

模糊优选方法及其在月集合预报中的应用

张永山 李维京 俞 汇

(国家气候中心, 北京, 100081)

俞 裕 国

(南京气象学院, 南京, 210044)

摘 要

根据 1996 年和 1997 年春季的 17 个 T 63L16 月集合预报个例, 利用模糊优选方法, 给出一个新的春季 500 hPa 月预报场的集合方案, 该方案中的权重系数比较客观地反映了预报水平的时空变化, 预报技巧有明显提高。根据春季 500 hPa 高度场误差的时空分布特征, 利用所得权重系数对预报场进行误差订正, 效果更为显著。

关键词: 模糊, 权重, 集合。

1 引 言

王向东、李维京等^[1]用算术平均方案、线性权重方案和抛物型权重方案分别对近几年来 T 63L16 的 500 hPa 高度预报场进行集合对比试验, 发现抛物型权重方案集合技巧最高, 算术平均方案集合技巧最差, 说明集合方法对提高 T 63L16 月动力延伸预报水平有一定的技巧。

根据模式预报中系统误差的分布特征^[2], Deque^[3]、张道民和纪立人等^[4,5]通过消除平均系统误差, Brankovic C. 和 Palmer T. N. 等^[6]在预报方程中附加一强迫项的方法, 使得误差明显减小。可见, 在目前模式本身难有大的改进的基础上, 误差订正也是提高预报水平的一种途径。

模糊优选方法能够揭示某一决策集中各决策因子的优劣程度, 具有较强的理论依据, 已被广泛应用于水文、化工、土建等方面的系统决策。本文用此方法寻求更好的集合方案和误差订正函数, 希望进一步改进 T 63L16 月动力延伸预报的集合技巧。

2 模糊优选方法简介^[7]

对某个集合中各因子的评价, 除了两个极端外, 往往还存在中介过渡因子, 这些中介过渡因子呈现“亦此亦彼”的特性, 具有模糊性。模糊优选方法就是根据某决策函数, 确定各因子的相对优劣、好坏程度。

设集合预报有 m 个因子, 每个因子预报时效为 n 天, 组成目标特征值矩阵

* 初稿时间: 1998 年 10 月 12 日; 修改稿时间: 1999 年 11 月 26 日。

资助课题: 国家“95”攻关项目 96-908-02-01 专题。

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{bmatrix} = (x_{ij}) \quad (1)$$

其中 $i = 1, 2, \dots, m; j = 1, 2, \dots, n$ 。这样的系统称为单元系统。

将目标特征值矩阵变换为目标相对优等程度阵

$$R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ r_{21} & r_{22} & \dots & r_{2n} \\ \dots & \dots & \dots & \dots \\ r_{m1} & r_{m2} & \dots & r_{mn} \end{bmatrix} = (r_{ij}) \quad (2)$$

其中 r_{ij} 称为优属度, 并且 $0 < r_{ij} < 1$, 表示第 j 天第 i 个预报结果的优等程度。它有多种确定方法, 可根据实际情况选出最能体现客观实际的一种表达方式。

定义 m 个集合因子的最大优属度和最小优属度分别为

$$g = (g_1, g_2, \dots, g_m)^T \quad (3)$$

$$b = (b_1, b_2, \dots, b_m)^T \quad (4)$$

设第 j 天预报结果的相对优等程度为 u_j , 它表示第 j 天预报结果在月预报中的权重系数, 与之相对应的相对劣等程度为 u_j^c , 并有

$$u_j^c = 1 - u_j \quad (5)$$

m 个集合因子的权重为

$$w = (w_1, w_2, \dots, w_m)^T, \quad \sum_{i=1}^m w_i = 1 \quad (6)$$

则第 j 天预报结果与最优预报的差异(称为距优距离)

$$d_{jg} = \sqrt[p]{\sum_{i=1}^m [w_i (g_i - r_{ij})]^p} \quad (7)$$

与最劣预报的差异(称为距劣距离)

$$d_{jb} = \sqrt[p]{\sum_{i=1}^m [w_i (r_{ij} - b_i)]^p} \quad (8)$$

p 为距离参数, $p = 1$ 为海明距离, $p = 2$ 为欧氏距离。则第 j 天的加权距优距离为

$$D_{jg} = u_j \quad d_{jg} = u_j \sqrt[p]{\sum_{i=1}^m [w_i (g_i - r_{ij})]^p} \quad (9)$$

