

# **BCC 二代季节预测模式历史回报性能 综合检验评估报告**

国家气候中心 气候研究开放实验室

完成人：任宏利、吴捷、刘颖、左金清、陆波、田奔、赵崇博

2015 年 10 月 20 日

## 摘 要

按照《气候预测模式的检验评估方案》，针对 BCC 二代季节预报模式，从确定性预报和概率预报两个方面，利用距平相关系数(ACC)、时间相关系数(TCC)、相对操作特征(ROC)和 Brier 技巧评分(BSS)等评估方法，对 2m 温度、降水量、大气环流的模式历史回报结果进行了综合检验评估。此外，还进一步对全球和我国气候具有高影响力的重要气候现象(如 ENSO、IOD、AO、北大西洋三级子、西太副高和东亚冬、夏季风等)代表性指标的预报技巧进行评估，从而全面考察模式预测性能，探讨可预报性来源，并为预报会商和模式改进提供参考。主要得到以下几点结论：

1) BCC 二代模式对全球 2m 温度、降水、环流的预报性能均明显优于 BCC 一代模式，特别是热带地区降水预报有明显改进，在技巧的空间分布上，TCC/MSSS/ROCA/BSS 等检验方法的评估结果在很大程度上具有一致性，即模式对热带地区的预报技巧更高；从不同季节上看，模式对冬季和春季的预报技巧较高，夏季和秋季偏低。

2) 与一代模式相比，二代模式在对于低纬度地区的预报改进最为显著，热带太平洋 NINO 区预报技巧改进明显，模式在热带外地区预报技巧及其改进也基本与热带强迫源(如 ENSO)所激发的遥相关(如 PNA 型)有关；在 ENSO 事件发生的当年秋冬和次年春夏季，模式整体的预报技巧明显高于其他季节。

3) 从概率预报上看，模式通常对负异常事件的预报技巧略高于正异常事件，二者都明显高于接近正常事件；从可靠性上看，模式普遍存在低概率低估和高概率高估的特点；BSS 的各项分析表明，随着预报时间的延长，模式概率预报技巧的下降主要来源于模式分辨能力的下降。

4) 对东亚大陆地区而言，模式的预报能力较弱，其中对各个季节平均降水的 TCC 技巧较低，冬季略好；但是二代模式对夏季 EAP 型等遥相关型有一定的把握能力，因此对夏季东亚地区环流(如西太副高)预报技巧较高；在冬季，模式对贝湖到鄂海的中高纬度环流预报技巧偏差，导致模式对冬季我国东北气温预报技巧偏低。

5) 从全球气候现象预报来看，二代模式对 ENSO (NINO3.4 指数)提前 6 个月预报的相关技巧接近 0.7，与 CFSv2 基本相当，但比国际上其他先进模式略偏低。以 0.5 相关为标准，对热带印度洋海温偶极子 (IOD) 的预报技巧约为 2-3 个月，略低于国际先进水平 (3-4 个月)；对北大西洋海温三极子指数 (NASTI) 的预报技巧约为 3 个月，其中春季会商时模式预报的夏季三级子指数与实况相关系数达到 0.50；模式只对 12 月起报的北极涛动 (AO) 指数存在一定预报技巧 (约 3 个月，以 0.2 相关为标准)，其它起报月份 AO 指数几乎无明显预报技巧；模式对东亚夏季风有一定预报技巧，对西太副高指数和东亚夏季风指数的预报技巧可达 6-7 个月，但对东亚冬季风预报技巧较低，约为 3 个月左右。

6) 从对中国 160 站温度和降水检验来看，二代模式对春季我国很多地区有正的预报技巧，特别在长江中下游及其以南部分地区能够通过显著性检验，对夏季气温和降水预报技巧总体偏低，秋季对气温技巧比夏季要好些，相比之下，冬季气温技巧整体偏低、但对南方降水有一定预报效果。总体上，二代模式与一代模式相比，对我国气候预报性能提升明显。

# 目 录

1 引言 .....	5
2 资料和方法.....	5
2.1 资料和预处理.....	5
2.2 确定性预报评估方法.....	5
2.2.1 距平相关系数 (Anomaly Correlation Coefficient, ACC) .....	6
2.2.2 时间相关系数 (Temporal Correlation Coefficient, TCC) .....	6
2.2.3 平均方差技巧评分 (Mean Square Skill Score, MSSS) .....	7
2.3 概率预报评估方法.....	7
2.3.1 相对操作特征 (Relative Operating Characteristics, ROC) .....	7
2.3.2 可靠性图表 (Reliability Diagrams, RD) .....	8
2.3.3 Brier 技巧评分 (Brier Skill Score, BSS) .....	8
2.4 评估区域.....	9
3 基本要素和环流要素评估结果.....	9
3.1 2m 温度 (t2m) .....	9
3.1.1 距平相关系数 (ACC) .....	9
3.1.2 时间相关系数 (TCC) .....	10
3.1.3 平均方差技巧评分 (MSSS) .....	12
3.1.4 相对操作特征 (ROC) .....	14
3.1.5 可靠性图表.....	15
3.1.6 Brier 技巧评分 .....	16
3.1.7 二代模式对 2m 温度预报小结.....	18
3.2 降水 (PREC) .....	18
3.2.1 距平相关系数 (ACC) .....	18
3.2.2 时间相关系数 (TCC) .....	18
3.2.3 平均方差技巧评分 (MSSS) .....	21
3.2.4 相对操作特征 (ROC) .....	22
3.2.5 可靠性图表.....	24
3.2.6 Brier 技巧评分 .....	24
3.2.7 二代模式对降水预报小结.....	26
3.3 500hPa 位势高度场 (Z500) .....	26
3.3.1 距平相关系数 (ACC) .....	26
3.3.2 时间相关系数 (TCC) .....	27
3.3.3 平均方差技巧评分 (MSSS) .....	29
3.3.4 相对操作特征 (ROC) .....	31
3.3.5 可靠性图表.....	32
3.3.6 Brier 技巧评分 .....	33
3.3.7 二代模式对 500hPa 位势高度场预报小结.....	35
4 全球气候现象评估.....	35
4.1 厄尔尼诺和南方涛动 (ENSO) .....	35
4.2 印度洋海温异常主模态 (IOD 等) .....	37
4.3 北大西洋海表温度三级子 (NAST) .....	38
4.4 北极涛动 (AO) .....	40

4.5 夏季西太副高和东亚冬、夏季风.....	41
5 我国台站气温和降水预报评估.....	45
5.1 气温.....	41
5.2 降水.....	41
6 总结 .....	45
参考文献.....	47

# BCC 二代季节模式预测性能综合检验评估

## 1 引言

BCC 二代季节预报模式作为国家气候中心新一代短期气候预测模式,已完成 1991 年起全部历史回算工作,并于 2014 年投入业务试运行。中国气象局预报与网络司和国家气候中心业务科技处委托开放室对 BCC 二代季节模式历史回报进行综合检验评估,形成评估报告,以此作为 BCC 二代季节模式业务化的重要依据。为全面检验模式预报性能,特别是为 3 月汛期会商和 10 月冬季会商提供参考,开放室参照 WMO 国际推荐标准和中国气象局标准,对二代季节模式的常规要素(气温、环流、降水)和重要气候现象(ENSO、IOD、AO 等)的预报技巧进行检验评估,探讨可预报性的来源,为模式进一步改进和发展提供思路。

报告的第 2 部分介绍了评估所用的观测资料和检验方法,第 3 部分介绍全球基本要素和环流要素的评估结果,第 4 部分介绍该模式对重要气候现象的预测技巧,第 5 部分对比 BCC 一代和二代模式对我国 160 台站的气温和降水预报性能,第 6 部分进行总结。

## 2 资料和方法

### 2.1 资料和预处理

BCC 二代季节模式的的回报和预报资料从 1991 年保留至今,考虑模式和观测资料的完整性,选取 1991-2013 共 23 年的资料进行评估。模式从每个月 1 号起报,共有 24 个集合成员,输出未来 13 个月的月平均等压面环流要素和地表常规要素,分辨率为 $1^{\circ} \times 1^{\circ}$ 。在进行确定性预报技巧评估时,对 24 个成员进行集合平均。考虑到模式的系统误差,取 1991-2010 共 20 年的平均作为模式自身的气候态,它既是时间的函数,也是预报 lead time 的函数。

本文侧重于对不同 lead time 的四个季节的预报性能进行评估,如对夏季而言(JJA),对 6 月 1 日预报结果的前 3 个月进行平均即得到 lead time 为 0 个月的预报结果,对 5 月 1 日的第 2-4 个月的结果进行平均即得到 lead time 为 1 个月的预报,依次类推得到每个季节的预报。由于通常在每个月的中旬才能得到该月的预报结果,因此考虑预报时效主要对 lead time 为 1 个月结果进行评估。按照 WMO 标准,模式数据均插值到 $2.5^{\circ} \times 2.5^{\circ}$ 的网格上。

对全球,检验主要对 2m 气温(t2m)、降水(PREC)和 500hPa 位势高度(Z500)3 个要素进行评估。其中,2m 气温和位势高度的观测资料均采用 NCEP/DOE 的第二套再分析资料(<http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis2.pressure.html>);降水观测数据使用 CMAP 提供的逐月降水资料(<http://www.esrl.noaa.gov/psd/data/gridded/data.gpcp.html>),分辨率均为 $2.5^{\circ} \times 2.5^{\circ}$ 。中国台站资料使用的是 160 个标准站逐月 2m 气温和降水资料。

此外,本文还对 BCC 一代季节模式的结果(图中记为 BCC1)进行了类似评估,与二代模式结果(图中记为 BCC2)进行对比分析。其中,一代模式的回报和预报资料从 1983 年至今。为保证评估结果的一致性,在模式资料预处理时均取 1991-2010 共 20 年的平均作为模式自身的气候态,选取 1991-2013 共 23 年的资料进行评估。

### 2.2 确定性预报评估方法

下面介绍本文所用检验评估方法,为了表述严谨,首先对所用符号及其含义进行统一,

令  $x_{i,j}$  代表观测值,  $f_{i,j}$  代表预测值, 其中  $i = 1, 2, 3, \dots, M$  代表评价区域的格点数,

$j = 1, 2, 3, \dots, N$  代表时间序列, 在进行区域平均时, 需要乘以系数  $w_i$ , 使用站点资料

时,  $w_i = 1$ , 使用格点资料时,  $w_i = \cos(\varphi_i)$ , 其中  $\varphi_i$  为格点所在纬度。

### 2.2.1 距平相关系数 (Anomaly Correlation Coefficient, ACC)

距平相关系数, 即 ACC, 主要反映的是预报值与实况值空间型的相似程度, 也可称为空间相似系数, 每次预报均可对预报场计算空间相似系数, 是 WMO 于 1996 年确定并建议使用的指标 (WMO, 1996), 需要首先计算各个格点的时间平均和异常:

$$\overline{x_i} = \frac{1}{N} \sum_{j=1}^N x_{i,j}, \quad \overline{f_i} = \frac{1}{N} \sum_{j=1}^N f_{i,j}$$

$$\Delta x_{i,j} = x_{i,j} - \overline{x_i}, \quad \Delta f_{i,j} = f_{i,j} - \overline{f_i}$$

计算 ACC 的公式如下

$$ACC_j = \frac{\sum_{i=1}^M (\Delta x_{i,j} - \overline{\Delta x_j}) \times (\Delta f_{i,j} - \overline{\Delta f_j})}{\sqrt{\sum_{i=1}^M (\Delta x_{i,j} - \overline{\Delta x_j})^2 \times \sum_{i=1}^M (\Delta f_{i,j} - \overline{\Delta f_j})^2}}$$

式中,  $N$  为评分总格点数或站点数, 其中  $\overline{\Delta x_j}$  和  $\overline{\Delta f_j}$  分别为观测和预测所有格点时间距平的空间平均值。

### 2.2.2 时间相关系数 (Temporal Correlation Coefficient, TCC)

时间相关系数能够在统计意义上较好地表征模式对各个格点异常的预报能力, 得到一个完整的相关技巧空间分布。计算 TCC 时需要求出每个格点的均方差和协方差, 公式如下

$$Sx_i^2 = \frac{1}{N} \sum_{j=1}^N (x_{i,j} - \overline{x_i})^2, \quad Sf_i^2 = \frac{1}{N} \sum_{j=1}^N (f_{i,j} - \overline{f_i})^2$$

$$Sxf_i = \frac{1}{N} \sum_{j=1}^N (x_{i,j} - \overline{x_i})(f_{i,j} - \overline{f_i})$$

$$TCC_i = \frac{Sxf_i}{Sx_i \times Sf_i} = \frac{\sum_{j=1}^N (x_{i,j} - \overline{x_i})(f_{i,j} - \overline{f_i})}{\sqrt{\sum_{j=1}^N (x_{i,j} - \overline{x_i})^2} \times \sqrt{\sum_{j=1}^N (f_{i,j} - \overline{f_i})^2}}$$

进行区域平均时, 同样需要考虑不同纬度的影响

$$TCC = \frac{\sum_{i=1}^M w_i TCC_i}{\sum_{i=1}^M w_i}$$

TCC 范围在-1 到 1 之间，越接近于 1 表明技巧越高，通常取 0.5 的相关技巧作为有预报意义的标准。

### 2.2.3 平均方差技巧评分（Mean Square Skill Score, MSSS）

2006 年，WMO 提出了标准评估系统，推荐模式的气候预测采用 MSSS 法进行评估（WMO, 2006）。MSSS 评分法主要是用于不分类的确定性预报检验和评估（WMO, 2009）。首先计算预测均方误差为：

$$MSE_i = \frac{1}{N} \sum_{j=1}^N (f_{i,j} - x_{i,j})^2$$

“气候学”预测的均方误差差为：

$$MSE_{c_i} = (N / (N - 1))^2 S_{x_i}^2$$

得到 i 点处均方差技巧评分为：

$$MSSS_i = 1 - MSE_i / MSE_{c_i}$$

同样的，进行区域平均时，需要考虑不同纬度的影响

$$MSSS = \frac{\sum_{i=1}^M w_i MSSS_i}{\sum_{i=1}^M w_i}$$

由计算公式可知，在理想预报情况下，MSSS=1，MSSS 评分越高，预报技巧越高，通常取 MSSS=0 作为有预报意义的标准

## 2.3 概率预报评估方法

### 2.3.1 相对操作特征（Relative Operating Characteristics, ROC）

BCC 二代季节预报模式的每次预报均有 24 个预报成员，因此虽然模式没有直接输出概率预报结果，但仍可通过各个集合成员构造 ROC 列联表（Contingency table）对概率预报技巧进行评估（表 1）。

表 1 集合成员构成的列联表

Bin number	Member distribution	Observed occurrences	Observed non-occurrences
1	F=0, NF=N	O <sub>1</sub>	NO <sub>1</sub>
2	F=1, NF=N-1	O <sub>2</sub>	NO <sub>2</sub>
3	F=2, NF=N-2	O <sub>3</sub>	NO <sub>3</sub>
...	...	...	...
n	F=n-1, NF=N-n+1	O <sub>n</sub>	NO <sub>n</sub>

...	...	...	...
N+1	F=N, NF=0	O <sub>N+1</sub>	NO <sub>N+1</sub>

在表 1 中, N 为集成员数, F 为预报事件发生的成员数, NF 为预报事件不发生的成员数。对区域进行累加时,

$$O_n = \sum W_i(O)_i, NO_n = \sum W_i(NO)_i$$

由此可进一步得到命中率 (Hit Rate, HR) 和空报率 (False Alarm Rate, FAR):

$$HR_n = \sum_{i=n}^N O_i / \sum_{i=1}^N O_i$$

$$FAR_n = \sum_{i=n}^N NO_i / \sum_{i=1}^N NO_i$$

以误报率作为横坐标、命中率作为纵坐标即可得到 ROC 曲线, 由上述两式可知 ROC 曲线通过 (0,0) 和 (1,1) 两点, ROC 曲线下的面积 (ROCA) 常用作代表预报技巧的统计指数, 其值介于 0-1 之间, 对角线 (ROCA=0.5) 代表无预测技巧。ROC 曲线越向上凸起, 即 ROC 面积越大代表预报技巧越高, 将各个变量按照气候百分位分为高值、常值和低值三类事件进行检验。

### 2.3.2 可靠性图表 (Reliability Diagrams, RD)

作为 ROC 的重要补充, WMO 同样推荐构造可靠性图表来评估概率预报技巧。众所周知, ROC 曲线更适用于大样本长序列的预报评估, 但 RD 可以很好弥补 ROC 曲线在预报可靠性评价上的缺陷。

可靠性图表的横坐标为预报概率, 纵坐标为命中率 HR, 但与 ROC 曲线不同, 此处 HR 的定义为:

$$HR_n = O_n / (O_n + NO_n)$$

因此, RD 曲线越接近对角线, 预报效果越好, 模式可靠性也越高。若曲线在对角线之下代表过高估计, 即预报概率太高; 而曲线在对角线之上则代表过低估计, 即预报概率太低。

### 2.3.3 Brier 技巧评分 (Brier Skill Score, BSS)

Brier 技巧评分 (BSS) 的定义基于 Brier 评分 (BS) 之上, 而 BS 也是一种定量估计概率预报准确率的方法, 它的定义为:

$$BS = \frac{1}{n} \sum_{i=1}^n (f_i - o_i)^2$$

其中 n 为预报次数,  $f_i$  为第 i 次预报的事件发生的概率,  $o_i$  为第 i 次观测到的概率, 若事件发生则  $o_i$  为 1, 否则  $o_i$  为 0。BS 的范围在 0-1 之间, 值越小表明模式技巧越高。但 BS 的缺陷在于它对事件的气候概率非常敏感。进一步 BS 评分可以分解为不确定性 (uncertainty)、可靠性 (reliability) 和分辨能力 (resolution) 三项, 即:

$$\begin{aligned} BS &= \bar{o}(1 - \bar{o}) + \frac{1}{n} \sum_{k=1}^m n_k (f_k - \bar{o}_k)^2 - \frac{1}{n} \sum_{k=1}^m n_k (\bar{o}_k - \bar{o})^2 \\ &= BS_{unc} + BS_{rel} - BS_{res} \end{aligned}$$



其中  $\bar{o}$  代表事件发生的气候概率,  $m = N + 1$  代表概率的范围,  $f_k$  代表预报概率,  $\bar{o}_k$  代表预报概率为  $f_k$  时事件实际发生的概率, 与可靠性图表中的  $HR_n$  一致。气候预报的 BS 为  $BS_{cli} = BS_{unc}$ , 因此 Brier 技巧评分的定义为

$$BSS = 1 - \frac{BS}{BS_{cli}} = \frac{BS_{res} - BS_{rel}}{BS_{unc}}$$

因此 BSS 的取值范围为负无穷到 1, BSS 越高代表概率预报技巧越高, 气候预报的 BSS 为 0 分。因此 BSS 可以定量估计模式概率预报相对于气候预报的技巧的提高, 同时能通过各个项的分解定量给出模式可靠性和分辨能力的贡献, 但它的局限性在于不适用于少量数据样本和罕见事件的评估。

## 2.4 评估区域

综合 WMO 推荐和我国预报员重点关注的地区, 对以下 6 个区域的平均值也进行了计算, 区域范围如表 2 所示

表 2 评估所用区域范围和简称

区域	范围	备注
全球 (Global)	0°-360°E, 90°S-90°N	
热带 (Tropics)	0°-360°E, 20°S-20°N	WMO 推荐
北半球热带外 (NET)	0°-360°E, 20°N-90°N	WMO 推荐
南半球热带外 (SET)	0°-360°E, 20°S-90°S	WMO 推荐
东亚 (EA)	90°-150°E, 20°N-50°N	
南亚 (SA)	60°-130°E, 10°S-30°N	

## 3 全球基本要素和环流要素评估结果

### 3.1 2m 温度 (t2m)

#### 3.1.1 距平相关系数 (ACC)

距平相关系数 ACC 是短期气候预测中最常用的检验方法之一, 反映的是模式的总体预测性能。图 1 给出了 BCC 一代和二代季节模式提前一个月起报的全球 2m 温度的逐年 ACC 技巧序列。由图可知, 二代模式回报的气温各年的 ACC 评分都明显高于一代模式, 其中冬季多年平均值为 0.27, 明显高于一代模式的 0.04, 其他各个季节二代模式的 ACC 的多年平均也都超过 0.1; 从 ACC 的年际变化上看, 二代模式 ACC 在 ENSO 年评分普遍较高, 如 1998 年的冬季和春季都超过 0.4。

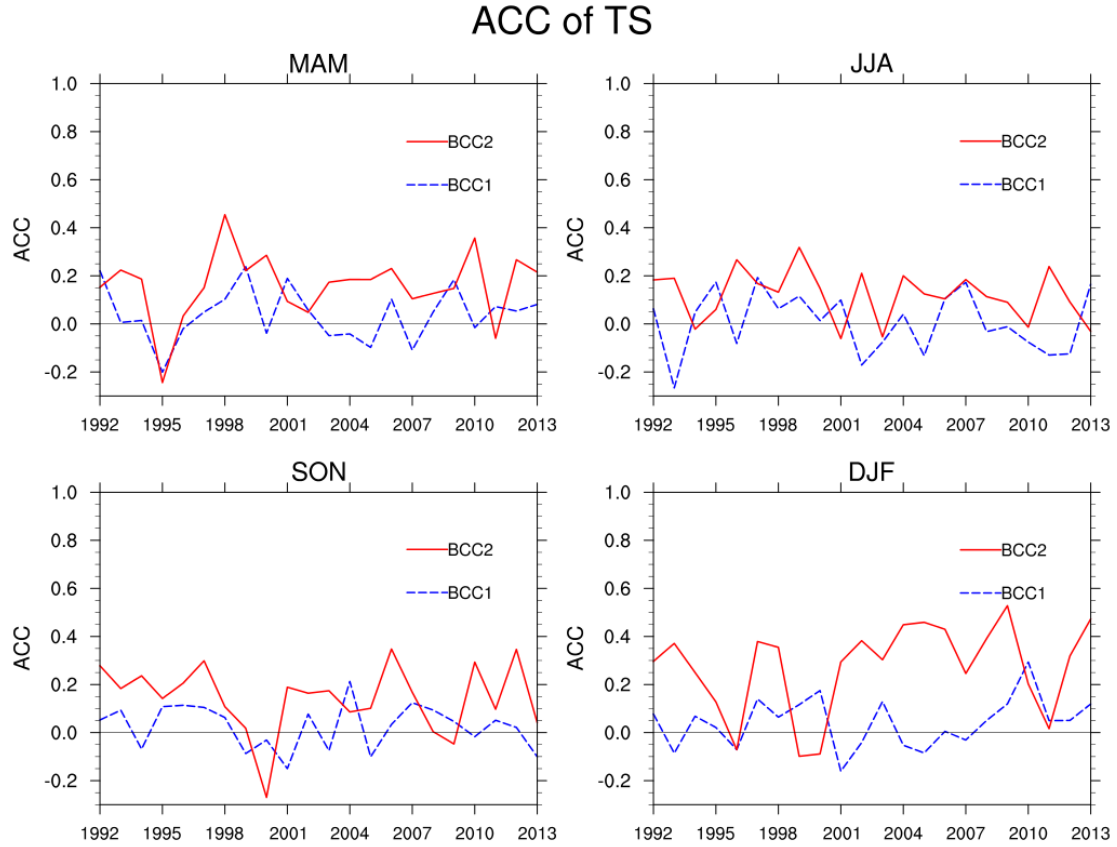


图 1 提前 1 个月 BCC 一代和二代季节模式对全球 2m 温度预报 ACC 技巧的四个季节时间序列

### 3.1.2 时间相关系数 (TCC)

考虑到预报时效的问题，这里主要 BCC 二代季节模式提前 1 个月对四个季节 2m 温度的预报技巧进行检验，各个格点的时间相关系数 (TCC) 技巧如图 2 所示，等值线分别代表通过显著性为 0.05 和 0.01 统计检验的区域。图 3 给出了 BCC 二代模式与一代模式 TCC 评分的差值分布。不同区域平均的 TCC 及其随 lead time 的变化则由图 4 给出。如图 2 所示，2m 温度预报技巧较高的区域主要位于热带海洋地区，特别是热带太平洋地区。这一方面与海表温度自身持续性较好有关，另一方面得益于热带海洋存在 ENSO 等具有较高可预报性的气候模态。在热带外地区，具有较高预报技巧的地区则类似于马蹄形分布，从热带海洋大陆地区分别向东北太平洋和东南太平洋延伸，这主要受到 ENSO 遥相关（例如 PNA 等）的影响。与一代模式相比（见图 4），二代模式 2m 温度的预报技巧提高明显，特别是冬季在赤道中东太平洋的 ENSO 区域，由此在 ENSO 遥相关影响地区内的预报技巧也有所提高。

