

文章编号: 2096-1618(2018)05-0517-08

气候模式数据集在区域性高温事件延伸期预报中的应用

周辰光¹, 杜良敏², 高伟¹, 王宏记³

(1. 中国地质大学(武汉)信息工程学院,湖北 武汉 430074;2. 湖北省气象局武汉区域气候中心,湖北 武汉 430074;3. 湖北省气象局气象信息与技术保障中心,湖北 武汉 430074)

摘要:针对近年来湖北省夏季区域性高温事件的频发,利用湖北省历史气温资料,采用两种高温判定方法统计了历史区域高温事件发生的时间和逐日平均最高温度;利用美国 NCEP-CFSv2 气候模式历史回报资料,采用集合平均预报方法开展逐日基于多时间序列样本的湖北省区域性高温事件的延伸期预报。设计并使用两种方案对区域性高温事件的延伸期预报结果开展检验,分别为模式输出结果直接使用方案,以及对比模式资料和湖北省历史区域高温事件发生的时间、持续过程以及逐日平均最高温度基础上的系统误差订正方案。在设计多样本集合延伸期预报高温事件有无的转换函数的基础上,开展了湖北省区域高温事件概率预报的检验。结果表明:湖北省区域性高温事件主要发生在7月中、下旬以及8月上旬,其中以7月下旬发生频次最多;采用系统误差订正方案对区域性高温的预报值和高温事件的预报准确率要优于模式输出结果直接使用方案,其订正后与订正前预报平均值误差相比要减少2.93℃,以及提前10、20、30天预报高温事件的准确率分别可达0.81、0.76和0.75。使用的气候模式数据以及研究方法对区域性高温事件的延伸期预报以及概率预报有较好的预报技巧。

关键词:区域高温事件;延伸期预报;集合平均预报;转换函数;概率预报

中图分类号: TP274+.2

文献标志码: A

doi: 10.16836/j.cnki.jcu.2018.05.006

0 引言

气候变化及其伴生的极端天气气候事件变化对经济社会发展有诸多不利影响。特别是近年来中国夏季高温热浪频发,极端高温事件严重影响了人们的生活和身心健康,并带来重大的经济损失^[1-2]。极端高温事件的发生通常具有一定的影响范围和持续时间,即是区域性高温事件,区域性高温事件已成为极端事件研究的一个新领域。近几年国内学者已经在区域性高温事件研究上取得了一定的成果,黄丹青等^[3]运用累积频率的阈值统计方法,分析了中国东部地区极端高温、极端低温以及全国的温度日较差极端事件的区域性特征。Ren等^[4]提出了一种区域性极端事件客观识别法。徐振亚^[5]利用区域性极端事件客观识别法对中国区域性高温事件进行模拟和预估。王艳姣等^[6]采用综合考虑高温事件温度强度、持续时间和影响范围等因子的区域持续性极端高温事件判别方法和指标体系,分析了中国近50年区域持续性极端高温事件的时空变化特征。

延伸期预报主要是指10~30 d的气象要素预报,对开展防灾减灾、工农业生产、决策等工作具有极其重要的意义^[7],但在预报准确性、稳定性和技巧性上又

具有相当的难度。本文采用延伸期集合平均预报方法对湖北省区域性高温和区域高温事件进行预报,通过结合历史资料,设计并采用了两种方案开展延伸期集合平均预报的检验。最后在设计多样本集合延伸期高温事件有无的转换函数基础上,开展了湖北省区域高温事件延伸期概率预报的检验。

1 资料说明

文中使用的资料包括历史最高气温资料以及NCEP-CFSv2气候模式历史回报资料,其中气温资料由湖北省1960~2017年76个观测站点逐日最高气温数据组成;气候模式资料由矩形范围(5°N~55°N,70°E~135°E)内气温格点数据组成,相邻气温格点经纬度间隔约为1°,模式原始的数据所占空间大小约为4TB。图1为气温资料观测站点的地理位置分布图。

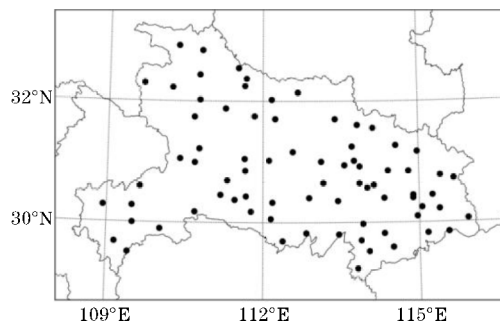


图1 气温资料观测站点地理位置分布

收稿日期: 2018-04-27

基金项目: 国家重点研发计划资助项目(2017YFC1502306、2016YFE0102400);中国气象局核心业务发展专项资助项目(YBGJXM(2018)04);中国气象局预报员专项资助项目(CMAYBY2018-048)

