

# 制作汛期降水集成预报的分区权重法\*

周家斌 张海福 杨桂英

(中国科学院大气物理研究所, 北京 100029)

彭 瑶

(中国科学院软件研究所, 北京 100080)

## 提 要

该文提出一种不受历史预报档案资料多少限制的集成预报方法. 这个方法的基本思想是以各种不同预报的历史评分资料为基础, 确定各个预报方法的权重. 根据若干原则设计了16种预报方案. 计算结果表明, 这个方法具有较高的准确度和可行性, 可以用于业务预报.

关键词: 集成预报 分区权重法 汛期降水

## 引 言

每年3月底4月初, 国家气候中心都要召开汛期预报会. 参加会议的各单位在会上发言, 展示各自的预报结果. 会议经过讨论, 最后得出当年的预报. 这个经过讨论得出的预报, 主要是由主班预报员根据会议的发言综合做出的, 难免有一定主观性. 能否设计一种客观方法将各单位的预报综合在一起呢? 这就是集成预报问题.

我国科学家于1960年发表了关于综合预报的论文<sup>[1]</sup>, 并提出应对准确率相差悬殊的预报采用不同的权重. 国外关于集成预报的研究始于70年代<sup>[2]</sup>. 后来主要集中于数值预报结果的集成, 其方法有综合平均预报(ensemble average forecast)<sup>[3]</sup>、滞后平均预报(lagged average forecast)<sup>[4]</sup>、Monte Carlo 预报<sup>[4]</sup>、调和预报(tempered forecast)<sup>[4]</sup>等. 周家斌于1975年发表了关于预报集成问题的论文<sup>[5]</sup>, 近年来国内这方面的研究见文献[6, 7].

集成预报还有很多方法<sup>[5]</sup>, 其基本思路都是将各种预报结果做统计处理. 统计需要有一定的样本, 然而国家气候中心现存各单位预报的资料除本中心有20多年外, 多数单位只有3、4年资料, 因此无法用统计方法加以处理. 为解决这一问题, 我们提出一种不受样本多少限制的分区权重法.

\* 本文得到“九五”国家重中之重项目“中国短期气候预测系统的研究”第四课题(96-908-04)的资助.  
1998-04-17收到, 1999-03-15收到再改稿.

## 1 预报公式

分区权重法的基本预报公式是

$$Y_k^*(t) = \sum_{j=1}^{j_k} \beta(t)_{j,k} X_{j,k}(t) \quad (1)$$

式中  $\beta(t)_{j,k}$  表示第  $j$  种预报方法  $k$  站  $t$  年的权重,  $J_k$  为对  $k$  站做预报的方法数.  $k=1, 2, \dots, 160$  表示台站, 不同的站处于不同的区域.  $t$  表示年序号,  $Y_k^*(t)$  表示  $k$  站  $t$  年的集成预报值.  $X_{j,k}(t)$  表示第  $j$  种预报方法对  $k$  站  $t$  年的预报.

权重  $\beta(t)_{j,k}$  的表达式为:

$$\beta(t)_{j,k} = P_{j,k}(t) S(t)_{j,k} / \sum_{i=1}^{j_k} P_{i,k}(t) S(t)_{i,k} \quad (2)$$

式中  $S(t)_{j,k}$  为第  $j$  种预报方法  $k$  站截止  $t-1$  年的评分的平均值, 即

$$S(t)_{j,k} = \frac{1}{n_t} \sum_{t'=t-n_t}^{t-1} S_{j,k,t'} \quad (3)$$

其中  $S_{j,k,t'}$  为  $t$  年以前(不含  $t$  年)第  $j$  种方法  $k$  站  $t'$  年的评分.  $n_t$  为截止  $t-1$  年的评分资料长度.

$P_{j,k}(t)$  为第  $j$  种方法对  $k$  站  $t$  年预报的稳健性程度. 其计算公式为

$$P_{j,k}(t) = 1 + N_{j,t} \rho_{j,k} \quad (4)$$

式中  $N_{j,t}$  为  $j$  种方法至  $t-1$  年的评分资料年份数,  $\rho_{j,k}$  为预报评分年平均增长率.

设计上式的基本指导思想是鼓励长期做预报, 肯定长期评分资料的稳健性.