总体而言，二代模式对冬季和春季 2m 温度的预报技巧较高，夏季技巧偏低，随着预报时间的延长，预报技巧下降较快，特别是前 2 个月，说明初值的重要影响。此外，模式对东亚地区的预报技巧偏低，甚至低于北半球平均水平，这可能与东亚季风区的复杂性有关。

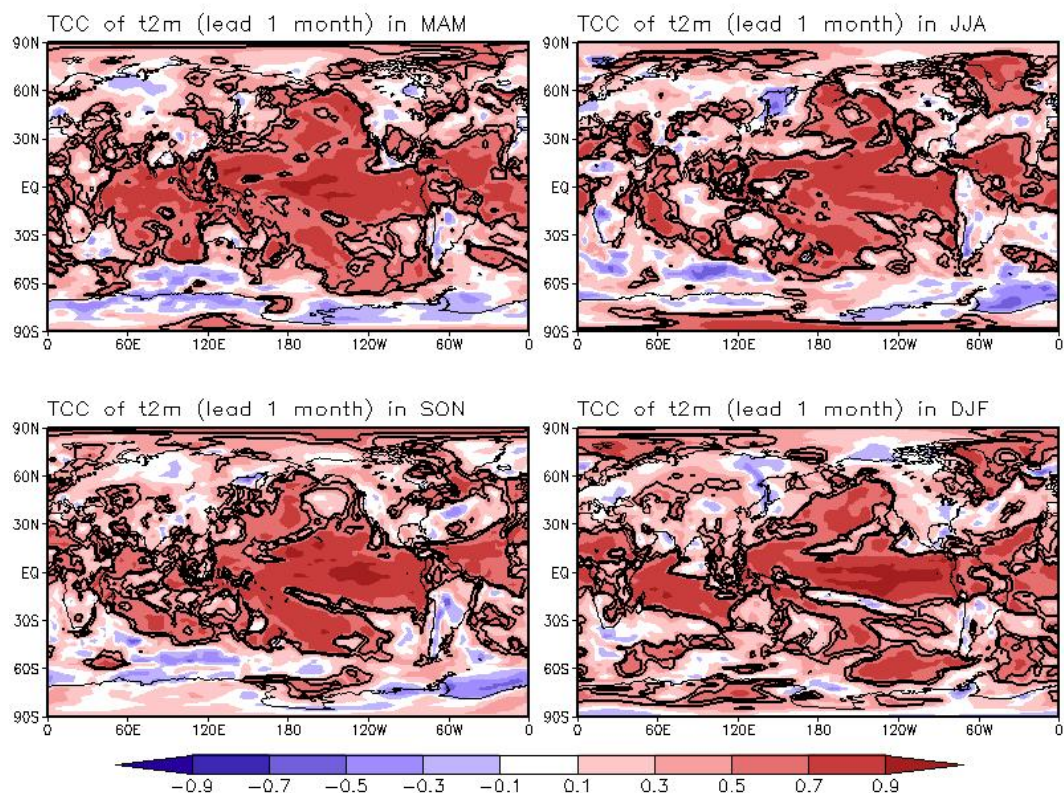


图2 提前 1 个月 BCC 二代季节模式对 2m 温度预报的时间相关系数(TCC)技巧的四个季节的空间分布图，等值线分别代表通过显著性为 0.05 和 0.01 检验的区域

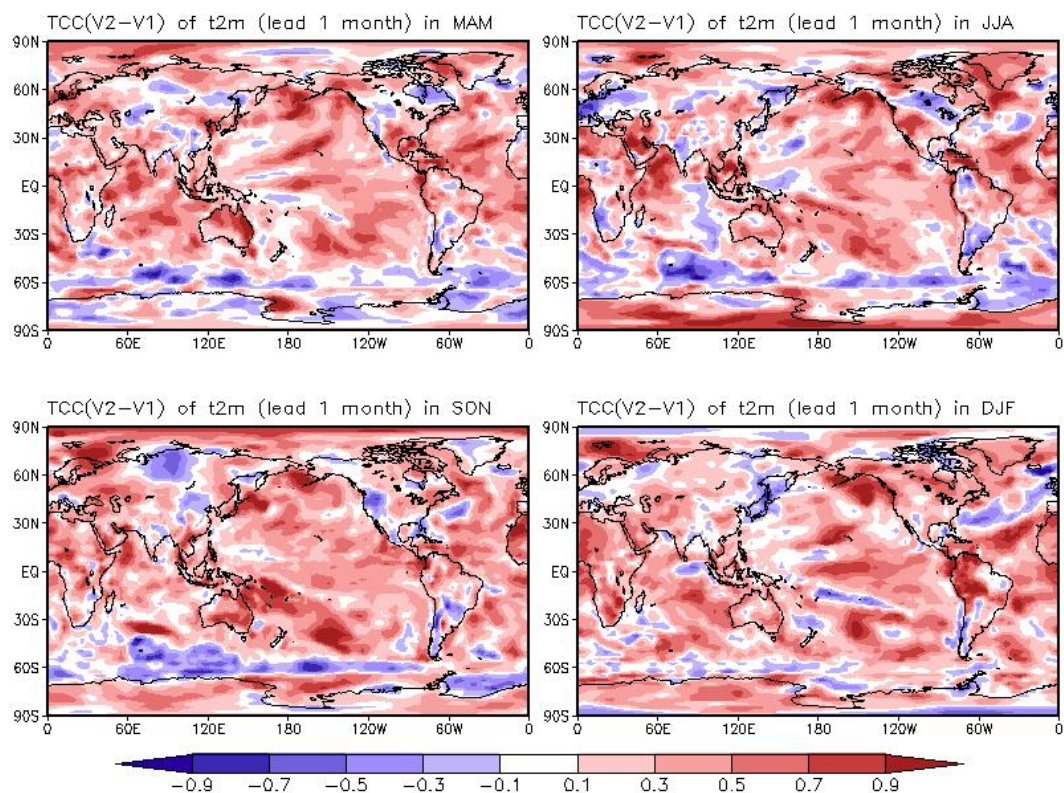


图3 同图2，但为二代模式与一代模式 TCC 技巧的差值图



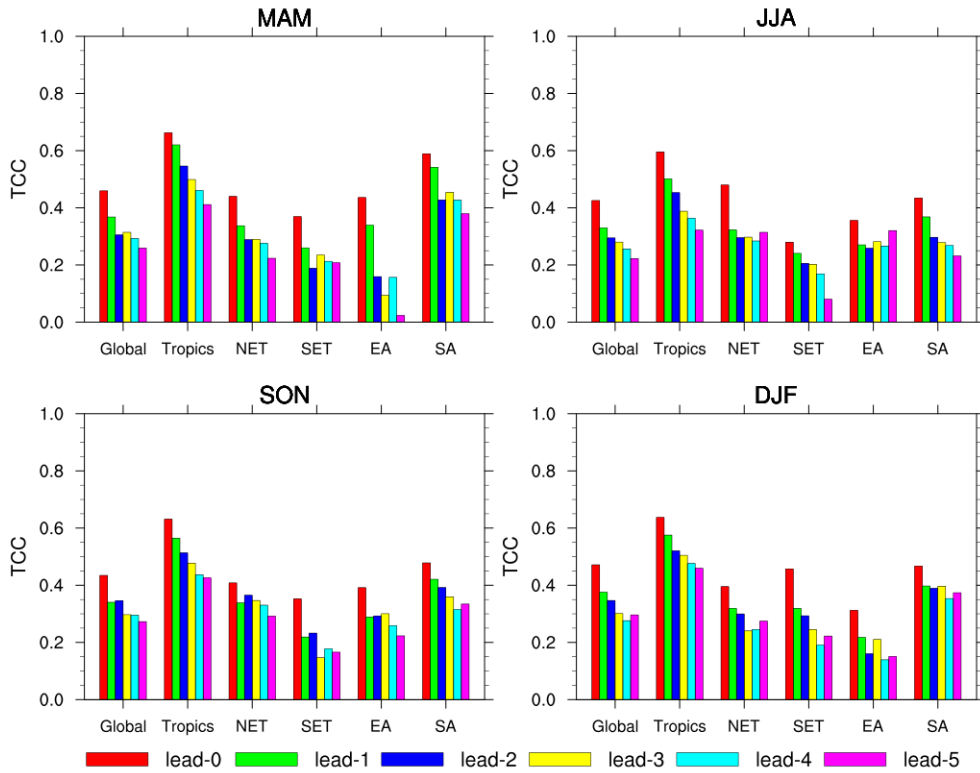


图 4 BCC 二代季节模式在不同 lead time 起报的四个季节 2m 温度的时间距平相关系数（TCC）技巧在不同地区的平均

此外，国家气候中心通常会在每年的 3 月和 10 月进行当年的夏季和冬季预报会商，因此考虑实际业务需要，图 5 专门给出了在 3 月和 10 月起报的夏季和冬季 TCC 技巧空间分布图。如图所示，对夏季而言，模式除了对 ENSO 区域具有较高的预报技巧，对东南亚、我国长江地区和东北区域的温度也具有一定的预报技巧，事实上这些地区也主要是受到 ENSO 次年夏季激发的菲律宾反气旋及 EAP 遥相关型的影响；而对冬季而言，该模式在东亚地区的表现反而偏差，技巧仅存在于华南一带，对华北一直到贝湖一带基本无技巧或为负技巧，这可能与模式对 AO 的预报能力较差有关。

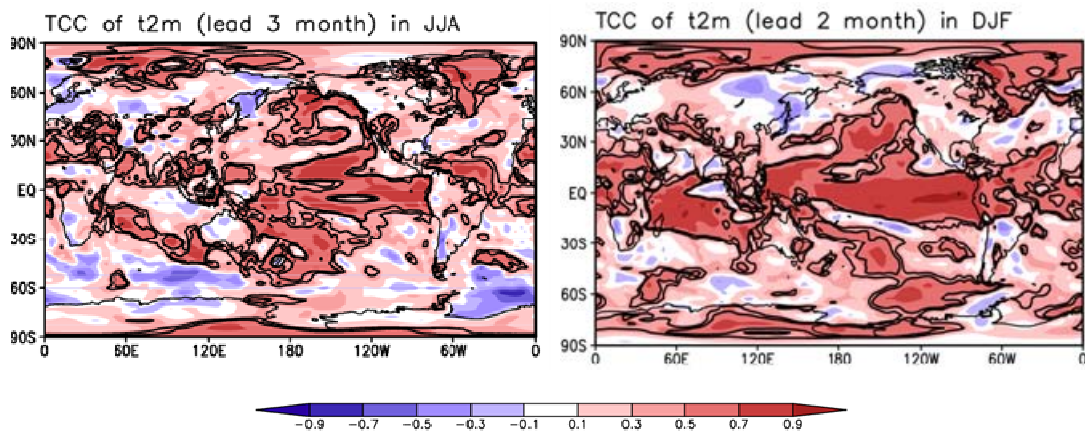


图 5 BCC 二代季节模式在 3 月起报的夏季平均和 10 月起报的冬季平均的 2m 温度的时间相关系数（TCC）技巧空间分布图，等值线分别代表通过显著性为 0.05 和 0.01 检验的区域

### 3.1.3 平均方差技巧评分（MSSS）

MSSS 为 WMO 推荐的确定性预报的主要评估方法，主要考虑预报误差和“气候”预报误

差的比值，MSSS 值越接近 1，表明模式预测越准确，大于 0 表明有预报意义。此处需要注意，由于预报中最关注对异常的预报而不是气候态的偏差，因此计算预报误差时未考虑气候态的漂移，仅考虑对距平预报的偏差。

全球平均而言，MSSS 总体得分在 0 值上下，其中秋冬季大于 0，同样是夏季技巧最低（图 6）。在空间上看，MSSS 与 TCC 的评估结果较为类似，技巧较高的地区均位于热带海洋和 ENSO 遥相关影响区域。需要注意的是，模式对夏季的热带印度洋预报技巧同样偏低，对东亚地区的预报基本无技巧。

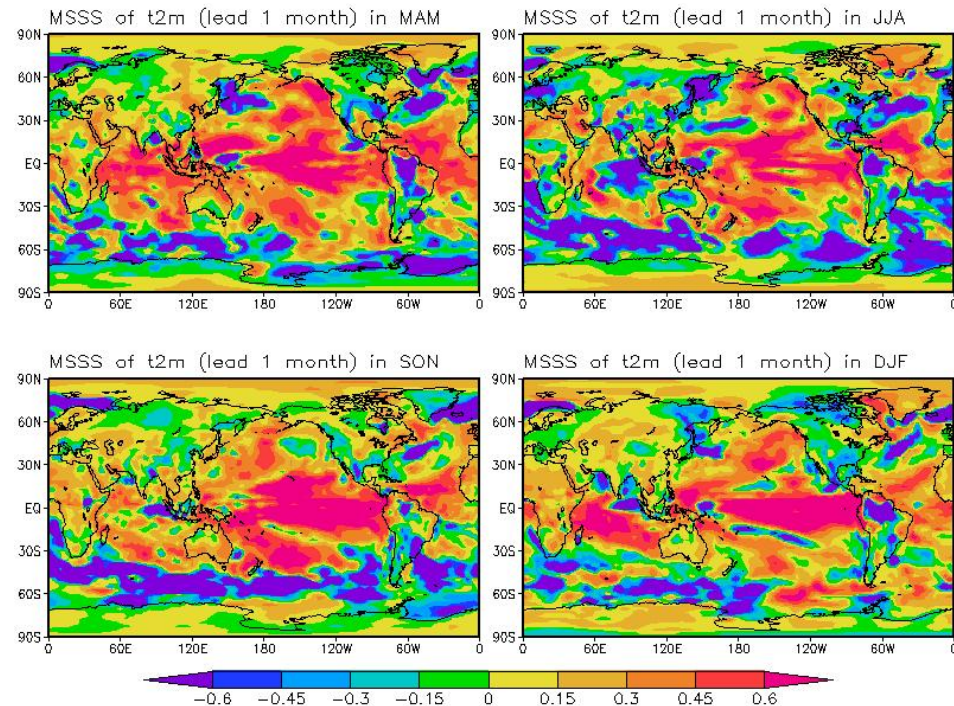


图 6 同图 2，但为均方技巧评分 (MSSS) 的空间分布图

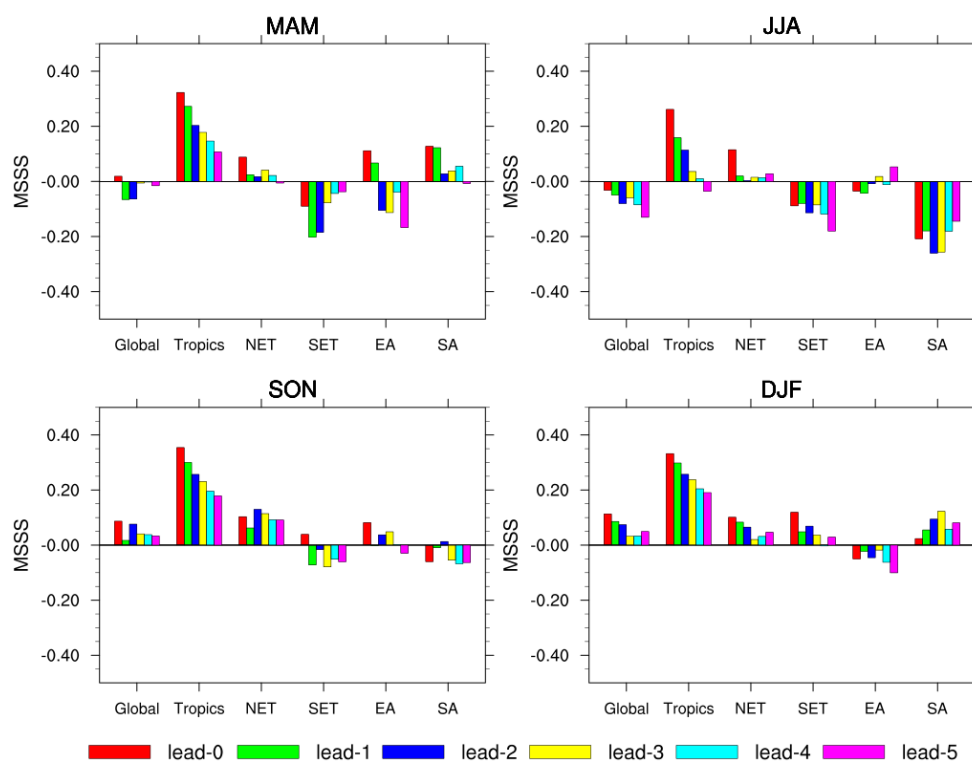


图7 同图4，但为均方技巧评分（MSSS）的区域平均图

### 3.1.4 相对操作特征（ROC）

下面采用相对操作特征（ROC）、可靠性图表（RD）和 Brier 技巧评分（BSS）等三种概率评估方法对按照气候概率百分位区分的正异常（AN）、负异常（BN）和接近正常（NN）三种事件的概率预报技巧进行评估。图8为二代模式提前一个月对四个季节2m温度正异常事件的ROC面积分布图，通常认为ROC面积大于0.5的区域模式具备预报技巧。总体而言，ROCA与TCC分布类似，二者接近线性相关（Wang等，2009），模式有技巧的区域集中在热带海洋和ENSO的遥相关影响区域，在亚欧大陆技巧较低。由三种事件ROC曲线可知（图9），模式对异常低温事件的预报能力稍好于异常高温事件，但它们的技巧都远高于对正常事件的预报。不同季节比较也表明，模式对春季和冬季的预报技巧较高，对夏季预报技巧较低。

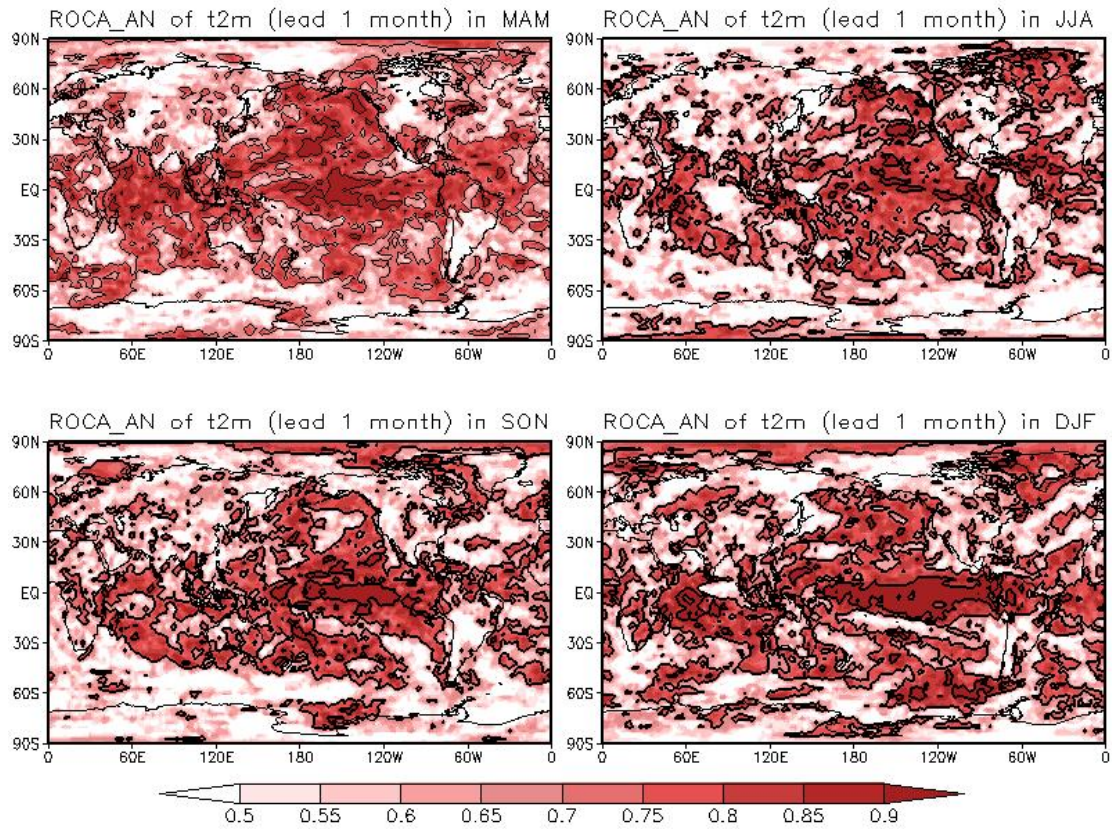


图8 同图2，但为正异常事件（Above Normal）的ROC面积的空间分布



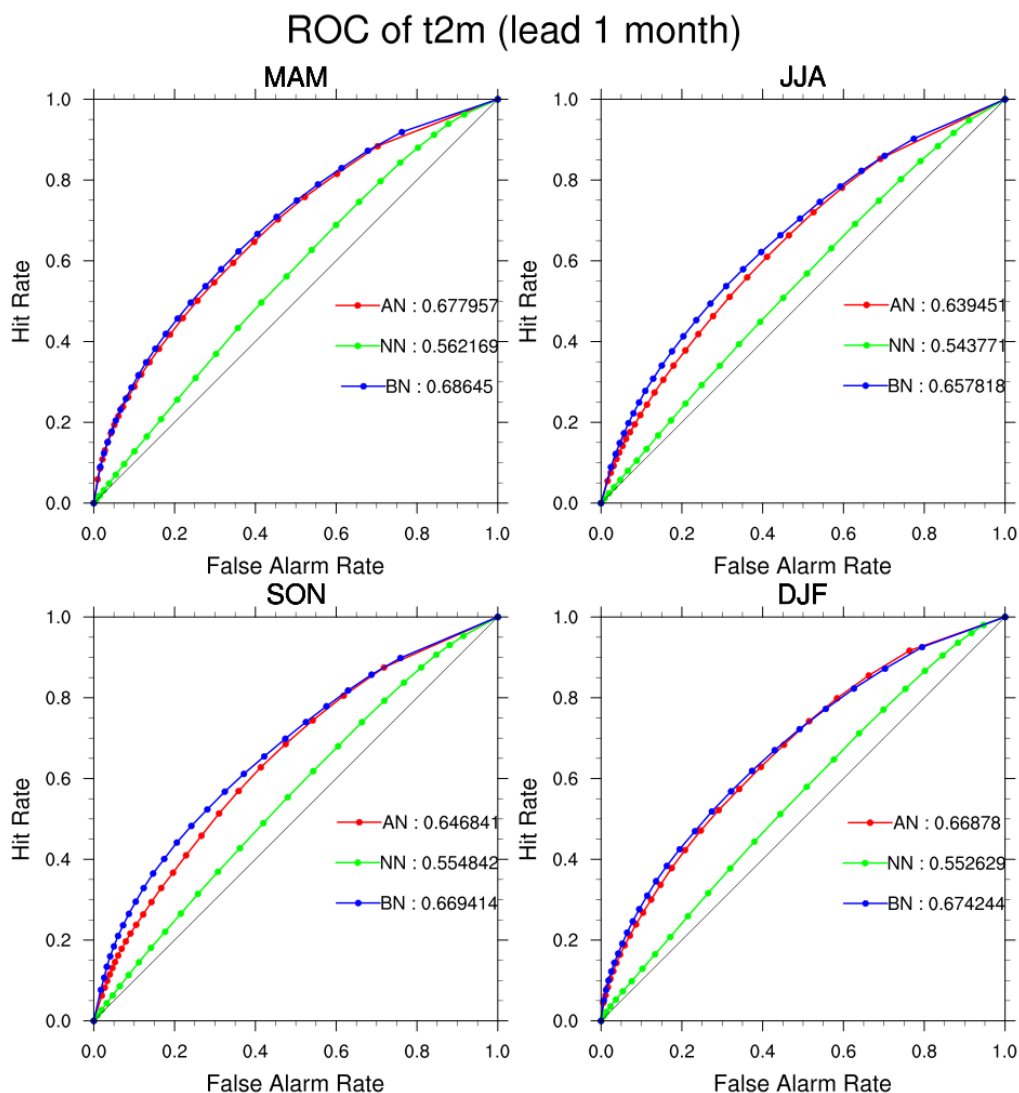


图9 为图8所示的全球平均的 ROC 曲线

### 3.1.5 可靠性图表

图10进一步给出二代季节模式对气温预报的可靠性图表(RD)。RD曲线的横轴代表预报概率,纵轴代表观测到的命中率,曲线越接近对角线,预报效果越好,模式可靠性也越高。由图10可知,在冬季和春季,RD曲线更接近对角线,模式对气温预报的可靠性也越高,在夏季曲线最平;模式同样对正异常(AN)和负异常(BN)事件的预报可靠性好于接近正常事件(NN)。此外,对于模式预报的低概率事件,实际观测中发生的频率偏高;而对预报的高概率事件,观测中发生的频率偏低,即BCC二代模式同其他模式类似,同样存在低概率低估,高概率高估的特征。

