

文章编号: 1004-4965(2010)05-0614-06

基于 BP 神经网络技术的西北太平洋热带气旋年频数预测

尹宜舟^{1,2}, 罗勇², GEMMER Marco², 曾小凡³, 李修仓²

(1. 中国科学院大气物理研究所, 北京 100029; 2. 中国气象局国家气候中心, 北京 100081;
3. 中国科学院南京地理与湖泊研究所, 江苏 南京 210008)

摘 要: 利用 NCEP 月平均的 500 hPa 高度场和海表温度场分析资料、中国气象局台风年鉴(1949—1988 年)和热带气旋年鉴(1989—2008 年)资料, 采用 BP(神经网络模式)对发生在西北太平洋上的热带气旋年频数进行预测。对隐层节点数进行普查, 采用试错法, 并综合考虑模式精度及迭代次数, 得到了相对较优的神经网络预测模式。利用选用的模式来预测热带气旋年频数, 给出了具体的预测值及具有范围的预测值, 8 年独立样本检验结果表明有 6/8 的预测值处于或接近相应的预测范围内。

关 键 词: 短期气候预测; 热带气旋; 年频数; BP 神经网络; 预测模式

中图分类号: P435

文献标识码: A

Doi: 10.3969/j.issn.1004-4965.2010.05.014

1 引 言

西北太平洋是热带气旋(TC)发生的主要源地之一^[1], 1949—2000 年平均每年有 28 个 TC 生成(不包括热带低压), 而中国等西北太平洋沿岸的国家是受其影响最大的区域。尤其是目前海上航行、海洋调查、海洋资源开发活动日益增多, 西北太平洋 TC 的发生发展越来越受到国际社会的关注, 而 TC 的年频数预测也是相关国家重视的内容之一。近年来针对西北太平洋 TC 年频数的预测取得了一些研究成果^[2-6], 但还需要进一步加强。

BP 神经网络算法是心理学家 McClelland L. L. 和 Rumelhart D. E. 提出并改进的算法, 是迄今为止一种应用最为广泛的神经网络学习算法。该算法主要设计思路是, 将输入信号通过隐层和输出层节点的处理计算得到的网络实际输出进一步与期望输出相比较, 并计算实际输出与期望输出的误差, 将误差作为修改权值的依据反向传播至输入层, 再修正各层的权系数, 并且反复这一过程, 直到实际输出与期望输出的误差达到预先设定的误差收敛标准, 从而获得最终的神经网络权值^[7]。由于 BP 算法解决了多

层网络的学习算法问题, 因此在许多学科的实际应用中获得了成功, 并从实践上证明了神经网络确实具备强大的功能, 从而大大推动了神经网络的研究热潮^[7-8]。在大气科学中神经网络方法的应用也是非常广泛的^[9-12], 在热带气旋的研究方面也有应用^[13-17]。

BP 神经网络模式(Back Propagation Neural Network Model, 简称 BP 模式)的网络结构及参数设定对其预报效果有着非常重要的意义, 目前许多研究在预报因子、隐节点数的选择以及网络权值、阈值确定方面采用了一些改进的措施, 如遗传算法、模糊数学方法、混沌理论等等与神经网络相结合, 优化了模式的结构^[18-22]。但是模式以什么条件作为结束标志值得注意, 实际应用中, 通常是给定一个收敛误差, 当模式误差达到收敛误差精度时, 模式运行结束。但是模式误差是由模式对历史样本的拟合值与实测值的偏离程度来确定, 而与模式的泛化能力无关, 即模式误差越小并不一定对应泛化能力越好。因此采用一个确定的收敛精度作为模式学习的结束标志并不一定对应有好预测结果。针对这两个问题, 本文将采用试错法, 即采用 4 年独立样本作为前期 BP 模式结构选择的试验样本, 对隐节