第 j 天的加权距劣距离为

$$D_{jb} = u_j^c \quad d_{jb} = (1 - u_j) \sqrt[p]{\sum_{i=1}^m [w_i (r_{ij} - b_i)]^p} \quad (10)$$

为了求解第 j 天预报结果相对优等程度 u_j 的最优值, 建立如下的优化准则: 第 j 天的加权距优距离平方与加权距劣距离平方之和为最小, 即目标函数为

$$\min\{F(u_j) = D_{jg}^2 + D_{jb}^2\} \quad (11)$$

求目标函数的导数, 且令导数为零, 解得

$$u_j = \frac{1}{1 + \left[\frac{d_{jg}}{d_{jb}} \right]^2} \quad (12)$$

由上式可知,当第 j 天的距优距离大于距劣距离,即 $d_{jg} > d_{jb}$ 时, $u_j < 0.5$,表明第 j 天的预报优的程度小于劣的程度, u_j 越小,预报水平越低,特别 $d_{jb} = 0, u_j = 0$ 时,第 j 天的预报为最劣预报;相反,当第 j 天的距优距离小于距劣距离,即 $d_{jg} < d_{jb}$ 时, $u_j > 0.5$,表明第 j 天的预报优的程度大于劣的程度, u_j 越大,预报水平越高,特别 $d_{jg} = 0, u_j = 1$ 时,第 j 天的预报为最优预报。可见,模糊优选方法具有清晰的数学、物理意义,用 u_j 作为第 j 天的预报结果在月预报中权重系数是比较客观的。

本文 r_{ij} 用下式表示

$$r_{ij} = 1 - 2 \frac{x_{ij} - o_j}{\sup(x_{ij}) - \inf(x_{ij})}, (i = 1, 2, \dots, m; j = 1, 2, \dots, n) \quad (13)$$

其中

$$\sup(x_{ij}) = o_j + \sigma_j p_{1-\alpha} \quad (14)$$

$$\inf(x_{ij}) = o_j - \sigma_j p_{1-\alpha} \quad (15)$$

式中 x_{ij} 为预报值, o_j 为第 j 天的观测值(分析值), σ 为均方根误差¹⁾, $p_{1-\alpha}$ 表示均值为 o_j 、均方根误差为 σ 的正态分布序列当信度为 α 时的标准化值。如果 $x_{ij} > \sup(x_{ij})$ (或 $x_{ij} < \inf(x_{ij})$), 则以 $\sup(x_{ij})$ (或 $\inf(x_{ij})$) 替代 x_{ij} 代入式(13)计算。显然, $0 \leq r_{ij} \leq 1$, 且满足预报场越接近客观场, r_{ij} 越接近于 1。

与 x_{ij} 相对应的 m 个集合因子的最大优属度和最小优属度分别为

$$g = (g_1, g_2, \dots, g_m)^T = (1, 1, \dots, 1)^T \quad (16)$$

$$b = (b_1, b_2, \dots, b_m)^T = (0, 0, \dots, 0)^T \quad (17)$$

则式(7), (8) 写成

$$d_{jg} = \frac{1}{\sqrt[m]{\prod_{i=1}^m [w_i(1 - r_{ij})]^p}} \quad (18)$$

$$d_{jb} = \frac{1}{\sqrt[m]{\prod_{i=1}^m [w_i r_{ij}]^p}} \quad (19)$$

3 模糊优选方法在月动力延伸集合预报中的应用

3.1 模糊优选方法对月动力延伸预报集合方案的改进

文献[1]对不同初始场的预报结果分析表明,对于月尺度预报,以时间间隔较短(6 h)的初始场做预报,并不一定初始时间最近的预报结果最好,因此,在预报集成时,一般不考虑初始时间的影响,认为各集合因子同一天的权重相同,即 $w_i (i = 1, 2, \dots, m)$ 的取值相同,只考虑某一天的预报结果在月趋势预报中的权重大小,即 u_j 的大小。

计算相对优等程度 u_j 时采用欧氏距离($p = 2$)^①, 并且取 $\alpha = 10\%$ ^②, 从 1996 年和 1997 年

1) 均方根误差定义为 $\sqrt{\frac{1}{n} \sum_{i=1}^n (x_{ij} - o_{ij})^2}$, x_{ij}, o_{ij} 分别为预报值、分析值, n 为所有个例的因子数。