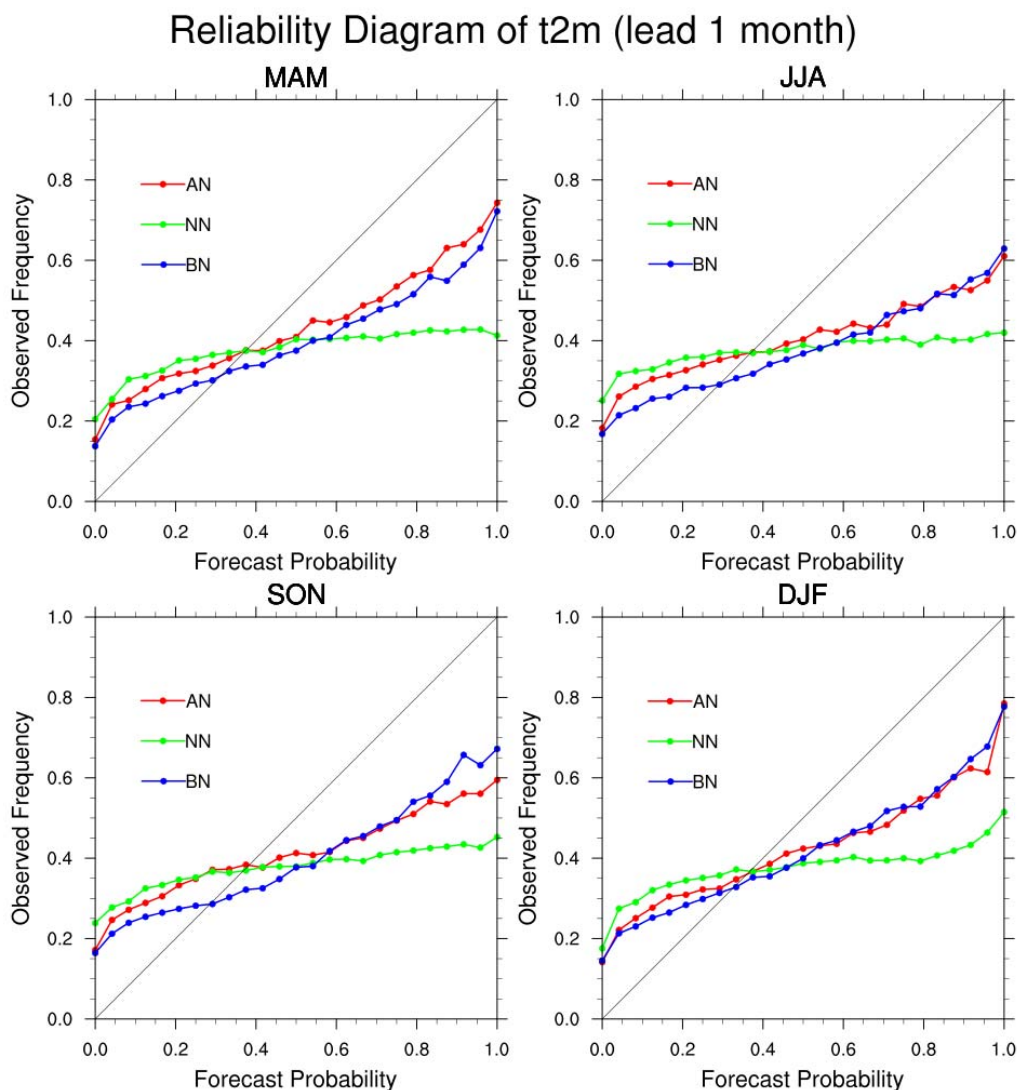


图 10 提前 1 个月起报时 BCC 二代季节模式对全球 2m 温度预报的可靠性图表 (Reliability Diagram)

### 3.1.6 Brier 技巧评分

与图 8 类似，图 11 给出二代模式提前一个月对四个季节 2m 温度正异常事件的 BSS 空间分布图。通常认为，模式在 BSS 大于 0 的区域具备预报技巧。由图可见，与 ROCA 类似，模式同样在热带中东太平洋以及西印度洋地区具有一定预报技巧，冬季技巧最高而夏季最低。图 12 进一步给出了随预报 lead time 的增加，BSS 以及各项得分的变化。由图可见，对全球平均而言，提前 1 个月以上预报的 BSS 基本都变为负值，而随着 lead time 的延长，BSS 的下降主要是由于模式分辨能力（与距平相关系数展现的特征类似）的下降造成的，而模式概率预报提供的可靠性信息变化不大。对比来看，模式对正、负异常概率预报的 BSS 没有明显差别，后者较前者略好。



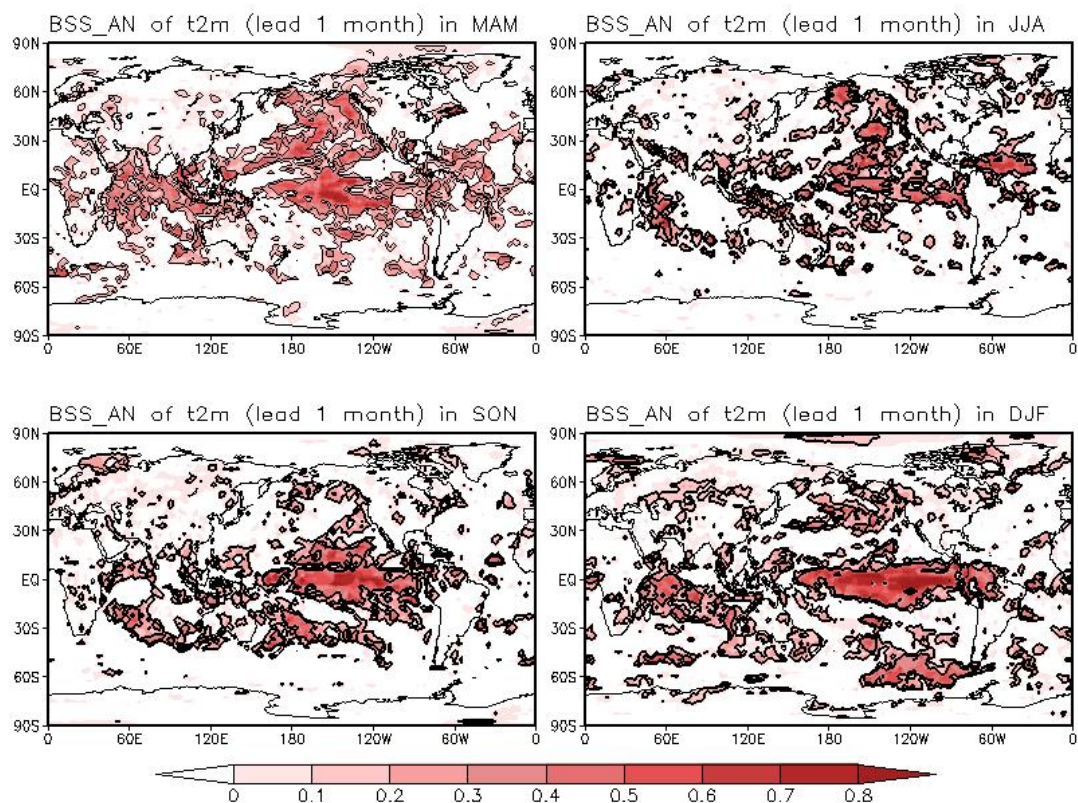
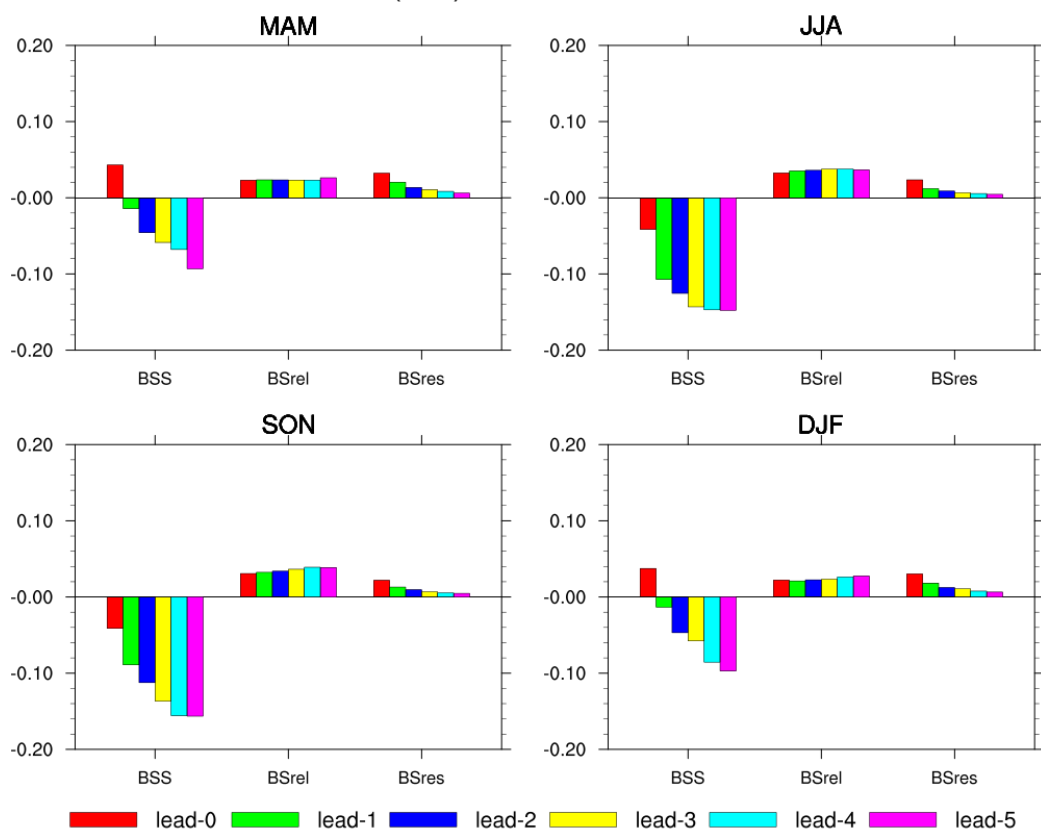


图 11 同图 8，但为正异常事件 (Above Normal) 的 BSS 空间分布

### BSS(AN) of t2m - Global



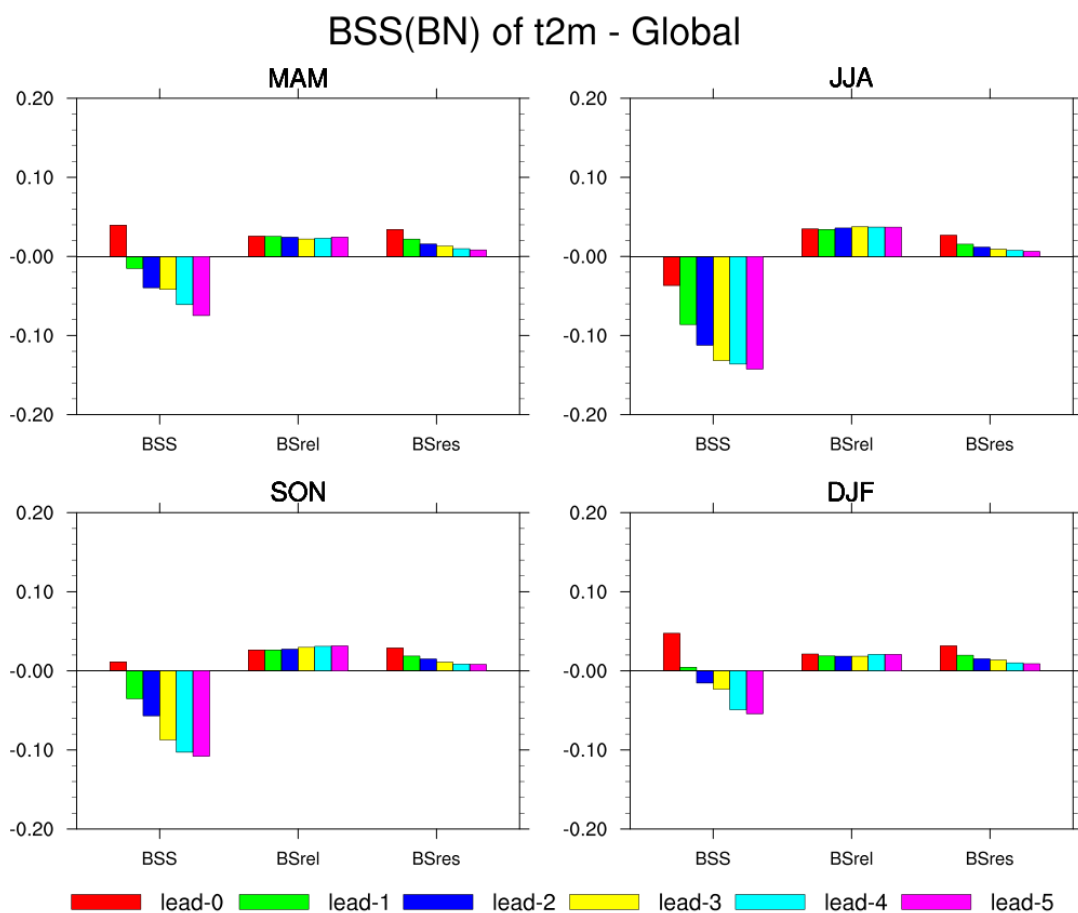


图 12 全球平均的正异常事件（AN，上图）和负异常事件（BN，下图）的 BSS 及可靠性和分辨能力得分

### 3.1.7 二代模式对 2m 温度预报小结

通过对二代季节模式对 2m 温度预报性能的系统评估及与一代模式的对比说明，二代模式的预报技巧较一代模式有了明显提高，其中冬季全球 ACC 多年平均值为 0.27，其他季节均超过 0.12；预报技巧的空间分布表明，模式在热带地区的预报技巧明显高于热带外，二代模式对 ENSO 区域气温预报的改进明显，由此模式热带外 ENSO 的遥相关地区也具备一定的预报能力，但模式对夏季热带印度洋地区没有预报能力；从不同季节来看，模式在北半球冬季和春季的预报技巧较高，夏季预报技巧偏差；具体到东亚地区而言，模式的整体预报性能偏差，但需要注意的是，在东亚地区模式在夏季预报性能高于冬季；在冬季，模式对我国东北地区的气温没有预报技巧；概率预报性能检验（ROCA, BSS）的结果与 TCC 分布非常类似，BSS 方程各项分析表明，随着预报 lead time 预报技巧的下降主要是由于模式的分辨能力下降造成的。

## 3.2 降水（PREC）

### 3.2.1 距平相关系数（ACC）

下面给出二代和一代季节预报模式对降水的预报性能检验。与 2m 气温相比，二代模式对降水预报性能的改进更为明显，其中二代模式提前一个月预报的全球降水多年平均的 ACC 在春夏秋冬四个季节分别达到 0.31、0.18、0.28 和 0.38，远远高于一代模式的结果（图

13)。这主要得益于二代模式对 ENSO 预报性能(详见本报告对全球气候现象预测评估部分)的提升带来的对热带降水把握能力的大幅提升。特别是在历史上典型的 ENSO 年,例如 1997 年的秋冬季和 1998 年的春季,模式对降水预报的 ACC 评分能够超过 0.6。

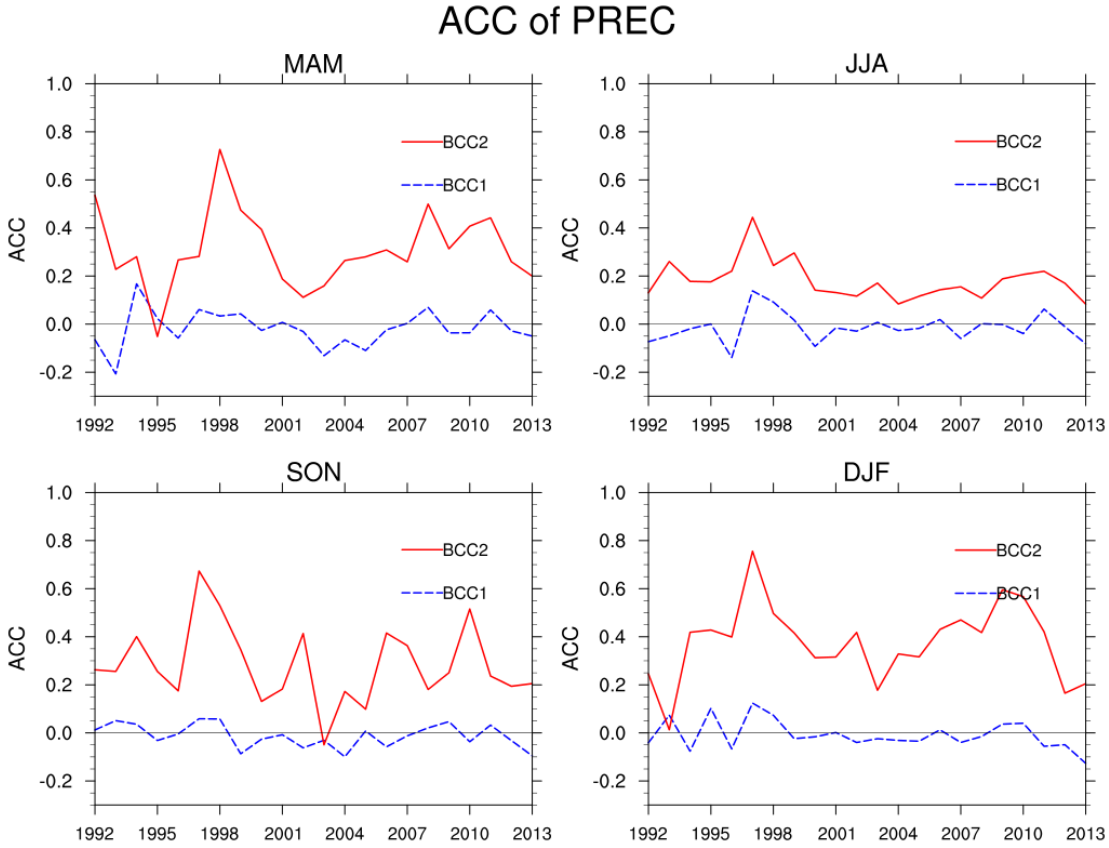


图 13 提前 1 个月 BCC 一代和二代季节模式对全球降水预报的 ACC 技巧的四个季节的时间序列

### 3.2.2 时间相关系数 (TCC)

图 14 和图 15 分别给出了提前 1 个月起报的二代模式对降水预报的 TCC 技巧以及二代与一代相比 TCC 技巧差值的全球空间分布图。由图可知,二代模式对降水的预报技巧主要位于热带中东太平洋、印度洋和西大西洋地区;在热带外地区,该模式主要对冬季 ENSO 的遥相关影响区域,如东北太平洋和北美部分地区、以及我国华南和长江地区,具有一定的预报技巧。与一代模式相比,二代模式在赤道中太平洋的预报技巧有明显的提高,而对东亚地区降水预报能力的提升比较有限。



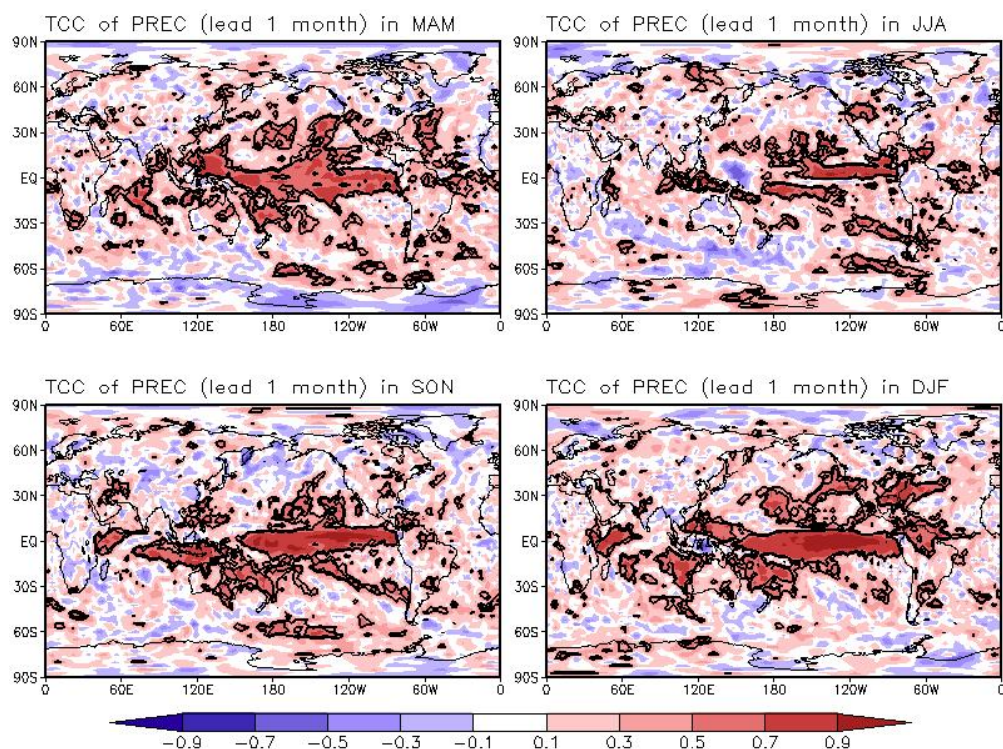


图 14 提前 1 个月 BCC 二代季节模式对降水预报的时间相关系数 (TCC) 技巧的四个季节的空间分布图，等值线分别代表通过显著性为 0.05 和 0.01 检验的区域