## 2 计算方案

在具体计算时对评分资料和预报图进行了一些处理, 并采用了不同的评分方法, 因而形成了不同的方案. 首先用下式对评分进行极差标准化:

$$S_n(t) = \frac{S(t) - S_{\min}(t)}{S_{\max}(t) - S_{\min}(t)} \quad (5)$$

式中  $S(t)$  为  $t$  年的评分值,  $S_n(t)$  为其标准化值,  $S_{\min}$  为全部评分资料截止  $t$  年的评分的最小值,  $S_{\max}(t)$  为截止  $t$  年的评分的最大值.

预报评分标准化后,  $0 \leq S_n(t) \leq 1$ , 有利于拉开不同评分结果的档次. 其次是对预报图进行平滑, 共分四步进行.

第一步, 用下式计算两站之间降水距平百分率的欧氏距离:

$$D_{kl} = \left[ \sum_{t=1}^T (R_{kt} - R_{lt})^2 \right]^{1/2} \quad (6)$$

式中  $R_{kt}$  和  $R_{lt}$  分别为  $k$  站和  $l$  站  $t$  年的降水距平百分率,  $T$  为资料年份数.  $D_{kl}$  表示两站间降水差异. 对  $k$  站, 选出  $D_{kl}$  小于某一阈值的若干站, 认为  $k$  站对这些站有代表性.

第二步, 计算不同台站间的几何距离, 选出与  $k$  站几何距离小于某一阈值的若干站.

第三步, 选出同时满足以上两步要求的台站中欧氏距离最小的若干个台站(如 5 个)作为某站可以代表的站.

表 1 计算方案

方案	评分	标准化否	平滑否	备注
111		×	×	
121	$P_c$	√	×	
112		×	√	
122		√	√	
211		×	×	
221	$T_{cc}$	√	×	
212		×	√	
222		√	√	
311		×	×	
321	$T_{cr}$	√	×	
312		×	√	
322		√	√	
411		×	×	
421	$Acc$	√	×	
412		×	√	
422		√	√	
000				等权重

第四步,若同时满足前两步要求的台站数较少(如不足3个),则适当放宽对欧氏距离和(或)几何距离的阈值要求,使该站可以代表的站适当增多.这种站很少,且均在西部台站稀少地区.

确定各站所代表的台站后,将预报图对各站用其被代表的站做一次平均计算.这样做的目的在于适当消除因人为读数所产生的各站降水距平百分率间的不协调性.

对历史评分资料,我们选取了四种,即预报准确率  $P_c$ ,与随机预报比较的技巧分  $T_{cr}$ ,与气候预报比较的技巧分  $T_{cc}$ 和距平相关系数  $Acc$ .这些评分计算方法见文献[8].评分资料由国家气候中心提供.根据不同评分方案评分是否标准化、预报图是否平滑组成表1中16种方案.为了比较起见,又计算了各种预报取等权重的投票法(代号为000),因此共有17种方案.

### 3 计算结果及实际预报效果

我们做了1995,1996,1997年的集成预报试验.参加集成的共有8个单位.1995年的预报用截止1994年的评分资料.1996年的预报用截止1995年的评分资料,1997年的预报用截止1996年的评分资料.每年计算17种方案,因此共有51张集成预报图.为了评价这51张图的预报效果,对每张图又计算了4种预报评分.表2就是评分结果.这些评分是用引文<sup>1)</sup>中的公式自行编制程序计算的.所有的图都是6~8月降水距平百分率图.

由表2可见,17种方案中某些方案获得了较高的评分.由各种评分高低综合考虑,选出方案311和121作为推荐方案.

下面对几个问题作一些讨论.

(1) 评分标准化问题 从表2看,标准化的  $P_c$  评分对预报改进比较明显.方案121的4种评分均高于111,方案122的12个评分值有11个高于112.这是很容易理解的.未经标准化的评分一般都是60~75分,由此计算的权重各单位间差别不大.而经过标准化后,差距明显拉大,相应的权重也就有了大的差别,因此在集成预报中历史上预报效果较好的单位占了更大的权重.例如1995年方案111中各单位权重最高为0.141,最低的为0.108,而采用标准化评分的方案121中权重最大者为0.177,最小者为0.083,差距明显

1) 周家斌.关于集成预报的几个问题.短期气候预测业务系统集成研讨会文集.96-908-04,96-908-05 课题组,1997,7~8.