春季 17 个个例的平均相对优等程度 u_j 的变化曲线(图 1) 可以看出: 随着时间增加, u_j 呈现减小的趋势, 最后一候, 又稍有上升。

不同地区的时间权重系数有所差异。以春季为例, 高纬度地区前 20 d 的权重值一般比平均值稍偏小, 后 10 d 又稍偏大; 中纬度地区前半月的权重值接近平均值, 后半月偏大; 低纬度地区前半月的权重值一般比平均值偏大, 后半月偏小较严重。

用模糊优选法得到的纬向平均权重系数做集合方案(模糊优选方案), 预报水平有明显提高。例如, 对 1996 年和 1997 年春季的 17 个个例统计表明: 前 10 d 距平相关系数基本没有改变, 最后 10 d 的预报水平大大提高了, 欧亚地区提高了近 0.15, 北半球地区提高了 0.05 以上。但是, 中间 10 d 的距平相关系数稍有下降。从月平均水平来看, 欧亚地区的距平相关系数提高了将近 0.12, 17 个个例中有 14 个都提高了, 最大的提高了 0.6 以上; 北半球地区提高了 0.05, 一半以上的个例都提高了, 最大的提高了将近 0.4(表 1)。

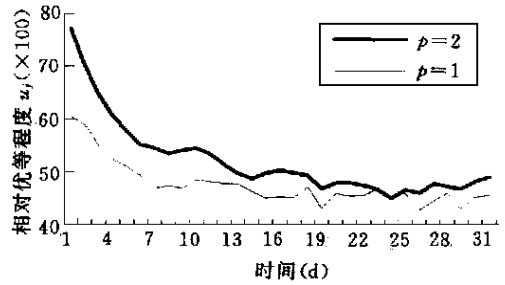


图 1 春季平均相对优等程度 u_j 的时间变化曲线($\alpha = 10\%$)

表 1 1996 年和 1997 年春季预报场 3 种方案的距平相关系数表

初始场时间	欧亚地区			北半球地区		
	算术平均方案	误差修正方案	模糊优选方案	算术平均方案	误差修正方案	模糊优选方案
1996 年 3 月 09 ~ 10 日	0.77	0.91	0.89	0.48	0.46*	0.45
1996 年 3 月 19 ~ 20 日	- 0.13	- 0.01	0.06	0.16	0.16	0.17
1996 年 3 月 30 ~ 31 日	0.38	0.45	0.43	0.33	0.30*	0.29
1996 年 4 月 09 ~ 10 日	0.48	0.65	0.59	0.32	0.23*	0.21
1996 年 4 月 19 ~ 20 日	- 0.33	0.29	0.31	- 0.13	0.26	0.25
1996 年 4 月 29 ~ 30 日	- 0.23	0.14	0.14	0.34	0.45	0.45
1996 年 5 月 09 ~ 10 日	0.55	0.62	0.58	0.46	0.35*	0.31
1996 年 5 月 30 ~ 31 日	0.39	0.18*	0.12	0.53	0.43*	0.41
1997 年 3 月 09 ~ 10 日	0.46	0.66	0.60	0.34	0.51	0.49
1997 年 3 月 19 ~ 20 日	0.64	0.86	0.85	0.44	0.66	0.67
1997 年 3 月 30 ~ 31 日	0.53	0.74	0.72	0.34	0.59	0.59
1997 年 4 月 09 ~ 10 日	0.85	0.95	0.95	0.65	0.63*	0.63
1997 年 4 月 19 ~ 20 日	0.69	0.76	0.78	0.34	0.36	0.36
1997 年 4 月 29 ~ 30 日	0.54	0.52*	0.48	0.25	0.47	0.43
1997 年 5 月 09 ~ 10 日	- 0.53	- 0.51	- 0.46	0.20	0.30	0.30
1997 年 5 月 19 ~ 20 日	- 0.14	- 0.03	- 0.02	0.17	0.24	0.24
1997 年 5 月 30 ~ 31 日	0.19	0.06*	0.09	0.46	0.32*	0.28
平均	0.300	0.426	0.418	0.334	0.395	0.384

* 号表示经修正预报水平没有提高。

- ① 效果优于海明距离($p = 1$)。
- ② 取 $\alpha = 10\%$ 时, 大于 0.5 和小于 0.5 的 u_j 个数几乎各占一半, 比较合理。