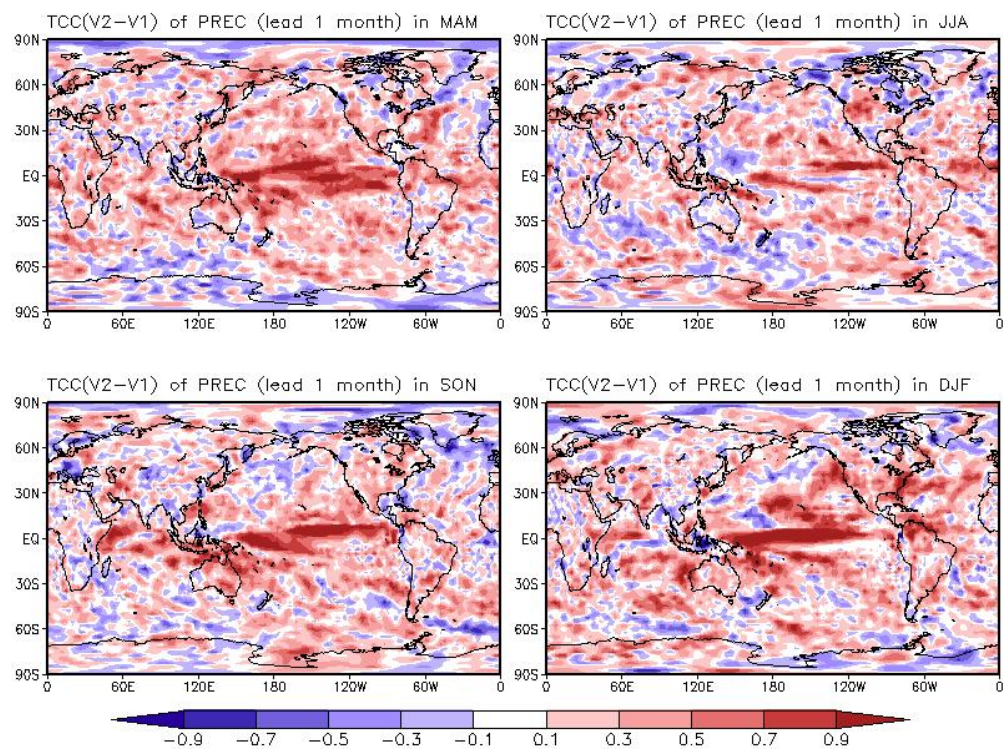


图 15 同图 14，但为二代模式与一代模式 TCC 技巧的差值图

从各区域平均来看，二代模式对降水的预报技巧在冬季和春季较高、夏季较低。此外，模式预报结果对初值比较敏感，提前 1 个月较提前 0 个月的预报技巧下降明显，特别是在东亚地区，之后随 lead time 的延长，预报技巧下降不大。

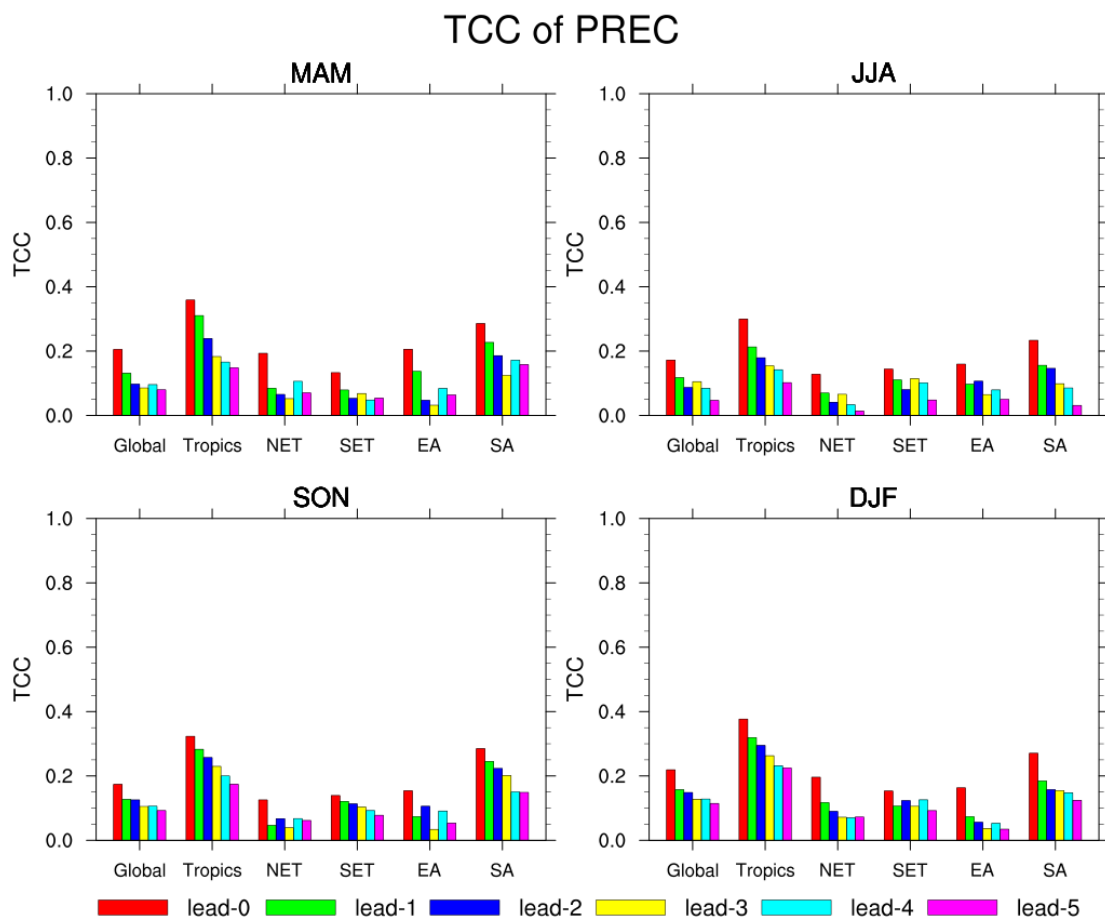


图 16 BCC 二代季节模式在不同 lead time 起报的四个季节降水的 TCC 技巧在不同地区的平均

若从业务预报会商的角度，二代模式在 3 月给出的当年夏季预报技巧偏低，特别是对东亚大陆区域降水的预报技巧较低，而 10 月份模式对当年冬季降水预报整体明显偏高，尤其是热带海洋区域，与夏季相比，对东亚和我国区域降水有一定的预报技巧。

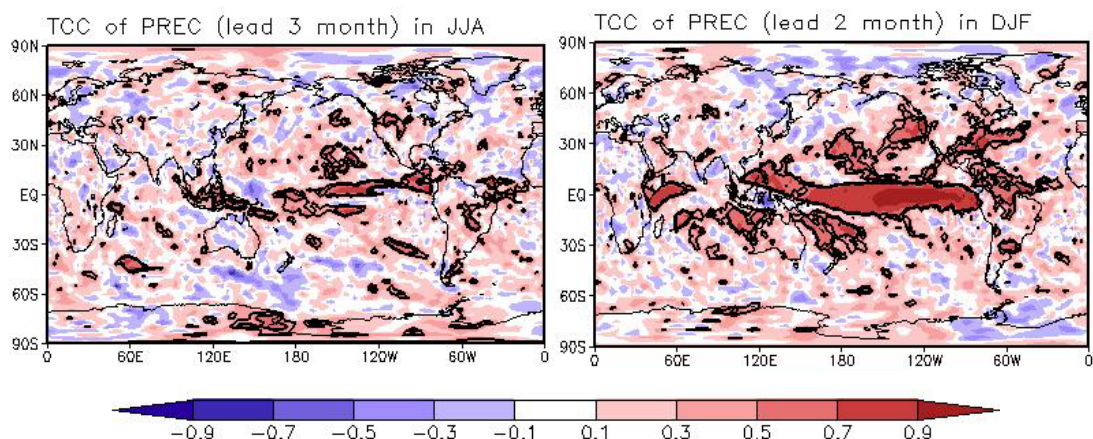


图 17 BCC 二代季节模式在 3 月起报的夏季平均和 10 月起报的冬季平均降水的 TCC 技巧空间分布图，等值线分别代表通过显著性为 0.05 和 0.01 检验的区域

### 3.2.3 平均方差技巧评分 (MSSS)

MSSS 的空间分布与 TCC 类似 (图 18)，除夏季外，二代模式对赤道中东太平洋、西北太平洋暖池和 ENSO 的遥相关地区降水的预报技巧较高。在夏季，模式对降水基本不具备



预报技巧（图 19）。相比之下，对我国东部区域，模式冬春技巧高于夏秋技巧，但均偏低。

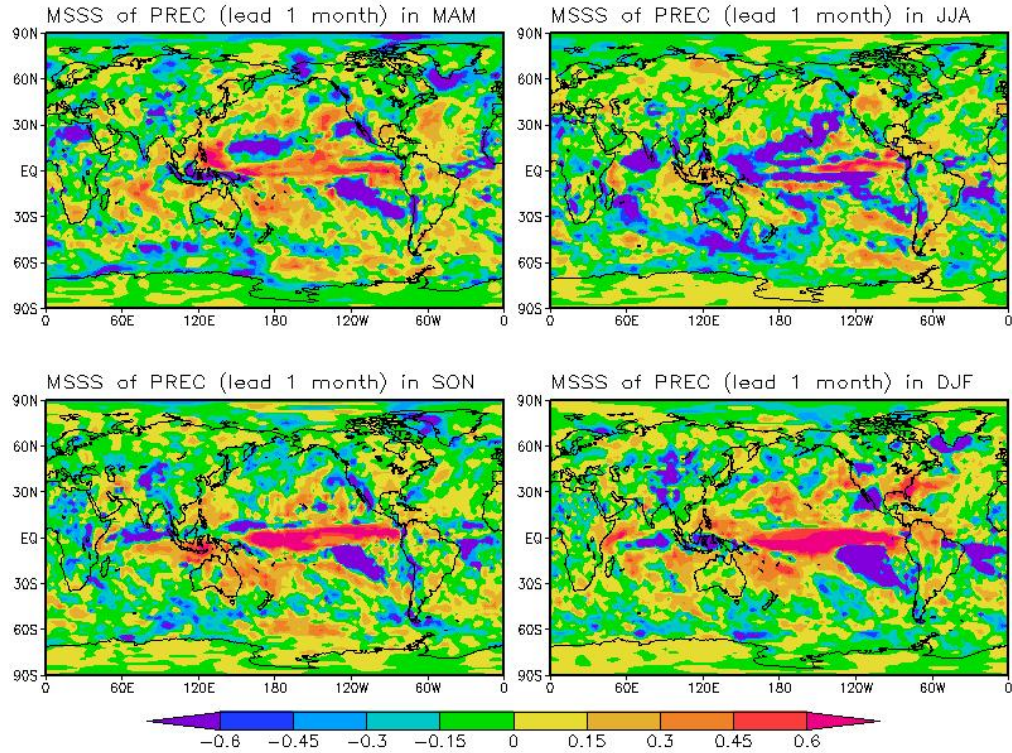


图 18 同图 14，但为均方技巧评分（MSSS）的空间分布图

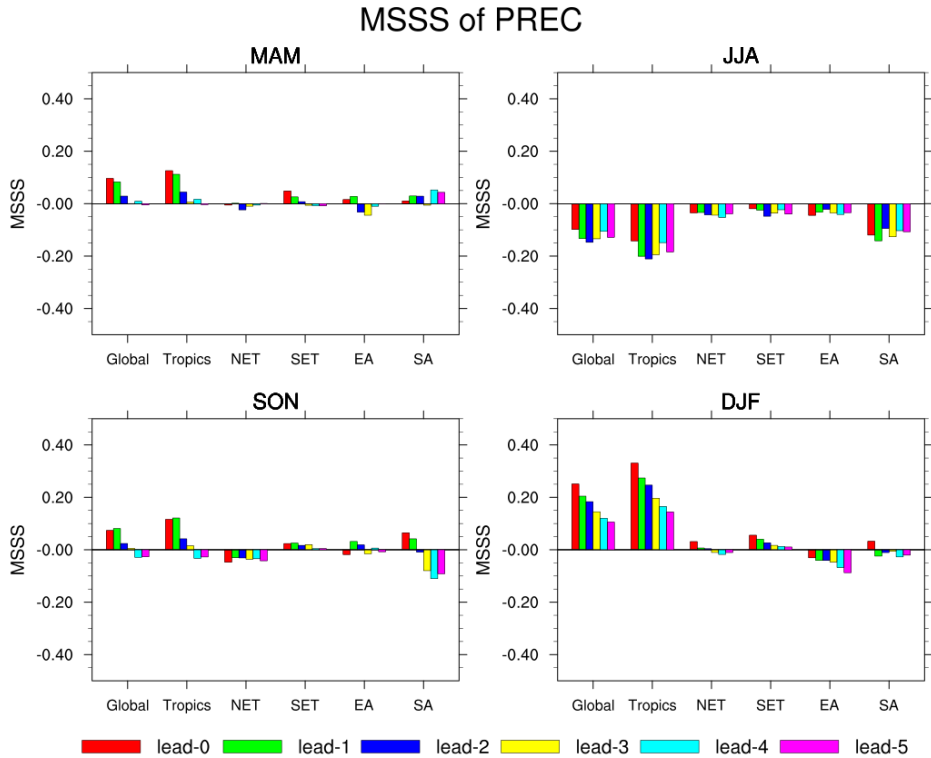


图 19 同图 16，但为 MSSS 的区域平均图

### 3.2.4 相对操作特征（ROC）

由图 20 可见，对降水正异常事件，ROC 面积的空间分布与 TCC 和 MSSS 比较接近，其中在北半球冬春季表现为明显的从西北太平洋暖池到墨西哥湾的沿大圆路径传播的波列

状技巧分布，而模式对东亚地区的降水基本不具备预报能力。ROC 曲线图进一步说明（图 21），模式对降水正异常事件（AN）的预报技巧略好于负异常事件（BN），而模式对接近正常值得降水基本无预报技巧。

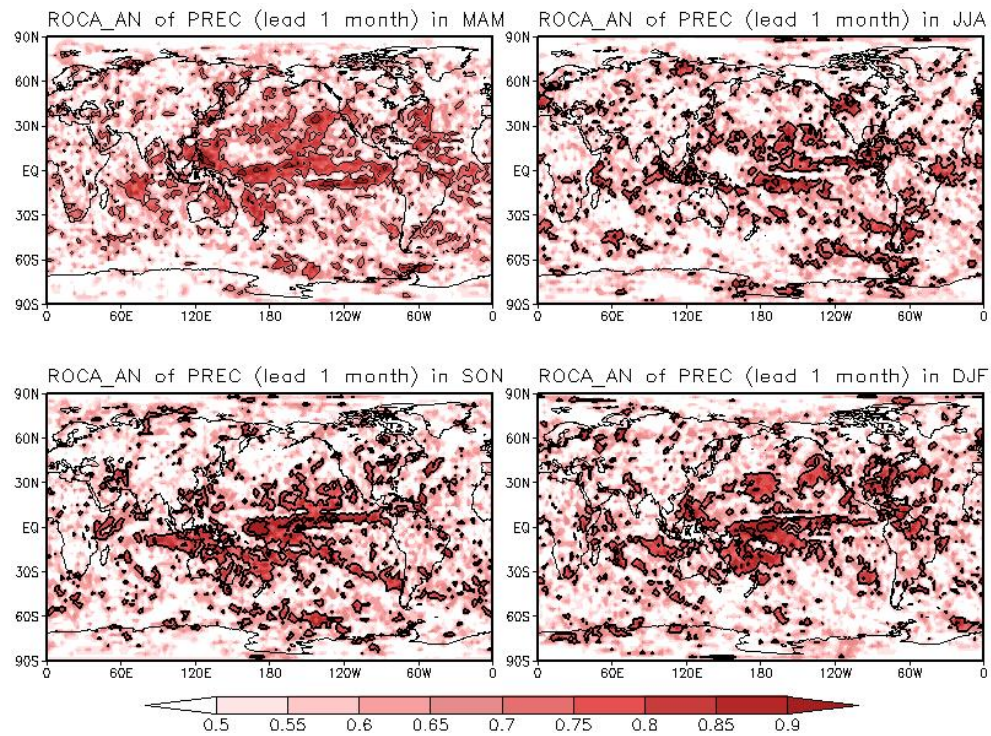


图 20 同图 14，但为正异常事件（Above Normal）的 ROC 面积的空间分布

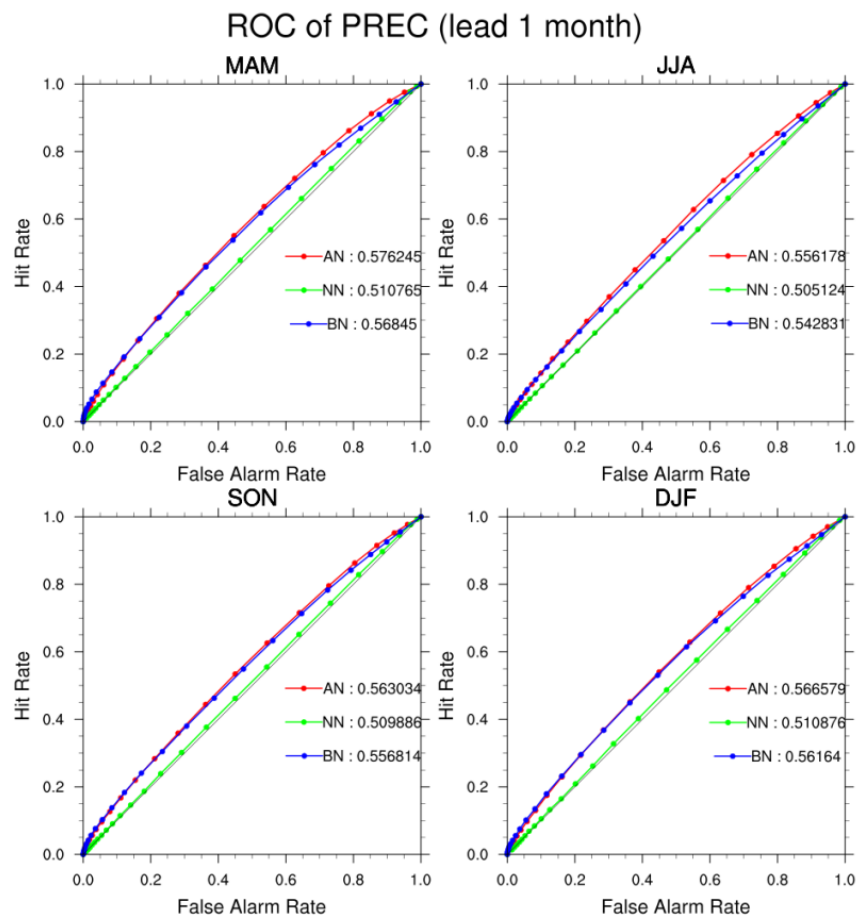


图 21 为图 20 所示的全球平均的 ROC 曲线

### 3.2.5 可靠性图表 (Reliability Diagram)

通过可靠性图表也可看出 (图 22), 模式对降水负异常事件预报的可靠性更高, 好于降水正异常事件, 特别是对预报较大概率时的可靠性更高。

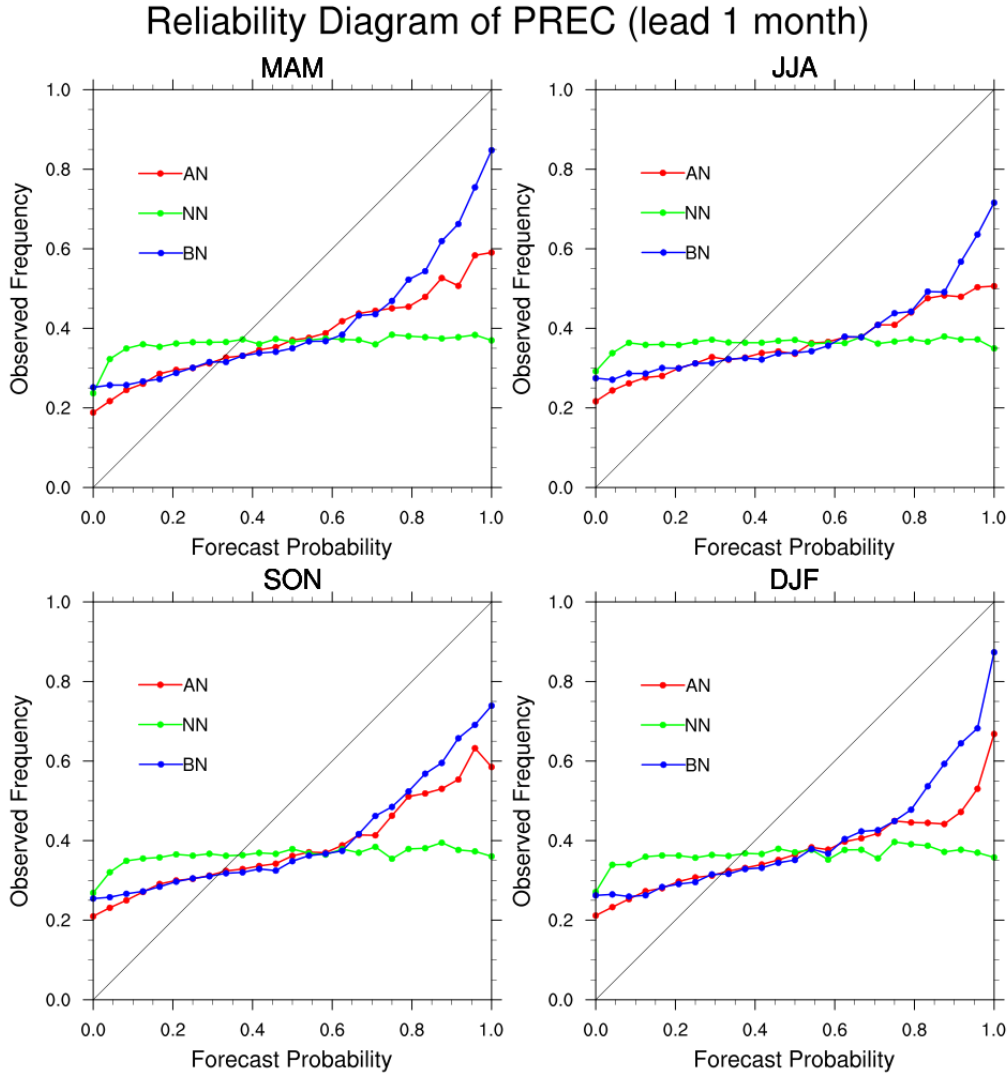


图 22 提前 1 个月起报时 BCC 二代季节模式对全球降水预报的可靠性图表

### 3.2.6 Brier 技巧评分

对降水 BSS 的空间分布与 TCC 和 ROC 面积均具有较好的一致性, 在北半球春季和冬季都有明显的波列状的分布存在 (图 23)。但全球平均的 BSS 评分在各个季节均为负值, 模式在 lead time 超过 2 个月后基本失去分辨能力。相比之下, 模式对正、负异常概率预报的 BSS 无明显差别, 后者略微高于前者。



