

区域气候模拟研究中的物理集合技术*

钟 科

(中国科学院大气物理研究所, LASG, 北京, 100029)

王汉杰

(中国科学院东亚区域气候-环境重点实验室, 北京, 100029)

摘 要

文中介绍了基于 MM5 有限区域模式的物理集合构造方法, 通过对模式中的 5 种物理过程加以扰动、组合而得到一个有 19 个成员的物理集合系统, 并用其对 1998 年夏季发生在中国东部地区的异常洪涝天气气候特征进行了模拟研究。以集合预报的 3 种基本释用方法, 对集合模拟产生的大量输出结果作了细致的分析, 发现在区域气候模拟中, 引入物理集合是可行的; 它为区域气候数值研究提供了更多的手段和信息; 集合系统存在的不足之处在于: 集合成员间离散程度不足、集合模拟范围较狭窄, 这使得离散对误差的实际预报能力低于潜在预报能力, 也使集合概率预报的精度降低。

关键词: 区域气候, 集合模拟, 集合平均, 成员离散, 概率预报。

1 引 言

集合预报方法起源于 Epstein^[1] 和 Leith^[2] 的工作, 从 20 世纪 90 年代初期开始, 中、短期数值天气预报中开始广泛应用这种方法。尽管集合预报方法在中、短期数值天气预报中得到了广泛应用, 然而在区域气候数值研究中, 集合预报方法的相关研究还很少。针对这种情况, 本文引入一个区域气候物理集合模拟系统, 对 1998 年夏季发生在中国东部的异常天气气候特点进行模拟, 并用 3 种基本释用方法对模拟结果进行分析。对区域气候模拟中, 如何生成、释用集合模拟结果做一些有益的探索, 同时也对集合系统的模拟误差进行分析, 提出对集合系统的改进意见。

2 集合系统及模拟个例

2.1 物理集合系统的构造

与广泛使用的初值集合不同, 本文选用物理集合, 这主要是考虑到气候积分对初值不敏感, 以及实现的简易性。数值模式采用了 NCAR/PSU 的

MM5V3^[3], 因为它提供了众多可以选择的物理过程, 便于用不同的物理配置来实现集合。

对 5 种物理过程(积云参数化方案、边界层方案、辐射方案、显式水汽方案和陆面过程方案)进行扰动, 每种物理过程有两种扰动态, 对各种扰动态进行组合之后共得到 19 个集合成员, 各个成员的物理方案配置如表 1 所示。在讨论中选用第一个成员作为控制试验。

2.2 集合模拟个例选取

研究对象选用了 1998 年夏季发生在中国东部的连续性强降雨过程^[4]。模拟积分时段为 1998 年 6 月 1 日 20 时至 8 月 31 日 20 时, 模式共积分 92 d。各集合成员的网格距均为 60 km, 东西和南北方向各有 101 个格点。模拟区域为 15~ 60°N, 60~ 160°E, 中心是 38°N, 120°E。

3 集合平均与模拟不确定性研究

用集合平均消除模拟不确定性是集合预报的第一种基本释用方法^[5]。分别用模拟结果对实况值的距平相关系数 (ACC) 和均方根误差 (RMSE), 来

* 初稿时间: 2003 年 3 月 15 日; 修改稿时间: 2003 年 4 月 29 日。

资助课题: 中国科学院创新工程 (ZK CX 2-SW-210)、国家重大基础发展规划 (G1999043400) 及国家自然科学基金项目 (90302015)。

