

陈德花, 陈创买, 周学鸣, 等. 2013. 福建汛期降水主分量逐步回归预测模型研究. 气象, 39(9): 1190-1196.

福建汛期降水主分量逐步回归 预测模型研究^{*1}

陈德花¹ 陈创买² 周学鸣¹ 孙琼博¹ 韦 晋¹

¹ 厦门市气象台, 厦门 361012

² 中山大学大气科学系, 广州 510275

提 要: 建立福建汛期降水量气候场的主分量逐步回归预测模型, 通过预测福建汛期降水量气候场的主分量来实现对福建汛期降水量场的预测。选取北半球 500 hPa 高度场、太平洋海温场、北半球海平面气压场等含不同区域不同季节的诸多因子场的前若干主分量因子作为预测因子。计算分析了福建汛期降水场的主分量分布特征和分型特点。利用相关筛选和双重检验逐步回归方法, 建立福建汛期降水场与多个因子场的关系, 建立主分量的预测方程。通过对 2012 年福建汛期降水的预报, 发现该模型对福建汛期降水的趋势和分布有较好的预测能力。

关键词: 气候统计预测, 主分量分析, 双重检验逐步回归, 福建汛期降水

中图分类号: P456

文献标志码: A

doi: 10.7519/j.issn.1000-0526.2013.09.015

PC Stepwise Regression Forecasting Model of Climatic Field of Fujian Flood Period Rainfall

CHEN Dehua¹ CHEN Chuangmai² ZHOU Xueming¹ SUN Qiongbo¹ WEI Jin¹

¹ Xiamen Meteorological Observatory, Xiamen 361012

² Department of Atmospheric Sciences, Sun Yat-Sen University, Guangzhou 510275

Abstract: A principal component stepwise regression prediction model of climatic field for the flood period rainfall in Fujian Province is established. The model predicts the precipitation of flood period in Fujian through predicting the principal component of the climatic field of that period. This study takes the 500 hPa height and the sea surface temperature of Pacific Ocean and the sea-level pressure in the Northern Hemisphere and many principal component factors in various seasons and regions as predictors, calculating and analyzing the distribution and types of the principal components of the flood period rainfall in Fujian. Through correlation screening and double test stepwise regression, a prediction equation is developed for the principal component of the Fujian flood period rainfall. By applying this model to predicting the flood period precipitation in Fujian in 2012, it is found that the model is capable to predict the precipitation and its distribution in the flood period in Fujian.

Key words: climatic statistical prediction, principal component analysis, double test step-wise regression, flood period rainfall in Fujian

* 公益性行业(气象)科研专项(GYHY20110133-02)和福建省气象局开放式基金(2012K4)共同资助
2012 年 10 月 20 日收稿; 2013 年 5 月 9 日收修定稿
第一作者: 陈德花, 主要从事预报研究工作. Email: dehua_chen@163.com