表 2 不同集成预报方案评分

方案 代号	1995 年				1996 年				1997 年				平均值			
	<i>P<sub>c</sub></i>	<i>T<sub>cr</sub></i>	<i>T<sub>cc</sub></i>	<i>Acc</i>												
111	72.2	0.07	0.17	0.16	62.8	-0.10	0.02	-0.08	65.4	-0.03	0.08	0.03	66.8	-0.02	0.09	0.04
121	74.4	0.12	0.22	0.18	63.8	-0.05	0.07	-0.04	67.3	-0.02	0.09	0.05	68.5	0.02	0.13	0.06
112	68.1	0.05	0.07	0.14	62.4	-0.05	0.07	-0.09	60.6	-0.17	-0.05	0.03	63.7	-0.09	0.03	0.01
122	67.7	0.05	0.07	0.15	63.4	-0.03	0.08	-0.06	63.1	-0.14	-0.01	0.01	64.7	-0.07	0.05	0.03
211	70.8	0.02	0.13	0.17	67.5	-0.10	0.02	0.02	70.1	-0.07	0.05	0.01	69.5	-0.05	0.07	0.06
221	73.0	0.07	0.17	0.17	62.8	-0.12	0.00	-0.09	64.2	-0.05	0.07	0.02	66.7	-0.03	0.08	0.03
212	69.9	0.02	0.09	0.20	61.5	-0.22	0.09	-0.00	68.7	0.03	0.14	0.03	66.7	-0.07	0.05	0.08
222	67.7	0.05	0.07	0.15	62.4	-0.07	0.05	-0.10	61.3	-0.16	-0.03	0.03	63.8	-0.09	0.04	0.01
311	73.3	0.07	0.17	0.21	62.8	-0.05	0.07	-0.05	71.1	0.03	0.14	0.01	69.1	0.02	0.13	0.06
321	73.0	0.06	0.16	0.19	63.2	-0.07	0.05	-0.08	65.4	-0.02	0.09	0.02	67.2	-0.01	0.10	0.04
312	68.1	0.01	0.10	0.21	61.6	-0.15	0.02	-0.08	60.6	-0.16	-0.03	0.04	63.4	-0.11	0.02	0.03
322	67.9	0.03	0.08	0.17	62.4	-0.07	0.05	-0.10	61.9	-0.14	-0.01	0.04	64.1	-0.08	0.04	0.01
411	73.8	0.03	0.14	0.23	65.7	-0.03	0.08	0.10	64.5	-0.08	0.03	0.07	68.0	-0.03	0.08	0.09
421	73.0	0.04	0.15	0.20	63.4	-0.07	0.05	-0.04	66.7	-0.02	0.09	0.00	67.7	-0.02	0.10	0.05
412	70.6	0.01	0.10	0.25	66.9	0.02	0.13	0.08	60.5	-0.12	-0.00	0.09	66.0	-0.04	0.08	0.08
422	68.9	0.02	0.09	0.18	62.4	-0.06	0.06	-0.06	61.3	-0.16	-0.03	0.06	64.2	-0.08	0.04	0.02
000	72.5	0.06	0.16	0.03	64.4	-0.08	0.03	-0.01	66.9	-0.03	0.08	0.00				

拉大. 相应的预报图的评分也由 72.2 提高到 74.4,  $T_{cr}$ 、 $T_{cc}$ 、 $Acc$  也都得到提高.

(2) 集成预报方案的可行性问题 为了讨论集成预报方案的可行性, 我们在表 3 中列出了 1995~1997 年各单位预报的评分和方案 311 和 121 历年预报的评分以及等权方案评分, 对 1997 年还列出方案 111 的评分. 这些评分是我们用文献[8]中的公式计算的. 在表中, 每年都对不同的预报评分进行排序. 由表可见, 集成预报的各种评分值都比较高, 而且在不同的年份保持稳定. 我们认为, 评分保持稳定这一点很重要, 因为各单位的评分在不同的年份里波动很大(参见表 3). 因此这种集成预报方案是可以考虑在业务预报中试用的. 表中列出的等权预报的评分不高, 可见用简单的投票法做预报效果是不好的, 多数情况下处于中下水平.

