

热带气旋频数的二次型预测模型^{*1}

梁 健 林永堂 谢定升

(广州中心气象台, 广州 510080)

摘 要

使用 1951—2005 年北半球 500 hPa 高度场格点资料、1949—2005 年海温场(SST)格点资料,计算与后期热带气旋发生频数的相关系数,分析两个相关场显著相关区的统计特征,进一步分析其天气气候学意义和物理意义。选取若干相关系数高的格点,构成组合因子,建立二次型曲线预测方程,进行西北太平洋、南海及登陆我国、登陆广东的热带气旋年月频数的预测。预测试验和检验表明,二次型预测模型有较高的拟合能力,在业务应用中有较好的效果。

关键词: 大气环流; 海温; 热带气旋频数; 二次型预测模型

引 言

西北太平洋及我国南海是全球热带气旋发生最多的地区之一,每年有 28 个左右的热带气旋生成,7 个热带气旋在我国登陆。广东是我国热带气旋登陆最多的省份,平均每年有 5 个,造成的直接经济损失年平均为人民币 83 亿元,1996 年经济损失高达 200 亿元,居广东各种自然灾害之首。随着我国进入 WTO,海洋航运、海洋捕捞养殖、南海石油开发、海上救捞等海上经济活动的增加,对海洋气象预报要求越来越高,特别是登陆台风往往造成国民经济和生命财产的重大损失。因此热带气旋活动是短期气候变化研究重要课题之一。

许多学者对热带气旋频数与大气环流^[1-2]、海温场的关系^[3-5]进行分析,得到许多研究成果。研究的方法有两种:一种是以时间序列为主的周期分析方法^[6-7];另一种是以相关分析为主的预报方法,如刘春霞对广东热带气旋频数异常成因进行相关分析^[8],谢炯光用海温建立了广东初始、终止台风的预报方程^[9]。但是这些研究多数只是分析海温及 EN-SO 与热带气旋的相关关系,以及定性的预测。本文根据业务实践经验,以周期和韵律为线索,深入分析热带气旋频数与海温场和大气环流的相关关系,找出具有天气气候意义和物理意义的因子,并重点讨论建立热带气旋频数的二次型预测模型的理论和方法,及其在实际业务中的应用效果。

1 资 料

预报因子文件包括① 海温场格点资料:使用 1949—2005 年太平洋 12 个月海温场(10°S~50°N, 120°E~80°W)5°×5°的 286 个格点资料作为预报因子场,共 57 年样本资料;② 北半球高度格点资料:使用 1951—2005 年(11 月、12 月、1—5 月)北半球 10°~85°N 500 hPa 东、西半球 10°×10°的 576 个格点资料作为预测因子场,共 55 年样本资料。

根据服务需要建立西北太平洋热带气旋年、月个数,南海热带气旋年、月个数,登陆我国热带气旋年、月个数,登陆广东等热带气旋年、月个数等,共 4 个预测对象文件。

2 热带气旋频数与前期高度场和海温场的相关分析

2.1 热带气旋频数与高度场相关分析

通过近 10 年的试验和相关分析表明,前期各月份的高度场格点与后期预报对象热带气旋分别存在着若干相关系数比较大(通过 0.05 或 0.01 检验)的关键区,且关键区区域稳定。

以 12 月 500 hPa 高度场与次年登陆我国的热带气旋年频数相关系数图为例,从 1995—2004 年共 10 年的相关系数图中,可以找出分别位于北大西洋的正相关关键区,格陵兰海、日本海附近的负相关关

* 广东省气象局预报员专项项目(2006001)资助。

2005-08-28 收到,2006-06-09 收到再改稿。

键区等 3 个关键区(图 1)。可以使用关键区的格点资料进行登陆我国的热带气旋年频数预测。

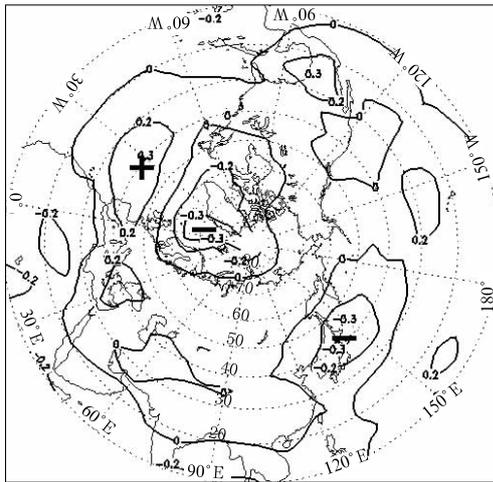


图 1 12 月 500 hPa 高度场与登陆我国热带气旋年频数的相关关键区

Fig.1 Spatial distribution of correlation coefficient between the geopotential height of 500 hPa in December and the landed tropical cyclone annual frequency in China