根据上述 1996 年和 1997 年春季 17 个个例得到的权重系数, 对 1998 年春季(3 ~ 4 月份)的 6 个个例进行集合, 欧亚地区和北半球地区的平均距平相关系数分别提高了 0.03 和 0.06 以上(表 2)。

表 2 1998 年春季预报场几种方案的距平相关系数表

初始场时间	欧亚地区			北半球地区		
	算术平均方案	误差修正方案	模糊优选方案	算术平均方案	误差修正方案	模糊优选方案
1998 年 3 月 09 ~ 10 日	0.31	0.38	0.43	0.23	0.46	0.49
1998 年 3 月 19 ~ 20 日	0.48	0.42*	0.39	0.26	0.27	0.25
1998 年 3 月 30 ~ 31 日	0.62	0.67	0.64	0.35	0.45	0.42
1998 年 4 月 09 ~ 10 日	0.55	0.48*	0.46	0.55	0.39*	0.36
1998 年 4 月 19 ~ 20 日	0.47	0.66	0.66	0.38	0.54	0.53
1998 年 4 月 29 ~ 30 日	0.54	0.62	0.59	0.25	0.44	0.43
平均	0.495	0.538	0.528	0.337	0.425	0.397

* 号表示经修正预报水平没有提高。

3.2 相对优等程度 u_j 在预报误差订正中的应用

3.2.1 纬向平均误差的分布特征

纬向平均误差定义为 $\frac{1}{m} \sum_{j=1}^m \frac{1}{n} \sum_{i=1}^n (x_{ij} - o_{ij})$, 其中 x_{ij} , o_{ij} , m 和 n 分别为预报值、观测值、纬向格点数和个例数。

根据 1996 年和 1997 年春季 17 个个例的统计结果, 春季纬向平均误差的分布是: 低纬度地区(20°N 以南)和中高纬地区(45° ~ 70°N 之间), 预报值一般低于分析值, 低纬度地区的误差相对稳定; 中低纬地区(20° ~ 45°N 之间), 预报值一般高于分析值。随着时间的增加, 上述现象越明显; 高纬地区(75°N 以北), 前 10 d 和后 10 d, 预报值一般高于分析值, 中间 10 d, 预报值一般低于分析值。总的看来, T63L16 月动力延伸模式, 大多数地区的预报值低于分析值(图 2)。

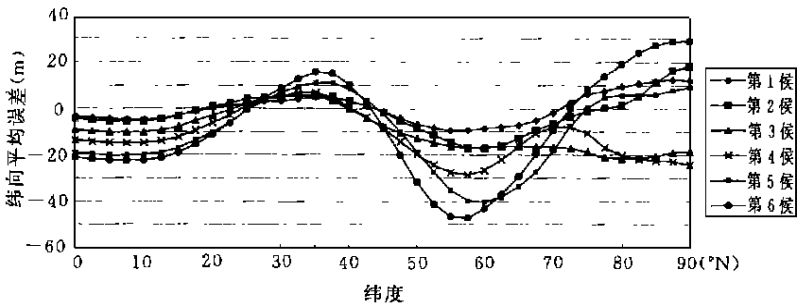


图 2 春季纬向平均误差的时空变化曲线

3.2.2 相对优等程度 u_j 在误差修正中的应用

设误差场 $F(\theta, \lambda, t)$ 可以写成空间函数 $H(\theta, \lambda)$ 和时间函数 $T(t)$ 的乘积

$$F(\theta, \lambda, t) = H(\theta, \lambda) T(t) \quad (20)$$

式中, θ , λ 和 t 分别为纬度、经度和时间。时间函数用劣的相对隶属度 u_j^c 表示。不考虑误差

的经向变化, 根据春季纬向平均误差的分布特征, 春季预报场修正值的拟合函数 $F(\theta, \lambda, t)$ 近似写成

$$F(\theta, \lambda, t) = \begin{cases} 5(1 - u_j) & (\theta < 25 \text{ N}) \\ -10(1 - u_j) \sin[9(\theta - 25)] & (25 \text{ N} \leq \theta < 45 \text{ N}) \\ 30(1 - u_j) \sin[6(\theta - 45)] & (45 \text{ N} \leq \theta < 75 \text{ N}) \\ 15(1 - u_j) & (\theta \geq 75 \text{ N}) \end{cases} \quad (21)$$