表 4 给出方案 311 和各单位预报的平均评分以及等权方案评分 3 年总平均的比较. 由表可知, 方案 311 与 8 单位的预报相比, 预报准确率高 5.1 分, 与随机预报比较的技巧分、与气候预报比较的技巧分和距平相关系数均高 0.05. 与投票法相比, 预报准确率高 2 分, 与随机预报比较的技巧分、与气候预报比较的技巧分和距平相关系数分别高 0.06、0.05、0.05.

表 3 不同单位预报与集成预报评分比较

	1	2	3	4	5	6	7	8	000	311	121	111	
1995 年	$P_c$	72.2	71.0	58.4	50.0	77.4	62.1	66.9	68.3	72.5	73.3	74.4	
	$R$	5	6	10	11	1	9	8	7	4	3	2	
	$Tr$	0.12	0.06	- 0.06	0.02	0.17	- 0.15	0.09	- 0.12	0.06	0.07	0.12	
	$R$	2	6	9	8	1	11	4	10	7	5	3	
	$T_c$	0.22	0.16	0.00	0.13	0.26	- 0.03	0.19	- 0.00	0.16	0.17	0.22	
	$R$	2	6	9	8	1	11	4	10	7	5	3	
	$Ac$	0.20	- 0.01	- 0.05	- 0.02	0.18	- 0.13	0.16	0.17	0.03	0.21	0.18	
	$R$	2	8	10	9	4	11	6	5	7	1	3	
	1996 年	$P_c$	63.8	60.1	71.7	65.6	62.8	62.1	62.9	62.6	61.9	62.8	63.8
$R$		4	11	1	2	7	9	5	8	10	6	3	
$Tr$		- 0.03	- 0.15	0.20	- 0.08	- 0.08	- 0.05	0.01	0.07	- 0.14	- 0.05	- 0.05	
$R$		4	11	1	9	8	5	3	2	10	6	7	
$T_c$		0.08	- 0.02	0.28	0.03	0.03	0.07	0.11	0.17	- 0.01	0.07	0.07	
$R$		4	11	1	9	8	5	3	2	10	6	7	
$Ac$		0.05	- 0.13	0.08	- 0.23	0.06	- 0.12	- 0.09	0.05	- 0.01	- 0.05	- 0.04	
$R$		3	10	1	11	2	9	8	4	5	7	6	
1997 年		$P_c$	58.4	72.6	57.3	58.5	66.3	67.6	44.6	72.3	66.9	71.1	67.3
	$R$	10	1	11	9	8	4	12	2	7	3	5	6
	$Tr$	- 0.10	0.22	- 0.21	- 0.20	0.06	- 0.06	- 0.48	0.02	- 0.03	0.03	- 0.02	- 0.02
	$R$	9	1	11	10	2	8	12	4	7	3	5	6
	$T_c$	0.02	0.31	- 0.08	- 0.07	0.16	0.06	- 0.32	0.13	0.08	0.14	0.09	0.09
	$R$	9	1	11	10	2	8	12	4	7	3	5	6
	$Ac$	- 0.15	0.30	- 0.03	- 0.06	0.13	0.12	- 0.28	0.04	0.00	0.01	0.05	0.03
	$R$	11	1	9	10	2	3	12	5	8	7	4	6

注:  $P_c, Tr, T_c, Ac$  分别表示预报准确率、与随机预报比较的技巧分、与气候预报比较的技巧分和距平相关系数,  $R$  为该方案在不同预报方案中评分的排名。

图 1 给出 1995 年方案 311 的预报图, 这张图在 11 种预报中  $Acc$  列第一位,  $P_c$  列第三位,  $Tcr, Tcc$  列第五位。图 2 是实况。易见, 方案 311 正确地报出该年全国两条雨带的特点。

图 3 是 1997 年方案 111 的预报图, 这张图是我们在 1997 汛期预报会上发布的。图 4

是实况图。图 3 的 4 种评分在 12 个预报结果中均排名第六, 属中上水平。易知, 方案 111 对我国当年南方的多雨和北方的干旱的预报是比较好的。发布预报时我们只做了 111 这一种方案的计算, 其预报效果是令人高兴的。事后计算的结果表明, 还可以计算出评分更

表 4 方案 311 的评分和各单位预报的平均评分以及等权方案评分 3 年总平均比较

	311	8 单位	000
$P_c$	69.1	64.0	67.1
$Tr$	0.02	- 0.03	- 0.04
$T_c$	0.13	0.08	0.08
$Acc$	0.06	0.01	0.01



