定量预报模型时,由于模型是直接要求能在数值上精确,这可能超过了这些统计相关预报因子实际所具有的预报能力。因为实际上由逐步回归方法选入的5个预报因子(见表1),与预报量只是统计上一定程度的线性相关关系,而且实际的预报量与预报因子并不一定只是线性相关,同时多因子对预报量的综合影响只要存在非线性关系时,其预报就变得更为复杂。而定量预报模型在方法上(无论预报因子的实际预报能力如何)都是要尽可能地追求预报模型拟合值与实际资料在数值上更一致。这种超过预报因子实际能力的数值上一致性的过高要求可能产生趋势类别上的错误识别。进一步对定性、定量两种预报模型的3a独立样本预报结果(表略)分析发现,定性预报模型对1995、1996和1997年3a的趋势预报均正确,而定量预报结果同样按正负距平划分,1996年预报错,1995、1997年预报对。

从以上两种神经网络的定性和定量预报模型的预报和拟合情况分析看到,两种预报方法所依据的预报因子完全相同,但是定性预报模型无论是历史样本的拟合还是独立样本的预报效果均显著好于定量预报模型,这种差别是否因为是由于神经网络定量预报模型本身的拟合和预报效果比较差,而显得定性预报更优越呢?为此,进一步用这同样的5个预报因子建立回归预报方程。结果表明,神经网络定量预报模型历史样本的拟合平均相对误差为8.98%,而相应的回归方法为15.75%,并且如果同样将回归方法历史样本拟合值按正负距平划分与实况对照,其趋势准确率为 $32/43=74.4\%$,明显低于神经网络的定量预报模型。在表2中也给出了回归方法3a独立样本的预报结果,其平均预报相对误差为36.6%,也大于神经网络的定量预报模型。因此,从神经网络的定量预报与相应的逐步回归预报相对比,再与神经网络的定性预报相比较可以看到,在相同的条件下,神经网络的定性预报模型通过“降低预报难度”,使得预报模型“解域”中解的个数减少,从而使定性预报结果比较准确和稳定是十分明显的。这种分析的合理性也比较容易理解,其基本思想也比较符合目前依据统计相关因子建立统计长期预报模型的实际情况。因此,如果在实际预报工作中,通过建立这样的两种神经网络定性和定量预报模型。并且当定性定量预报结果相悖时,可以对定性预报结果给予更多的“信任”,或者再采用其它预报方法进行预报,则有可能使得预报出错的可能性减小。而如果神经网络的定性预报与定量预报趋势一致,这显然可以增强我们对预报结论的信心。

利用以上两种神经网络的定性和定量预报模型, 对 1998 年江淮 6~7 月平均雨量进行了实际预报。定性预报模型预报结果是 0 1 型(正距平), 定量预报模型的预报值为 448.5 mm(平均值 339.5 mm)。两种模型预报的意见一致, 均为正距平, 增强了对预报结果可靠性的信心。而 1998 年实际降水量为 479.2 mm, 与预报结果十分相近。

6 小 结

人工神经网络方法与传统的统计预报方法相比具有很多优良的性能, 根据相关统计预报的特点, 利用神经网络方法灵活可变的拓扑结构, 构造定性和定量的预报模型并进行综合分析, 这是一种新的尝试。将定性和定量预报模型的配合使用比较符合目前长期预报工作中, 以统计相关因子为依据的预报建模实际情况。这种从定性到定量的逐步推进的预报分析方法, 具有比较容易理解的分析依据, 值得进一步深入研究。但是, 为了比较定性和定量预报模型的差异, 在建立两种预报模型时采用了相同的预报因子。然而在实际预报工作中, 一般预报因子个数较多, 入选预报方程的预报因子仅是少数。因此, 如果在选定几个预报因子建立定性预报模型, 对未来趋势作出预报后, 再用这几个预报因子建立定量预报模型作出的定量预报结论如果与定性预报结果不一致, 则能否重新选择另外的预报因子建立定量预报方程, 如果可以建立, 则当该定量预报结果与前面的定性趋势预报结果一致时, 这样的定性到定量预报结果是否也能增加预报结果的可靠性, 对这个问题值得进一步讨论。

参考文献

- 1 周家斌, 黄嘉佑. 近年来中国统计气象学的新进展. 气象学报, 1997, 55(3): 297~304
- 2 Apostols N R. Stock performance modeling using neural networks: A comparative study with regression models. Neural Network, 1994, 7(2): 375—388
- 3 Lee K Y, Cha Y T, Park J H. Short-term load forecasting using the artificial neural networks IEEE transaction on power system, 1992, 7: 124—132
- 4 宋新华, 陈茁, 俞汝勤. 人工神经网络用于对位取代苯酚定量构效关系的研究. 中国科学(B 辑), 1993, 23(3): 245~251
- 5 Jin Long, Luo Ying, Lin Zhenshan. Comparison of long-term forecasting of June-August rainfall over Changjiang-Huaihe Valley. Adv Atmos Sin, 1997, 14(1): 87—92
- 6 Mccann D W. A neural network short-term forecasting of thunderstorms. weather and forecasting, 1992, 7: 525—534

COMPREHENSIVE ANALYSIS ON QUALITATIVE AND QUANTITATIVE MODEL OF LONG- TERM FORECASTING

Jin Long

(Jiangsu Research Institute of Meteorological Sciences, Nanjing, 210008)

Miao Chunsheng

(Nanjing Institute of Meteorology, Nanjing, 210044)

Chen Ning

(Jiangsu Research Institute of Meteorological Sciences, Nanjing, 210008)

Luo Ying

(Jiangsu Research Institute of Applied Climatology, Nanjing, 210009)

Abstract

In terms of the same predictors of previous monthly mean 500 hPa height and sea surface temperature, a qualitative and quantitative models of long-term forecasting of rainfall over Changjiang-Huaihe Valley are constructed aid of the variable topological structure of neural network. A stepwise regression model is established using same predictors. The contrast analysis shows the comprehensive forecast method of the qualitative and quantitative is superior in improving predictive accuracy and stability. This shows great promise in the research of precipitation prediction on long-range basis due mainly to the reasonable analysis for prediction method and predictors.

Key words: Comprehensive analysis, Neural network, Qualitative, Quantitative.