引言

福建省地处欧亚大陆东南边缘(鹿世谨,1999),属副热带南部的季风气候区。不少学者对福建汛期降水的气候特征和预测问题进行了探索。吴滨等(2005)用典型相关分析方法对福建前汛期降水进行预测,这是一种用两个场相关作预测的方法;王艳姣(2003)对福建省汛期降水作了特征分析并用人工神经网络建立预测模型,又采用基于BP和Elman神经网络对福建省汛期旱涝作预测;许金镜等(2004)对福建伏早期的旱涝的环流特征作分析及预测;张容焱等(2007)对福建旱涝用4种方法作短期气候预测试验;王艳姣等(2005)用RBF神经网络对福建省汛期旱涝作预报方法研究。以上所述各种分析和预测方法,除了吴滨等(2005)和王艳姣等(2005)的方法涉及场的特征分析外,其余都是采用对降水指标序列进行预测。陈创买等(1999;2009)充分利用各种统计方法的优越性,提出“气候场的主分量逐步回归预测模型”,该模型将气候场的预测变成对气候场的前若干重要主分量的预测,预测因子可取自对气候变化有重要影响的500 hPa高度场、太平洋海温场和北半球海平面气压场等含不同区域不同季节因子场的前若干个主分量,它们集中了各因子场的最重要信息。彭端等(2007)将该方法用于预测粤中西部汛期降水,黄露菁等(2001)用于预测广东低温阴雨年景,彭武坚(2005)用于预报桂东南低温阴雨,刘艳群等(2007)用于预测珠江流域年降水量,黄刚等(2006)用于预报华北夏季降水气候场,陈慧娴等(2007)用于研究广东春季水资源的时空分布特征和预测,曾琮等(2008)用于广东冬季气温的预测,周文等(2002)用于预测南海西南季风的爆发及广东春雨的预报,刘利平等(2003a;2003b)用于预测广东韩江、北江和全省各大江河年最大流量,陈慧娴等(2004;2008)用于预报番禺单站年降水量及分析广东冬季水资源的时空分布特征,向甸等(2008)利用旋转主分量分析等对热浪指数和暖夜指数的时空变化特征进行分析,曾琮等(2006)利用主分量方法对广东冬季气温时空变化特征开展研究,这些分析和预测例子均表明,用气候场的主分量逐步回归预测模型进行预测均取得了较好的效果。本文将采用该方法研究福建省汛期(4—9月)降水的气候变化及预测。

1 资料说明

取福建省分布较均匀的69个气象观测站自建站年至2011年的1—12月月降水量作为分析的基本资料。为主分量计算的需要,利用气候序列订正法将建站较迟的测站一律向前订正到1954年。

2 计算方法的简介

2.1 主分量逐步回归预测模型的建立思路

主分量分析能够把随时间变化的气候要素场分解成为空间函数(特征向量矩阵)和时间函数(主分量矩阵)两部分。空间函数部分概括了场的不随时间变化的地域分布特点,而随时间变化的前若干个主分量占有原空间点变量的总方差的很大部分(黄嘉佑,2004)。根据这些特点,首先可将气候场的预测变成对场的前若干个主分量的预测。同样可从各区分季节的因子场的主分量分析中,提取其前若干个主分量作为预测因子。通过相关筛选和双重检验逐步回归对气候场的主分量进行预测,及至利用相对少变的空间函数的回算,最后预测气候场。

福建汛期降水场主分量逐步回归预测模型的建立分三步进行。第一步对汛期降水场作主分量分析,第二步对降水场的前25个主分量逐个进行双重检验逐步回归预测,第三步利用降水场的预测结果与该场的特征向量进行回算,得到气候场的预测结果。

2.2 降水场的主分量分析方法

取福建69站1954—2011年4—9月降水量作为福建汛期降水场,2012年资料作预测。其主分量计算采用标准化变量方法,得到其前25个主分量和相应的特征向量及特征值等。

2.3 相关筛选和逐步回归预测

对降水场前25个主分量逐个与因子群进行相关系数计算,设第 k 个预测量 y_k 与第 l 个预测因子 f_l 的相关系数为:

$$r_{kl} = \frac{1}{m} \sum_{i=1}^m \frac{y_{ki} - \bar{y}_k}{s_k} \frac{f_{li} - \bar{f}_l}{s_l} \quad (1)$$

式中, \bar{y}_k 与 \bar{f}_l 为平均值, s_k 与 s_l 为标准差, m 为因子和预测量的时间序列长度。给定显著性水平 α ,

相应显著相关系数为 r_a , 若 $r_{kl} \geq r_a$, 则该因子与预测量相关显著, 从显著因子中选出 70 个左右供逐步回归计算。然后再对前 25 个主分量逐个进行双重检验的逐步回归计算。