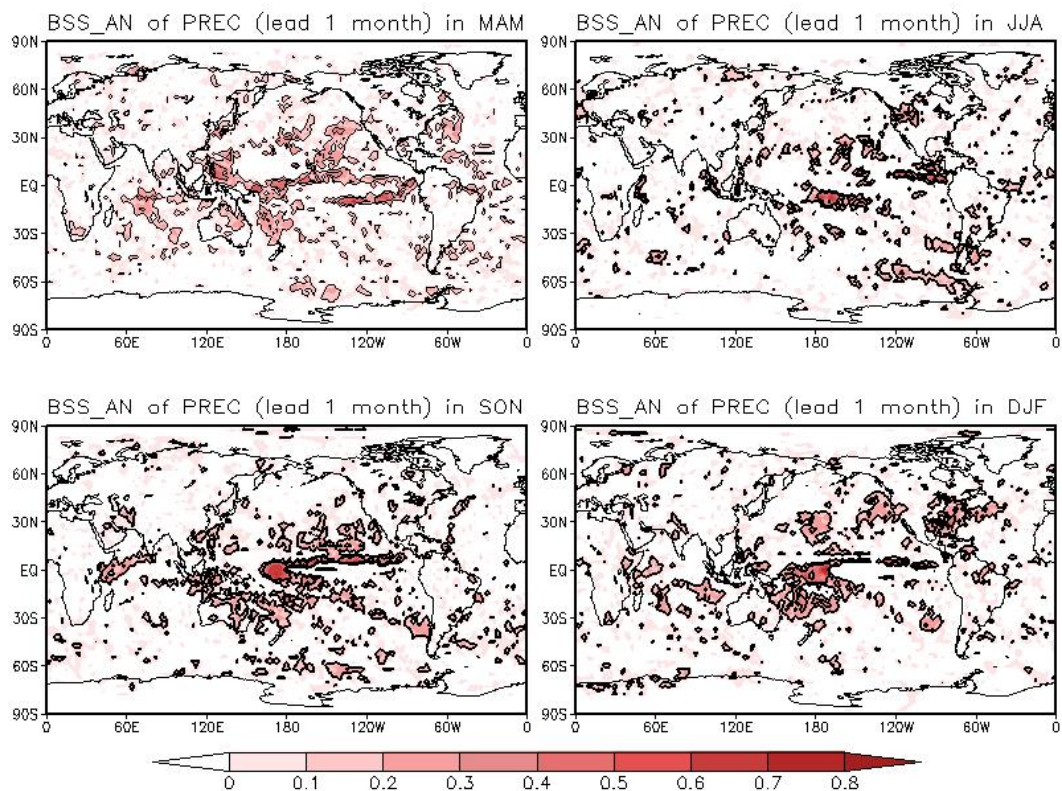
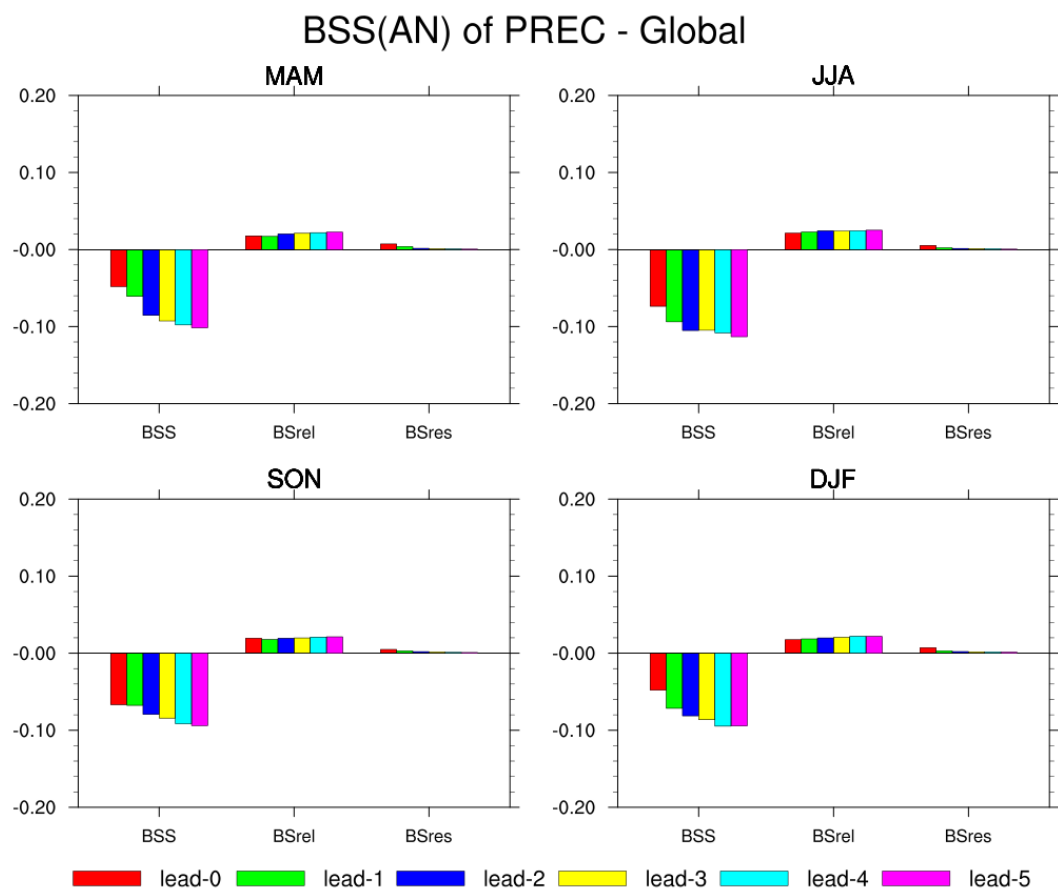


图 23 同图 20, 但为正异常事件 (Above Normal) 的 BSS 空间分布



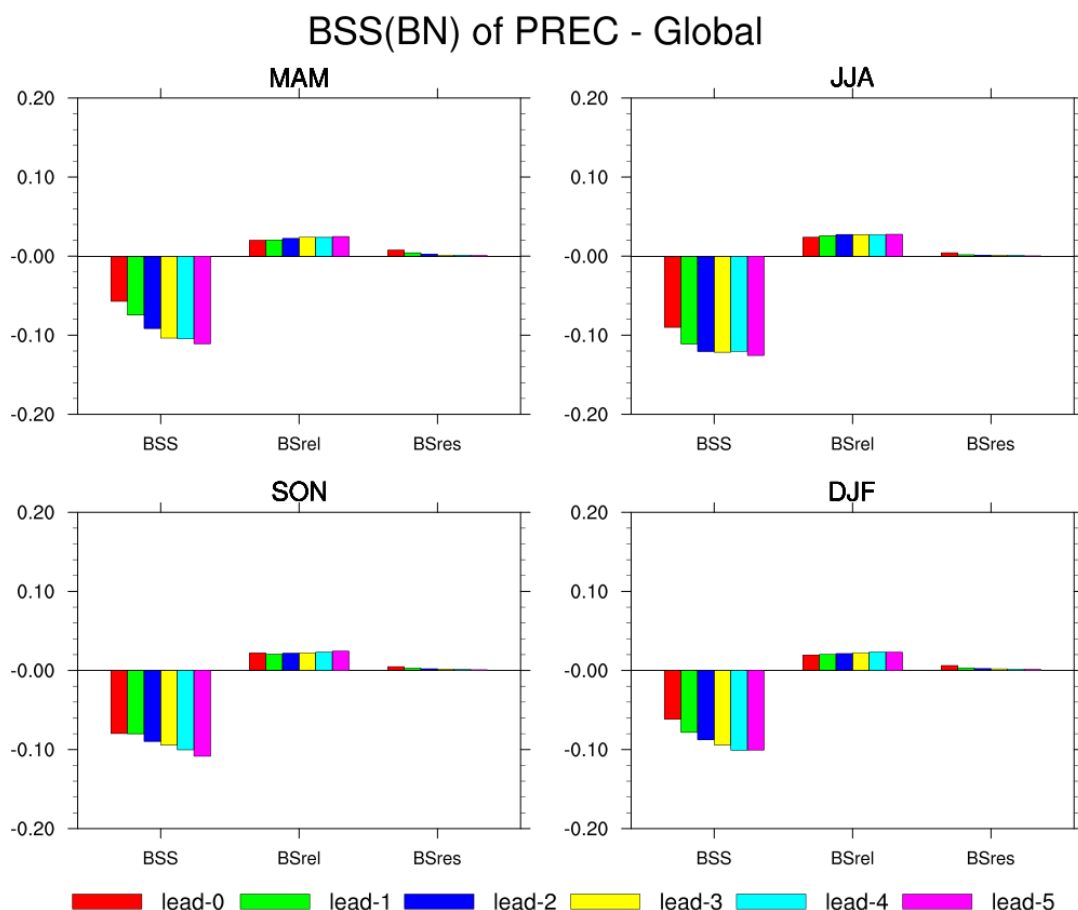


图 24 全球平均的正异常事件（AN，上图）和负异常事件（BN，下图）的 BSS 及可靠性和分辨能力得分

### 3.2.7 二代模式对降水预报小结

通过与一代模式的对比，BCC 二代季节模式对全球降水预报性能的改进较为明显，特别是在热带地区，其提前一个月预报的全球降水多年平均的 ACC 在春夏秋冬四个季节分别达到 0.31、0.18、0.28 和 0.38，技巧远高于一代模式的结果。此外，二代模式在 ENSO 年的预报技巧较高；从预报技巧的空间分布上看，除夏季（JJA）外，二代模式对赤道中东太平洋、西北太平洋暖池及北半球热带外 ENSO 的遥相关区域的降水具有一定的预报能力；从概率预报评分上看，二代模式对降水负异常事件的预报能力略高于正异常事件。与其他地区相比，模式对东亚陆地区域的降水异常的预报能力仍然较低。

## 3.3 500hPa 位势高度场（Z500）

### 3.3.1 距平相关系数（ACC）

对 500hPa 位势高度场而言，目前 BCC 二代模式对春夏秋冬四个季节的 ACC 预报技巧分别为 0.17、0.10、0.11 和 0.26，均高于一代模式的预报技巧（图 25）。然而，与 2m 温度和降水预报不同，500hPa 位势高度场的 ACC 年际变化很大，在主要 ENSO 年(97/98 和 09/10) 预报技巧较高，反映出二代模式对 ENSO 事件预报性能有了比较明显的改进。

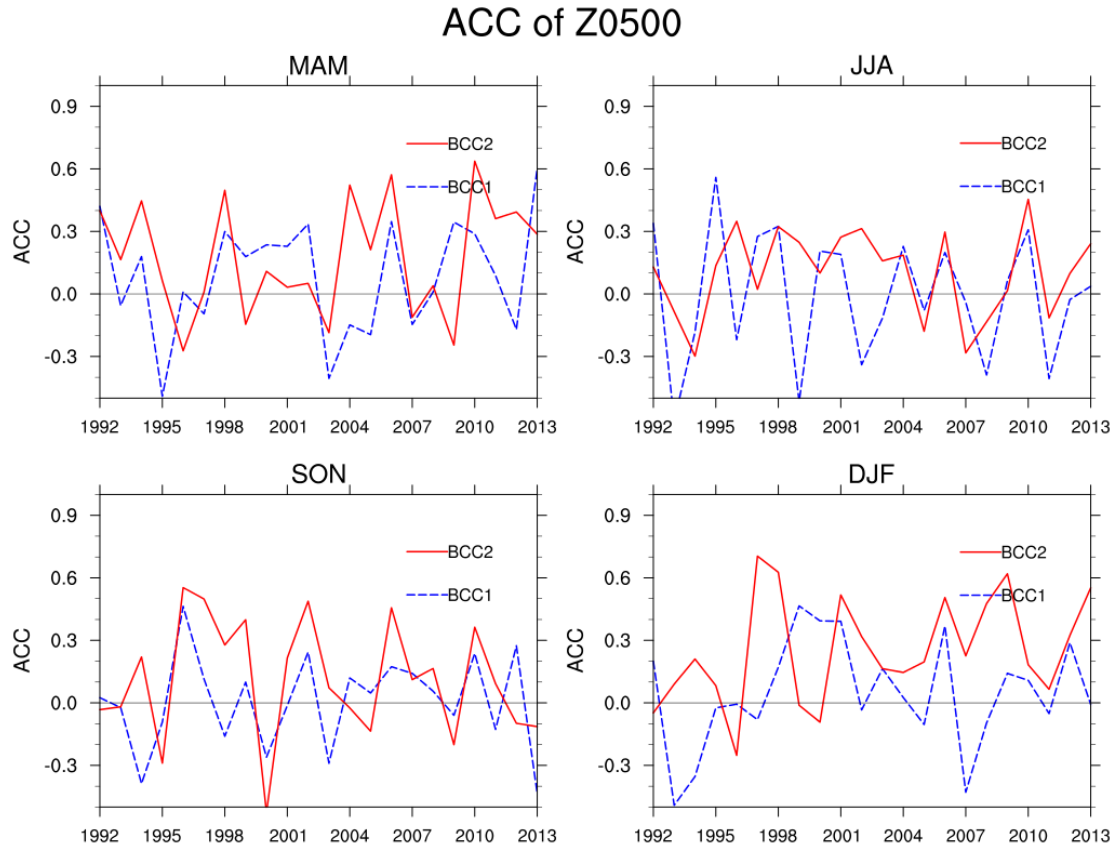


图 25 提前 1 个月 BCC 一代和二代季节模式对全球 500hPa 位势高度场预报的距平相关系数 (ACC) 技巧的四个季节的时间序列

### 3.3.2 时间相关系数 (TCC)

从时间相关系数的空间分布上看 (图 26), 二代模式对热带地区的 500hPa 位势高度场的预报技巧最高, 在中纬度的高技巧区域成波列状分布, 与 EU 型和 ENSO 激发的 PNA 型等遥相关型类似。对东亚地区而言, 模式在夏季对西太副高有一定的预报技巧, 但在冬季对贝加尔湖以北到鄂霍茨克海地区等中高纬阻塞形势易发生地区的预报技巧偏低。与一代模式相比较 (图 27), 二代模式对热带地区、特别是热带印度洋和热带中太平洋地区的预报技巧有明显改进, 而对中高纬地区的改进则与热带地区明显相关, 与热带预报改进明显的地区形成类似大圆路径的波列状分布, 例如, 夏季形成类似 EAP 型波列, 冬季形成 PNA 型波列, 说明预报技巧的来源仍然来自热带地区。



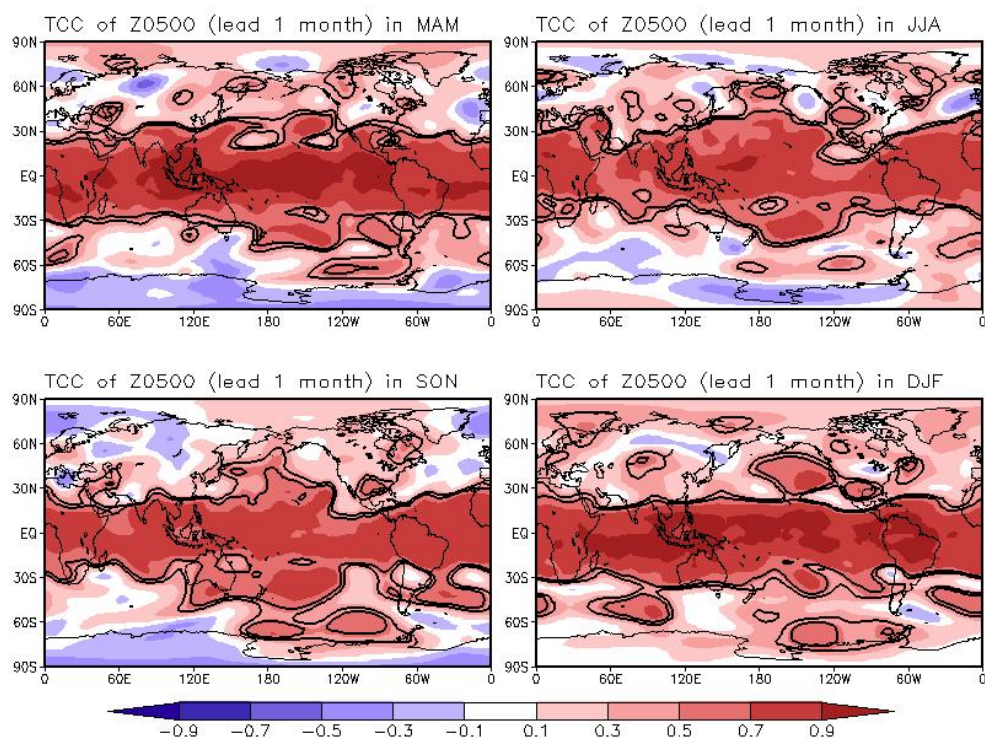


图 26 提前 1 个月 BCC 二代季节模式对 500hPa 位势高度预报的时间相关系数 (TCC) 技巧的四个季节的空间分布图, 等值线分别代表通过显著性为 0.05 和 0.01 检验的区域

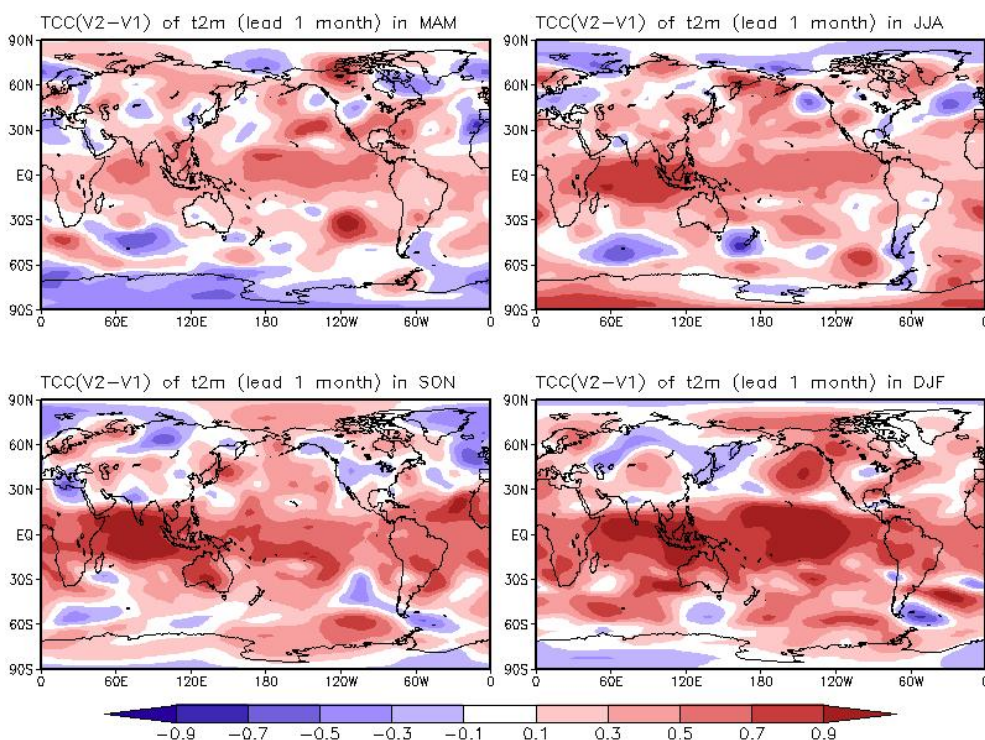


图 27 同图 26, 但为二代模式与一代模式 TCC 技巧的差值图

按照预报会商的需要, 考虑模式 3 月和 10 月给出的夏季和冬季 500hPa 位势高度场预报, 由图 28 可见, 在 3 月模式给出的当年夏季东亚地区的 500hPa 位势高度场具备一定的预报技

巧，特别是西太副高所在区域，然而 10 月份模式对贝湖到鄂霍次克海地区的预报技巧为负值，这也可以从一定程度解释模式对我国冬季东北地区的气温预报性能不佳。

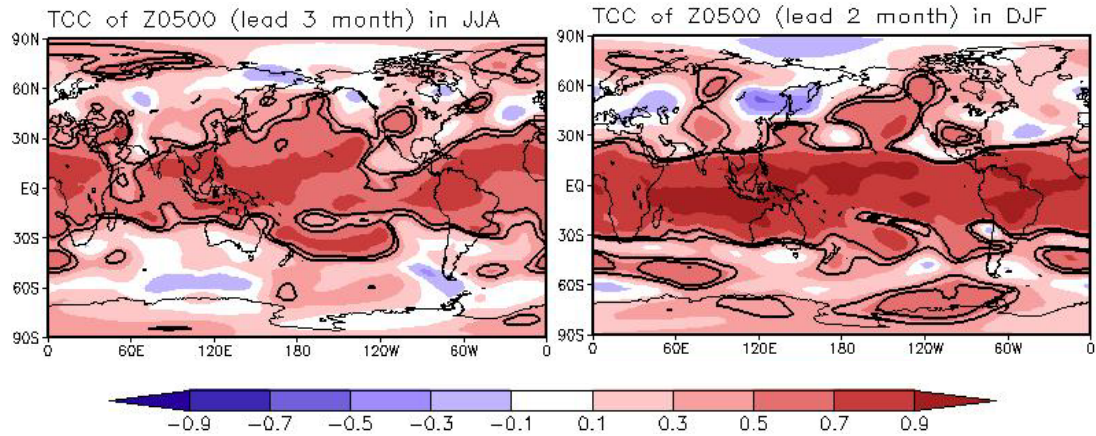


图 28 BCC 二代季节模式在 3 月起报的夏季平均和 10 月起报的冬季平均 500hPa 位势高度场的 TCC 技巧空间分布图，等值线分别代表通过显著性为 0.05 和 0.01 检验的区域

从不同季节的区域平均上看（图 29），全球平均而言，二代模式对冬季 500hPa 位势高度场的预报技巧最高，对秋季最低。而且，模式预报技巧对初值比较敏感，从 lead-0 延长到 lead-1 个月模式技巧下降较快，之后随着 lead time 的延长模式预报技巧下降不明显。此外，与 2m 温度和降水不同，在夏季，该模式对东亚地区预报技巧高于北半球热带外，可能与模式对 EAP 遥相关型预报能力较好有关；但在冬季，模式对东亚地区的预报技巧偏低。

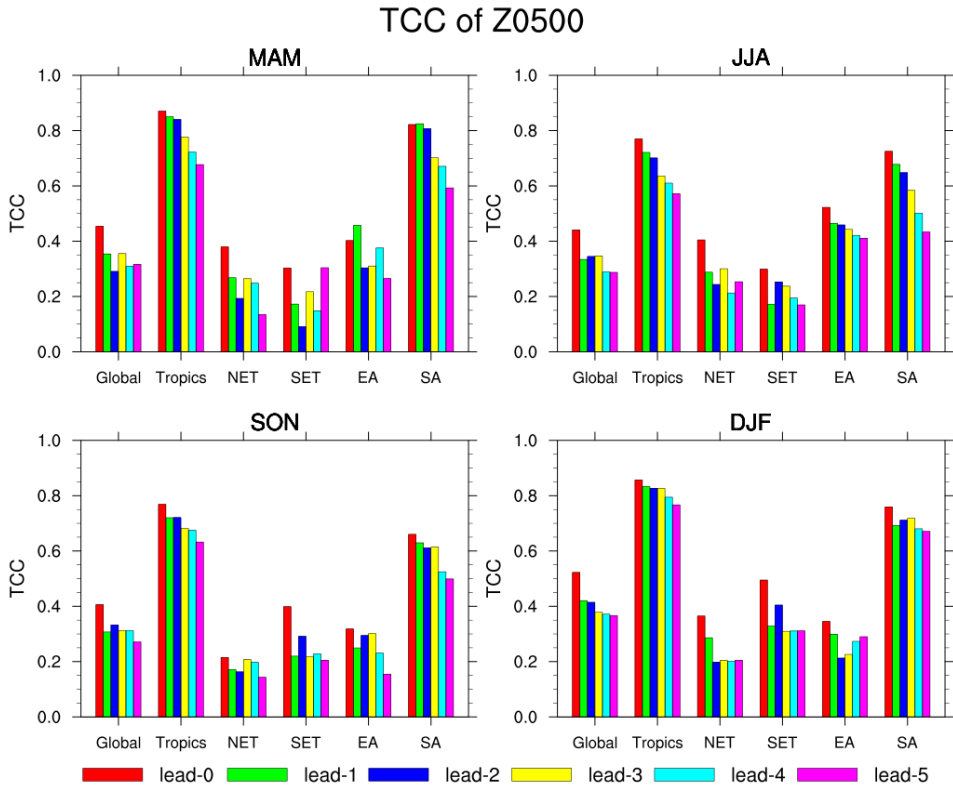


图 29 BCC 二代季节模式在不同 lead time 起报的四个季节 500hPa 位势高度的 TCC 技巧在不同地区的平均

### 3.3.3 平均方差技巧评分（MSSS）

二代模式的平均方差技巧评分的空间分布与 TCC 非常接近，热带地区的预报技巧最高，特别是在冬季。热带外预报技巧主要来自 ENSO 等所激发的遥相关的贡献，但是在冬季，