收稿日期: 2009-03-23; 修订日期: 2009-09-28

资助项目: 中德合作“气象指数灾害保险”项目; 德国 CIM 专家项目共同资助

通讯作者: 尹宜舟, 男, 安徽省人, 在读博士研究生, 研究方向: 短期气候预测。E-mail: ship2001691@sohu.com

点数进行普查,并综合考虑模式精度及迭代次数,得出相对较优的预测模式,并利用该模式对西北太平洋 TC 年频数进行预测。

2 资 料

本文采用的预报因子是前一年 1—12 月及当年 1—2 月各月的 NCEP 月平均 500 hPa 高度场(HGT500)以及海表温度场(SST)。HGT500 场资料分辨率为 $2.5^{\circ} \times 2.5^{\circ}$, SST 场分辨率为 $2^{\circ} \times 2^{\circ}$ 。试验中 TC 的即中心最大风速必须大于 17.2 m/s ,不包括热带低压,数据来源于中国气象局台风年鉴(1949—1988 年)和热带气旋年鉴(1989—2008 年)。

3 预报因子筛选

首先对预报因子进行相关因子普查,求出全球格点上的各月预报因子与相应的西北太平洋 TC 年频数的相关系数,时间序列为 1949—2000 年共 52 个,即 1948 年 1—12 月以及 1949 年 1—2 月对应 1949 年 TC 年频数。然后从中选出成片高相关区(相关系数大于 0.3)中的最高相关格点(需通过 0.01 的显著性检验),以此格点上的因子作为主要预报因子,共得到 79 个预报因子,其中高度场 46 个, SST 场 33 个。

但是 BP 模式本身不提供因子挑选的功能,79 个因子都加入模式,势必会造成模式冗繁,计算速度缓慢。为此须采用逐步回归方法挑选最终的预报因子。在逐步回归分析中,取 F_{α} 检验信度值为 2.5,得到 16 个因子,其中高度场 10 个, SST 场 6 个,复相关系数达到 0.976,表 1 给出了这些因子的说明,因子序号为这些因子在 79 个因子中的编号。

表 1 逐步回归分析所得到的 16 个因子说明

因子序号	因子属性(与 TC 相关系数)	格点位置(经度, 纬度)	因子序号	因子属性(与 TC 相关系数)	格点位置(经度, 纬度)
2	前一年 1 月 HGT50 (0.442)	(37.5 °W, 20.0 °N)	36	前一年 12 月 HGT500 (-0.426)	(107.5 °E, 25.0 °S)
4	前一年 2 月 HGT500 (-0.396)	(117.5 °W, 40.0 °S)	44	当年 2 月 HGT500 (-0.474)	(62.5 °W, 17.5 °S)
7	前一年 3 月 HGT500 (-0.373)	(107.5 °W, 37.5 °S)	47	前一年 1 月 SST (-0.463)	(50.0 °E, 46.0 °S)
14	前一年 5 月 HGT500 (-0.453)	(75.0 °W, 62.5 °S)	54	前一年 4 月 SST (0.342)	(20.0 °W, 20.0 °S)
17	前一年 7 月 HGT500 (-0.437)	(60.0 °E, 35.0 °N)	67	前一年 8 月 SST (-0.499)	(78.0 °E, 52.0 °S)
23	前一年 7 月 HGT500 (0.433)	(55.0 °W, 57.5 °S)	74	前一年 12 月 SST (-0.425)	(168.0 °E, 14.0 °S)
25	前一年 9 月 HGT500 (-0.389)	(152.5 °E, 47.5 °N)	75	前一年 12 月 SST (-0.423)	(12.0 °W, 38.0 °N)
32	前一年 10 月 HGT500(-0.4301)	(177.5 °W, 20.0 °S)	77	前一年 12 月 SST (-0.442)	(78.0 °W, 22.0 °S)