表 1 19 个集合成员的物理方案配置

Table 1 The physical schemes of 19 ensemble members

成员序号	积云参数化	显式水汽方案	辐射方案	边界层方案	陆面过程方案
1	Anthes-Kuo	simple ice	CCM2	Eta	OSU/Eta LSM
2	Anthes-Kuo	simple ice	CCM2	MRF	OSU/Eta LSM
3	Anthes-Kuo	simple ice	云辐射-RRTM 长波	Eta	OSU/Eta LSM
4	Anthes-Kuo	simple ice	云辐射-RRTM 长波	MRF	OSU/Eta LSM
5	Anthes-Kuo	mix phase	CCM2	Eta	OSU/Eta LSM
6	Anthes-Kuo	mix phase	CCM2	MRF	OSU/Eta LSM
7	Anthes-Kuo	mix phase	云辐射-RRTM 长波	Eta	OSU/Eta LSM
8	Anthes-Kuo	mix phase	云辐射-RRTM 长波	MRF	OSU/Eta LSM
9	Betts-Miller	simple ice	CCM2	MRF	OSU/Eta LSM
10	Betts-Miller	simple ice	云辐射-RRTM 长波	Eta	OSU/Eta LSM
11	Betts-Miller	simple ice	云辐射-RRTM 长波	MRF	OSU/Eta LSM
12	Betts-Miller	mix phase	CCM2	Eta	OSU/Eta LSM
13	Betts-Miller	mix phase	CCM2	MRF	OSU/Eta LSM
14	Betts-Miller	mix phase	云辐射-RRT 长波	Eta	OSU/Eta LSM
15	Betts-Miller	mix phase	云辐射-RRTM 长波	MRF	OSU/Eta LSM
16	Anthes-Kuo	simple ice	CCM2	Eta	5层土壤模式
17	Anthes-Kuo	simple ice	云辐射-RRTM 长波	Eta	5层土壤模式
18	Anthes-Kuo	mix phase	CCM2	Eta	5层土壤模式
19	Anthes-Kuo	mix phase	云辐射-RRTM 长波	Eta	5层土壤模式

表征形势场(500 hPa 高度场的 3 个月平均)和要素场(3 个月累计降水)的模拟精度。

在计算了各成员、集合平均形势场与观测值间的 ACC(图 1)后,发现成员间具有一定的离散,并且 ACC 最大相差达 0.1,这表明各成员的模拟结果之间差异较大,存在明显的不确定性,使用集合方法消除不确定性是必要的。

图 1 还显示,集合平均后,得到了大于成员 ACC 平均值的 ACC,这说明集合平均消除了不确定性,使模拟精度得到一定程度的提高。另外,集合平均并不是优于所有成员,但它能使模拟精度得到稳定的提高。对于累计降水量场,计算了 RMSE,得到了类似的结果(图略)。

4 成员间离差对模拟误差的预报

用集合成员间离差对预报误差进行预报,是集合预报的第 2 种基本释用方法^[5]。基本做法是:首先检查在理想状态下,成员间离差是否具有对误差的预报能力(理想状态指的是集合成员的模拟结果能够包含所有的观测值)。理想状态的集合一般称为“完全集合”,把完全集合具有的预报能力称为“潜在预报能力”,在具体实现中随机选择一个成员作为观测值来实现完全集合;其次,如果集合系统有潜在预报能力,那么在实际情形下这种预报能力的表现如何?

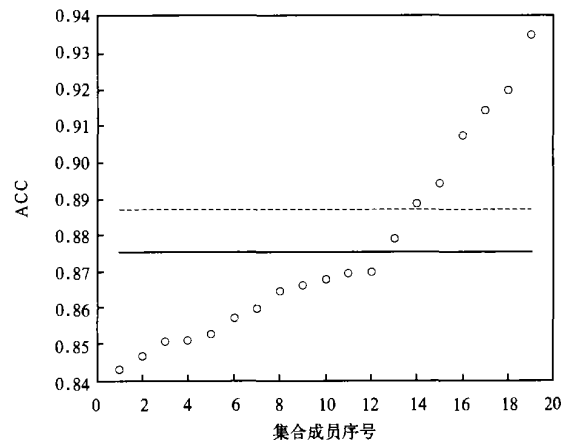


图 1 1998 年 6~8 月 500 hPa 高度场平均(gpm)的 ACC(19 个集合成员)

(圆圈:集合成员 ACC;实线:集合成员 ACC 的平均值;虚线:集合平均 ACC)