模式对鄂海地区预报技巧偏低；在秋季则对整个东亚大陆地区基本没有预报技巧。

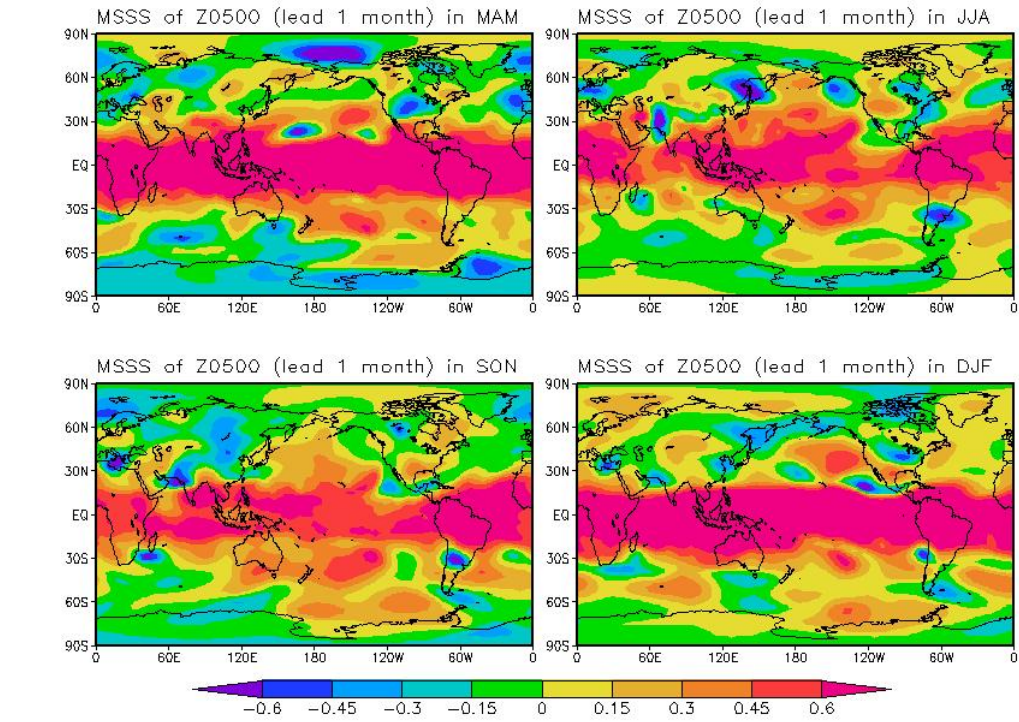


图 30 同图 26，但为均方技巧评分（MSSS）的空间分布图

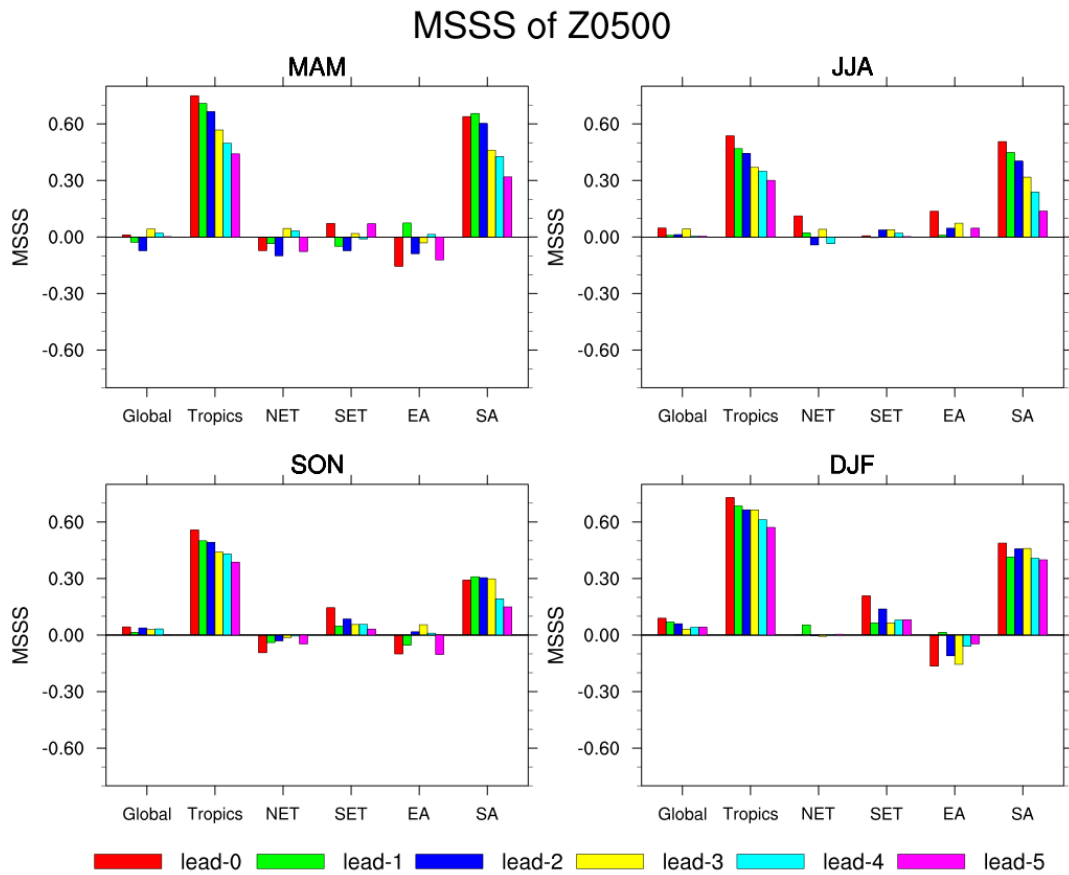


图 31 同图 28，但为均方技巧评分（MSSS）的区域平均图

### 3.3.4 相对操作特征 (ROC)

对 500hPa 位势高度场正异常事件, 从 ROC 面积的空间分布上看(图 32), 除热带地区, 在春季和夏季, 二代模式对我国南方一直到副热带西太平洋的副高活跃区的位势高度有较高的预报技巧; 从秋季到冬季, 模式对巴尔喀什湖到贝加尔湖之间的中纬度地区有一定的预报技巧。从全球平均的 ROC 曲线来看(图 33), 除冬季外, 模式对 500hPa 位势高度异常偏低事件的预报技巧高于异常偏高事件, 这在秋季最为明显。

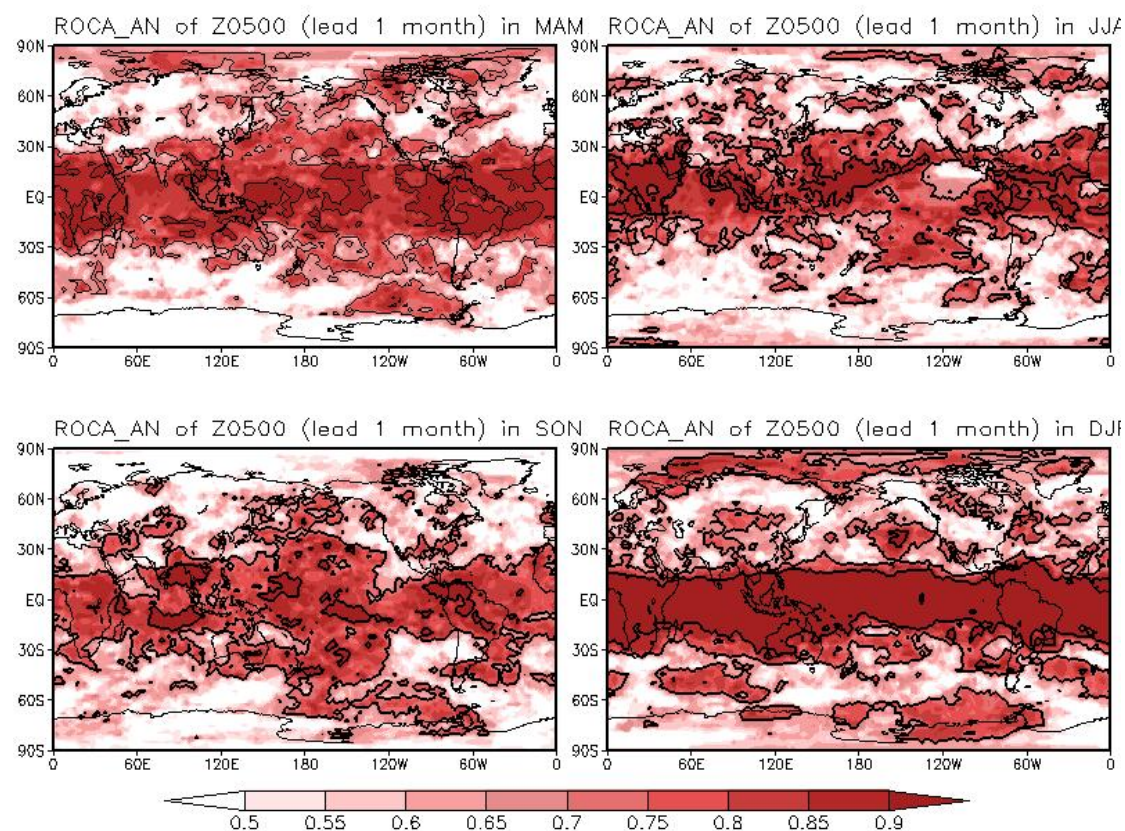


图 32 同图 26, 但为正异常事件 (Above Normal) 的 ROC 面积的空间分布

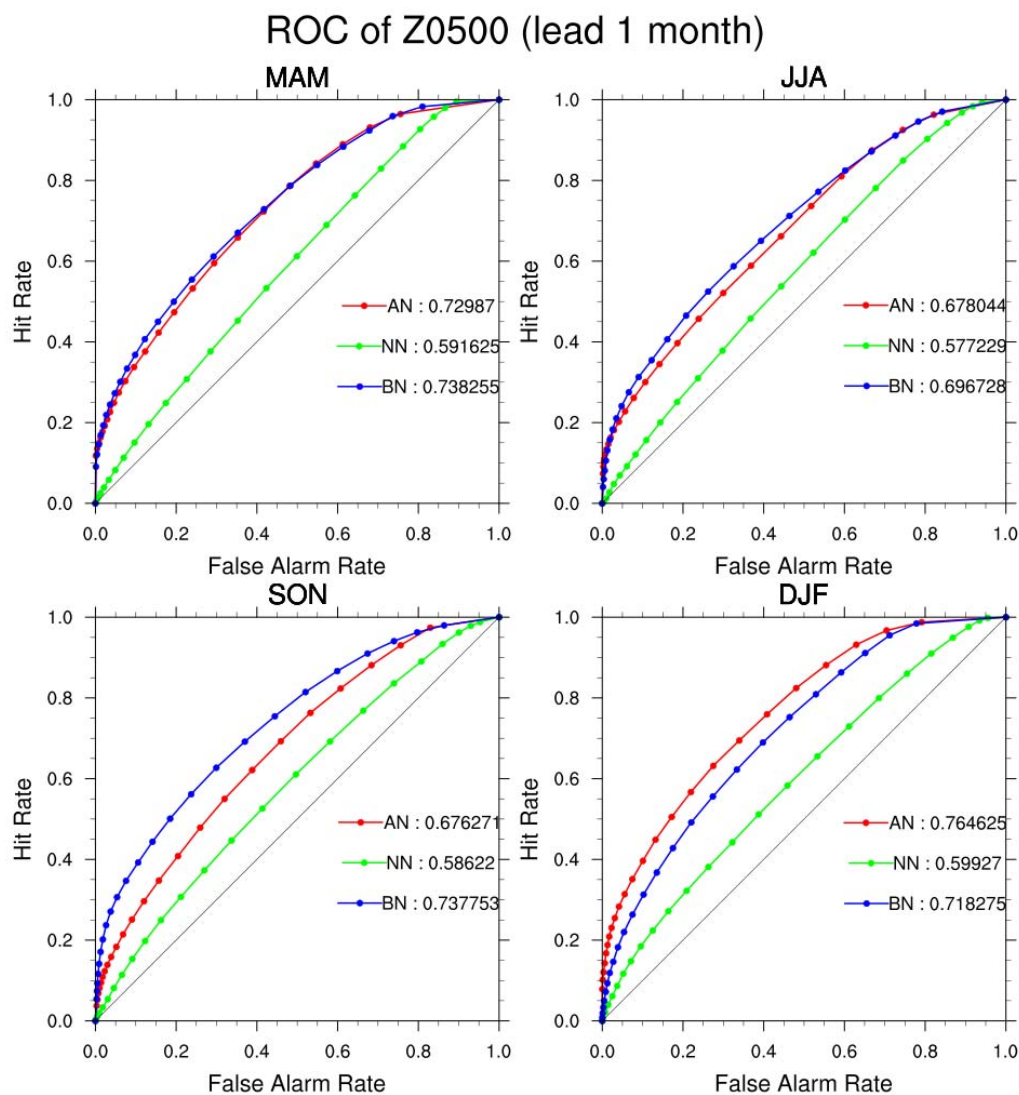


图 33 为图 32 所示的全球平均的 ROC 曲线

### 3.3.5 可靠性图表 (Reliability Diagram)

从可靠性图表上看, 相较于 2m 温度和降水, 模式对 500hPa 位势高度场的预报可靠性更高, 虽然仍然存在预报低概率时低估和预报高概率时高估的问题, 但整体与对角线更加接近, 特别是在冬季。此外, 在春、夏、秋季模式对负异常事件预报的可靠性好于正异常事件。



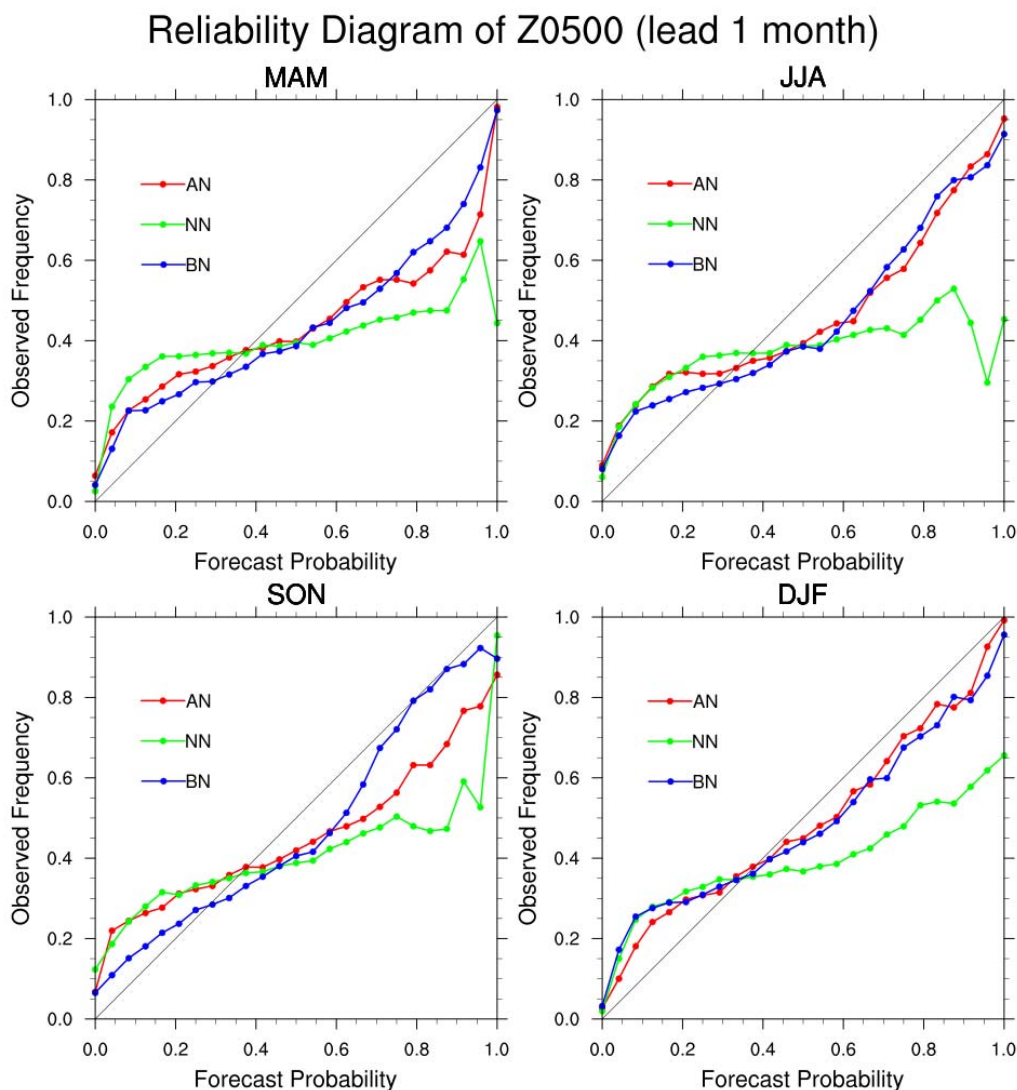


图 34 提前 1 个月起报时 BCC 二代季节模式对全球 500hPa 位势高度预报的可靠性图表

### 3.3.6 Brier 技巧评分

图 35 给出二代模式对 500hPa 位势高度场正异常事件预报的 BSS 的空间分布。与 TCC 类似，该模式对春季和冬季的预报技巧更高，对夏季和秋季预报技巧较低。模式主要在热带地区预报技巧较高。与其他要素类似，随预报 lead time 的延长，模式分辨能力的下降导致 BSS 降低（图 36）。

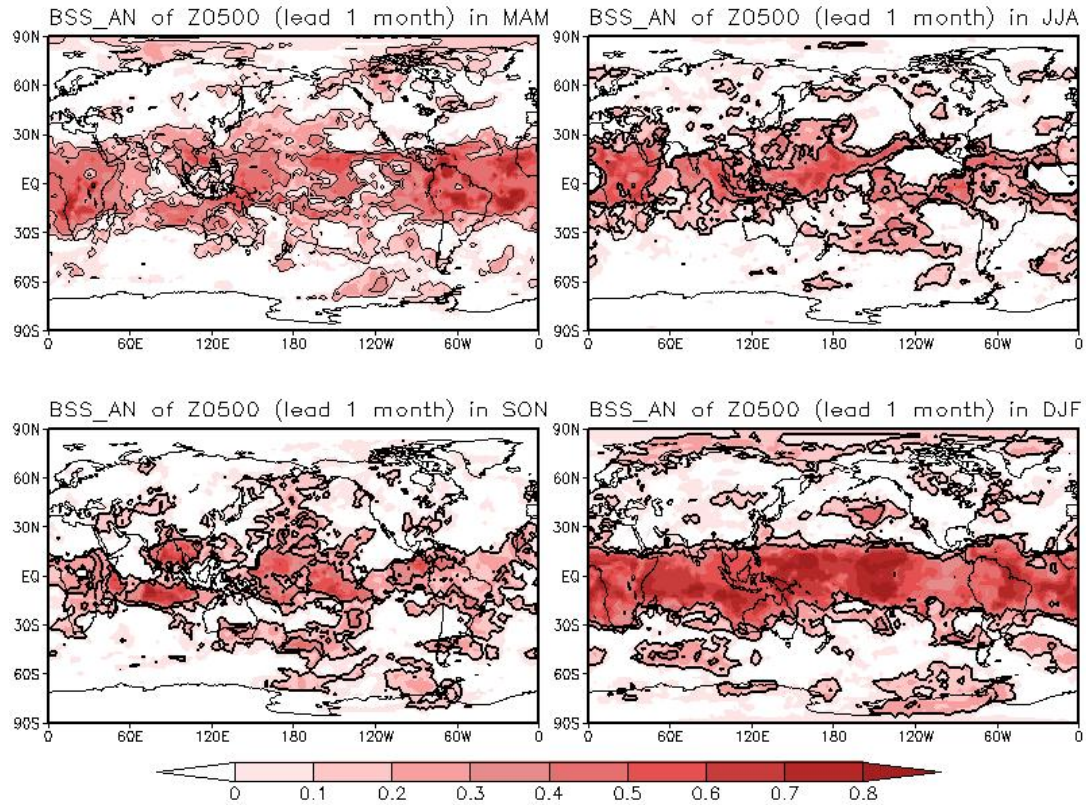
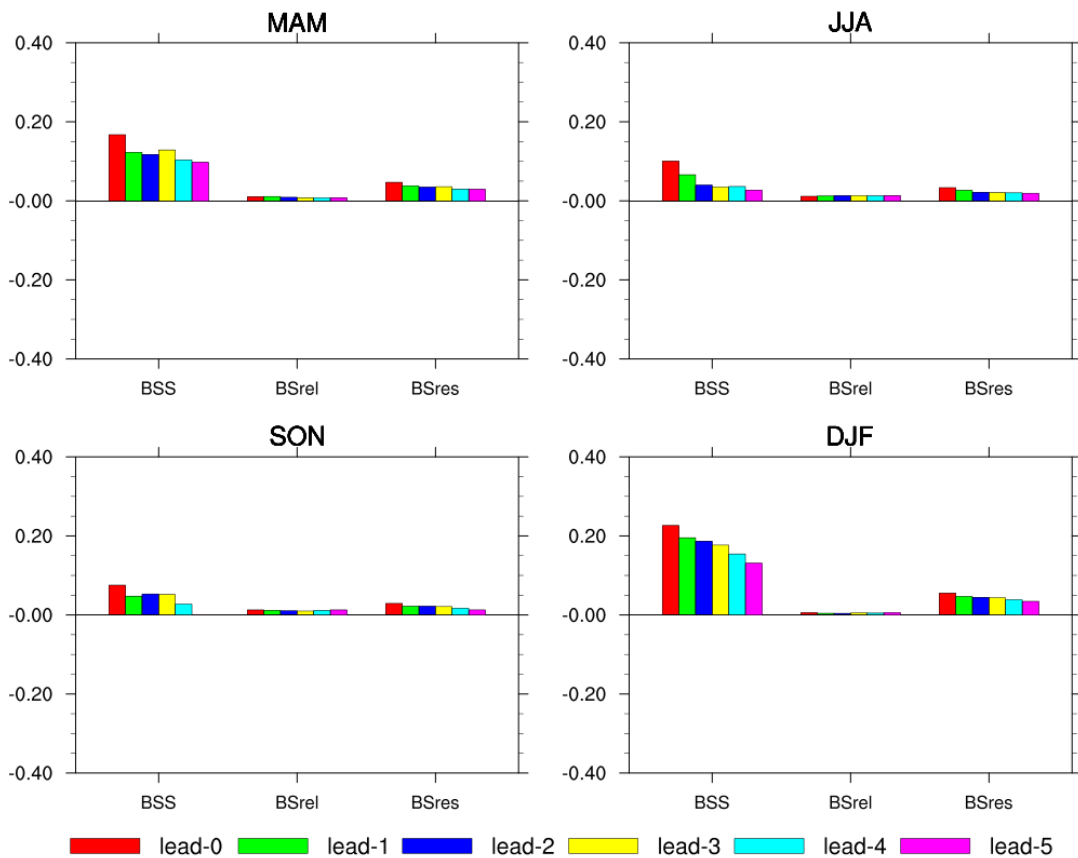