4 归一化方案

进入 BP 模式学习之前,预报因子要进行归一化处理,处理方案为^[7],

$$K_i = \frac{z_i - s}{t - s} \quad (1)$$

K_i 为 BP 模式的标准化输入数据, z_i 为原始观测数据, s 、 t 为待定系数。 s 、 t 由以下公式计算,

$$\begin{cases} a-s=0.1(t-s) \\ b-s=0.9(t-s) \end{cases} \quad (2)$$

其中 a 、 b 分别为各样本序列中的最小值和最大值。训练目标样本即 TC 实际观测值也需要进行归一化处理,为体现 TC 数的大小比照,归一化方案为,

$$T_i = T / (1.1 * T_{\max}) \quad (3)$$

T_i 为归一化后的目标数据, T 为原始观测数据, $T_{\max}$ 为资料序列中的最大值, $T_{\max}$ 乘以 1.1 主要是因为

未来可能会出现比 $T_{\max}$ 还要大的值。最终训练结束后, BP 模式输出层输出的数据需要根据式(3)进行反向演算来得出最终的模拟值。

5 BP 模式结构选择

本文 BP 模式输入层节点数为 16,采用 1 个隐层,隐节点数在后面进行普查,输出层为 1 个节点,学习因子和动量因子分别取为 0.3、0.4。学习样本为 1949—2000 年共 52 年的资料,独立检验样本为 2001—2008 年共 8 年资料。有关 BP 神经网络的详细说明请参照文献[7],本文不作介绍。

隐节点数是影响 BP 网络模式泛化能力的因子之一。关于隐节点数,一般认为要小于输入样本总数,不少学者进行相关的研究后得到了一些计算公式^[7]。但是不同的学科领域之间差别很大,而且选择不同的公式会得到不同的隐节点数,因此在实际

应用中需要根据具体情况来确定隐节点数。

本文对 8—51 个隐节点数进行普查,并在各种隐节点数的模式下考虑不同的迭代次数(最大为 30 000 次)及相应的模式精度,并用 2001—2004 年共 4 年的独立样本对各种模式进行预测检验,最终确定出各种隐节点数下预测效果相对较优的 BP 模式,选择过程中考虑了模式预测效果随迭代次数变化的稳定性以及排除局部最小值。表 2、3 是通过对比得到的预测效果排名前 5 位的 BP 模式及相关数据。可见隐节点数为 18,迭代次数为 11 642 次的 BP 模式预测效果较优。对比可知,这 5 个模式的预测结果均优于逐步回归预测结果。纵观这 5 个模式的模式误差及对应的预测效果,可以证实模式误差并不一定是模式预测能力的度量。图 1 是模式 1 的历史拟合图,可见拟合效果较好。图 2 是隐节点数为 18,具有不同迭代次数的各种 BP 网络模式对 4 年(2001—2004 年)独立测试样本进行预测检验结果的平均相对误差(Mean Relative Error, 简称 MRE),从图对应的数据可知,当迭代次数为 11 642 次时有

最小的 MRE 为 6.753 8。因此确定进行西北太平洋 TC 年频数预测的 BP 模式的隐节点数为 18,迭代次数为 11 642 次。

表 2 前 5 位预测效果最优的 BP 模式

模式编号	隐节点数	迭代次数	模式误差
1	18	11 642	0.003 07
2	23	13 068	0.002 75
3	21	12 401	0.002 75
4	22	16 246	0.002 61
5	19	12 760	0.002 09

表 3 前 5 位预测效果最优的 BP 模式及逐步回归预测检验 I (独立样本: 2001—2004 年)

年份	TC 观测值	模式 1 预测值	模式 2 预测值	模式 3 预测值	模式 4 预测值	模式 5 预测值	逐步回归预测值
2001	25	22.78	23.28	22.82	22.06	23.07	22.88
2002	26	25.85	25.43	25.49	24.60	25.39	26.60
2003	21	24.36	24.29	24.23	23.39	24.25	26.05
2004	30	29.53	29.14	29.27	29.61	28.70	29.03
MRE/%		6.75	6.90	7.12	7.46	7.47	9.52