Fig. 1 The time mean ACC(19 ensemble members) of 500 hPa height during summer(June, July, and August) in 1998. The X-axis corresponding to ACC, while the Y-axis to the order number of ensemble members. The circle points, slash line, point line denote ACC of ensemble members, the mean of ensemble members ACC, and the ACC of ensemble mean

仍使用 Buizza^[6]的 3 个标准来判别成员间离差对误

差的预报能力, 即:

- (1) 成员间离差在数值上与预报误差可比;
- (2) 小的成员间离差对应小的预报误差;
- (3) 观测值包含在集合成员构成的取值区间内。

上述 3 个标准满足程度越高, 离差对误差的预报能力越高。用模拟结果对实况值的距平相关系数 ACC 来定义形势场(500 hPa 高度场的 3 个月平均)的模拟误差 S_k 和成员间离差 S_p ^[6]。

4.1 完全集合的潜在预报能力

计算完全集合在各个输出时次(每日输出两次, 间隔 12 h)对应时间的 S_k 和 S_p , 发现 S_k 和 S_p 的比值在积分时段内的平均值为 1.006, 两者几乎相等, 这较好地满足标准 1。

图 2 是模拟时段内, S_p 相对于 S_k 的点聚图。图中的竖线是 S_p 的高、低值区域分界线, 其量值是模拟时段内 S_p 平均值, 而横线是 S_k 的高值、低值区域分界线, 其量值是模拟时段内的 S_k 平均值, 横线和纵线对应的 ACC 都为 0.94, 这也较好地满足标准 1; 横线和纵线把平面分为 4 个象限, 如果第 2 个标准得到满足, 则图 2 中的大部分点应该落在第 1, 3 象限。经过统计, 在图 2 中, 落在第 1, 3 象限的点数占总点数的 78.7%。而落在第 2, 4 象限的点数仅占总点数的 21.3%。绝大多数点落在 1, 3 象限, 较好地满足标准 2。

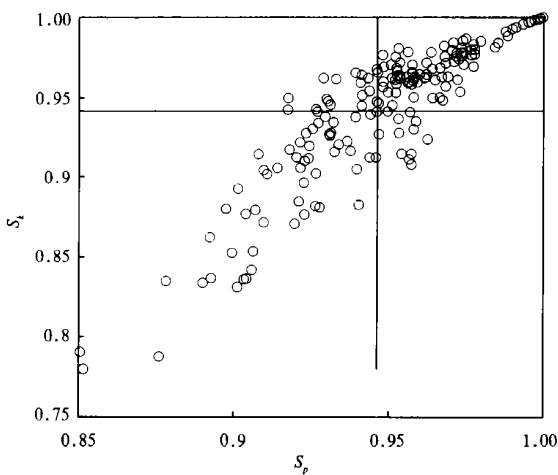


图 2 完全集合的各个输出时次的 S_p 对 S_k 的点聚分布

Fig. 2 the scatter plot of (the X-axis) to (the Y-axis) during all output times of perfect ensemble

由于是完全集合, 观测值被视为成员之一, 第 3 个标准必然是满足的。

总之, 完全集合较好地满足了 3 个标准, 成员间离差具有对控制试验误差的潜在预报能力, 这表明只要集合系统能达到理想状态, 就能够用成员间离差对误差进行预报。

4.2 实际集合的预报能力

将随机选取的成员换成实际观测值后, 就可以讨论集合系统对模拟误差的预报能力。计算实际集合在各个输出时次(每日输出两次, 间隔 12 h)对应时间的 S_k 和 S_p , 在完全集合中, 两者比值在积分时段内平均为 1.006(图 2), 而在实际集合中, 两者比值在积分时段内的平均值为 1.22, 与完全集合相比, 实际集合中离差明显地小于误差, 满足标准 1 的程度不如完全集合高。另外积分时段内 S_k 大于 S_p , 这也表明成员与控制试验相关程度较高, 物理扰动没能有效地使成员离散开来。