图 35 同图 32，但为正异常事件 (Above Normal) 的 BSS 空间分布

### BSS(AN) of Z500 - Global



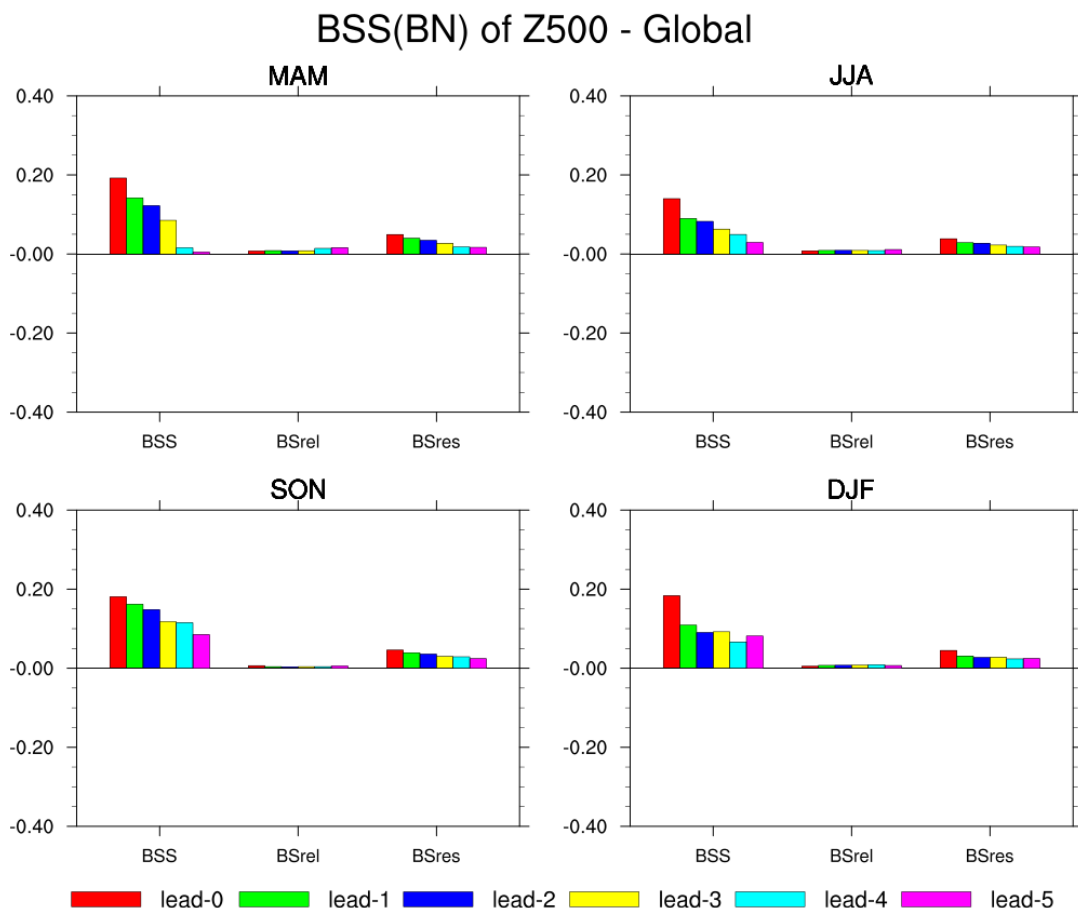


图 36 全球平均的正异常事件（AN，上图）和负异常事件（BN，下图）的 BSS 及可靠性和分辨能力得分

### 3.3.7 二代模式对 500hPa 位势高度场预报小结

通过将 BCC 二代季节模式对 500hPa 位势高度场的预报性能评估和与一代模式的比较，得到以下主要结论：与一代模式相比，二代模式对 500hPa 位势高度场的预报技巧提升明显，但 ACC 技巧评分的年际变化很大，在 ENSO 年预报技巧较高；从 TCC 的空间分布上看，二代模式在热带地区的预报技巧提高最明显，热带外技巧提升的区域在冬季为 PNA、夏季为 EAP 型遥相关波列，表明模式能抓住大气中显著的遥相关波列；从全球平均上看，模式对冬季和春季预报技巧较高，夏季和秋季预报技巧较偏低，但就东亚地区而言，模式在夏季对西太副高影响区域有一定的预报技巧，但在冬季对东亚地区、特别是贝湖到鄂海地区预报技巧偏低；概率预报评估表明，除冬季外，模式对负异常事件的预报技巧高于正异常事件。

## 4 全球气候现象评估

### 4.1 厄尔尼诺和南方涛动（ENSO）

热带太平洋海表温度是衡量气候变率的重要指标，是 ENSO 监测和诊断的基本统计量。本节将首先评估 BCC 二代季节模式（BCC-CSM1.1m）对 ENSO 的预报能力。赤道中、东太平洋被划分为以下 4 个 ENSO 监测区，各海区海表面温度的距平值即表示各海区海边温度特征的指数。ENSO 监测区包括：Niño3 区（5N-5S, 160W-90W），Niño4 区（5N-5S,

160E-150W), Niño3.4 区 (5N-5S, 170W-120W), Niño1+2 区 (0-10S, 90W-80W)。中国气象局国家气候中心业务上以往主要以 NiñoZ 区 (亦称 Niño 综合区, 即 Niño1+2+3+4 区) 的海温距平指数作为判定 ENSO 事件的依据。近期, ENSO 监测业务新标准采用了 Ren 和 Jin(2011) 提出的暖池指数 (NiñoWP) 和冷舌指数 (NiñoCT) 来判别两类 ENSO 事件, 计算公式如下:

$$\begin{cases} N_{CT} = N_3 - \alpha N_4 \\ N_{WP} = N_4 - \alpha N_3, \end{cases} \quad \alpha = \begin{cases} 2/5, & N_3 N_4 > 0 \\ 0, & \text{otherwise.} \end{cases}$$

新的 ENSO 监测业务规范采用上述定义的两类 ENSO 指数(业务上统称为 NiñoT 指数, 即 Niño indices for Two types of ENSO) 作为事件判别指标, 将  $N_{CT}$  指数和  $N_{WP}$  指数 (也曾简记为 CTI 和 WPI) 分别改称为 Niño Eastern-Pacific Index 和 Niño Central-Pacific Index, 简称为 NEPI 和 NCPI, 可称为东部型指数和中部型指数。即 NiñoT 指数 = (NEPI, NCPI)。

由图 37 可见, 对于五个传统的 ENSO 指数 (NiñoZ, Niño3.4, Niño3, Niño4 和 Niño1+2 指数), 从统计的角度可以认为该模式对于 ENSO 预报, 尤其是超前 6-8 个月月的预报结果具有较高的可信度, 预测结果与观测的相关系数维持在 0.6 以上, 相关系数在 0.8 之上的预报时效为 2-4 个月。其中, NiñoZ 指数、Niño3 和 Niño4 指数预报技巧相当, 均略高于 Niño1+2 指数的预报效果。就近 20 年的 Niño3.4 指数预报效果来看, BCC-CSM1.1m 模式 (超前 6 个月相关系数约为 0.68, 接近 0.7) 与国际上 ENSO 预测效果较好的其他模式 (如 NCEP CFSv2, Kumar 等, 2015) 水平接近, 低于多模式集合预报技巧 (Jin 等, 2008)。相比而言, 与 2020 年现代化目标 (超前 6 个月相关系数 0.8) 仍存在一定差距。

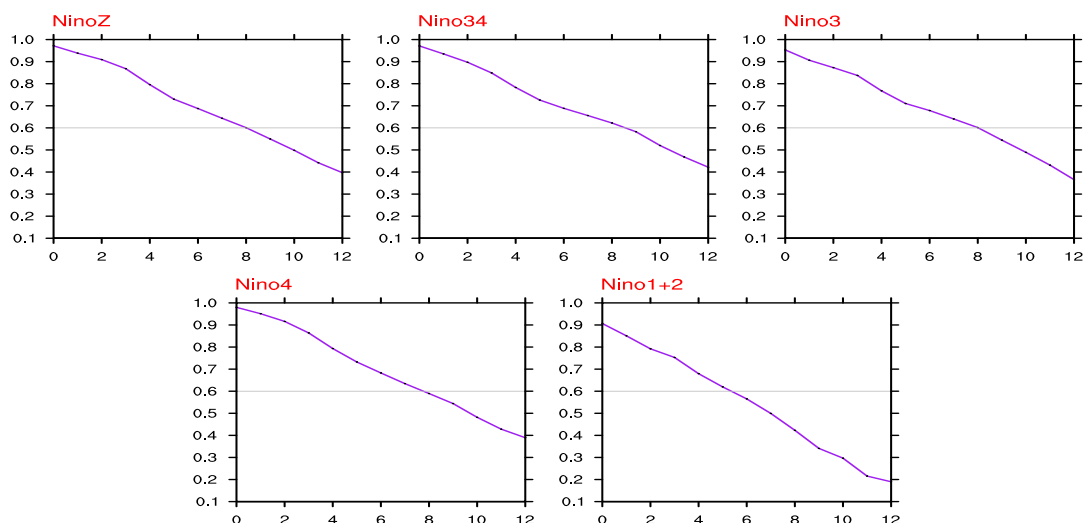


图 37 BCC-CSM1.1m 模式预测的 ENSO 指数与对应观测 (取自 OISST 资料) 的相关系数图。图中, 横坐标对应不同的超前预报时效, 纵坐标为相关系数。其中, 相关系数为 0.6 的等值线用灰线标注。

图 38 中依次给出了 NiñoZ 指数, Niño3.4 指数, 冷舌指数和暖池指数 (本次评估报告依然沿用 NiñoCT 和 NiñoWP 叫法) 的模式预报技巧。可以看出, 对应不同的预测起始月份, 模式表现出不同预报技巧。该模式对 NiñoZ 指数的预报技巧与 Niño3.4 指数相当, 且开始于



北半球 5 月-9 月的 ENSO 预报往往有更高的预报技巧。相比之下，起始于夏季 7 月、8 月的 ENSO 预报对 ENSO 冷舌指数和暖池指数呈现出较好的预报效果，对暖池指数预报效果略低于冷舌指数。

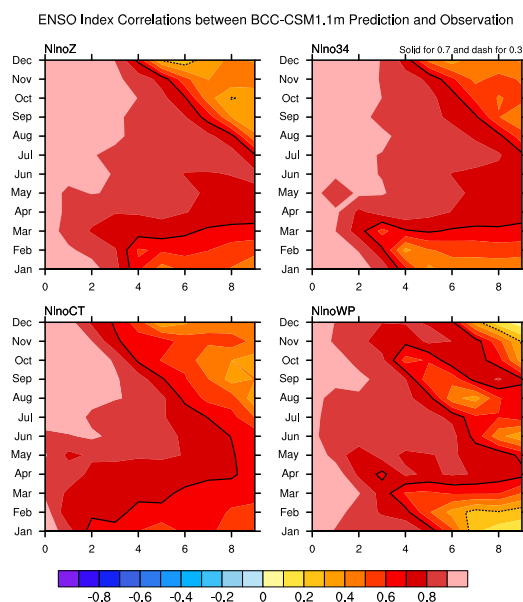


图 38 二代模式不同起报月份预测的 ENSO 指数与对应观测（取自 OISST 资料）的相关系数图。图中纵坐标对应不同起始预报月份，横坐标为预报时长。相关系数为 0.7 和 0.3 的等值线分别用黑实线与黑虚线标注。

## 4.2 印度洋海温异常主模态（IOD 等）

本节将评估 BCC-CSM1.1m 对印度洋海温的预报能力，预报对象为热带印度洋全区一致海温模态（Klein 等，1999）、热带印度洋偶极子（Saji 等，1999）以及副热带南印度洋偶极子（Behera 和 Yamagata, 2001）等三个主要模态。其中，热带印度洋全区一致海温模态（IOBW）定义为热带印度洋（20°S-20°N，40°-110°E）区域平均的海温距平；热带印度洋海温偶极子（TIOD）定义为热带西印度洋（10°S-10°N，50°-70°E）的海温距平与热带东南印度洋（10°S-0°N，90°-110°E）的海温距平差；副热带南印度洋偶极子（SIOD）定义为西南印度洋（45°S-30°S，45°-75°E）与东南印度洋（25°S-15°S，80°-100°E）区域平均海温距平的差值。

IOBM 模态是印度洋海温变率的最主要模态，它通常在冬季开始发展，在春季达到最强。研究表明，IOBM 的发生发展与 ENSO 的强迫关系密切（Klein 等，1999），因此预报的前期信号较强。BCC 二代模式对 IOBM 模态的预报能力非常高，总体上来看（图 39），IOBM 指数的 12 个月预报技巧均在 0.5 以上。对于不同的起报月份而言（图 40），在北半球冬季起报的预报技巧最高，这与 IOBM 通常在冬季发展、春季达到最强的特征是一致的。

IOD 模态是热带印度洋海温变率的第二模态，表现为印度洋海温的东西差异。IOD 表现出显著的季节锁相特征，即在夏季开始发展，秋季达到峰值，冬季快速衰减。总体上看（图 39），BCC 二代模式对 IOD 模态的预报能力一般，在预报 3 个月以后技巧就下降至 0.5 以下；

而国际上的 IOD 预报的预报水平也比较有限，通常对 IOD 指数的预报技巧可以达到 3-4 个月（Li 等，2012）。对于不同的起报月份而言（图 40），BCC 二代模式在北半球夏季起报的预报技巧最高，6 个月的预报技巧超过 0.5，这与 IOD 通常在夏季发展、秋季达到峰值是紧密联系的；另外，BCC 二代模式对 IOD 的预测还呈现出“冬季预报障碍”的特点，国际上的 IOD 预报尝试也指出了同样的特征（Luo 等，2007）。SIOD 模态也有明显的季节位相锁定特征，通常在初冬开始发展，次年 1-3 月达到最强，4 月明显减弱。BCC 二代模式对 SIOD 的模拟水平十分有限，在预报 1-2 个月后相关系数即下降至 0.5 以下。

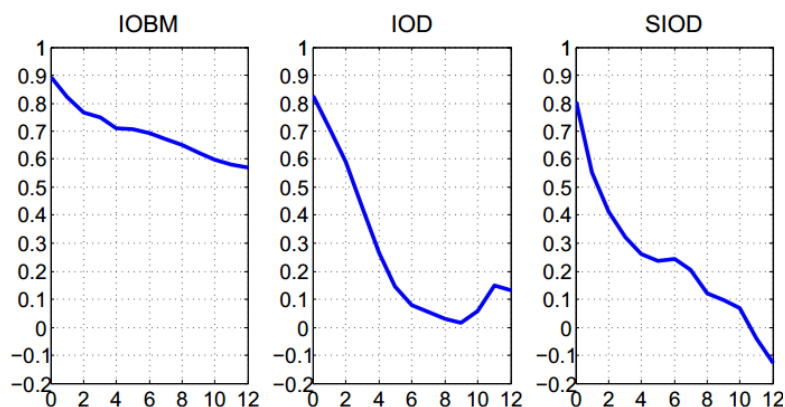


图 39 BCC 二代季节模式针对 IOBW、IOD、SIOD 的预报相关系数总体技巧图

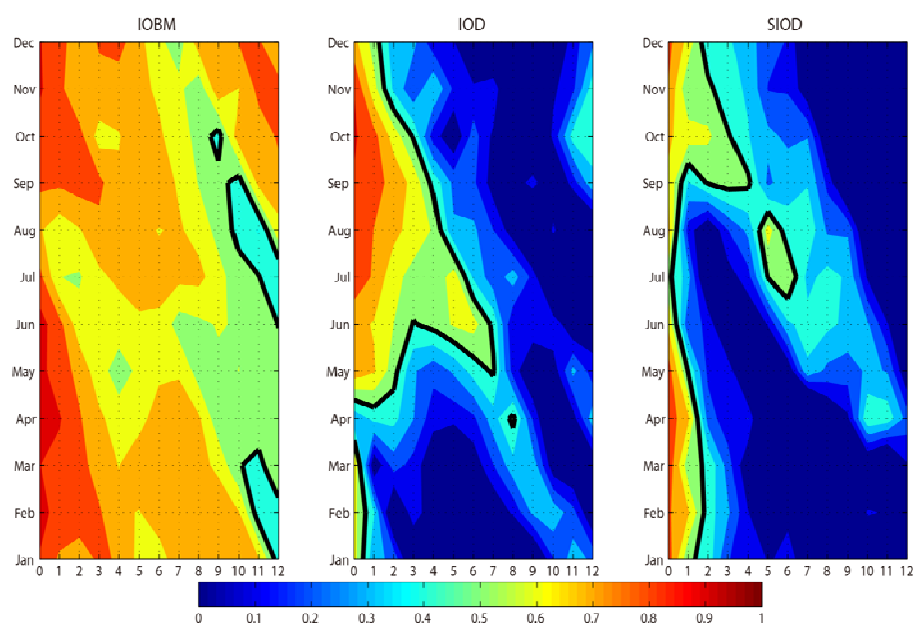


图 40 二代模式不同月份预报的 IOBW、IOD、SIOD 相关系数技巧图，其中粗实线为 0.5 相关系数

### 4.3 北大西洋海温三级子（NAST）

下面对二代季节模式对北大西洋的海温显著模态的预报能力进行评估，选取了北大西洋海温三极子指数（NASTI），其定义为大西洋的中部（34°S-44°N，72°-62°W）海温距平与热带（0°-18°N，46°-24°W）和副极地（44°S-56°N，40°-24°W）海温距平差值（参考 Zuo 等，2013）。这一指数与东亚夏季风指数近 20 年具有较好的同位相关关系。在二代模式对北大西洋海表温度三级子的回报结果中，模式超前一个月的预报结果与观测的相关系数为 0.75，超前两个月时降低为 0.58，起报九个月后其相关系数将无法通过 99% 信度检验（图 41）。

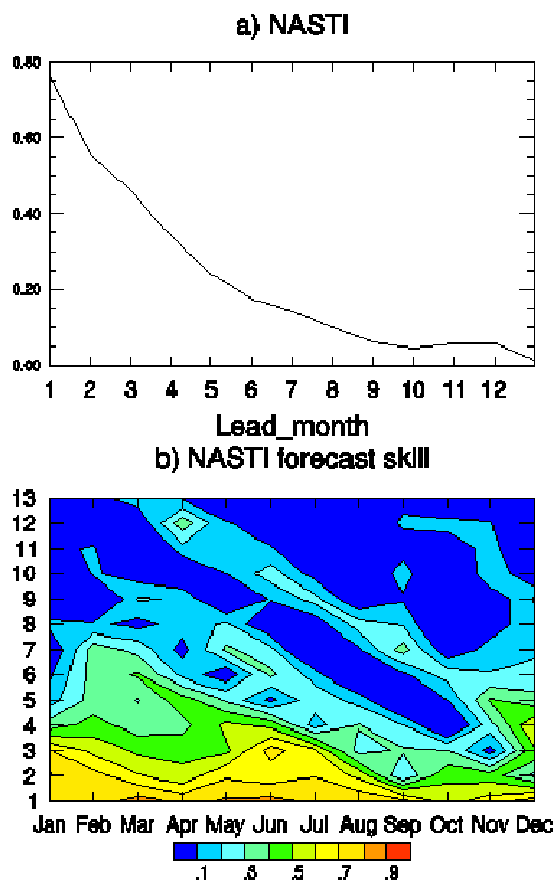


图 41 (a) BCC 二代模式对 NASTI 的预报技巧, (b) 不同起报月对 NASTI 的回报技巧

按照春季预报业务会商的需求,我们还对比了历年 3 月份模式预报和观测的春季和夏季 NASTI 指数。由图 42 和图 43 可见,模式预报的春季四五月份 NASTI 指数与观测相关达到 0.74, 在夏季达到 0.50, 基本能够反映出主要异常事件, 为汛期预测提供了有用参考信息。

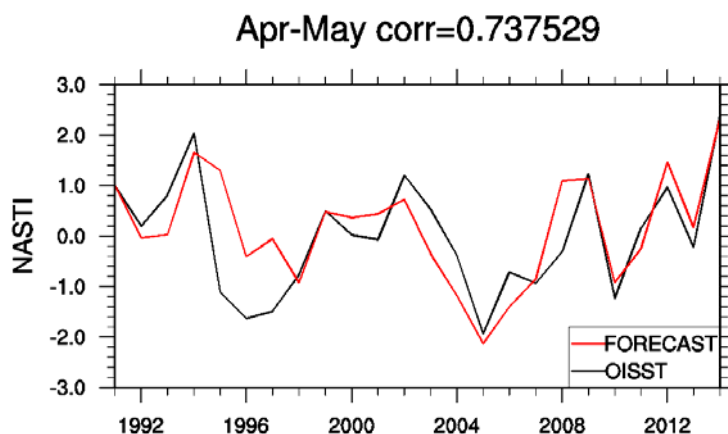


图 42 4-5 月平均的 NASTI 指数和 BCC 二代模式 3 月起报结果时间序列图

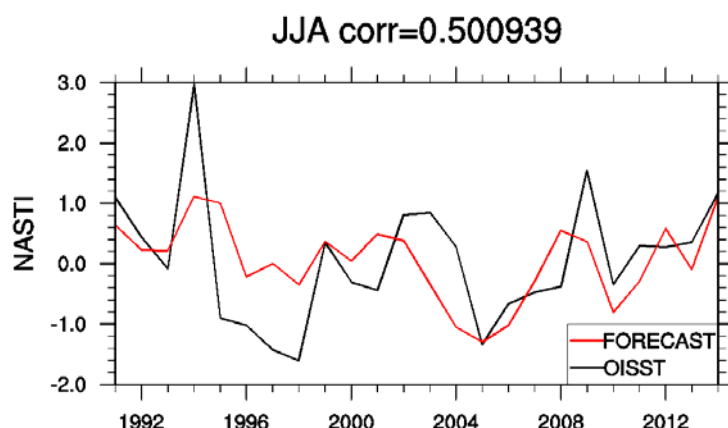


图 43 6-8 月平均的 NASTI 指数和模式 3 月起报结果时间序列图

## 4.4 北极涛动 (AO)

北极涛动 (AO) 是北半球热带外大气低频变率的主导模态。AO 模态通常定义为 20°N 以北逐月海平面气压 (SLP) 距平自然正交函数 (EOF) 分解的第一主模态 (Thompson 和 Wallace, 2000)。在进行 EOF 分解之前, 对各格点数据进行了面积加权处理。EOF 第一模态所对应的标准化时间系数则定义为 AO 指数。为了评估模式对 AO 指数的预报技巧, 首先将模式回报的 SLP 距平投影到观测的 EOF 模态上, 从而得到模式预报的 AO 指数。然后, 计算观测和预报的 AO 指数之间的相关系数, 以此衡量 BCC-CSM1.1m 模式的预报技巧。

图 44 给出了该模式对逐月 AO 指数的预报性能。Lead=0 时, 最大相关技巧为 0.75, Lead=1 时, 最大相关技巧为 0.41。模式只对 12 月 AO 指数存在一定预报技巧, 如以 0.2 为标准, 可用的有技巧预报长度约为 3 个月, 其它月份则几乎无明显预报技巧, 如以 0.5 为标准, 则该模式基本上无技巧。从图 45 可以看到, 该模式对冬季平均 AO 指数也几乎无明显预报技巧。目前, 国际上多个先进的季节动力预报系统对冬季平均 AO 指数已表现出有用的预报技巧 (高于观测的持续性预报) (Riddle 等, 2013; Kang 等, 2014; Scaife 等, 2014)。

