

文章编号:1673-8411(2018)04-0001-04

延伸期暴雨过程的神经网络预报技术应用初探

覃卫坚¹, 廖雪萍², 陈思蓉¹

(1. 广西壮族自治区气候中心, 广西 南宁 530022; 2. 广西壮族自治区气象减灾研究所, 广西 南宁 530022)

摘要:利用DERF2.0延伸期环流预报数据资料,首先使用暴雨过程信号指标就一般降水和暴雨过程进行分类,结果延伸期逐日降水分类预报准确率为65%。最后利用逐步回归和粒子群-神经网络方法就延伸期暴雨综合强度进行建模预报,逐步回归方法在F=3条件下对广西暴雨综合强度预报误差最小;粒子群-神经网络预报误差均小于逐步回归方法,相对误差较逐步回归方法预报效果最好的方程减小了32.5%,可见粒子群-神经网络在延伸期量化预报中具有很好的应用前景。

关键词:延伸期暴雨过程;粒子群-神经网络;暴雨综合强度

中图分类号:P466

文献标识码:A

Preliminary Study on Neural Network Forecasting Technology Application for Extended Rainstorm Process

Qin Weijian¹, Liao Xueping², Chen Sirong¹

(1. Guangxi Climate Center, Nanning Guangxi 530022;

2. Guangxi Institute of Meteorological Disaster Reduction, Nanning Guangxi 530022)

Abstract: This paper used DERF 2.0 circulation forecasting data in the extended period. Firstly, the general precipitation and rainstorm were classified by using the signal indices of rainstorm process. The accuracy of daily precipitation classification forecast in extended period is 65%. Finally, the stepwise regression method and Particle Swarm Optimization-Artificial Neural Network (PSO-ANN) were used to model and forecast the comprehensive intensity of rainstorm in the extended period. Under the condition of F=3, the stepwise regression method has the smallest error in the comprehensive intensity prediction. The prediction error of PSO-ANN is less than that of stepwise regression, and the relative error of PSO-ANN is 32.5% less than that of equations with the best forecasting effect by stepwise regression, which shows that PSO-ANN has a good application prospect in quantitative prediction of extended period.

Keywords: extended rainstorm process; Particle Swarm Optimization-Artificial Neural Network; comprehensive rainstorm intensity

引言

天气的可预报性一般认为在2周以内,超过2周天气预报由于计算误差剧增使预报准确率显著下

降。延伸期预报指未来10~30d的天气预报,介于天气和月季气候尺度之间,是气象预报的薄弱区。随着经济社会的发展,人们对气象预报的要求越来越高,迫切地需要提高延伸期天气预报准确率,以满足人

收稿日期:2018-7-15

基金项目:中国气象局预报员专项“延伸期暴雨过程的神经网络预报技术应用”(项目编号:CMAYBY2018-057)、“广西延伸期气候预测创新团队”项目资助。

作者简介:覃卫坚(1971-),男,广西上林县人,在读博士,正高,主要从事气候变化与气候预测研究,E-mail:qinweijian2008@126.com。

们对气象服务日益增长的迫切需求,这对防灾减灾工作具有重要的意义。在气象服务业务中,延伸期暴雨过程等灾害性天气的预报方法不多,预报准确率不高,是气象工作者急需解决的问题。当前延伸期过程预报多基于模式的数值模拟,在可预报信息方面,很多把大气低频振荡作为预报信号^[1-6]。气象预报水平的提高离不开数值模式的发展,国家气候中心第二代月动力延伸预测模式业务系统(DERF2.0)是基于BCC_AGCM2.2模式发展建立的,可提供未来52d逐日滚动预报,从模式的直接的降水预报来看,其延伸期逐日降水预报准确率还不高,尤其暴雨过程的预报能力低,但模式的环流预测还是具有一定的参考价值。因此,本研究拟从环流预报场中提取出可靠的信号,既考虑这些信号跟广西暴雨有很高的相关性,又参考模式对这些信号预报技巧,最后利用逐步回归和粒子群-神经网络方法建模进行延伸期逐日暴雨综合强度预报,为延伸期暴雨过程预报提供一种新方法。

1 资料与方法

1.1 资料

使用1961~2016年广西90个气象观测站20:00~20:00的日降水资料和ECMWF ERA-Interim逐日再分析资料。模式资料使用国家气候中心月动力延伸集合预测二代产品DERF2.0每日滚动输出的未来1~52d $1^\circ \times 1^\circ$ 分辨率5d滑动集合平均的高度场、风场预测数据资料。