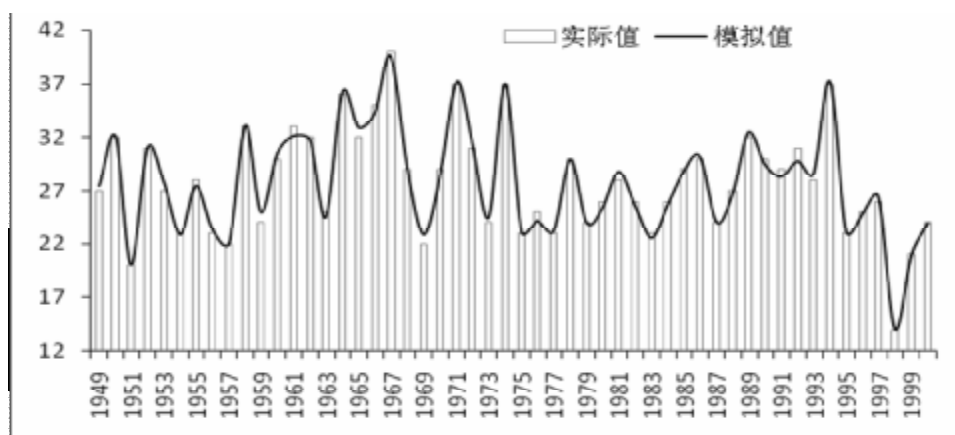


图 1 隐节点数为 18, 迭代次数为 11 642 次的 BP 模式的历史拟合图(1949—2000 年)

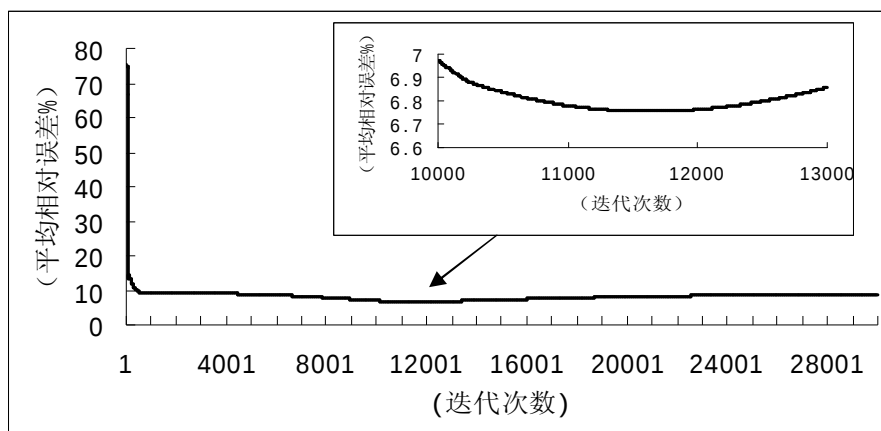


图 2 不同迭代次数的 BP 模式下(隐节点数为 18)对 4 年(2001—2004 年)独立样本检验结果的平均相对误差

6 西北太平洋 TC 年频数预测

上节已得知隐节点数为 18, 迭代次数为 11 642 次的 BP 模式的预测效果较优, 本节将用此模式, 利用另外 4 年(2005—2008 年)独立样本来预测这 4 年的 TC 年频数, 并分别对另外 4 个 BP 模式进行预测检验, 以及给出 2001—2008 年 8 年的各模式预测结果的 MRE, 以此来与模式 1 进行预测比较。检验结果见表 4, 比较 5 个模式及逐步回归方法, 模式 1(隐节点 18, 迭代次数 11 642)的预测效果相对较优。

表 4 前 5 位预测效果最优的 BP 模式及逐步回归预测检验 II (独立样本: 2005—2008 年)

年份	TC 观测值	模式 1 预测值	模式 2 预测值	模式 3 预测值	模式 4 预测值	模式 5 预测值	逐步回归预测值
2005	23	22.25	22.48	22.25	21.44	22.56	21.19
2006	24	21.12	21.39	21.16	20.45	21.30	20.04
2007	25	25.98	26.96	26.69	24.63	26.50	23.24
2008	22	22.63	22.68	22.59	21.95	22.79	21.29
4 年*		5.51	6.03	6.14	5.82	5.68	8.66
MRE/%							
8 年**		6.13	6.47	6.63	6.64	6.58	9.09
MRE/%							

注: *4 年是指 2005—2008 年, **8 年是指 2001—2008 年。

作为预测, 一方面预测值与实际值之间的偏差

表 5 模式 1 具有范围的预测结果(2001—2008 年)

年份	2001	2002	2003	2004	2005	2006	2007	2008
预测值	22.78	25.85	24.36	29.53	22.25	21.12	25.98	22.63
实际值	25	26	21	30	23	24	25	22
绝对误差	2.22	0.15	3.36	0.47	0.75	2.88	0.98	0.63
预测范围	21—24	24—27	23—26	28—31	21—24	20—23	25—27	21—24