图 3 是模拟时段内, S_p 对 S_k 的点聚分布。落在第 1, 3 象限的点数占总点数的 73.3%; 而落在第 2, 4 象限的点数占总数的 26.7%。与完全集合相比, 标准 2 的满足程度略差一些, 但仍能满足误差预测的要求。

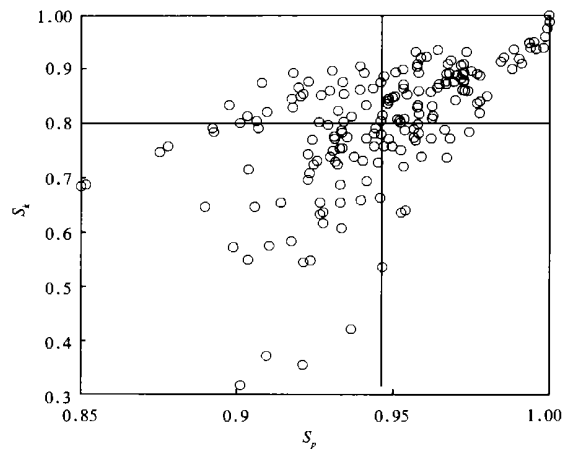


图 3 实际集合的各个输出时次的 S_p 对 S_k 的点聚分布

Fig. 3 the scatter plot of (the X-axis) to (the Y-axis) during all output times of actual ensemble

在积分时段内, 一共有 25 个时次观测值落在集合模拟范围的左边, 28 个时次观测值落在集合模拟范围的右边, 二者之和达到观测总数的 29.0%。与

完全集合相比, 标准 3 的满足程度下降比较明显。

总之, 在实际集合中, 由于无法保证观测值总能包含在集合成员的取值区间内, 直接影响了标准 3 的满足程度, 与完全集合相比, 总有某些观测点落在取值区间之外。另一方面, 与完全集合相比, 实际集合的离差小于误差, 且二者差值较大, 这影响了标准 1 的满足程度, 导致实际预报能力低于潜在预报能力。但从标准 2 来看, 实际集合 1, 3 象限总点数只比完全集合下降 5%, 仍然达到 73.3%, 说明本集合系统仍具有一定的预报能力, 能够预报误差的相对变化趋势。

5 集合概率预报的误差分析

用集合成员构成的概率密度函数制作概率预报, 是集合预报的第三种基本释用方法^[5]。

5.1 连续分级概率评分(C_{RPS})的定义及分解

C_{RPS} ^[7]反映了概率预报与实况间的差距, C_{RPS} 的值越小概率预报的误差越小。此外, C_{RPS} 的空间或者时间平均 $\overline{C_{RPS}}$ 可以进行分解: $\overline{C_{RPS}} = R_{ELI} + U - R_{ESO}$, 其中 R_{ELI} 是可靠性分量, 通过它可以更为细致地分析概率预报误差的生成原因; U 是不确定性分量, 它指示概率事件本身的复杂度, 即正确做出概率预报的难度; R_{ESO} 是分辨率分量, 用于比较概率预报与气候预报的优劣。

记 N 个成员把集合模拟范围分成 $N + 1$ 个小区间: $X_0 = \{x | x < x_1\} \dots, X_i = \{x | x_i \leq x < x_{i+1}\} \dots, X_{N-1} = \{x | x_{N-1} \leq x < x_N\} \dots, X_N = \{x | x \geq x_N\}$, 其中 $x_1, x_2 \dots, x_N$ 表示各个成员。根据 R_{ELI} 的计算公式^[7], 当没有观测值落在集合模拟范围之外, 并且观测值均匀地分布在 $N - 1$ 个小区间 $X_i (i = 1, 2 \dots N - 1)$ 内时, $R_{ELI} = 0$, 而当上述条件不成立时, $R_{ELI} \neq 0$ 。所以, 下列情形都是 R_{ELI} 分量误差的来源: (1) 有观测值落在集合预报范围之外; (2) 观测值在小区间 $X_i (i = 1, 2 \dots N - 1)$ 内分布不均匀; (3) 观测值在区间内均匀分布, 但 $X_i (i = 1, 2 \dots N - 1)$ 内的观测点数与参与平均的总点数之比不等于 $1/(N - 1)$ 。