2 区域高温事件的定义及历史统计结果

对于极端高温事件阈值的确定,刘绿柳等^[8]以最高气温 $\geq 35^{\circ}\text{C}$ 和 38°C 作为极端高温的绝对阈值指标,以及基于百分位数的高温阈值和基于概率统计分布方法计算的多少年一遇作为相对阈值指标。此外,其他研究学者也对基于百分位数的高温阈值指标划分和计算方式有不同的解释^[1, 9-11],这一阈值指标也广泛应用于全球极端气候事件的研究^[11-13]。文中统计历史区域高温事件的阈值指标采用的是2012年国家气象中心起草的《区域高温天气过程等级划分》气象行业标准,根据标准规定:当某区域内同一天有20%以上的站点日最高气温达到或超过 35°C ,并且是成片出现,即为一个区域高温日(也称区域高温事件)。

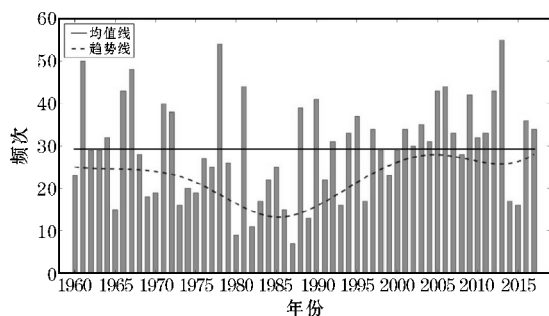


图2 湖北省区域高温事件日数历史频次统计

图2为1960–2017年湖北省区域高温事件发生的频次统计结果,其中实线为多年均值线,虚线为多年结果拟合趋势线。由图2可知,1960–2017年发生区域高温事件的频次以2013年最多,1987年最少。在这58年里发生区域高温事件的频次超过多年均值的共有28年,占有年份的51%。从拟合曲线可得,20世纪60年代湖北省区域高温事件发生频次较多;从70年代初到80年代中期,区域高温事件发生频次呈减少趋势;从1987年后又开始逐渐增多并在2000年左右趋于稳定,平均每年大约发生30次;自2000年以来,除了2014年和2015年以外高温频次数都大于或接近多年均值,可以明显看出区域高温事件发生频次较20世纪末至21世纪初有了增多的趋势,气候变暖与区域高温事件不断发生有一定的联系。图3为1960–2017年湖北省5–9月份区域高温事件分旬频次统计图,区域高温事件主要发生在7月中、下旬以及8月上旬,其中以7月下旬发生频次最多。

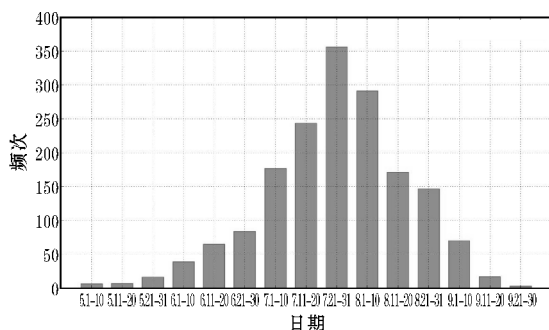


图3 区域高温事件分旬频次统计

3 区域高温事件延伸期预报方法

3.1 Grib2 数据预处理

在文中,通过wgrib2将模式输出的原始grib2文件转换为netcdf文件,将原始文件的格式转换为netcdf格式以便后期使用Python编程软件进行读取、判断等操作。转换格式后的netcdf文件变量名和经纬度不变,其中4个变量分别为“latitude”,“longitude”,“time”和“TMP_surface”。使用湖北省边界大小的矩形边框作为矢量文件对转换后的netcdf格式数据进行裁剪,裁剪后的数据共包含32个气温格点,每个格点的数值代表该位置(经纬度)的气温值,将气温值(开尔文温度)转化为摄氏温度。

3.2 确定历史气温资料的高温阈值

数值天气预报在预报过程中存在误差,且不能完全消除。如何减小预报误差是延伸期预报的关键。通过引入历史资料,结合历史资料对预报误差进行订正。将预报值与历史资料做对比,求出多年累计误差,最后将误差累加在预报值上,且有效提高了预报效果。