2.4 回算

因为预测的是降水场的主分量, 因此, 还需要利用降水场的特征向量和预测的前 25 个主分量进行回算, 最后得到原始降水场的预测。

3 福建汛期(4—9月)降水的时空分布特征分析

对福建汛期降水量场进行主分量计算, 以下分析该降水场的时空分布特征。

3.1 福建汛期(4—9月)平均降水量的分布

福建汛期平均降水量分布如图 1。图 1 可见存在 3 个大降水中心分别位于宁德地区(周宁—福安)、泉州地区(德化—永春)、漳州地区(南靖—平和), 其中心降水量均超过 1300 mm。还存在 2 个次大值中心分别位于南平东部地区、三明中北部地区, 降水量达到 1200 mm 以上。小值中心出现在福建沿海北部地区、中部(莆田—泉州)、南部(厦门), 总降水量在 1000 mm 以下。从汛期平均降水的总体分布情况来看, 山区大于沿海, 但是沿海地区的山区大于内陆地区的山区; 大值中心位于沿海地区的靠山地带, 小值中心位于沿海地区的靠海地带; 内陆山区介于两者之间。

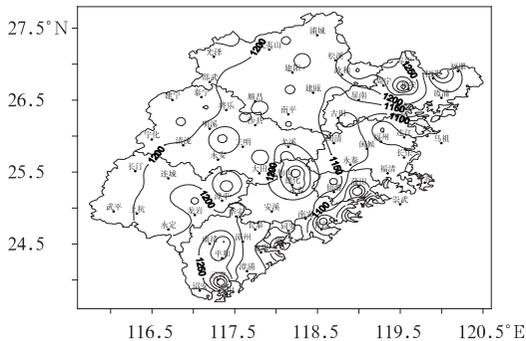


图 1 福建汛期(4—9月)平均降水量分布图(单位:mm)

Fig. 1 Distribution of average rainfall in Fujian flood period (April to September) (unit: mm)

3.2 福建汛期(4—9月)降水的标准差分布

福建汛期降水的标准差分布如图 2, 它与平均降水量场的分布较相似, 也存在多个大小值中心。300 mm 以上的大值中心位于宁德—福安一带以及东山—云霄一带。小值中心也与平均降水量场的分布较一致。

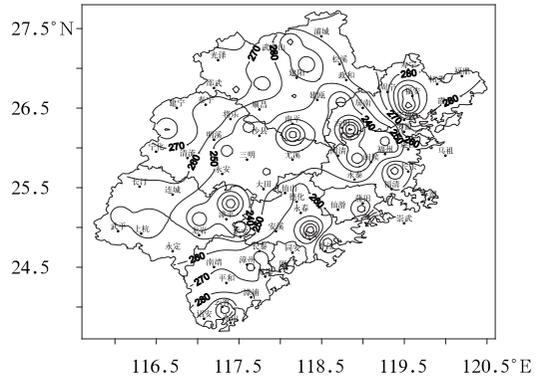


图 2 福建汛期(4—9月)降水的标准差分布图(单位:mm)

Fig. 2 Standard deviation distribution of precipitation in Fujian during flood period (April to September) (unit: mm)

3.3 福建汛期降水的主分量计算

对福建汛期(4—9月)降水进行主分量计算得到特征向量矩阵和主分量矩阵, 其前 25 个主分量所占总方差百分数如表 1。由表可见前 4 个主分量已占总方差约 75%, 前 25 个达到近 96%。以下将对前 4 个主分量和特征向量做分析。

3.4 福建汛期降水的时空分布特征

为了解释预测结果的需要, 以下分析了福建汛期第一、第二特征向量和主分量的分布特征。

3.4.1 第 1 特征向量及主分量分布情况

图 3 为福建汛期降水第一特征向量值的分布。全省均为正值, 绝对大值中心与汛期平均降水量分布的大值中心类似。分布在福建的宁德北部山区、福州的中部地区及泉州的山区(德化—南安)。次大值中心位于龙岩西北部(漳平)、三明南部(大田)及漳州的南部地区(平和—云霄)。由于第 1 特征向量全省均为正值, 代表全省性涝(旱)分布(I)型。第 1 主分量的正(负)对应全省性的涝(旱)年份。从表