经过误差修正, 1996 年和 1997 年春季 17 个个例的欧亚地区的距平相关系数平均提高了 0.12 以上, 其中 14 个个例都提高了, 最大的提高了 0.6 以上; 北半球地区平均提高了 0.06 以上, 且一半以上的个例都提高了, 最大的提高了将近 0.4(与模糊优选方案很相似, 表 1)。

根据上述 17 个个例得到的权重系数, 对 1998 年春季(3 ~ 4 月份) 的 6 个个例进行误差订正, 欧亚地区和北半球地区的平均距平相关系数分别提高了近 0.04 和 0.09(表 2)。

4 环流场对比分析

和算术平均方案相比, 模糊优选方案和误差修正方案不仅从总体上提高了距平相关系数, 对环流场也确实有所改进。以初始时间为 1997 年 4 月 19 日 20 日的第 3 个 10 d 的 500 hPa 预报场为例, 图 3c 和图 3b 相比, 除了大西洋北部的低压槽加深了, 欧亚大槽的位置也东移了, 使得这两个地区的环流形势更接近于客观分析场图 3a。图 3d 和图 3c 的环流形势基本相似, 只是低纬地区位势高度稍有增大, 更接近于客观分析场图 3a, 这说明误差修正方案确实稍优于模糊优选方案。

表 3 列出了几种集合方案对初始时间为 1997 年 4 月 19 日 20 日的 500 hPa 预报场处理后的均方误差, 进一步说明了模糊优选方案和误差修正方案确实比算术平均方案优越。

表 3 几种集合方案得到的 500 hPa 预报场均方误差的比较*

	前 10 d(m)	中 10 d(m)	后 10 d(m)	30 d(m)
算术平均方案	41.4	45.2	61.8	50.3
模糊集合方案	37.4	45.5	60.1	48.6
误差订正方案	39.5	44.1	58.2	47.9

注: 初始时间为 1997 年 4 月 19 ~ 20 日, 统计区域为 0 ~ 85 N, 分辨率为 $2.5^\circ \times 2.5^\circ$ 。

5 结论与讨论

通过上面的分析我们发现:

1) 用模糊优选方法可以得出比较客观地反映预报水平时空变化的量度, 以此作为集合预报的权重系数, 集合效果优于算术平均方案。

* 均方根误差定义为 $\sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - o_i)^2}$, x_i, o_i 分别为预报值、分析值, n 为格点数。