图 1 1995 年 6~8 月降水距平百分率预报图(方案 311)(图例表示降水距平百分率  $\rho$  的不同档次, 即  $\rho \geq 30\%$ ,  $15\% \leq \rho < 30\%$ ,  $0 \leq \rho < 15\%$ ,  $-15\% \leq \rho < 0$ ,  $-30\% < \rho < -15\%$ ,  $\rho \leq -30\%$ )

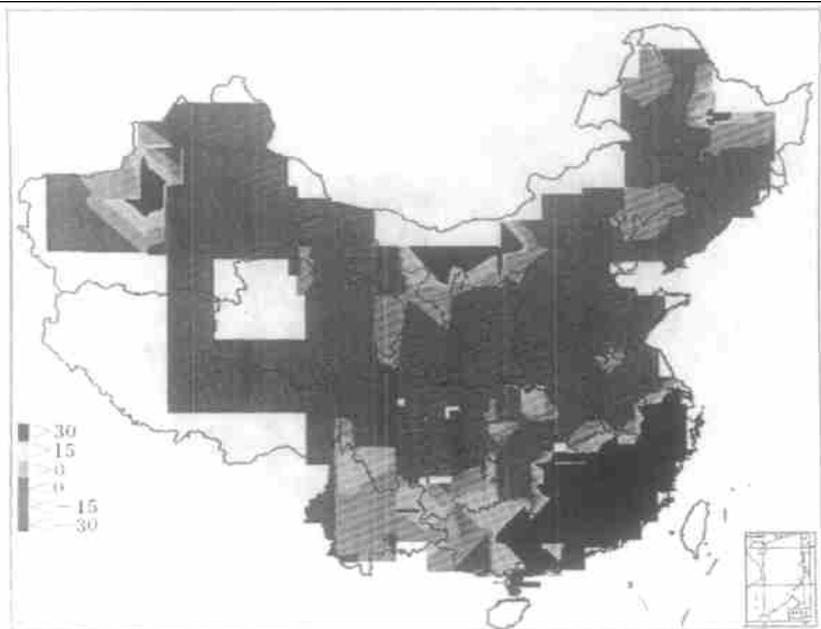


图 2 1995 年 6~8 月降水距平百分率实况图(说明同图 1)

高的方案 311 的预报图. 我们用本文的方法制作的 1998 年汛期降水集成预报已经在当年汛期旱涝趋势预报会商会上发布.

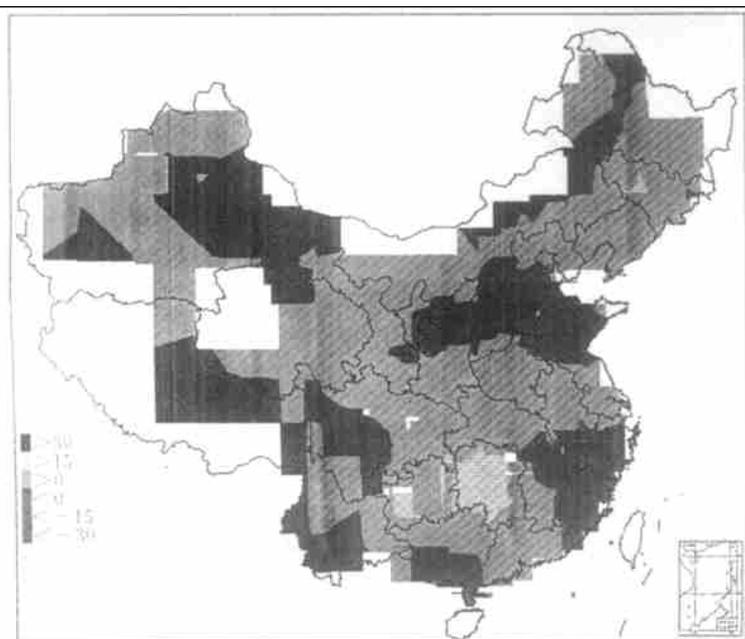


图3 1997年6~8月降水距平百分率预报图(方案111)(说明同图1)