2.2 热带气旋频数与海温场相关分析

通过近 10 年的试验和相关分析表明,前期各月的海温场与后期预测对象热带气旋分别存在着若干相关系数比较大(通过 0.05 或 0.01 检验)的关键区,且关键区区域稳定。

以 12 月海温场与次年南海的热带气旋年频数的相关系数图为例,从 1995—2004 年共 10 年的相关系数图中,可以找出分别位于 28°~30°N,148°~152°E 附近的负相关关键区,38°~42°N,187°~192°E 附近的正相关关键区以及 225°~235°E 的赤道附近正相关关键区等 3 个关键区(见图 2)。可以使用关

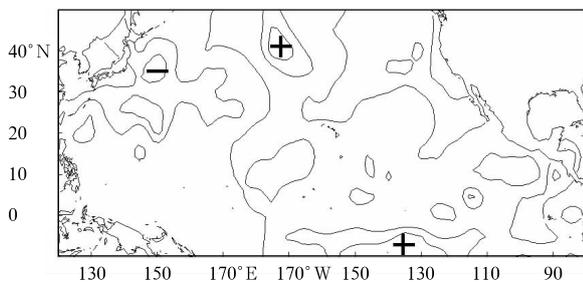


图 2 12 月海温场与南海热带气旋年频数的相关关键区

Fig.2 Spatial distribution of correlation coefficient between the sea-surface temperature (SST) of Pacific Ocean in December and the tropical cyclone annual frequency over South China Sea

键区的格点资料进行南海的热带气旋年频数预测。

3 热带气旋频数的二次型预测模型

20 世纪 70 年代和 80 年代热带气旋活动的预测多采用方差分析、多元线性回归、逐步回归等线性预测模型,90 年代后期,随着统计方法的发展,非线性预测模型有了进一步的广泛应用。本文将介绍二次型预测模型^[10]作热带气旋频数的预测原理和预测方法。

3.1 组合因子的建立

本方法是计算某月 500 hPa 高度场(或某月 SST 海温场)与热带气旋的年、月个数的相关系数,挑选 6 到 12 个通过相关系数信度检验为 0.05 的格点,每 3 个格点值组合为一个因子,计算组合因子与预测对象的相关系数。下面以 11 月 500 hPa 高度场格点资料为预测因子,次年 8 月登陆我国(热带风暴以上级别的)热带气旋个数为预测对象为例,说明预测原理及操作方法。

经普查,11 月高度场有如下 12 个格点 21,22,23,56,57,58,110,111,112,435,458,475,与预测对象的相关系数分别是 -0.340, -0.335, -0.347, -0.360, -0.355, -0.344, 0.304, 0.319, 0.282, 0.334, 0.276, 0.361。每 3 个格点值组合为一个因子,即:

$$X_A = X_{21} + X_{22} + X_{23} \quad (1)$$

$$X_B = X_{56} + X_{57} + X_{58} \quad (2)$$

$$X_C = X_{110} + X_{111} + X_{112} \quad (3)$$

$$X_D = X_{435} + X_{458} + X_{475} \quad (4)$$

X_A, X_B, X_C, X_D 与预测对象 Y 的相关系数分别是: $R_A = -0.342, R_B = -0.359, R_C = 0.308, R_D = 0.504$ 。

显然组合因子的相关系数与原来 3 个格点的相关系数比较,大多数绝对值大。也就是拟合和预测能力有明显提高,而且较稳定。

3.2 用组合因子建立权重回归方程

如上述 4 个组合因子 X_A, X_B, X_C, X_D , 其相关系数为 R_A, R_B, R_C, R_D 。

令相关系数绝对值之和为 E , 则 $E = |R_A| + |R_B| + |R_C| + |R_D| = 1.513$, 权重回归系数为 $W_A = R_A/E = -0.2262, W_B = R_B/E = -0.2371, W_C = R_C/E = 0.2032, W_D = R_D/E = 0.3334$ 。