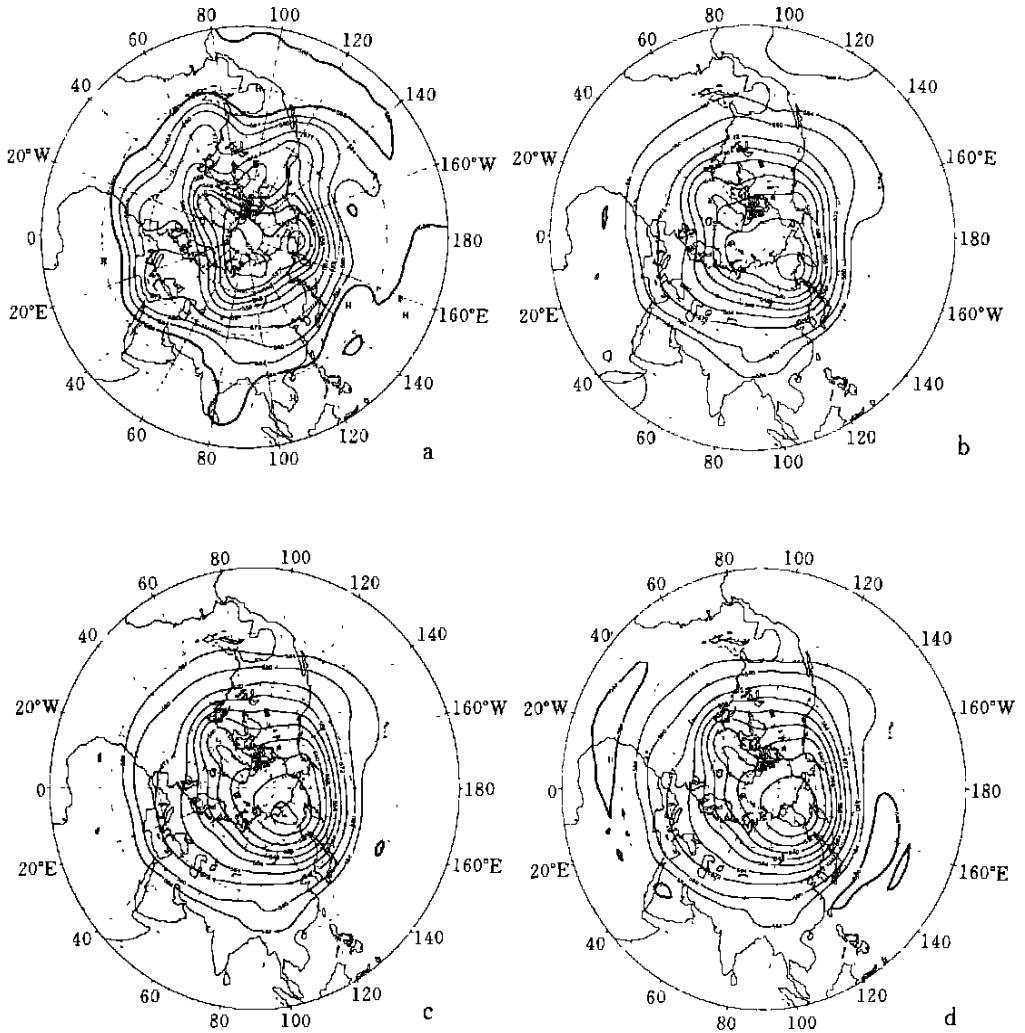


图3 初始时间为1997年4月19日20日的第3个10 d的预报场

(a. 为客观分析场; b. 为用算术平均方法集合得到的环流场;

c. 为用模糊优选方法集合得到的环流场; d. 为在b的基础上经误差订正后得到的环流场)

2) 用模糊优选方法得到的权重系数, 用来拟合误差场, 修正预报场, 效果又稍好于模糊集合方案。

此外, 我们也发现, 用模糊优选方法得到的权重系数, 无论用于集合方案, 还是用于误差修正, 总体预报水平虽然提高了, 但是并没有改变预报的优劣属性, 即好的预报还是好, 差的预报还是差。这除了与优属度 r_{ij} 确定方法有关, 与模式的性能亦有很大关系。

如果优属度 r_{ij} 能更加准确反映预报技巧的时空变化, 用模糊优选方法确定权重系数做集合预报, 并用来修正预报场, 会取得更好效果的。这种方法值得进一步深入研究。

参考文献

- 1 王向东, 李维京等. 1996 年月动力集合预报进展报告. 气候通讯, 1996, 3: 29 ~ 36
- 2 Brankovic C. Systematic errors in deterministic extended-range forecasts: [Ph. D. Thesis]. University of Zagreb, 1989, 61—68
- 3 Deque. Removing the model systematic error in extended-range forecasting. Geophysics, 1991, 9: 242—251
- 4 张道民, 纪立人等. 月数值天气预报试验研究. 大气科学, 1996, 20(4): 429 ~ 438
- 5 张道民, 纪立人等. 动力延伸(月) 预报的季节变化和误差分析. 大气科学, 1997, 21(2): 141 ~ 150
- 6 Brankovic C, Palmer T N, et al. Extended-range predictions with ECMWF models: Time-lagged ensemble forecasting. Quart J Roy Meteor Soc, 1990, 116: 867—912.
- 7 陈守煜. 系统模糊决策理论与应用. 大连: 大连理工大学出版社, 1994. 5 ~ 22

FUZZY SET METHODS AND ITS APPLICATION IN MONTHLY DYNAMICAL EXTENDED- RANGE ENSEMBLE FORECAST

Zhang Yongshan Li Weijing Ding Yihui

(National Climate Center, Beijing, 100081)

Ding Yuguo

(Nanjing Institute of Meteorology, Nanjing, 210044)

Abstract

According to 17 forecast cases of Model T63L16 in spring 1996 and 1997, a group of new weight coefficients are calculated by use of Fuzzy Set Theories and Methods. The new weight coefficients can significantly improve ensemble forecast skill and reflect distributed features of forecast results.

On the basis of error features in spring, the results are better through error correction using the new weight coefficients.

Key words: Fuzzy, Weight, Ensemble.