1.2 粒子群-神经网络方法

粒子群算法是1995年Kennedy等^[7]最早提出的一种基于群体智能方法的计算技术,具有全局搜索、通过个体之间的互动协作来搜寻全局最优解的特点,对神经网络的连接权和网络结构进行了优化,解决了以往神经网络计算中的问题^[8-11]。

1.3 暴雨过程的综合强度指数

依据有关广西暴雨气候特征研究^[12-13],暴雨过程的综合强度指数计算公式:

$$Q = \text{Pre} \cdot \text{Sn}^{0.5} \quad (1)$$

式中:Q为暴雨过程的综合强度指数;Sn为广西90个气象观测站的日雨量 $\geq 50\text{mm}$ 的总站数。暴雨雨强指数(Pre)是指在暴雨过程中广西90个气象观测站暴雨量级以上的平均雨量,其计算公式:

$$\text{Pre} = \frac{\sum_{t=1}^{\text{day}} \sum_{i=1}^{S_t} R_{t,i}}{\sum_{t=1}^{\text{day}} S_t} \quad (2)$$

式中:day为暴雨过程持续天数; S_t 为暴雨过程的第t天日降水量 $\geq 50\text{mm}$ 的站数; $R_{t,i}$ 为暴雨过程中第t天的第i站日降水量 $\geq 50\text{mm}$ 的日降水量。

2 暴雨过程指标信号的选取

在预报建模之前,利用以下方法把降水过程分为一般降水过程和暴雨过程两类,如果是一般过程,那么当日的暴雨综合强度为零。从历年降水资料中筛选出95个持续性区域暴雨过程,首先根据公式(1)计算每日的暴雨综合强度。对暴雨过程综合强度和前期模式延伸期环流预报场求相关,结果发现暴雨综合强度与赤道东太平洋南部区域200hPa纬向风速相关显著,相关显著区域同时为模式预报技巧通过显著性检验的区域,其中位于 25°S 附近区域A为正相关,而位于靠近赤道的区域B为负相关。前期研究指出,当预报量与模式预报的相关场中存在两个相邻的、符号相反、相关比较高的区域时,将这两个区域平均值相减,得到组合预报因子,较单个因子相关更加显著。本研究定义当模式预报出现区域A和区域B200hPa纬向风速差达到 42m/s 时,判断当日广西出现暴雨过程,反之为一一般过程。

利用以上暴雨过程指标对2018年6月8日起报未来10~30d降水预报进行分类,结果漏报了6月21、22、26日共3d暴雨过程,其中26日只有1站暴雨,空报了6月19、27、28日和7月5日共4d,其余13d一般降水和暴雨过程分类预报均正确(表2)。

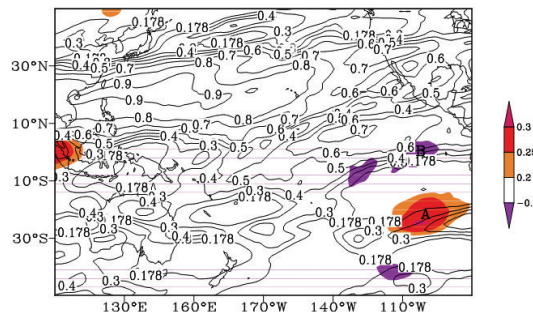


图1 广西暴雨综合强度与DERF2.0延伸期200hPa纬向风速预报值的相关(阴影区域为相关通过0.05水平的显著性检验,样本数=131)和模式预报效果(曲线,画线部分通过0.05水平的显著性检验)

3 预报建模及试验对比分析

3.1 查找气候影响因子

历年降水资料中筛选出 95 个持续性区域暴雨过程, 针对 5~7 月暴雨过程得到 131d 暴雨日, 计算每日暴雨综合强度, 统计暴雨综合强度与前期 DERF2.0 延伸期各层高度及风场预报的相关系数, 查找高相关区域, 并结合模式预测技巧, 筛选得到关键因子共 11 个(相关通过了 0.05 水平的显著性检验), 作为预报因子。相关情况如表 1 所示。

表 1 广西暴雨综合强度与各预报因子的相关系数(n=131)