与上文所提到的区域高温事件判定方法不同,这里结合历史资料,首先通过对历史气温资料逐日76个站点的最高气温值求平均,作为湖北全省每日的高温值,然后设定一个阈值,通过判断高温值是否大于或等于该阈值来判定该日是否为区域高温日,该阈值的设定影响区域高温日的预报结果。每日76个站点最高气温的平均值为

$$T_{\text{obs}} = \sum_{i=1}^{76} T_i \div 76 \quad (1)$$

其中, T_i 为逐日每个站点的最高气温。根据该公式可求得每日的最高气温平均值 T_{obs} ,通过判断 T_{obs} 是否大于或等于高温阈值来确定该日是否为区域高温日。根据《区域高温天气过程等级划分》已经统计出区域高温事件发生的具体日期,通过 T_{obs} 与区域高温事件发生的具体日期对比,可以得出高温阈值大致分

布在33 ℃ ~ 34 ℃ ,为了将该高温阈值更好地在业务中应用,取33 ℃ ~ 34 ℃ 间隔为0.1 ℃ 的 11 个阈值,通过将阈值判定的区域高温日与实际区域高温日做对比,采用混淆矩阵误差检验方法,准确率最高的阈值作为最终的高温阈值。准确率为

$$Acc=\frac{a+d}{a+b+c+d}$$

(2)

式中 a, b, c 和 d 的含义如表 1 所示, a 为阈值判定的区域高温事件准确数, b 为空报数, c 为漏报数, d 为

阈值判定的区域高温事件不发生准确数。准确率计算结果如表 2 所示,当阈值为33.5 ℃ 时,通过阈值判定的区域高温日与实际区域高温日基本一致。

表 1 阈值判定混淆矩阵

阈值判定的区域 高温事件	实际区域高温事件	
	有	无
有	a	b
无	c	d

表 2 准确率结果

年份	33 ℃	33.1 ℃	33.2 ℃	33.3 ℃	33.4 ℃	33.5 ℃	33.6 ℃	33.7 ℃	33.8 ℃	33.9 ℃	34 ℃
2013	0.935	0.935	0.954	0.961	0.948	0.961	0.954	0.954	0.954	0.941	0.928
2014	0.974	0.974	0.974	0.974	0.98	0.98	0.98	0.98	0.98	0.974	0.98
2015	0.9	0.908	0.922	0.928	0.941	0.961	0.961	0.961	0.961	0.961	0.967
2016	0.93	0.935	0.941	0.948	0.954	0.954	0.948	0.954	0.954	0.967	0.974
准确率	0.935	0.938	0.948	0.953	0.956	0.964	0.961	0.962	0.962	0.961	0.962

3.3 延伸期集合预报

集合预报的目的主要有 3 个:提高预报质量;提高预报的可靠性;三是为概率预报提供定量基础^[14]。集合预报会过滤掉预报中不确定成分,而保留集合成员中一致的部分倾向。气候模式回报资料每天共生成 16 个预报样本,每次生成 4 个预报样本,时间周期为 6 小时。对转换后的 netcdf 数据样本读取包括预报日在内 30 天的预报数据,表 3 为延伸期集合预报样本的格式,“○”代表该时间点有预报数值。行数据为单个预报样本的预报数值,列数据为 16 个预报样本预报同一

时间点的预报数值。

采用集合平均预报的方法求出同一时间点所有样本的集合预报值,计算公式如下

$$T_{pre}=\sum_{i=1}^aT_i\div a$$

(3)

T_i 为每列的预报数值, a 为每列的样本数。计算后的集合预报值共有 120 个,根据每天时间点的不同伴随的气温不同,集合预报值呈有序排列,取每天同一时间点预报值的最高气温作为延伸期集合预报每日气温预报值。

表 3 延伸期集合预报样本

样本	第 1 天				第 2 天				……	第 30 天			
样本 1	○	○	○	○	○	○	○	○	……	○	○	○	○
样本 2	○	○	○	○	○	○	○	○	……	○	○	○	○
……	……	……	……	……	……	……	……	……	……	……	……	……	……
样本 5		○	○	○	○	○	○	○	……	○	○	○	○
……		……	……	……	……	……	……	……	……	……	……	……	……
样本 9			○	○	○	○	○	○	……	○	○	○	○
……			……	……	……	……	……	……	……	……	……	……	……
样本 13				○	○	○	○	○	……	○	○	○	○
……				……	……	……	……	……	……	……	……	……	……
样本 16				○	○	○	○	○	……	○	○	○	○

3.4 概率预报

为了更好地表达逐日发生区域性高温事件的可能性,使用概率预报方法预报每日发生高温事件的概率。

概率预报是用概率值表示预报对象出现可能性的大小,它所提供的不是区域高温事件的“有”或“无”,而是区域高温事件发生的可能性有多大。概率预报以延伸期集合预报为基础,采用多样本集合延伸期预报方

法,建立高温事件有无转换函数,函数公式为

$$P(x)=\sum_{i=1}^{16}F(x)\div16$$

(4)