1 可知(I)型占总方差的 47.4%。福建省最严重的旱、涝年份都与这一分布型有关。

图 4 是福建汛期第一主分量的逐年变化曲线。从图可见,2006、1973、1990、1997、1959 和 1961 年存在正值,采用比拟法可得对应出现的实况这 6 年福建属于严重的涝年。负的大值出现在 1967、

表 1 福建汛期降水场前 25 个主分量所占方差及其累积百分数

Table 1 The variance and its accumulated percentage of the first 25 principal components which influence flood period rainfall in Fujian

主分量	特征值	占总方差/%	累计方差/%
1	27.5	47.4146	47.4146
2	10.6	18.2695	65.2695
3	3.04	5.2432	70.9273
4	2.45	4.2251	75.1524
5	1.89	3.2507	78.4032
6	1.34	2.5304	80.9597
7	1.17	2.0193	82.9790
8	0.92	1.5803	84.5593
.....
24	0.23	0.3949	95.9977
25	0.21	0.3602	96.3579

2003、1991、1986 和 2011 年,对应实况以上 5 年福建属于严重的旱年,就是说从第一主分量的前若干个绝对最大值可以推断出严重的灾害气候年份。

3.4.2 第二特征向量及主分量分析

图 5 为福建汛期降水第二特征向量的分布。从图上可知,第二特征向量的分布呈现明显的正、负两种相态,零值等值线沿着寿宁—屏南—古田—尤溪—大田—漳平—长汀将福建分为两半,福建北部—西北部呈现正值区,福建东北—西南及龙岩西南部呈现负值区。表现为福建北部—西北部涝(旱),福建东北—西南部旱(涝)分布(II)型。当第二主分量为正(负)时,福建北部—西北部地区出现涝(旱)年,而东北部—西南部及龙岩西南部地区则相反。

图 6 为福建汛期第二主分量的逐年变化曲线。从图 6 可见,1990 和 2008 年存在最大值和次大值,这两年福建北部—西北部出现涝,东北—西南部出现旱。第二主分量的最小值和次小值出现在 1954 和 1962 年,这两年福建北部—西北部地区出现旱,

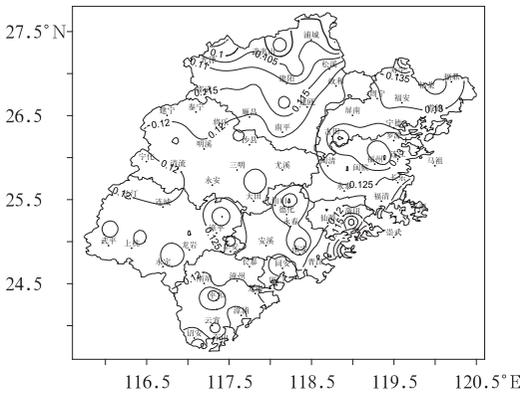


图 3 福建汛期降水第一特征向量分布
Fig. 3 The first eigenvector of flood period rainfall in Fujian

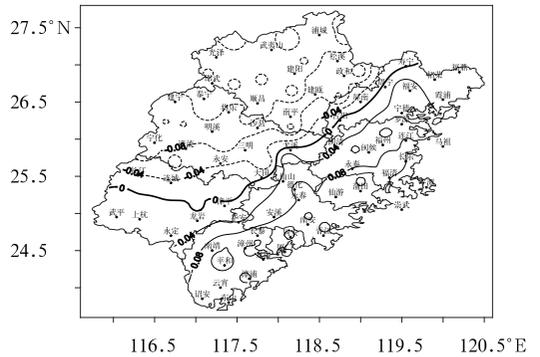


图 5 福建汛期降水量第二特征向量分布
Fig. 5 The second eigenvector of flood period rainfall in Fujian