表 6 一般 BP 模式预测结果(2001—2008 年)

收敛误差	2001 年预测值	2002 年预测值	2003 年预测值	2004 年预测值	2005 年预测值	2006 年预测值	2007 年预测值	2008 年预测值	8 年 MRE
0.01	22.58	26.17	25.11	28.34	20.91	21.32	23.79	21.33	7.94
0.05	22.89	24.43	23.99	26.15	21.65	20.94	21.55	21.61	9.47
0.001	21.72	25.35	24.14	29.69	22.82	20.60	26.78	23.05	7.30
0.005	22.23	25.24	24.67	28.94	21.63	21.31	25.16	22.14	6.68

7 夏季热带气旋频数的预测

为了进一步验证本文方法的有效性, 增加独立样本数, 将按照上面的方法对夏季(6—8 月)发生在西北太平洋上的热带气旋频数进行预测检验。预报

是必然存在的; 另一方面模式在不同的年份预测效果不同, 如模式 1 预测 2005 年的结果是 22.25, 而 2006 年结果是 21.12, 很明显 2005 年的预测更接近于实际观测值。为了使预测结果更具有参考价值, 将另外给出具有一个范围的模式预测结果作为参考。预测值 y 由模式 1 确定, 变化幅度值 V_y 由模式 1 对 8 年(2001—2008 年)独立样本检验结果的绝对误差的平均值确定, 因此预测范围的上下限将为 $y \pm V_y$ 。计算得到 $V_y = 1.43$, 则预测结果上下限为 $y \pm 1.43$, 并将最终结果进行四舍五入。表 5 给出了相应的 8 年(2001—2008 年)预测范围, 可见 5/8 的预测值处于预测范围内, 1/8 的非常接近于预测范围, 只有 2/8 的较预测范围相差较大。预测值和预测范围均可以作为实际预测参考。

为了与一般计算过程中以设定模式收敛误差作为模式训练结束标志的 BP 模式进行对比, 将以输入层节点数为 16, 隐节点数为 18, 收敛误差分别为 0.01、0.05、0.001、0.005 的 BP 模式对 2001—2008 年各年 TC 频数进行预测。表 6 为相应的结果。与表 4 的 8 年 MRE 对比可发现, 表 4 中的 5 个 BP 模式预测结果均优于表 6 中的 4 个一般模式。虽然设定的收敛误差精度有可能接近于相对最优模式最终的模式误差, 但这种设定具有很大的偶然性。

因子是前一年 6 月—当年 5 月的月平均 500 hPa 高度场及海温场。相关因子普查共得到 55 个高相关格点因子, 其中高度场 38 个, 海温场 17 个。取 $F=3.0$, 逐步回归得到 8 个因子, 其中高度场 6 个, 海温场 2 个(表 7)。

表 7 逐步回归分析所得到的 8 个因子说明

因子序号	因子属性(与 TC 相关系数)	格点位置(经度, 纬度)	因子序号	因子属性(与 TC 相关系数)	格点位置(经度, 纬度)
5	前一年 7 月 HGT500(0.384)	(60.0 °W, 60.0 °S)	21	当年 4 月 HGT500(-0.536)	(12.5 °E, 22.5 °S)
7	前一年 8 月 HGT500(-0.424)	(32.5 °W, 52.5 °S)	30	当年 4 月 HGT500(-0.439)	(45.0 °W, 45.0 °S)
11	前一年 12 月 HGT500(-0.436)	(70.0 °E, 10.0 °S)	42	前一年 9 月 SST(-0.511)	(80.0 °E, 56.0 °S)
15	当年 2 月 HGT500(-0.382)	(90.0 °W, 45.0 °N)	50	当年 3 月 SST(-0.512)	(104.0 °W, 0 °)

表 8 给出了经过普查得到的前 5 位预测效果最优的 BP 模式, 则相对最优模式为输入节点为 8, 隐节点为 10, 迭代次数为 12 304 的 BP 模式。表 9 给出了相应的预测结果。可以看出按照文中方法选择出的模式一具有相对较优的预测能力。