因子	相关系数
x ₁ 赤道哥伦比亚、巴拿马及沿海一带 200hPa 高度场	0.22
x ₂ 图瓦卢附近 500hPa 高度场	0.23
x ₃ 赤道东印度洋 200hPa U 风场	0.28
x ₄ 赤道东太平洋科隆群岛以南区域 200hPa U 风场	-0.23
x ₅ 秘鲁以西海洋 200hPa U 风场	0.29
x ₆ 赤道东太平洋附近区域 850hPa U 风场	0.27
x ₇ 赤道大西洋附近区域 850hPa U 风场	-0.29
x ₈ 朝鲜半岛附近区域 200hPa V 风场	-0.31
x ₉ 南美洲玻利维亚附近区域 200hPa V 风场	0.26
x ₁₀ 巴西东北部及沿海地区 200hPa V 风场	-0.29
x ₁₁ 120W 赤道南侧区域 850hPa V 风场	0.34

3.2 逐步回归预测

使用广西暴雨综合强度指数与模式高度场、风场预报资料相关显著的 11 个关键因子, 利用逐步回归方法建立预测方程, 建模拟合取 131 个预测样本, 采用不同的 F 值逐步回归方法从前面因子群中进行关键因子再提取, 以 F=2、F=3、F=4 三种不同 F 值建立逐步回归预报方程:

$$y = -17571.164 + 2.996x_2 + 8.866x_5 - 10.702x_8 - 9.819x_{10} \quad (3)$$

(F=2, σ= 125.008, R= 0.495)
$$y = 247.665 - 15.099x_8 - 12.685x_{10} + 34.697x_{11} \quad (4)$$

(F=3, σ=128.094, R= 0.448)
$$y = 295.02 - 17.911x_8 - 17.118x_{10} \quad (5)$$

(F=4, σ=129.496, R= 0.421)

利用以上三个方程对 2018 年 6 月 8 日起报延伸期 20d 进行独立样本预测, 把模式环流预报值代入

以上预报方程即可得到暴雨综合指数的预报值。从预测结果和实况对比来看, 各方程预报平均值较实况值偏大, 其中(4)方程绝对误差最小, 预报值较实况值偏大 98.12, 其次为(5)方程, (3)方程预报误差最大, 达到了 118.28(见表 2)。

表 2 逐步回归方程的预测检验结果

日期	实况值	F=2(4 个因子)		F=3(3 个因子)		F=4(2 个因子)	
		预报值	误差	预报值	误差	预报值	误差
18/6	75.6	508.26	432.66	474.71	399.11	469.53	393.93
19	0	480.12	480.12	436.63	436.63	429.25	429.25
20	123.85	448.21	324.36	425.92	302.07	428.47	304.62
21	171.54	0	-171.54	0	-171.54	0	-171.54
22	330.79	0	-330.79	0	-330.79	0	-330.79
23	318.57	414.42	95.85	375.39	56.82	378.34	59.77
24	177.9	415.44	237.54	416.12	238.22	419.68	241.78
25	255.09	385.22	130.13	375.12	120.03	380.19	125.1
26	59.8	0	-59.8	0	-59.8	0	-59.8
27	0	386.74	386.74	323.63	323.63	347.78	347.78
28	0	374.22	374.22	331.1	331.1	348.15	348.15
29	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0
1/7	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0
5	0	347.91	347.91	299.28	299.28	341.16	341.16
6	263.64	337.35	73.71	292.98	29.34	334.9	71.26
7	323.69	368.19	44.5	311.9	-11.79	356.93	33.24
平均	105.02	223.3	118.28	203.14	98.12	211.72	106.7

3.3 粒子群-神经网络建模预报

采用不同的 F 值逐步回归方法从 11 个关键因子群中进行因子提取, 利用计算得到的关键因子建立神经网络预报模型, 三个不同 F 值得到不同因子数, 并利用三个不同的因子数来做的神经网络预报, 得到三种不同的预报结果。预报模型的建模样本个数为 131, 独立预报样本个数为 20 个, 设定粒子位置的上限为 2(下限为-2), 隐节点为输入节点倍数上限为 1.9(下限为 0.3), 学习因子为 0.5, 训练次数为 200, 总体误差为 0.01。预报结果来看(见表 3), 粒子群-神经网络预报的绝对误差均小于逐步回归预报, 其中利用在 F=4 下逐步回归方法得到的因子粒子群-神经网络建模预报误差最小, 仅为 64, 相对误差较逐步回归方法预报效果最好的方程减小了 32.5%。