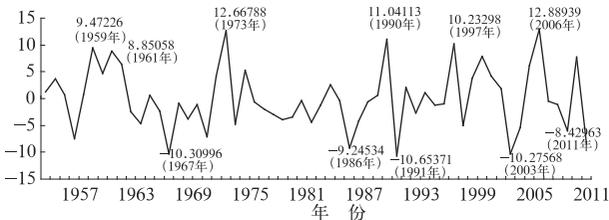


图 4 福建汛期降水第一主分量年变化曲线
Fig. 4 The annual variation of the first principal component of flood period rainfall in Fujian

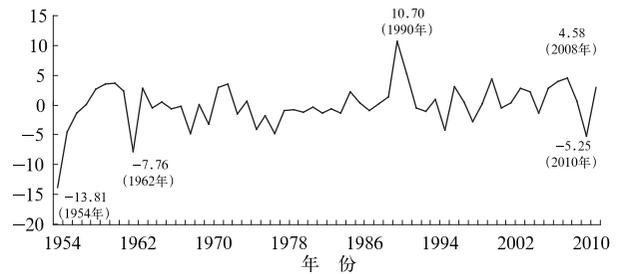


图 6 福建汛期降水量第二主分量年变化曲线
Fig. 6 The annual variation of the second principal component of flood period rainfall in Fujian

东北—西南部地区则涝。

3.5 福建汛期(4—9月)降水分类

将福建汛期降水的历年分布看成是由该年前 5 个主分量的方差贡献所组成,分布类型就可以用前 4 个主分量确定。将主分量的正负分别记为 a 和 b ,从而定出该年和历年的分布类型为 I_a 、 I_b 、 II_a 、 II_b 、 III_a 、 III_b 、 IV_a 、 IV_b 。为了区分那些汛期降水严重涝和严重旱的年份,特定义一个数据结构函数 YC ,它是气候场第一主分量中前第 k 绝对最大值其累积方差占第一主分量总方差的 30% 以上。

通过分型统计 1954—2011 年(具体分型表略),其中 $I_a=6$ 、 $I_b=5$ 、 $I_c=9$ 、 $I_d=14$ 、 $II=13$ 、 $III=3$ 、 $IV=4$ 、 $V=4$,以 T_1 表示第一主分量,若 $T_1 \geq YC$ 为 a , $T_1 \leq -YC$ 为 b , $0 \leq T_1 < YC$ 为 c , $-YC < T_1 < 0$ 为 d ,得 I_a 、 I_b 、 I_c 、 I_d ,分别表示福建汛期降水严重涝、严重旱、偏涝、偏旱 4 种状况。本文计算的 $YC=6.438$ 。统计结果中 I 型的有 34 年占 58.6%,其中 I_a 为 6 年占 10.3%, I_b 为 5 年占 8.6%, $I_a + I_b=11$ 年占 19.0%,这些是属于全省性严重涝和严重旱的年份,以 1959、1961、1973、1990、1997 和 2006 年为涝年代表,以 1967、1986、1991、2003 和 2011 年为旱年代表。 I_c 为 9 年占 15.5%, I_d 为 14 年占 24.1%,出现 II 、 III 、 IV 、 V 型的分别有 13、3、4 和 4 年。这 4 个型共占 41.4% 是属于涝和旱分布不均匀的年份。

4 福建汛期降水的预测模型的建立

以上进行了福建汛期降水场的主分量计算,完成建模的第一步。

4.1 福建汛期降水场前 25 个主分量的预测因子库

为了对降水场进行主分量逐步回归预测,预先建立北半球 500 hPa 高度场、太平洋海温场、北半球海平面气压场的主分量因子库以及高空环流物理量指数季节因子库。

因子场取 1953—2011 年太平洋海温场(286 个 $5^\circ \times 10^\circ$ 格点)月平均海温资料和北半球 500 hPa 高度场(576 个 $5^\circ \times 10^\circ$ 梅花格点)月平均高度资料。北半球海平面气压场(576 个 $5^\circ \times 10^\circ$ 格点)月平均气压。将太平洋海温场分为 5 个区,北半球 500 hPa 高度场分为 11 个区,北半球海平面气压场分为 11