$P(x)$ 为每日发生高温事件的概率, $F(x)$ 为每日有无发生高温事件,有为“1”,无为“0”。

4 区域高温事件延伸期预报检验

4.1 误差订正

数值预报过程中,存在着一定的误差,这些误差导致数值预报结果与实际存在偏差,采用逐月系统误差订正方法^[15],根据公式(5)–(7),将逐月逐日的延伸期集合预报值 T_{pre} 与实际高温值 T_{obs} 相减,计算出每月的平均值误差后进一步得到多年同月的平均值误差,不同月份得到不同的预报误差。将预报误差与对应月份预报日的预报值相加后得到订正后的预报值 T_{pred} 。将订正后的预报值再次通过公式(5)和(6)计算订正后的预报误差,订正后预报平均值误差比订正前要减少2.93℃,且预报值接近于实际高温值。因为区域高温日主要发生在每年的5–9月,所以误差订正研究日期为5–9月。

$$Error1=T_{\text{obs}}-T_{\text{pre}}$$

(5)

$$Error=\sum_{j=1}^c(\sum_{i=1}^bError1\div b)\div c$$

(6)

$$T_{\text{pred}}=T_{\text{pre}}+Error$$

(7)

其中, T_{obs} 为逐日实际高温值, T_{pre} 为逐日预报值(订正前), T_{pred} 为最终预报值(订正后), $Error1$ 为逐日预报误差, $Error$ 为逐月预报误差, b 为该月的天数(30或31), c 为年数。

4.2 延伸期预报评估检验

为了更清晰地描述预报时段内的高温事件,可以用“1”和“0”分别表示逐日是否发生区域性高温事件。延伸期预报有多种检验方法,为了客观地检验出高温事件预报的准确性,借鉴杜良敏等^[16]使用命中率和假警报率延伸期预报检验方法,通过混淆矩阵,使用准确率、假正率(假警报率)和假负率做延伸期预报评估检验。准确率反映了预报的准确性,假正率反映了达到这一预报能力包含空报的比例,假负率反映了达到这一预报能力包含漏报的比例。通过使用准确率、假正率和假负率来共同表征其预报技巧。准确率(Accuracy)、假正率(FP)和假负率(FN)的分别用式(8)–(10)表示,其中, A, B, C, D 的含义如表4所示, A 为预报发生准确数, B 为空报数, C 为漏报数, d 为预报不发生准确数。

$$Accuracy=\frac{A+D}{A+B+C+D}$$

(8)

$$FP=\frac{B}{A+B+C+D}$$

(9)

$$FN=\frac{C}{A+B+C+D}$$

(10)

表4 预报结果混淆矩阵

预报	观测	
	有	无
有	A	B
无	C	D

4.3 延伸期预报结果

由于区域高温事件主要发生在每年的5–9月中的6,7和8月,为了尽可能地减小预报准确率被预报和实况同为无高温事件的过高影响,延伸期预报准确率评估检验选择的时间序列为逐年6–8月共92 d。通过公式(8)分别计算每年逐日提前10 d,20 d和30 d的预报准确率,将多年所有日求平均值可得,提前10 d,20 d和30 d预报高温事件的准确率分别可达0.81,0.76和0.75,可以看出,所使用的延伸期预报方法对提前10~30 d预报高温事件有较好的预报性。通过公式(8)–(10)分别计算2013–2016年6–8月的预报准确率,假正率和假负率。图4为延伸期预报评估检验图像。其中,图4(a)为2013–2016年6–8月的总体预报准确率,假正率和假负率结果,根据图可得,从预报日起,预报1~30 d高温事件的准确率逐渐降低,但总体准确率为0.802,尽管假正率和假负率结果分别为0.084和0.113,综合三者共同特征,所使用的预报方法对延伸期高温事件预报有较好的预报技巧。图4(b)、(c)、(d)分别为2013–2016年6,7和8月的延伸期预报检验结果。可以看出8月的预报效果最好,6月次之,7月的预报准确率较低,但三者的预报的准确率均大于0.5。

图5是以历史上3个高温事件为例提前30 d,20 d和10 d预报高温事件有无发生的预报图,图中蓝色实线为集合平均预报高温值,红色虚线为高温阈值(33.5℃),蓝色柱状图表示预报有高温事件发生,红色柱状图表示实际有高温事件发生。表5为图5所代表3个高温事件的日期说明。图5(a)、(b)、(c),以7月13日起报,提前30 d没有预报出8月11日为高温日,虽然能较好地预报出其他高温日,但存在5次空报,4次漏报现象,预报准确率(见公式7)为0.7。以7月23日起报,提前20 d准确预报出8月11日为高温日,其他高温日全部预报准确,但存在1次空报现象,预报准确率为0.95。以8月2日起报,提前10 d准确