权重回归方程为:

$$Y_1 = W_A X_A + W_B X_B + W_C X_C + W_D X_D \quad (5)$$

上述回归方程综合了 4 个组合因子。计算 Y_1

与预测对象 Y 的相关系数 $R_{Y_1} = 0.459$ 。相关系数 R_{Y_1} 比各单因子和各组合因子的相关系数大多数有明显的提高。

3.3 多次型曲线模拟热带气旋频数的气候规律及预测

统计学理论和实践表明,非线性回归更能反映预测对象与因子的相关关系。多次型回归在统计中有着特殊的地位。因为任何函数(或复杂问题)都可用正交多项式回归进行分析、计算以及模拟。下面介绍二次型曲线模拟热带气旋频数的预测方法。

3.3.1 二次型方程的建立

由高度场格点资料代入式(1)到式(5)得到自变量 Y_1 和预测对象 Y 是按等时间取样的等间隔值。适用正交多项式。

$$Y_t = a_0 + a_1 X_t + a_2 X_t^2 + \dots + a_k X_t^k \quad (6)$$

式(6)中, Y_t 为预测对象(即热带气旋频数), X_t 为高

度场格点的综合预测因子。令 $X = Y_1$, 系数 $a_0, a_1, a_2, \dots, a_k$ 由正规方程组求得。

将式(5)计算得到的 Y_1 历史资料,作为预测因子,代入式(6),可求得二阶预报方程,也就是二次型方程:

$$Y = 11.45146126 - 0.27489326X + 0.00188835X^2 \quad (7)$$

式(7)中, Y 与 X 的相关系数 $R_{Y_2} = 0.578$, 对比相关系数 R_{Y_1} , 有明显提高。这符合非线性预测的突变理论。

3.3.2 二次型方程的拟合情况

一般来说,相关系数越高,预测对象与预测因子的拟合率越高,即 Y 与 X 有式(7)的分布规律。图3是1952—1994年11月500 hPa高度场与次年8月登陆我国热带气旋频数的拟合曲线和实况曲线的对比。以拟合值与实况的绝对误差值小于1评定为正确,评定1952—1994年的8月登陆我国热带气旋月个数拟合率是 $34/43 = 79\%$ 。

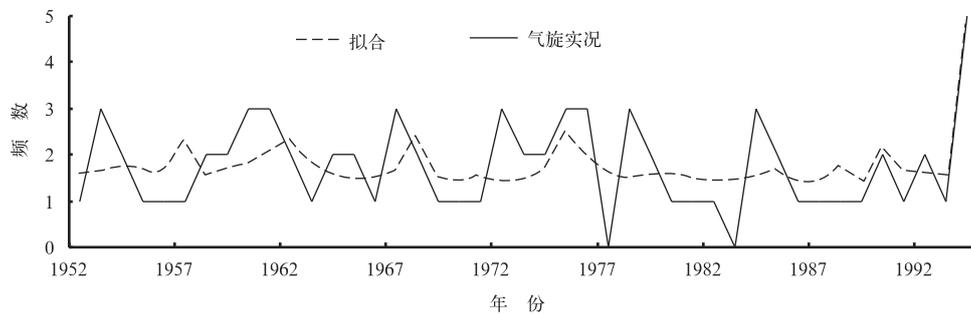


图3 1952—1994年8月登陆我国热带气旋月频数实况及利用前期11月500 hPa高度场的拟合曲线

Fig. 3 The landed tropical cyclone monthly frequency in August in China and the fitting curve using the preceding geopotential height of 500 hPa in November during 1952—1994

3.3.3 二次型方程的预测

将1994年11月12个高度场格点资料: -1, -2, -2, 1, 1, 0, 49, 49, 46, 85, 86, 88 分别代入式(1)~式(4)后,再代入式(5)可得 $X = 116.2766$ 。

将 $X = 116.2766$ 代入式(7),得到1995年8月登陆我国热带气旋频数预测值: $Y = 5.019$ 个。

8月登陆我国热带气旋数多年平均值为1.8个,而1995年8月登陆我国带气旋实况有5个,是历史的极大值。

4 热带气旋频数的气候预测系统

上述二次型预测模型其优点是能够简便地进行相关分析,容易选到通过信度为0.05的相关系数

检验的预测因子格点资料;二次型曲线拟合率高;多年试验预测实际效果好。因此建立热带气旋频数二次型气候预测系统,简称热带气旋预测系统。

热带气旋预测系统主要功能:普查500 hPa高度场格点资料、海温场(SST)格点资料与热带气旋年、月频数的相关系数;选取2~4个组合因子;建立二次型预测方程;显示预测过程和结果。