个区。对每个区各取 29 种季节。每一个区的每一种季节构成一个因子场,总共可形成 783 个因子场。对它们进行主分量分析,各取其前 8 个主分量作为预测因子。对高空环流物理量指数则按 29 种季节统计析出季节因子,形成季节因子库。以上共可产生 8318 个因子。

4.2 预测因子的相关筛选

为了适应逐步回归计算,将降水场的前 25 个主分量逐个与所有因子进行相关筛选。从大量的预测因子中选出高相关的 70 个因子供逐步回归计算。结果所有因子的单相关系数在 $\gamma=0.300$ 以上,最大达到 0.556。

4.3 双重检验逐步回归

利用双重检验逐步回归方法,对福建汛期降水场前 25 个主分量逐个进行回归计算,设定 F -信度检验的临界值 $p_a=95\%$,从而确定了前 25 个主分量的回归方程。其中第一主分量逐步回归方程:

$$y = 47.06218 - 0.03902x_1 - 0.11485x_2 - 0.02875x_3 + 0.35010x_4 - 0.14971x_5 + 0.41526x_6 - 0.48047x_7 + 0.09111x_8 - 0.16064x_9 + 0.14095x_{10} \quad (2)$$

式中, x_1 为大西洋欧洲区 2—4 月极涡强度指数、 x_2 为东北半球南部 6—7 月海平面气压场第七主分量、 x_3 为北半球区极涡面积指数、 x_4 为全北半球 8—9 月 500 hPa 高度场第五主分量、 x_5 为东北半球南部 5—6 月海平面气压场第一主分量、 x_6 为全北太平洋 2—3 月海温场第六主分量、 x_7 为东北半球南部 5—8 月海平面气压场第七主分量、 x_8 为印度副热带高压北界($65^\circ \sim 95^\circ E$)、 x_9 为北半球区极涡中心强度、 x_{10} 为大西洋欧洲夏季环流型 E。回归方程的复相关系数达 $r=0.863$ 。

图 7 是福建汛期降水第一主分量的预测曲线。由图 7 可见,第一主分量的预测值和实测值的拟合度相当高,平均预测残差为 2.89,最大残差为 -7。两线趋于重合,58 年中只有 7 年预测值与实测值符号相反。符号相关系数为 87.9%,2012 年预测值为 -1.16。

4.4 福建汛期降水分布预测分析

图 8 为 2012 年福建汛期降水预测值和实测值的分布图。对比可知,降水量大于 1250 mm 的区域

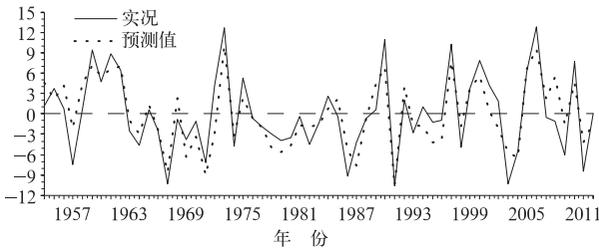


图 7 福建汛期降水第一主分量原值及模型预测曲线
 Fig. 7 The original value and model-predicted curve of the first principal component of flood period rainfall in Fujian

实况与预测对应较好,其中在三明和龙岩市的两个强降水预报准确,位于邵武的强中心预测值中心略偏北。漳平、德化、霞浦的次大值中心也对应较好,福安的大值区预报偏弱。降水量小值区域也对应较好,位于福建沿海一带。实况中古田—建瓯一带的小值中心,预测场中该中心偏北。