预报出 8 月 11 日为高温日,其他高温日全部预报准确,预报准确率为 1;图 5(d)、(e)、(f) 分别为以 2014 年 8 月 5 日为例的提前 30 d、20 d 和 10 d 的预报图,其

预报准确率分别为 0.8、0.8 和 1;图 5(g)、(h)、(i) 分别为以 2015 年 8 月 2 日为例的提前 30 d、20 d 和 10 d 的预报图,其预报准确率分别为 0.47、0.5 和 0.5。

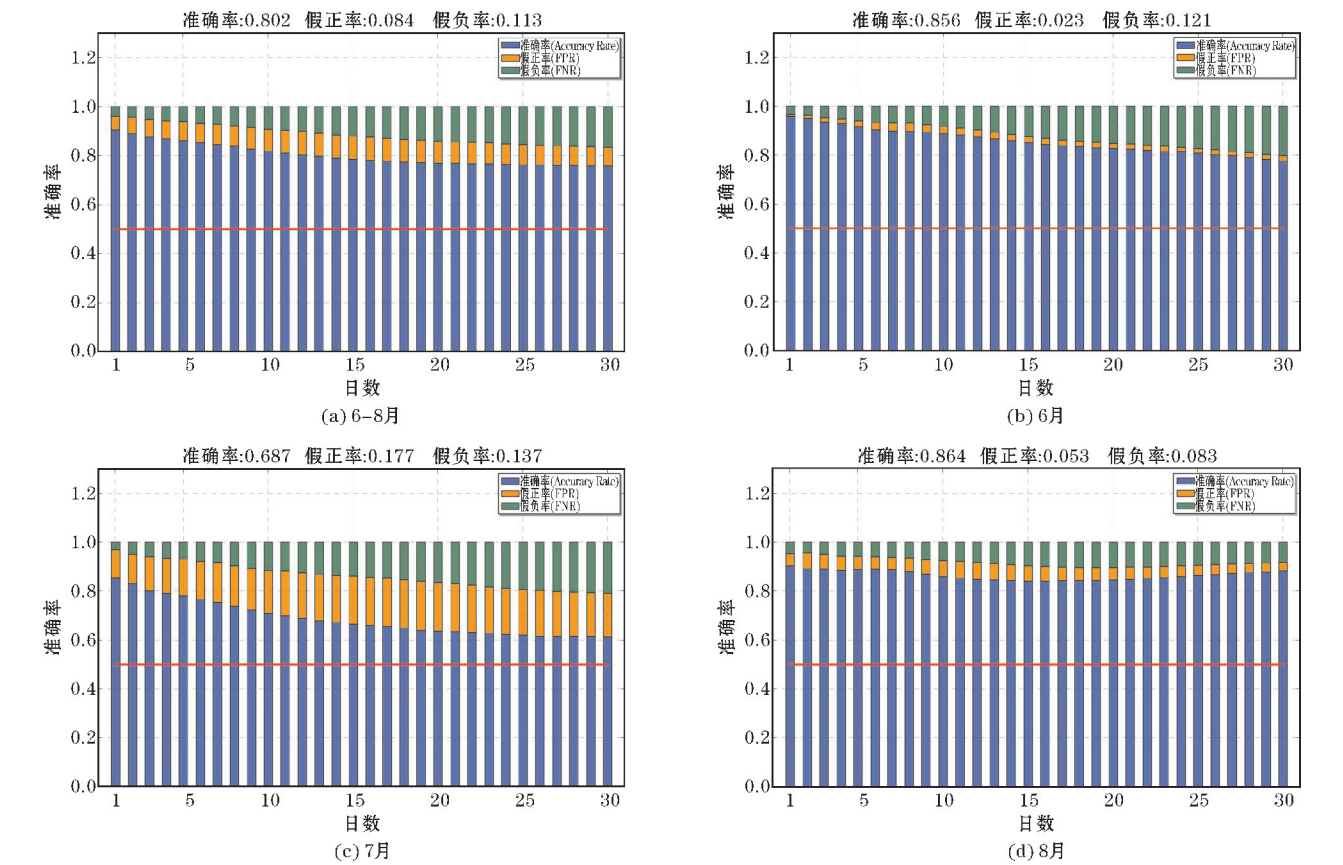


图 4 2013-2016 年 6-8 月延伸期预报检验结果

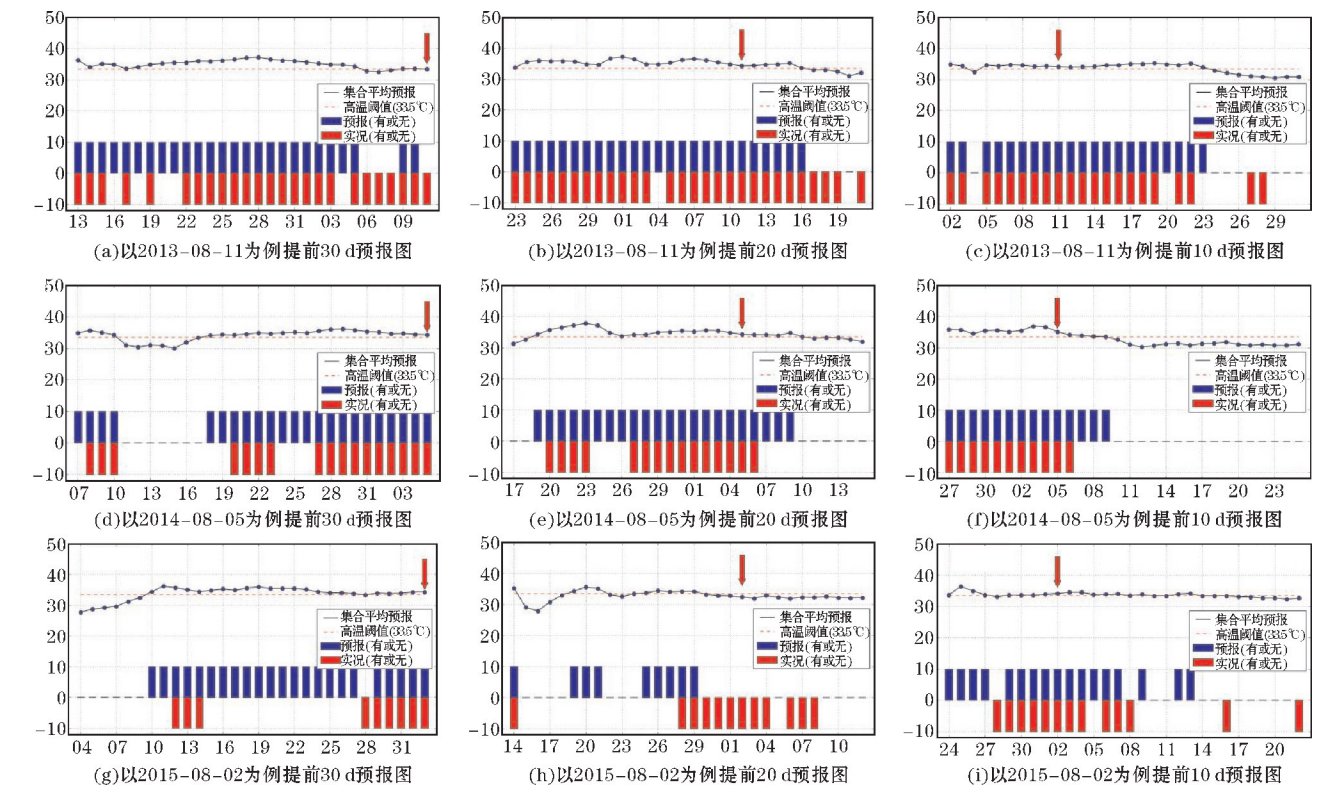


图 5 区域高温事件预报

表5 高温日日期

图5子图序	高温日
a, b, c	2013-08-11
d, e, f	2014-08-05
g, h, i	2015-08-02

图6是以历史上3个高温事件为例提前30 d, 20 d和10 d预报高温事件发生的概率预报图。其中, 每张图分别由两张小图组成, 上图中颜色较暗的16条线表示每日16个样本预报温度值, 蓝色虚线表示集合平均预报温度值, 红色实线表示实况温度值; 下图蓝色

柱状图表示预报有高温事件发生的概率(为了方便显示, 已放大10倍), 红色柱状图表示实际有高温事件发生。图6所使用的3个高温事件例子的日期与图5相同, 但与图5预报高温事件有无发生的结果相同的是, 图6概率预报中能够以较高概率预报出实况高温事件的发生; 不同的是, 概率预报能够以较低概率预报出图5预报中漏报的实况高温事件。虽然预报过程中存在空报现象, 但整体预报效果较为理想。综上可得, 本文所使用的概率预报方法对延伸期高温事件发生的概率有较好的预报技巧。