5 预测结果和验证

热带气旋预测系统分别用500 hPa高度场格点资料、海温场(SST)格点资料,制作(太平洋、南海、我国、广东)4个预报对象的年、月热带气旋频数预测。一般来说,当绝对误差小于该预测对象历年平

均值的 30%, 评定预测准确。表 1 中误差栏中误差值标有 * 的, 预测误差大, 评定当年预测为错。太平洋、登陆我国(热带风暴级别以上)、南海、登陆广东的历年热带气旋平均频数分别是 27.6, 7.1, 10.2, 5.3; 评定年频数标准(历年平均值的 30%)分

别是 4.0, 2.1, 3.0, 1.6。

5.1 用海温格点资料作热带气旋的年频数预测结果验证

使用 12 月海温资料做 1995—2005 年热带气旋年频数的预测结果验证, 如表 1 所示。

表 1 12 月海温预测热带气旋年频数结果验证

Table 1 Verification of the tropical cyclone annual frequency prediction using the preceding SST in December

年份	太平洋热带气旋			登陆我国热带气旋			南海热带气旋			广东热带气旋		
	预测值	实况值	误差	预测值	实况值	误差	预测值	实况值	误差	预测值	实况值	误差
2005	23.435	23	-0.4	6.1	8	1.9	8.9	10	1.1	4.4	4	-0.4
2004	26.427	30	3.5	6.314	8	1.7	11.58	12	0.4	4.738	4	-0.7
2003	24.995	21	-4.0*	6.29	6	-0.3	10.711	10	-0.7	4.419	5	0.6
2002	25.681	26	0.3	7.107	5	-2.1	10.699	8	-2.7	5.699	4	-1.7*
2001	25.693	25	-0.7	5.454	9	3.5*	8.997	12	3.0	5.322	5	-0.3
2000	27.18	24	-3.2	8.448	8	-0.4	10.107	13	2.9	6.109	5	-1.1
1999	23.686	21	-2.7	7.188	5	-2.2*	9.949	13	3.1*	6.09	7	0.9
1998	14.935 [△]	14 [△]	-0.9 [△]	5.532	3	-2.5*	9.424	9	-0.4	4.92	4	-0.9
1997	26.983	26	-1.0	6.397	4	-2.4*	9.181	5	-4.2*	4.625	2	-2.6*
1996	27.798	25	-2.8	7.207	7	-0.2	9.949	8	-1.9	4.333	3	-1.3
1995	29.092	23	-6.1*	7.961	9	1.0	11.884	13	1.1	6.249	7	0.8

注: 标有 * 的数值表示预报误差大; 标有 Δ 的数值表示气旋频数极小值情况。

评定表明, 用海温资料做热带气旋年频数预测, 11 年预测准确率为 $35/44=79.5\%$ 。这里特别指出 1998 年太平洋热带气旋年频数只有 14 个, 是极小值, 而预测系统报 15 个, 说明热带气旋频数的气候预测系统有预测极值的能力。

5.2 500 hPa 格点资料作热带气旋年频数预测结果验证

使用 11 月 500 hPa 资料做 1995—2005 年热带气旋年频数预测结果验证, 如表 2 所示。

评定表明, 用 500 hPa 格点资料做热带气旋年

表 2 11 月 500 hPa 预测热带气旋的年频数结果验证

Table 2 Verification of the tropical cyclone annual frequency prediction using the preceding 500 hPa in November