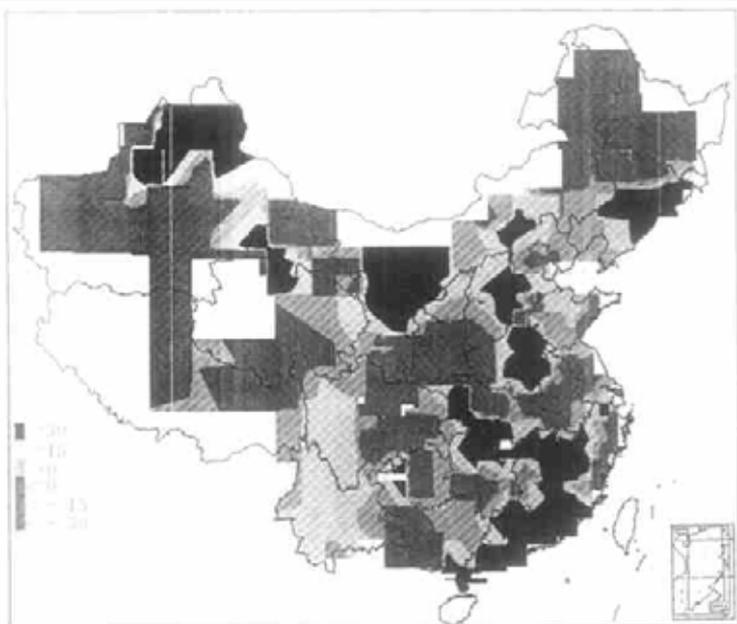


图4 1997年6~8月降水距平百分率实况图(说明同图1)

#### 4 结 语

集成预报是一个难题. 要想使集成预报每年对各种评分方案都达到最高的预报评分,

不是一朝一夕就可以做到的. 本文给出的分区权重法, 可以达到较高的评分, 且保持稳定. 我们认为这一方法已经可以在业务预报中试用了.

本文给出的  $\beta_{jk}$  和  $P_{jk}$  皆有下标  $k$ , 表示权重和方法的稳健度与测站有关, 这是为了适应不同区域预报结果而设计的, 也可以适应不同区域不同评分的情况. 至今我们没有得到不同省、市、自治区的预报, 也没有得到不同区域的评分, 因此本文给出的计算结果中权重和稳健度是全场一致的. 关于不同省、市、自治区的预报结果和不同区域的评分参加集成的问题, 留待以后进一步研究.

本文的降水集成预报方法同样可以用于其它要素的集成预报.

## 参 考 文 献

- 1 章少卿, 丁士晟. 预报综合问题的初步探讨. 气象学报, 1960, 31: 110~ 118.
- 2 Stael V, Holstein C A S. An experiment in probabilistic weather forecasting. *J. Appl. Meteor.*, 1971, 10: 635~ 645.
- 3 Tracton M S, et al. Dynamical Extended Range Forecast (DERF) at the National Meteorological Center. Proc. Workshop ECMWF, 16~ 18 May 1988. 153~ 198.
- 4 Hoffman R N, et al. Lagged average forecasting an alternative to Monte Carlo forecasting. *Tellus*, 1983, 35A: 100~ 113.
- 5 周家斌. 关于预报集成问题. 中国科学院大气物理研究所集刊, 第三号. 北京: 科学出版社, 1975. 62~ 71.
- 6 张道民, 纪立人, 李金龙. 月数值天气预报的试验研究. 灾害性气候的模拟和预测. 北京: 气象出版社, 1996. 1~ 11.
- 7 周家斌, 王允宽, 杨桂英. 华北降水分布集成预报试验. 灾害性气候的模拟和预测. 北京: 气象出版社, 1996. 1~ 11.
- 8 陈桂英, 赵振国. 短期气候预测评估方法和业务初估. 应用气象学报, 1998, 9(2): 178~ 185.

## REGIONALIZED WEIGHTED METHOD OF CONSENSUS FORECAST FOR RAINFALL DURING FLOOD SEASON

Zhou Jiabin Zhang Haifu Yang Guiying

(Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029)

Peng Yao

(Institute of Software, Chinese Academy of Sciences, Beijing 100081)

### Abstract

A regionalized weighted method of consensus forecast is proposed which is not confined by the number of sample. The assemble forecast is got based on the score data of different forecast in Previous Periods. 16 kinds of schemes are designed. The results indicate that the method possesses certain accuracy and is able to be used in operational forecast.

**Key words:** Consensus forecast Regionalized weighted method Rainfall during flood season