5.2 C_{RPS}空间平均值的分解结果

记 3 个月累计降水概率预报的 C_{RPS} 的空间平均值为 $\overline{C_{RPS}}$, 计算结果表明 R_{ELI} 占 $\overline{C_{RPS}}$ 的 1/3, 为了分析这部分误差产生的原因, 我们绘制了图 4, 5 和 6。

图 4 是 R_{ELI} 分量曲线, 实线在横坐标方向上的宽度表示 R_{ELI} 的量值。前面指出了产生 R_{ELI} 的几种情形, 这些情形都与观测点在 $X_i (i = 0, 1, 2 \dots N)$ 内的分布有关, 因此每个 $X_i (i = 0, 1, 2 \dots N)$ 都对应一个 R_{ELI_i} 分量, 图中纵坐标表示 $N + 1$ 个小区间 $X_i (i = 0, 1, 2 \dots N)$ 的序号, 横向上各段折线的长度就是与各个小区间 $X_i (i = 0, 1, 2 \dots N)$ 相应的 $R_{ELI_i} (i = 0, 1, 2 \dots N)$, $R_{ELI} = \sum_{i=0}^N R_{ELI_i}$ 。

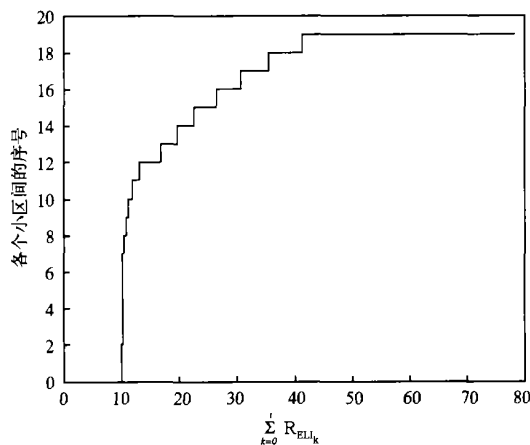


图 4 1998 年 3 个月(6~8 月) 累计降水 $\overline{C_{RPS}}$ 的 $R_{ELI}(\text{mm})$ 分量

$$\left(\sum_{k=0}^N R_{ELIk}, \text{ 中 } i \text{ 表示小区间 } X_i \text{ 的区间序号} \right)$$

Fig.4 The RELI(mm) part of for the accumulating precipitation during summer (June, July, and August) in 1998. The X-axis corresponding to the order number of $X_i (i = 0, 1, 2 \dots N)$, while the Y-axis to, where i denotes the order number of X_i

图 5 是所有参与平均的点对应的各个小区间 $X_i (i = 0, 1, 2 \dots N)$ 的平均长度 g_i ^[7]。横向的各段折线的长度表示 X_i 对应的 g_i , 其相对大小表示集合成员的离散情况; 另外, 考察 g_0 和 g_N 是否为 0, 可以看出集合模拟范围是否足够宽而能够包括所有的观测值。

图 6 是 o_i 和 p_i 曲线。 $o_i (i = 1, 2 \dots N - 1)$ 是小于 $(x_i + x_{i+1})/2$ 的观测点数与参与平均的总点数之比; o_0 和 o_N 分别表示小于等于 x_1 和大于 x_N 的观测点数与总点数的百分比。 p_i 是使 R_{ELI} 分量为 0 的理想 o_i 曲线。

图 4 中, R_{ELI_0} 和 R_{ELI_N} 构成了 R_{ELI} 的主要部分, 两者占 RELI 的 59%, 在图 6 中 o_0 和 o_N 分别为 0.1 和 0.75, 说明有 35% 的观测点超过了集合模拟