年份	太平洋热带气旋			我国热带气旋			南海热带气旋			广东热带气旋		
	预测值	实况值	误差	预测值	实况值	误差	预测值	实况值	误差	预测值	实况值	误差
2005	24.9	23	-1.9	5.3	8	2.7*	9.0	10	1.0	4.2	4	-0.2
2004	30.14	30	-0.1	8.41	8	-0.4	13.47	12	-1.5	5.91	4	-1.9*
2003	25.67	21	-4.7*	6.66	6	-0.7	11.02	10	-1.0	4.64	5	0.4
2002	25.62	26	0.4	6.66	5	-1.7	9.26	8	-1.3	4.82	4	-0.8
2001	25.77	25	-0.8	8.00	9	1.0	11.88	12	0.1	5.28	5	-0.3
2000	22.21	24	1.8	7.02	8	1.0	14.73	13	-1.7	4.80	5	0.2
1999	24.78	21	-3.8	7.16	5	-2.2*	9.9	13	3.1*	5.56	7	1.4
1998	24.09	14	-10.1*	5.05	3	-2.1	9.25	9	-0.3	4.71	4	-0.7
1997	29.16	26	-3.2	5.46	4	-1.5	10.38	5	-5.38*	4.50	2	-2.50*
1996	31.2	25	-6.2*	7.92	7	-0.9	9.03	8	-1.0	3.89	3	-0.9
1995	28.39	23	-5.4*	9.11	9	-0.1	14.58	13	-1.6	6.50	7	0.5

注: 标有 * 的数值表示预报误差大。

个数预测, 11 年预测准确率为 $34/44=77.3\%$ 。

5.3 热带气旋的月频数预测结果验证

下面是 1995—2005 年, 使用 11 月 500 hPa 高度场格点资料作为预测因子, 预测 7—9 月登陆我国(热带风暴以上级别的)热带气旋频数, 其预测结果

的验证如表 3 所示。

据表 3 资料, 若预测误差的绝对值 <1.5 时, 评定为预测正确。根据这个标准, 评定 1995—2005 年的 7—9 月, 登陆我国热带气旋月个数预测准确率是 $25/33=75.8\%$ 。

表3 7—9月份登陆我国热带气旋月频数预测验证

Table 3 Verification of the tropical cyclone monthly frequency prediction from July to September

年份	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	
7月	预测值	1.21	1.46	2.85	2.33	1.17	1.60	1.15	2.28	1.77	2.32	1.58
	实况值	1	3	0	0	0	1	5	1	2	3	2
	误差	-0.21	1.54	-2.85	-2.33	-1.17	-0.60	3.85	-1.28	0.23	0.68	0.42
8月	预测值	5.01 [#]	3.02	1.92	2.31	1.68	1.80	2.67	1.62	1.60	2.18	2.64
	实况值	5	2	4	2	1	2	1	2	3	2	2
	误差	-0.01 [#]	-1.02	2.08	-0.31	-0.68	0.20	-1.67	0.38	1.40	-0.18	-0.64
9月	预测值	1.81	2.11	2.17	0.97	2.10	2.05	2.44	1.49	1.89	1.90	1.51
	实况值	1	2	0	1	2	5	2	2	1	1	3
	误差	-0.81	-0.11	-2.17	0.03	-0.10	2.95	-0.44	0.51	-0.89	-0.90	1.49

注:标有#的数值表示1949—2005年中的极大值。

6 结束语

研究表明,已建立热带气旋年月频数的气候预测系统,在近几年实际应用中较好的参考意义。使用二次型预测系统有如下经验:

1) 选取2~4个组合因子时,要注意同一组合的因子,其符号要相同。

2) 通过高相关关键区分析选取组合预测因子法优于常用的分海区分析法,权重回归综合分析优于常用的经验正交函数(EOF)分析或主分量分析方法。因其方法可较大地提高拟合因子的相关系数,易于建立客观定量的预测方程。

3) 使用权重回归综合分析海温、高度场资料与热带气旋年月频数,建立二次型预报方程的方法,有较好的物理意义和统计意义,其方法方便易用。

4) 多次型预测方程对极值的预测能力优于常用的线性回归方程。经多年试验和预报实践,热带气旋的年月频数预测,以二次为宜,因此建立二次型预测模型做热带气旋频数的气候预测。

5) 二次型热带气旋年月频数气候预测系统可以根据不同的业务需要,做相应的客观定量预测,也可推广应用到其他类型的气候预测中,如短期气候降水预测等。

参考文献

[1] 丁一汇, 莱特 E R. 影响西北太平洋台风形成的大尺度环流

条件. 海洋学报, 1983, 5(5): 561-574.