表 8 前 5 位夏季预测效果最优的 BP 模式

模式编号	隐节点数	迭代次数	模式误差
1	10	12 304	0.003 12
2	9	12 862	0.003 26
3	14	14 291	0.002 85
4	11	11 433	0.003 65
5	13	13 247	0.002 99

表 9 前 5 位夏季预测效果最优的 BP 模式及逐步回归预测检验(独立样本: 2001—2008 年)

年份	TC 观测值	模式 1 预测值	模式 2 预测值	模式 3 预测值	模式 4 预测值	模式 5 预测值	逐步回归预测值
2001	13	14.18	14.06	14.13	14.17	14.19	14.28
2002	14	13.56	13.52	13.37	13.49	13.47	13.24
2003	8	9.10	9.21	9.15	9.23	9.29	9.61
2004	16	15.95	15.88	15.91	15.89	15.61	15.48
2005	11	11.20	11.13	11.25	11.21	11.23	12.01
2006	11	10.83	10.79	10.82	10.78	10.73	11.62
2007	8	8.87	8.92	8.94	8.96	9.04	9.31
2008	7	7.89	7.91	8.01	7.94	8.09	8.32
4 年*		6.57	6.86	7.03	7.18	7.88	9.66
MRE/%							
8 年**		6.65	6.88	7.28	7.26	8.08	11.09
MRE/%							

注: *4 年是指 2001—2004 年, **8 年是指 2001—2008 年。

8 讨论与结论

本文在应用 BP 模式的过程中首先对隐节点数进行一次有益的普查, 并确定出预测结果相对较优的迭代次数, 然后选出预测效果相对最好的 BP 模式(即确定出隐节点数及迭代次数)。进一步做检验及与其他方法进行对比的结果表明, 对于年频数预测及夏季频数预测, 所选出的模式具有较好的预测能力。最后针对实际预测中的偏差, 给出了具有范围的预测结果。此结果是在选用的模式预测结果的基础上, 加减去此模式对 8 年(2001—2008 年)独立样本预测结果的绝对误差平均值作为上下限。结果显示只有 2/8 的预测值与预测范围相差较大些, 1/8 接近, 5/8 处于预测范围内。因此预测范围也可以作为实际预测的参考。

制约 BP 模式预测效果的因素有很多, 除模式本身因素外, 预测因子的前期选择也很重要。本文只采用最基本的高度场和海温资料, 这可能对模式的预测效果有一些影响。研究表明^[3-6, 23-24]一些诊断量或指数, 如海冰指数、850 hPa 涡度、Niño 指数、涛动指数(如 NPO, AAO 等)、高低层纬向风异常等等与台风活动关系密切, 若采用这些因子用 BP 神经网络来预测西北太平洋 TC 年频数, 相信预测效果会进一步提高, 这些将在后期工作中展开。

致 谢: 中国气象局国家气候中心姜彤研究员在论文的撰写过程中提出了非常宝贵的建议。

参 考 文 献:

- [1] 朱乾根, 林锦瑞, 寿绍文, 等. 天气学原理[M]. 北京: 气象出版社, 2000: 507-509.
- [2] 王会军, 郎咸梅, 范可, 等. 关于 2006 年西太平洋台风活动频次的气候预测试验[J]. 气候与环境研究, 2006, 11(2): 133-137.
- [3] FAN Ke. North Pacific sea ice cover, a predictor for the Western North Pacific typhoon frequency[J]. Science in China Series D: Earth Sciences, 2007, 50(8): 1 251-1 257.
- [4] FAN Ke. New predictors and a new prediction model for the typhoon frequency over western North Pacific[J]. Science in China Series D: Earth Sciences, 2007, 50(9): 1 417-1 423.