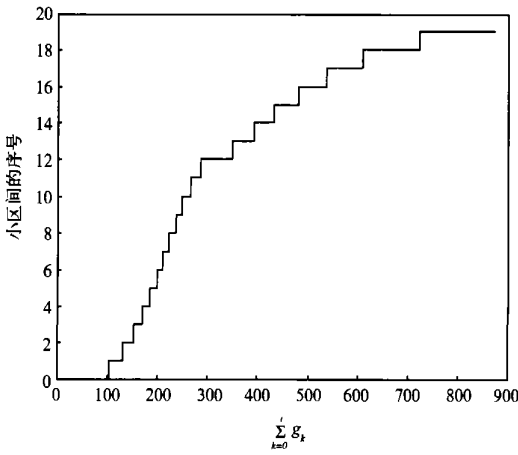


图5 1998年3个月(6~8月)累计降水 $\overline{C_{RPS}}$ 的 g_i (mm)

Fig. 5 the g_i (mm) part of for the accumulating precipitation during summer (June, July, and August) in 1998. The X-axis corresponding to the order number of X_i ($i = 0, 1, 2 \dots N$), while the Y-axis to, where i denotes the order number of X_i

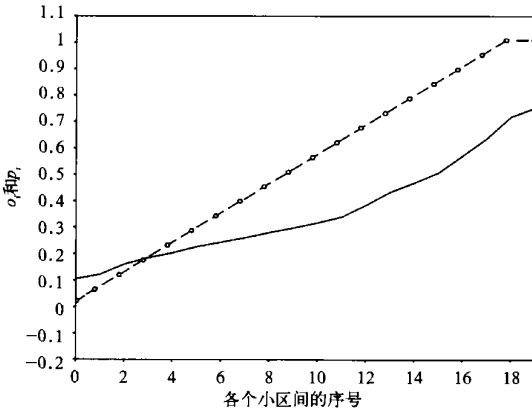


图6 1998年6,7和8月3个月累计降水 $\overline{C_{RPS}}$ 的 o_i (实线) 和 p_i (虚线)

Fig. 6 the o_i (the solid line) and p_i (the slash line) of for the accumulating precipitation during summer (June, July, and August) in 1998. The X-axis corresponding to the order number of X_i ($i = 0, 1, 2 \dots N$)

范围, 这些越界观测点的存在正是 R_{ELI0} 和 R_{ELIN} 的生成原因。

其余的 R_{ELI_i} ($i = 1, 2, 3 \dots N-1$) 之和 $\sum_{i=1}^{N-1} R_{ELI_i}$ 占 R_{ELI} 的 41%, 在图6中, o_i 曲线与 p_i 曲线相比, 大

部分线段的斜率偏小, 在多数小区间 X_i ($i = 1, 2, 3 \dots N-1$) 内, 观测点数都小于 $1/(N-1)$; 在 $i = 15, 16, 17, 18$ 时, o_i 与 p_i 近乎平行, 其中的观测点接近于 $1/(N-1)$, 而其余区间的点数都小于 $1/(N-1)$, 特别是那些处于曲线下端的 X_i 更是如此, 所以观测点在小区间 X_i ($i = 1, 2 \dots N-1$) 内分布不均匀, 并且观测点落在 X_i ($i = 1, 2 \dots N-1$) 的比例并不都等于 $1/(N-1)$, 这是 $\sum_{i=1}^{N-1} R_{ELI_i}$ 的生成原因。

消除 $\sum_{i=1}^{N-1} R_{ELI_i}$ 是困难的, 但只要增大集合模拟范围、减少越界观测点就可以消除 R_{ELI0} 和 R_{ELIN} 。在图5中, g_i 分布具有“下密上疏”的特点, 一半以上成员集中在集合模拟范围的下端, 相比之下, 如果要扰动效应更明显的成员替换原有成员来达到扩大集合模拟范围的目的, 这些集中的量值较小的成员是最合适的被替换对象。另外, 在图6中, 超过集合模拟范围上界与低于下界的点的比例约为 2.5:1, 说明模拟量值偏小区域比偏大区域广阔; 在图5中, $g_N \approx 1.8 g_0$, 这表明降水集合模拟误差中, 模拟值偏小现象比偏大更为严重。这两个误差特征表明, 为了扩大集合模拟范围, 应该优先选取那些模拟量值较大的配置作为替换成员。