图 9 为福建汛期降水预测值和实况值的距平百分率分布图。从图 9 可以看出,模型对于福建的西北部和北部地区,其正距平中心对应较好。预测值对福建东北部、西南部地区为负距平,其余为正距

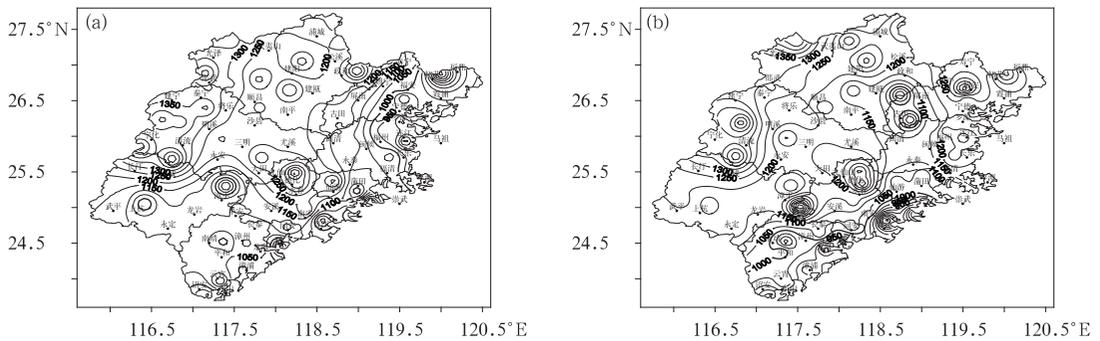


图 8 2012 年福建汛期降水预测值分布图(a)和实测值分布图(b)(单位:mm)
 Fig. 8 The rainfall predicted by model (a) and measured rainfall by rain gauge (b) in Fujian during flood period in 2012 (unit: mm)

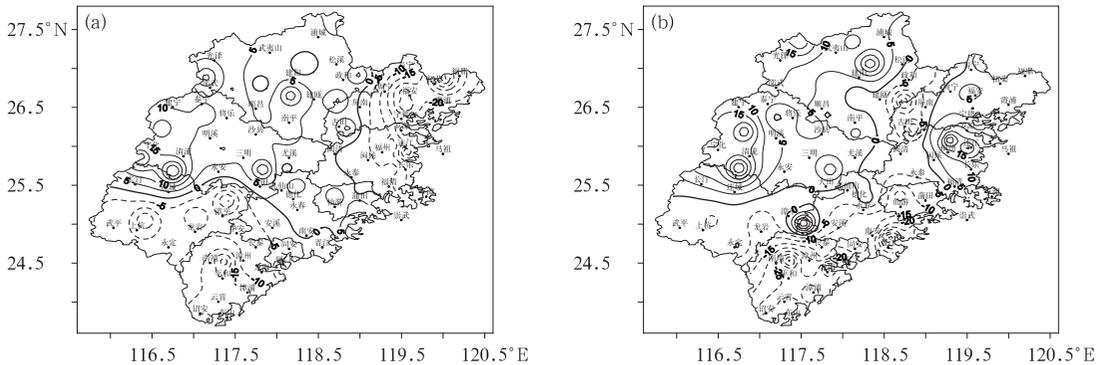


图 9 2012 年福建汛期降水预测值距平百分率分布图(a)和实测距平百分率分布图(b)(单位:%)
 Fig. 9 The anomaly percentage distribution of rainfall predicted by model (a) and measured rainfall by rain gauge (b) during flood period in Fujian in 2012 (unit: %)

平。在三明西北部地区汛期降水明显偏多 15%~20%,在漳州的中北部地区明显偏少 25%左右,这两个中心预测与实况对应较一致。而在福建的东北部、中部沿海地区预测结果与实况相反,并且负距平区域预报比实测区域更为偏南。从总体的预报结果

来看,中北部和西南部的预测结果基本正确,福建东北部和中部沿海结果相反。共有 70% 以上的站点其距平预报准确,主分量逐步回归预测模型对于福建省汛期降水的预测效果较好。