- [2] 何诗秀, 张宝严, 傅秀琴. 西北太平洋盛夏台风频数与大尺度环流条件的关系. 热带气象学报, 1986, 2(3): 251-256.
- [3] 邓自旺, 屠其璞, 冯俊茹, 等. 我国登陆台风频率变化与太平洋海表温度场的关系. 应用气象学报, 1999, 10(增刊): 55-60.
- [4] 吴国雄. 海温异常对台风形成的影响. 大气科学, 1992, 16(3): 322-332.
- [5] 江剑民, 林锦瑞. 全球月平均海表水温异常的区域性及其年际变化特征. 海洋学报, 1993, 15(2): 17-23.
- [6] 谢定升, 张晓辉, 梁凤仪. 热带气旋的年月频数预测. 海洋预报, 2000, 17(4): 66-68.
- [7] 刘春霞. 广东热带气旋年际变化的小波分析. 热带气象学报, 2001, 17(4): 381-390.
- [8] 刘春霞. 广东省登陆热带气旋活动异常成因分析. 热带气象学报, 2004, 20(1): 24-34.
- [9] 谢炯光. 登陆广东省的初、终热带气旋分析及预测. 海洋预报, 1999, 16(2): 62-67.
- [10] 谢定升, 翁向宇, 曾琮. 热带气旋气候数学模型的预报应用. 海洋科学, 2004, 28(4): 52-57.
- [11] 李崇银. 厄尔尼诺与南海台风活动. 热带气象学报, 1986, 2(2): 117-123.
- [12] 陈特固. 厄尔尼诺与南海台风活动的初步分析. 热带海洋, 1986, 5(2): 80-83.
- [13] 吴尚森, 梁建茵, 黄增明. 厄尔尼诺事件对华南天气气候的影响. 广东气象, 1998, 1: 8-10.
- [14] 施能, 罗伯良. 北半球 500 hPa 极地环流与西北太平洋台风活动的遥相关. 热带气象学报, 1991, 7(4): 341-347.

Binomial Climate Prediction Model of Tropical Cyclone Frequency

Liang Jian Lin Yongtang Xie Dingsheng

(Guangzhou Central Meteorological Observatory, Guangzhou 510080)

Abstract

There are about 28 tropical cyclones per year over the Northwest Pacific Ocean and 7 tropical cyclones per year which make landfall or have significant impacts on China. Tropical cyclones lead to huge damages and economic losses. Accurate prediction of the tropical cyclone (TC) in advance would be a powerful tool in disaster preparedness and prevention.

Based on the monthly mean 500 hPa geopotential height in Northern Hemisphere during 1951—2005 of $10^{\circ}\times 10^{\circ}$ grid points between 10°N and 85°N , monthly mean sea-surface temperature (SST) of the Pacific Ocean during 1949—2005 of $5^{\circ}\times 5^{\circ}$ grid points between 10°S and 50°N , monthly and annual numbers of TC over the Northwest Pacific Ocean, South China Sea, and TCs make landfalls or generate significant impacts on China or Guangdong Province during 1951—2005, the correlation coefficients between them are calculated. By analyzing the spatial distribution of the correlation coefficients between the TC frequency and the 500 hPa geopotential height and SST respectively, several key regions exist (over significant level of 5%). Those highly significant region's geographical locations are stable. The synoptic climatologically significance and the physical characteristics are investigated. Several high correlated factors are selected and combined and used to construct the binomial prediction equations to predict the TC's monthly and annual frequency in the Northwest Pacific Ocean, South China Sea, and landing ones in China and Guangdong Province respectively.

The operational binomial climate prediction system is constructed and verified using the last 11-year data. The accuracy of the TC annual frequency predicted using the 500 hPa and SST is 79.6% and 77.3% respectively. The accuracy is 75.8% by using the preceding 500 hPa in November to predict the TC monthly frequency that lands in China from July to September during last 11 years.

The results are as follows; Select the same sign factors to combine in the same group when two to four predictors are chosen. By analyzing high significance regions, it is found that the combined predictor method is preferable to the single sea analysis. The weighted regression ensemble analysis is superior to the general statistical methods, such as the stepwise regression analysis of principle components and empirical orthogonal function. The method has the statistical and physical significance. The prediction of the multinomial prediction equation is better than the linear prediction equation. In practice the binomial prediction equation is the best. The Binomial Climate Prediction Model can be used for many purposes and performs well.

Key words: atmosphere circulation; sea-surface temperature; the tropical cyclone frequency ; Binomial Climate Prediction Model