表3 粒子群-神经网络建模预报检验结果

日期	实况	F=2(4个因子)		F=3(3个因子)		F=4(2个因子)	
		预报值	误差	预报值	误差	预报值	误差
18/6	75.6	490.78	415.18	568.64	493.04	498.36	422.76
19	0	428.37	428.37	351.22	351.22	309.88	309.88
20	123.85	283.28	159.43	347.07	223.22	192.92	69.07
21	171.54	0	-171.54	0	-171.54	0	-171.54
22	330.79	0	-330.79	0	-330.79	0	-330.79
23	318.57	291.56	-27.01	310.29	-8.28	253.72	-64.85
24	177.9	214.4	36.5	328.46	150.56	242.64	64.74
25	255.09	300.35	45.26	310.25	55.16	307.07	51.98
26	59.8	0	-59.8	0	-59.8	0	-59.8
27	0	299.5	299.5	308.35	308.35	346.78	346.78
28	0	292.38	292.38	307.91	307.91	330.75	330.75
29	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0
1/7	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0
5	0	365.74	365.74	276.41	276.41	305.93	305.93
6	263.64	335.16	71.52	274.02	10.38	302.68	39.04
7	323.69	299.77	-23.92	267.78	-55.91	289.33	-34.36
平均	105.02	180.06	75.04	182.52	77.5	169	64

4 结论与讨论

通过分析得到以下结论:

(1)利用DERF2.0延伸期环流预报数据资料,使用暴雨过程信号指标就一般降水和暴雨过程进行分类,结果日降水分类预报准确率为65%。

(2)逐步回归方法在F=3条件下对广西暴雨综合强度预报误差最小。粒子群-神经网络预报误差均小于逐步回归方法,相对误差较逐步回归方法预报效果最好的方程减小了32.5%。

本研究利用DERF2.0环流预报数据和暴雨过程信号指标来做延伸期暴雨过程预报,具有一定的预报效果,但暴雨过程信号指标还有待进一步的应用

来检验和改进。

参考文献:

- [1] 丁一汇,梁萍.基于MJO的延伸预报[J].气象,2010,36(7):111-122.
- [2] 陈官军,魏凤英.基于低频振荡特征的夏季江淮持续性降水延伸期预报方法[J].大气科学,2012,36(3):633-644.
- [3] 孙国武,李震坤,信飞,等.延伸期天气过程预报的一种新方法—低频天气图[J].大气科学,2013,37(4):945-954.
- [4] 覃卫坚,李耀先,廖雪萍.MJO活动对广西6月阶段性降水的影响研究[J].气象研究与应用,2015,36(1):25-30.
- [5] 覃卫坚,李耀先,陈思蓉.同期MJO对广西汛期强降水过程的影响[J].气象研究与应用,2017,38(4):18-21.
- [6] 覃卫坚,李耀先,廖雪萍,等.大气低频振荡对广西持续性区域性暴雨的可能影响[J].气象研究与应用,2015,36(3):1-7.
- [7] Kennedy J, Eberhart R C. Particle Swarm Optimization[A]. In: Proc IEEE International Conference on Neural Networks [C]. IV Piscataway, NJ: IEEE Service Center, 1995, 1942-1948.
- [8] 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[J].气象学报,1999,57(2):198-207.
- [9] Hua-sheng Zhao, Long Jin and Ying Huang. An Objective Prediction Model for Typhoon Rainstorm Using Particle Swarm Optimization - Neural Network Ensemble [J]. Natural Hazards. 2014, 73(2):427-437.
- [10] 覃卫坚,李耀先,陈思蓉,等.粒子群-神经网络在华南夏季降水短期气候预测中应用研究[J].气象研究与应用,2015,36(2):1-7.
- [11] 覃卫坚,陆虹,黄志,等.粒子群-神经网络在广西寒露风日数预报中的应用[J].气象与环境学报,2015,31(6):158-162.
- [12] 覃卫坚,李耀先,廖雪萍.广西热带气旋的暴雨统计分析 & 数值模拟[J].气象研究与应用,2013,34(1):1-6.
- [13] 覃卫坚,李耀先,覃志年.广西暴雨的区域性和连续性研究[J].气象研究与应用,2012,33(4):1-4.