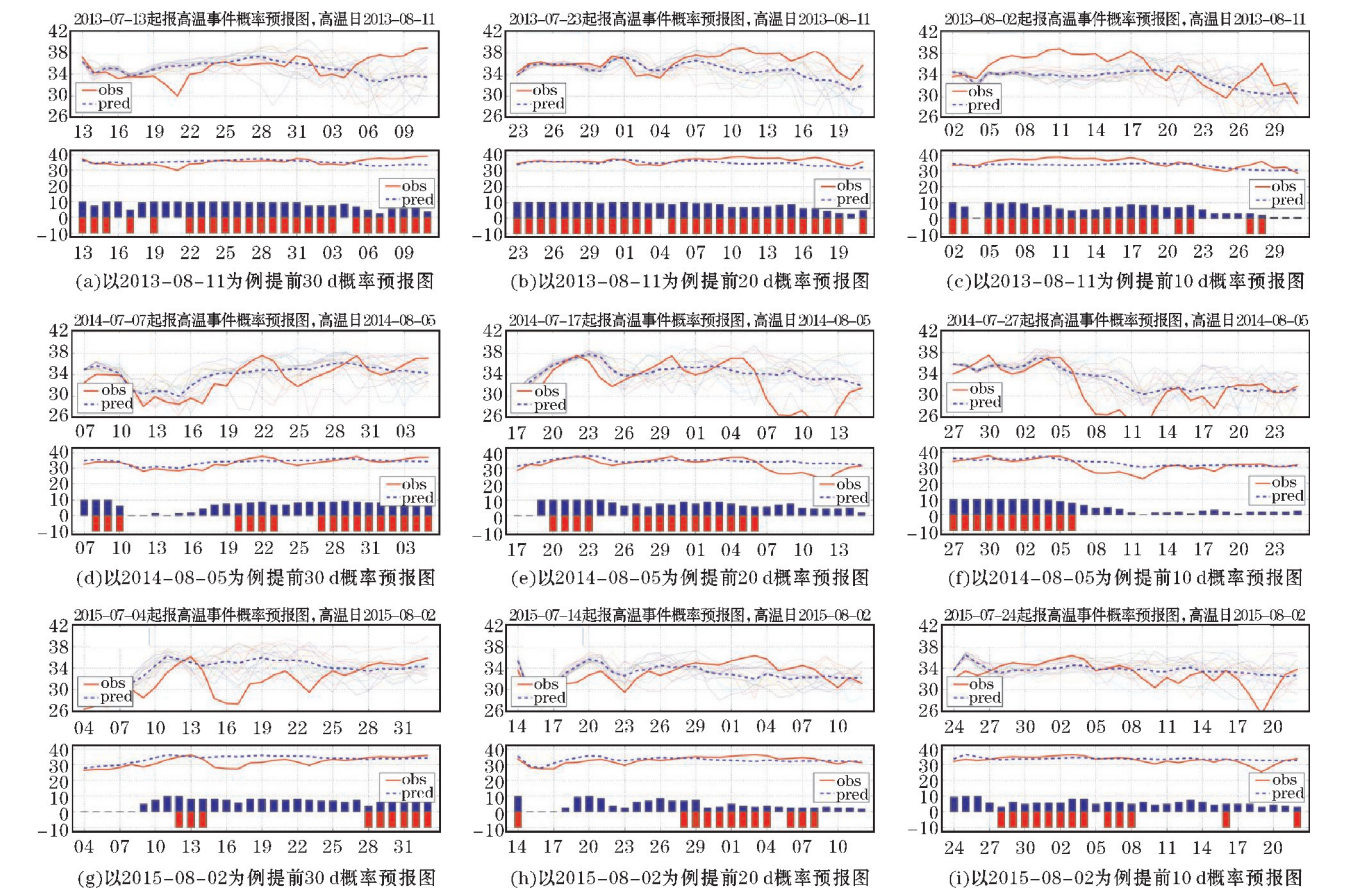


图6 区域高温事件概率预报

5 结论与讨论

利用历史气温资料分析得出湖北省区域高温事件主要发生在每年的7月中、下旬以及8月上旬, 其中7月下旬发生频次最多。利用 NCEP-CFSv2 气候模式回报资料, 分别使用两种高温判定方法统计了历史区域高温事件发生的时间和逐日平均最高温度, 使用集合平均预报方法预报出逐日最高气温, 在历史资料误差

订正的基础上, 进一步提高了预报准确率, 最后在设计多样本集合延伸期高温事件有无的转换函数基础上, 开展了湖北省区域高温事件延伸期概率预报的检验。得到如下结论:

- (1) 直接使用模式产品具有一定的预报效果, 提前10 d、20 d和30 d预报高温事件的准确率分别可达0.68, 0.67, 0.65。
- (2) 误差订正后对区域高温的预报值和高温事件预报准确率要好于订正前, 其预报平均值误差相比要

减少 $2.93\text{ }^{\circ}\text{C}$,以及提前10 d、20 d和30 d预报高温事件的准确率分别可达0.81、0.76和0.75。

(3) 概率预报检验结果中,虽然存在空报现象,但整体预报效果较理想,能以较高概率预报出延伸期高温事件的发生。

综上所述,文中所使用的气候模式数据和预报方法对延伸期高温事件的预报有较好的预报技巧,该数据可以应用于区域高温事件延伸期预报研究。

参考文献:

- [1] 李庆祥,黄嘉佑. 对我国极端高温事件阈值的探讨[J]. 应用气象学报,2011,22(2):138-144.
- [2] 周后福,王兴荣,翟武全,等. 基于混合回归模型的夏季高温日数预测[J]. 气象科学,2005,25(5):505-512.
- [3] 黄丹青,钱永甫. 极端温度事件区域性的分析方法及其结果[J]. 南京大学学报(自然科学),2009,45(6):715-723.
- [4] Ren Fumin, Cui Donglin, Gong Zhiqiang, et al. An Objective Identification Technique for Regional Extreme Events [J]. Journal of Climate, 2012, 25: 7015-7027.
- [5] 徐振亚. 中国区域性高温事件的模拟与预估[D]. 南京:南京大学,2012.
- [6] 王艳姣,任福民,闫峰. 中国区域持续性高温事件时空变化特征研究[J]. 地理科学,2013,33(3):314-321.
- [7] 杨秋明. 10-30 d 延伸期天气预报方法研究进展与展望[J]. 地球科学进展,2015,30(9):970-984.
- [8] 刘绿柳,孙林海,廖要明,等. 国家级极端高温短期气候预测系统的研制及应用[J]. 气象,2008,34(10):102-107.
- [9] 黄丹青,钱永甫. 我国极端温度事件的定义和趋势分析[J]. 中山大学学报(自然科学版),2008,47(3):112-116.
- [10] Alexander L V, Zhang X, Peterson T C, et al. Global observed changes in daily climate extremes of temperature and precipitation[J]. Journal of Geophysical Research, 2006, 111, D05109.
- [11] Zhang X, Hegerl G, Zwiers F W, et al. Avoiding inhomogeneity in percentile-based indices of temperature extremes [J]. Journal of Climate, 2005, 18: 1641-1651.
- [12] 江志红,丁裕国,陈威霖. 21 世纪中国极端降水事件预估[J]. 气候变化研究进展,2007,3(4):202-207.
- [13] 黄丹青,钱永甫. Community Climate Model 3 模拟夏季极端降水的初步分析[J]. 南京大学学报(自然科学版),2007,43(3):238-248.
- [14] 郑志海. 月动力延伸预报研究进展回顾[J]. 气象科技进展,2013,3(1):25-30.
- [15] 苏海晶,封国林,杨杰,等. 基于动力统计方法对中国夏季温度模式误差订正的研究[J]. 高原气象,2015,34(5):1345-1356.
- [16] 杜良敏,柯宗建. 一种适用于延伸期过程事件预报的检验方法[J]. 应用气象学报,2013,24(6):686-694.

Application of Climate Model Dataset in Extended-Range Forecast of Regional High Temperature Events

ZHOU Chen-guang¹, DU Liang-min², GAO Wei¹, WANG Hong-ji³

(1. Department of Information Engineering, China University of Geosciences, Wuhan 430074, China; 2. Wuhan Region Climate Center, Hubei Meteorological Bureau, Wuhan 430074, China; 3. Weather Information and Technology Support Center, Hubei Meteorological Bureau, Wuhan 430074, China)

Abstract: In view of the frequent occurrence of regional hightemperature events in Hubei Province in recent years, in this paper,utilizing the historical temperature data of Hubei Province, two high temperature determination methods were used to count the time of the historical regional high temperature eventsand daily average maximum temperature; With

the American NCEP-CFSv2 climate model data, an ensemble mean forecast method was used to carry out the extended-range forecast of regional high temperature events in Hubei Province based on daily multiple time series samples. Two schemes were designed and used to test the extended-range forecast results of regional high temperature events, they are the directly used scheme of the model output results, as well as the system error correction scheme based on the comparison model data and the time, duration, and daily average maximum temperature of the historical regional high temperature events in Hubei Province. Based on the design of the high temperature event extended-range ensemble forecast presence or absence transfer function, carrying out the probabilistic forecast test of regional high temperature events in Hubei Province. The results show that the regional high temperature events in Hubei Province mainly occurred in the middle and late July and early August, of which the frequency of occurrence was the most in late July; the forecast value of the regional high temperature and the forecast accuracy of the regional high temperature events through using the system error correction scheme was better than using the scheme of the model output results direct use, the corrected forecast average error than before correction is reduced by 2.93°C , and the forecast accuracy of high temperature events at 10, 20, and 30 days in advance can reach 0.81, 0.76, and 0.75, respectively. The research method of this paper has better prediction skills for extended-range forecast and probability forecast of regional high temperature events.

Keywords: regional high temperature events; extended-range forecast; ensemble mean forecast; transfer function; probability forecast