6 结论

本文通过对 MM5V3 的 5 种物理过程进行扰动, 得到了一个成员间有一定离散度的物理集合系统。分析表明, 在区域气候模拟研究中引入集合方法是必要的, 集合方法特有的释用方法为区域气候模拟研究提供了更多的手段和信息。集合平均使模拟精度得到稳定的提高; 利用成员间离差对误差的预报能力, 对模拟结果的信度可以预先做出评估; 用集合预报也可以方便地制作概率预报。

集合系统有待改进之处是成员间离散度不够充分和集合模拟范围较狭窄, 这使离差对误差的实际预报能力低于潜在预报能力, 也使集合概率预报的精度降低。进一步的工作是: 用扰动效应更强烈的成员来代替现有成员; 对物理集合而言, 可以从考察参数化过程对模拟结果的扰动效应入手, 在扰动效应明显的参数化过程中选取更多的扰动态。另外, 在计算条件允许的前提下, 增加集合成员个数也是改进模拟精度的有效方法。

参考文献

- 1 Epstein E S. A scoring system for probability forecasts of ranked categories. *J Appl Meteor*, 1969, 8, 985~ 987
- 2 Leith C E. Theoretical skill of Monte Carlo forecasts. *Mon Wea Rev*, 1974, 102: 409~ 418
- 3 Dudhia Jim y, Dave Gill, Kevin, Manning, et al. PSU/NCAR Mesoscale Modeling System Tutorial Class Notes and Users' Guide (MM5 Modeling System Version 3) (for MM5 release- 3- 5), 2002
- 4 孙淑清, 马淑杰. 西太平洋副热带高压异常及其与 1998 年长江流域洪涝过程关系的研究. *气象学报*, 2001, 59(6): 719~ 729
- 5 Wilson L J. Statistical interpretation methods applied to ensemble forecasts uses and interpretation of ensemble output. WMO, Proceedings of the Workshop on Ensemble Prediction, Beijing, 2001
- 6 Buizza R. Potential forecast skill of ensemble prediction and spread and skill distributions of the ECMWF ensemble prediction system. *Mon Wea Rev*, 1997, 125: 99~ 119
- 7 Hensbach H. Decomposition of the continuous ranked probability score for ensemble prediction systems. *Wea Forecasting*, 2000, 15: 559~ 570

THE PHYSICAL ENSEMBLE TECHNIQUE OF THE REGIONAL CLIMATE SIMULATION

Zhong Ke

(*LASG, Institute of Atmospheric Physics, CAS, Beijing 100029*)

Wang Hanjie

(*Key Laboratory of Regional Climate- Environment Research for Temperate East Asia,
Chinese Academy of Science, Beijing 100029*)

Abstract

The five physical processes used in the limited area model MM5 are disturbed to form a physical ensemble system of 19 members. The abnormal synoptic situation and heavy rain occurred in summer of 1998 in Eastern China was simulated with this physical ensemble system. The outputs from the ensemble simulation are analyzed using the three basic interpretational methods commonly used in the ensemble forecast; the consequence is that it is practicable to introduce ensemble technique in regional climate study. The ensemble technique appends new methods and additional information to regional climate simulation as well. However, there are still some flaws in the ensemble system, the ensemble members are not diverse enough and the range of ensemble simulation is rather narrow, which not only lowers the accuracy of probability forecast of deviation but also leads to the actual forecast capability being inferior to the potential one.

Key words: Regional climate, Ensemble simulation, Ensemble mean, Ensemble spread, Probability forecast.