5 结 论

(1) 建立福建汛期降水量气候场的主分量逐步回归预测模型,通过对福建汛期降水量气候场的主分量预测来实现对福建汛期降水量场的预测。

(2) 选取北半球 500 hPa 高度场和太平洋海温场、全球海平面气压场等含不同区域不同季节的诸多因子场的前若干主分量因子库,将它们作为预测因子。

(3) 利用相关筛选和双重检验逐步回归分析,建立福建汛期降水的预测模型。通过对 2012 年福建汛期降水的预报,发现该模型对福建汛期降水的趋势和分布有较好的预测能力。

参 考 文 献

陈创买,周文. 1999. 气候场的主分量逐步回归预报. 广东灾害性气候的分析和预测研究. 广州:中山大学出版社,1-30.

陈创买,周文,林爱兰. 2009. 气候场的主分量逐步回归预测模型及应用. 热带气象学报,25(2):196-204.

陈慧娴,陈创买,林爱兰. 2007. 广东春季水资源的时空分布特征和预测研究. 热带气象学报,23(6):601-610.

陈慧娴,陈创买,袁宇平,等. 2008. 广东冬季水资源的时空分布特征分析. 安徽农业科学,36(29):12826-12829.

陈慧娴,黄露菁,陈创买. 2004. 用逐步回归方法预报番禺年降水量. 广东气象,(4):7-9.

黄刚,周文. 2006. 华北夏季降水气候场的主分量逐步回归预报模型. 气候与环境研究,11(3):296-301.

黄嘉佑. 2004. 气象统计分析与预报方法(第三版). 北京:气象出版

社,66-67.

黄露菁,简裕庚. 2001. 广东低温阴雨年景预测,中山大学学报(自然科学版),40(6):91-94.

刘利平,陈创买,陈健,等. 2003a. 广东韩江年最大流量的预测研究. 广东水利水电,(6):48-50.

刘利平,陈创买,陈健,等. 2003b. 主分量与广东北江年最大流量预测. 水科学进展,14(4):485-488.

刘艳群,陈创买,郑勇,等. 2007. 珠江流域年降水的预测. 2007 年中国气象学会年会论文集,622-629.

鹿世谨. 1999. 福建气候. 北京:气象出版社,163-165.

彭端,何华庆,黄天文,等. 2007. 气候场主分量逐步回归模型预测粤中西部汛期降水. 广西气象,28(A02):81-83.

彭武坚. 2005. 主分量逐步回归法在桂东南低温阴雨预报的应用. 广西气象,26(A02):41-43.

王艳姣. 2003. 福建省汛期降水的特征分析及神经网络预测模型的研究. 南京:南京气象学院.

王艳姣,张鹰,邓自旺,等. 2005. RBF 神经网络的汛期旱涝预报方法研究. 热带气象学报,21(1):100-106.

吴滨,蔡学湛. 2005. 用典型相关分析预测福建前汛期降水. 气象科技,33(1):32-36.

向旬,王冀,王绪鑫,等. 2008. 我国极端气温指数的时空变化和分区研究. 气象,34(9):73-80.

许金镜,温珍洽,杨林,等. 2004. 福建伏早期旱涝环流特征及其预测探讨. 气象科学,24(2):233-239.

曾琮,陈创买. 2006. 主分量逐步回归在气温预测中的应用. 中山大学学报(自然科学版),45(4):107-110.

曾琮,陈创买,李晓娟,等. 2008. 广东冬季气温时空变化特征. 气象,31(3):56-59.

张容焱,李永尧,徐宗焕,等. 2007. 福建旱涝短期气候预测统计方法试验研究. 热带气象学报,23(6):636-642.

周文,温之平,陈创买. 2002. 南海西南季风爆发的预测研究. 中山大学学报(自然科学版),41(3):95-98.