- [5] 王会军, 范可, 孙建奇. 关于西太平洋台风气候变异和预测的若干研究进展[J]. 大气科学, 2007, 31(6): 1 076-1 081.
- [6] 周波涛, 崔绚, 赵平. 亚洲-太平洋涛动与西北太平洋热带气旋频数的关系[J]. 中国科学D辑: 地球科学, 2008, 38(1): 118-123.
- [7] 金龙. 神经网络气象预报建模理论与应用[M]. 北京: 气象出版社, 2004: 1-208.
- [8] YAO Xin. A review of evolutionary artificial neural network[J]. International Journal of Intelligent System, 1993, 8(3): 539-565.
- [9] 金龙. 人工神经网络技术发展及在大气科学领域的应用[J]. 气象科技, 2004, 32 (6): 385 -392.
- [10] 任志鸿. BP 神经网络在气象数据预测研究中的应用[J]. 农业网络信息, 2007, 11: 42-44.
- [11] WILLIAM Hsieh W, TANG Benyang. Applying neural network models to prediction and data analysis in meteorology and oceanography[J]. Bulletin of the American Meteorological Society, 1998, 79(9): 1 855-1 870.
- [12] 何慧, 金龙, 覃志年, 等. 基于BP神经网络模型的广西月降水量降尺度预报[J]. 热带气象学报, 2007, 23(1): 72-77.
- [13] 高建芸, 李永尧, 宋德众. 基于前馈网络的热带气旋短期气候预测模型[J]. 海洋预报, 1999, 16(4): 33-38.
- [14] 李祚泳, 邓新民. 人工神经网络在台风预报中的应用初探[J]. 自然灾害学报, 1995, 4(2): 86-90.
- [15] 陆虹, 金龙, 缪启龙, 等. 影响广西热带气旋年频次的神经网络预测模型[J]. 南京气象学院学报, 2003, 26(1): 56-62.
- [16] 谢玲娟. 热带气旋路径客观综合集成预报方法的研制[J]. 气象, 1995, 21(8): 30-33.
- [17] BAIK Long-Jin, HWANG Hong-sub. Tropical cyclone intensity prediction using regression method and neural network[J]. Journal of the Meteorology Society of Japan, 1998, 76(5): 711-717.
- [18] 吴建生, 金龙, 汪灵枝. 遗传算法进化设计BP神经网络气象预报建模研究[J]. 热带气象学报, 2006, 22(4): 411-416.
- [19] 农吉夫, 金龙. 基于MATLAB的主成分RBF神经网络降水预报模型[J]. 热带气象学报, 2008, 24(6): 713-717.
- [20] 金龙, 林熙, 金健, 等. 模块化模糊神经网络的数值预报产品释用预报研究[J]. 气象学报, 2003, 61(1): 78-84.
- [21] 陈业辉, 邢贞相, 付强. 基于模拟退火算法的神经网络模型在洪灾易损性分析中的应用[J]. 东北水利水电, 2005, 23(4): 45-60.
- [22] 曹杰, 谢应齐. 一种基于混沌理论的联想记忆神经网络模型[J]. 热带气象学报, 1997, 13(1): 82-87.
- [23] 胡邦辉, 卢秋珍, 刘群燕. 西北太平洋热带气旋气候特征及其变化研究的若干进展[J]. 解放军理工大学学报(自然科学版), 2006, 7(2): 181-188.
- [24] 何敏, 龚振淞, 徐明, 等. 高低层纬向风异常与西北太平洋热带气旋生成年频数关系的研究[J]. 热带气象学报, 2007, 23(3): 277-283.

FORECAST METHOD STUDY ON YEARLY TROPICAL CYCLONE FREQUENCY IN NORTHWEST PACIFIC BASED ON THE BP NEURAL NETWORK TECHNIQUE

YIN Yi-zhou^{1,2}, LUO Yong², GEMMER Marco², ZENG Xiao-fan³, LI Xiu-cang²

(1. Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China;

2. National Climate Center, China Meteorological Administration, Beijing 100081, China;

3. Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing 210008, China)

Abstract: The yearly frequency of tropical cyclone(TC) happened in the Northwest Pacific Ocean was forecasted based on the BP neural network model with monthly average geopotential field on 500hPa and sea surface temperature of NECP data and the typhoon year-book of China Meteorological Administration (1949—2008). In order to get the best forecast model, the number of hidden layer node was investigated and the model precision and iterative times were considered. Finally, the selected model was used to forecast the yearly frequency of TC and the results showed the model had better forecast ability.

Key words: short-term climate forecast; tropical cyclone; yearly frequency; forecast model