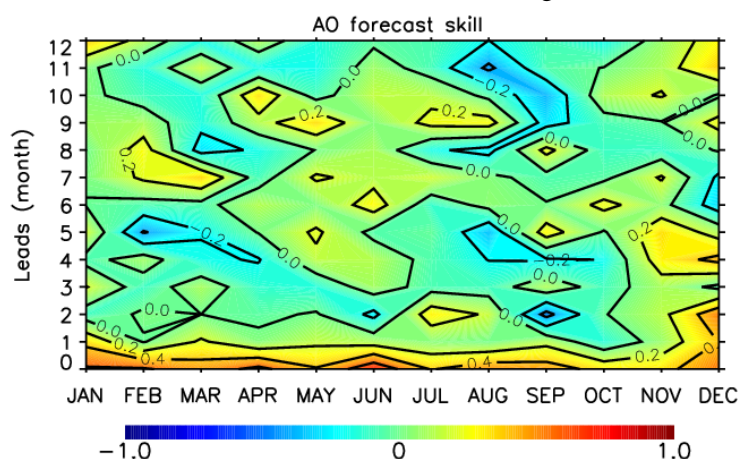


图 44 模式对逐月 AO 指数的预报技巧。图中, 横坐标对应不同起始预报月份, 横坐标为预报时长



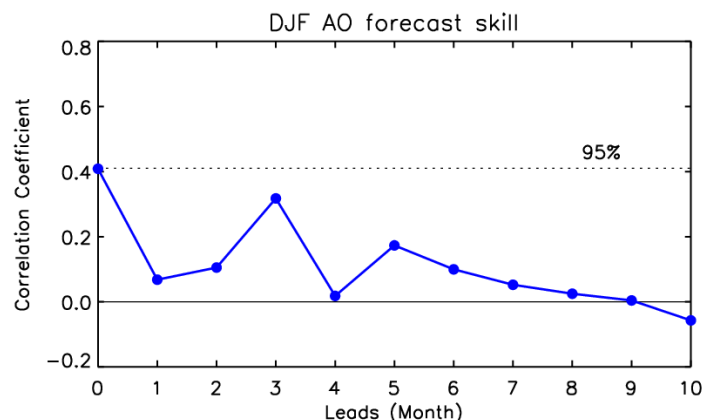


图 45 模式对冬季 AO 指数的预报技巧

#### 4.5 夏季西太副高和东亚冬、夏季风

首先，对 BCC-CSM1.1m 模式中西太平洋副热带高压指数的 1991-2014 年 24 年回报进行评估。这里选用的西太平洋副热带高压指数定义为夏季平均  $15^{\circ}\text{N}$ – $25^{\circ}\text{N}$ ,  $115^{\circ}\text{E}$ – $150^{\circ}\text{E}$  区域上空 850hPa 高度场的标准化值(Wang et al. 2013)。图 46 中横坐标为 6、5、4、3、2、1、12、11、10、9、8 月起报的对应未来 0-10 个月的 TCC。由图中结果可以看到，该模式对夏季西太副高预测的 TCC 较高，lead time=1 时 TCC 达到 0.8；lead time=6 时 TCC 达到 0.45 以上，通过 0.05 信度检验水平。随着 lead time 的增加，TCC 下降不明显，表现较稳定。

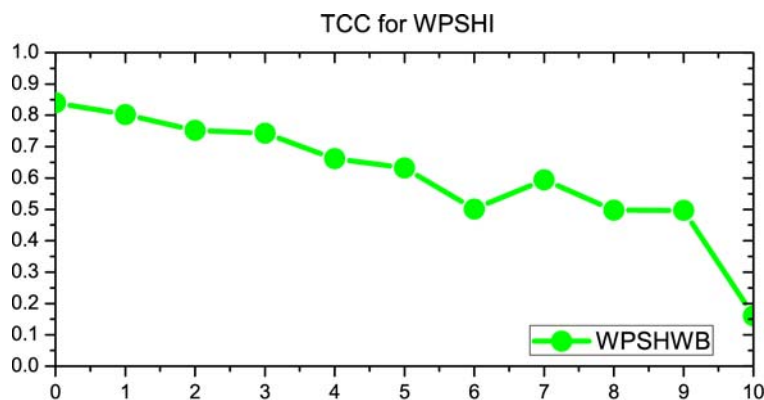


图 46 模式对未来 0-10 个月的西太副高指数预报的 TCC，横坐标为预报超前于夏季的 Lead month

其次，对该模式中东亚夏季风指数的回报进行评估。这里东亚夏季风指数定义为东亚热带季风槽区( $10^{\circ}$ – $20^{\circ}\text{N}$ ,  $100^{\circ}\text{E}$ – $150^{\circ}\text{E}$ )与东亚副热带季风槽区( $25^{\circ}$ – $35^{\circ}\text{N}$ ,  $100^{\circ}\text{E}$ – $150^{\circ}\text{E}$ )6-8 月平均的 850hPa 风场的纬向风距平差(张庆云等, 2003)。图 47 中横坐标为 6、5、4、3、2、1、12、11、10、9、8 月起报的对应未来 0-10 个月的 TCC。由图中结果可以看到，该模式对东亚夏季风的预测 TCC 随着 lead time 的增加表现较稳定，预测性能良好。lead time=1 时，TCC 达到 0.7 以上；lead time=6 时，TCC 达到 0.5，通过 0.01 信度检验水平。同时也给出了使用另外两个夏季风指数的评估结果，同样可以看到，二代模式总体上对于东亚夏季风指数强度的提前 3 个月仍具有一定把握能力，特别是对 Wang 和 Fan (1999) 指数提前 8 个月预报仍能达到 0.6 以上，但施能等 (1996) 指数预报技巧偏低。

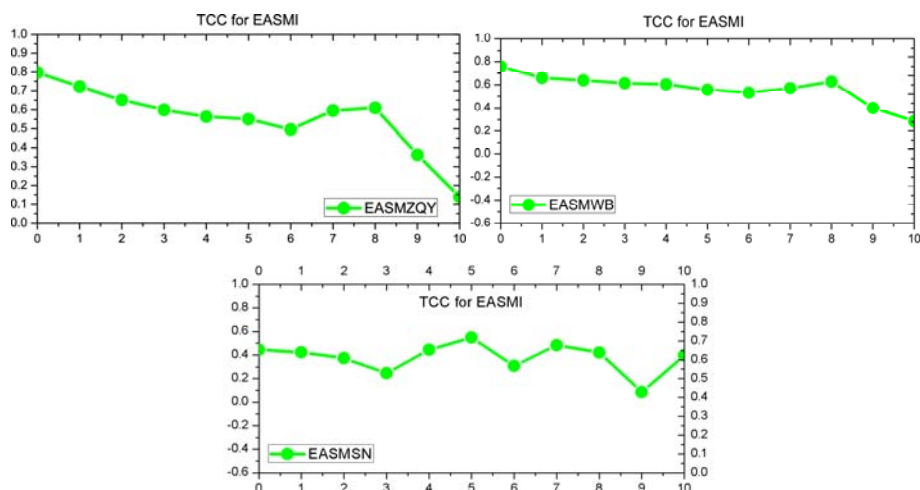


图 47 二代模式对于夏季风指数分别提前 0-10 个月预报的 TCC 评分，左上：张庆云等（2003）指数；右上：Wang 和 Fan（1999）；下：施能等（1996），横坐标为预报超前于夏季的 Lead month

最后，对该模式中东亚冬季风指数的 1991/92-2013/14 年共 23 个冬天（DJF 平均）历史回报进行评估。东亚冬季风指数定义选用用西伯利亚高压附近 3 点(60°N, 100°E、60°N, 90°E、50°N, 100°E)的海平面平均气压距平来表示（郭其蕴，1994）。图 48 中横坐标为 12、11、10、9、8、7、6、5、4、3、2 月起报的对应未来 0-10 个月的 TCC。该模式对东亚冬季风的预报技巧在 lead time 较小时量值很小，lead time=4 时，模式的 TCC 接近于 0。同时评估了另两个常用冬季风指数技巧，得到了较为一致的结论：该模式对冬季风基本无预报能力。

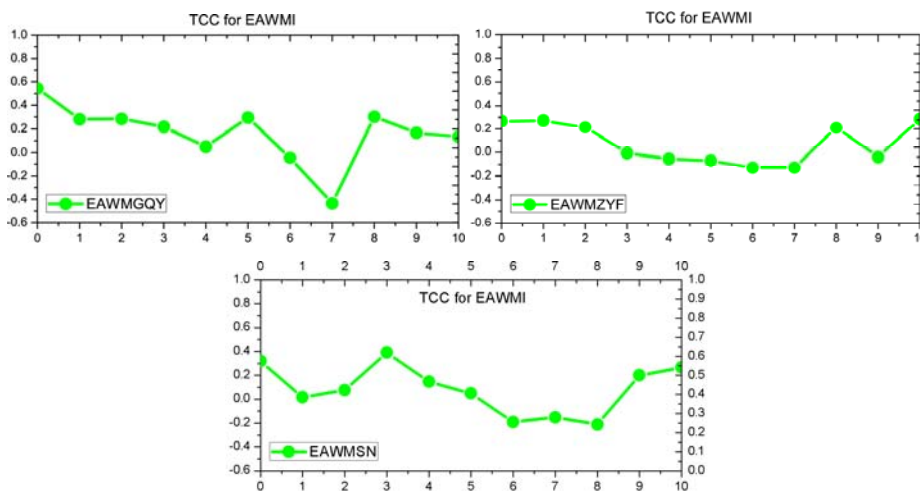


图 48 二代模式对于夏季风指数分别提前 0-10 个月预报的 TCC 评分，左上：郭其蕴（1983）指数；右上：朱艳峰等（2008）；下：施能等（1996），横坐标为预报超前于冬季的 Lead month

由于夏季西太副高指数、东亚冬夏季风指数的定义种类繁多，导致评估结果可能有较大差异，我们尽可能选择业务常用的指标做评估，这里结果也仅供参考。

## 5 我国台站气温和降水预报评估

### 5.1 气温

图 49 和图 50 分别给出 BCC 二代模式和一代模式对我国 160 站站点气温预报 TCC 技巧

的空间分布。由图可见，二代模式对春季新疆北部、黄河中游地区，秋季和冬季的河套地区以及秋季的长江中下游地区的气温异常的预报技巧能够通过显著性检验，对夏季气温预报技巧分布与春季类似、但数值明显偏低，相比之下，秋季正技巧的范围和数值都是最高的，并在华北北部到西北地区东部以及江南北部都能达到显著，但冬季技巧整体偏低。与一代模式相比，二代模式对我国东部地区的预报技巧有明显提升。需要注意的是，模式在冬季对我国东北地区气温的预报能力较弱，这也与第 3.1 节的评估结果相一致。

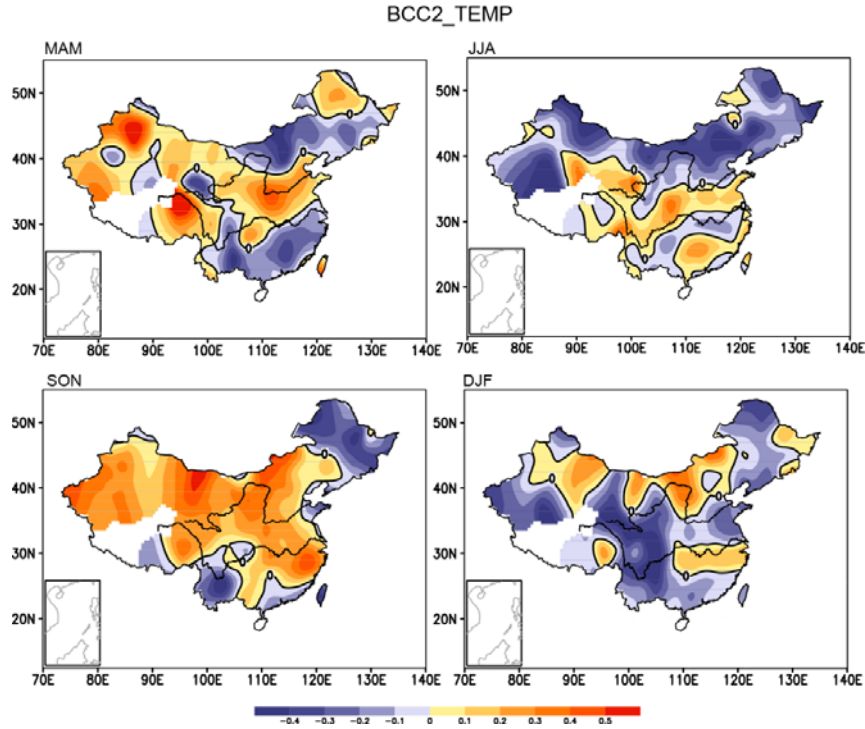


图 49 提前 1 个月 BCC 二代季节模式对我国 160 站站点温度预报的 TCC 技巧的四个季节的空间分布图，达到 95% 信度统计显著性的相关系数大约为 0.4

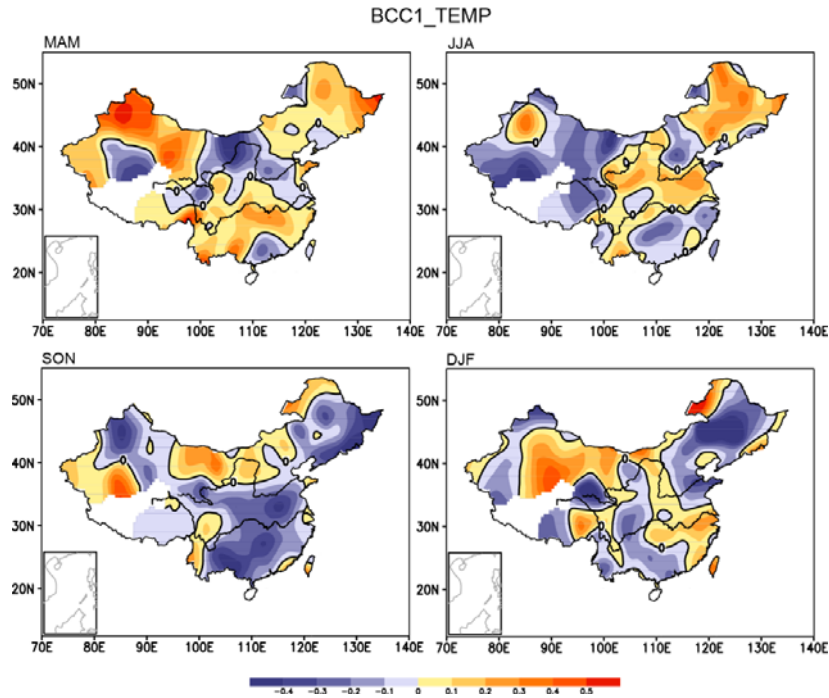


图 50 同图 49，但为 BCC 一代模式结果

## 5.2 降水

图 51 和图 52 分别给出 BCC 二代模式和一代模式对我国 160 站点降水预报 TCC 技巧的空间分布。由图可见，二代模式对春季我国长江以南和西北地区地区的降水有不错的预报技巧，但对夏季降水整体预报能力偏低，秋冬季在我国南方地区有一定预报效果，但显著性不理想。总体上，与一代模式对我国东部大部分地区的降水预报性能相比，有了明显提升。

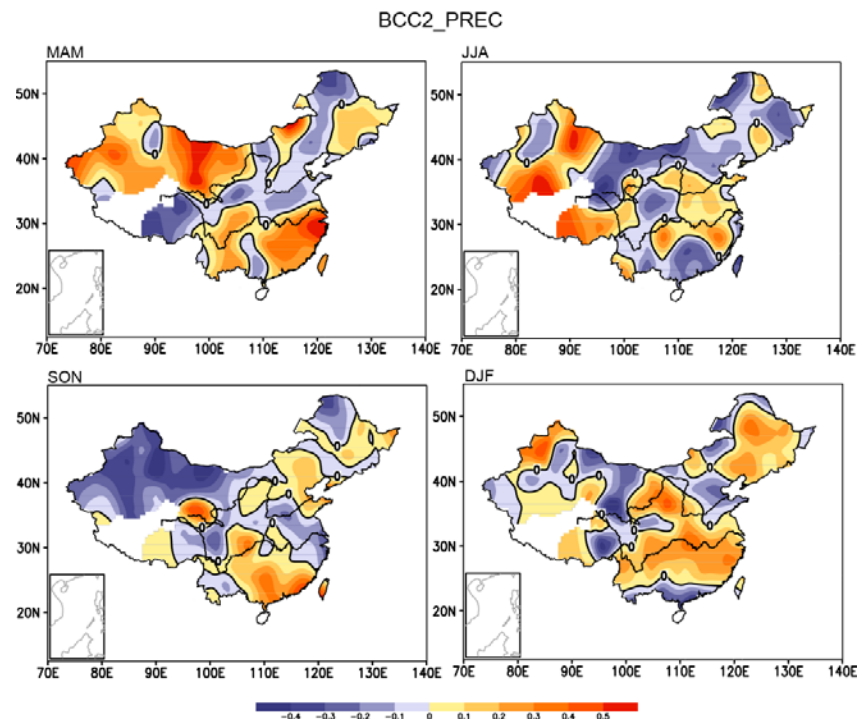


图 51 同图 49，但为我国 160 站点降水预报技巧分布

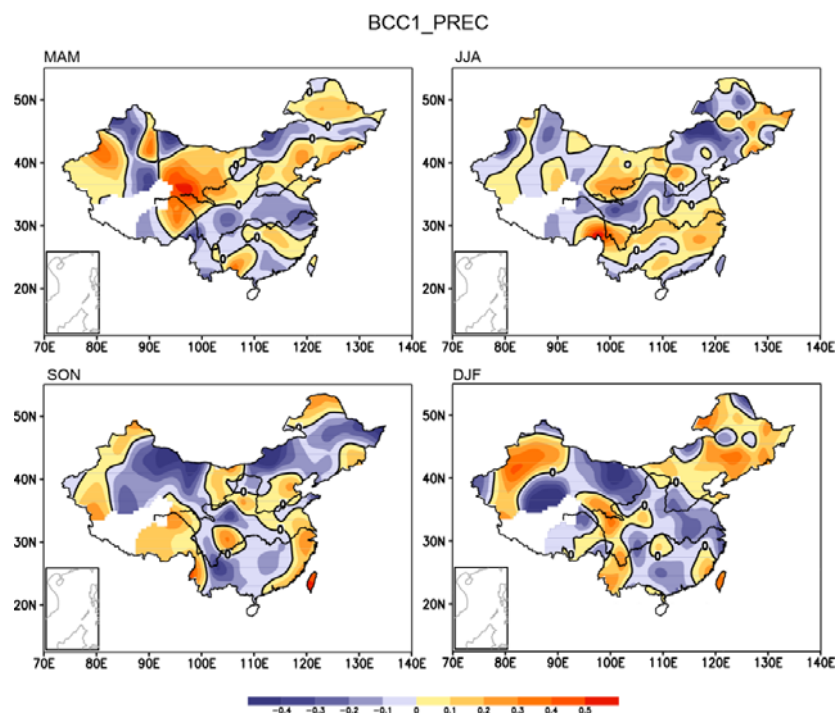


图 52 同图 51，但为 BCC 一代模式结果



## 6 总结

本报告完成了针对国家气候中心 (BCC) 二代季节预测模式对我国 160 站站点和全球格点 2m 温度、降水、环流以及全球主要气候现象代表性指标的历史回报综合检验评估。

得到以下主要结论:

1) BCC 二代模式对全球 2m 温度、降水、环流的预报性能均明显优于 BCC 一代模式,特别是热带地区降水预报有明显改进,在技巧的空间分布上,TCC/MSSS/ROCA/BSS 等检验方法的评估结果在很大程度上具有一致性,即模式对热带地区的预报技巧更高;从不同季节上看,模式对冬季和春季的预报技巧较高,夏季和秋季偏低。

2) 与一代模式相比,二代模式在对于低纬度地区的预报改进最为显著,热带太平洋 NINO 区预报技巧改进明显,模式在热带外地区预报技巧及其改进也基本与热带强迫源(如 ENSO)所激发的遥相关(如 PNA 型)有关;在 ENSO 事件发生的当年秋冬和次年春夏季,模式整体的预报技巧明显高于其他季节。这些反映出 ENSO 现象及其热带外遥相关型预报效果的改善是提升二代模式季节气候预测水平的根本所在。

3) 从概率预报上看,模式通常对负异常事件的预报技巧略高于正异常事件,二者都明显高于那些接近正常的事件;从可靠性上看,模式普遍存在低概率低估和高概率高估的特点;BSS 的各项分析表明,模式对负异常事件的概率预测效果略好于正异常事件;而且,随着预报时间的延长,模式概率预报技巧的下降主要来源于模式分辨能力的下降。这反映出该模式还需要在集合成员构建方面进行有针对性的改进。

4) 二代模式对东亚大陆地区预报能力较弱,其中对各个季节平均降水 TCC 技巧偏低,冬季略好;但是该模式对夏季 EAP 型等遥相关型有一定的把握能力,因此对夏季东亚地区环流(如西太副高)预报技巧较高;在冬季,模式对贝加尔湖到鄂霍茨克海的中高纬度环流预报技巧偏低,进而导致二代模式对冬季我国东北气温预报技巧偏低,这可能会严重影响到该模式对于的东亚冬季风的预测效果。

5) 从全球气候现象预报来看,二代模式对 ENSO (NINO3.4 指数)提前 6 个月预报的相关技巧接近 0.7,与 CFSv2 基本相当,但比国际上其他先进模式略偏低。以 0.5 相关为标准,对热带印度洋海温偶极子 (IOD) 的预报技巧约为 2-3 个月,略低于国际先进水平 (3-4 个月);对北大西洋海温三极子指数 (NASTI) 的预报技巧约为 3 个月,其中春季会商时模式预报的夏季三级子指数与实况相关系数达到 0.50;模式只对 12 月起报的北极涛动 (AO) 指数存在一定预报技巧(约 3 个月,以 0.2 相关为标准),其它起报月份 AO 指数几乎无明显预报技巧。仍以 0.5 为标准,模式对东亚夏季风有一定预报技巧,对西太副高指数和东亚夏季风指数的预报技巧可达 6-7 个月,但对东亚冬季风预报技巧较低,仅为 3 个月左右。

6) 从对中国台站温度和降水 TCC 检验来看,二代模式对春季我国很多地区有正的预报技巧,特别在长江中下游及其以南部分地区能够通过显著性检验;对夏季气温和降水预报技巧总体偏低,前者略高些;秋季对气温的预报技巧比夏季要好些;相比之下,对冬季气温的预报技巧整体不高、与夏季相当,但对南方降水有一定预报效果。总体上,二代模式与一代模式相比,对我国气候预报性能有明显提升。

最后,为了给出直观对比,我们制作了 BCC 一代和二代季节预报模式对中国温度、降水和环流的评分表格(见表 3)。与一代模式相比,BCC 二代模式对我国站点气温、降水及全球环流的预报性能在一定程度上优于一代模式(唯独夏季降水略有下降);整体上,二代模式对我国 160 站降水 PS 评分提高有限,但对我国气温 PS 评分则提升明显;相比之下,ACC 评分均由负值提高至 0.03,但依然数值较小;从全球来看,二代模式对 500hPa 位势高度 ACC 评分从 0.03 提高至 0.16,改进幅度较大;季节对比来看,冬季技巧提升最显著。

表 3 BCC 一代和二代季节预测模式对中国温度、降水和环流的预报评分

		平均	春季	夏季	秋季	冬季
全国 160 站降水 (PS 评分)	一代	64.2	65.6	66.4	62.9	62.0
	二代	65.8	68.2	65.3	64.8	64.8
全国 160 站温度 (PS 评分)	一代	66.7	66.5	65.1	68.1	66.9
	二代	71.3	68.1	73.1	71.6	72.3
全国 160 站降水 (ACC 评分)	一代	-0.02	-0.01	0	-0.02	-0.04
	二代	0.03	0.07	-0.04	0.04	0.04
全国 160 站温度 (ACC 评分)	一代	-0.02	0	-0.03	-0.05	0.01
	二代	0.03	0.03	0.02	-0.03	0.1
全球 500hPa 高度场 (ACC 评分)	一代	0.03	0.08	-0.01	0.01	0.05
	二代	0.16	0.17	0.10	0.11	0.26

## 参考文献

- Behera S K, Yamagata T. 2001. Subtropical SST dipole events in the southern Indian Ocean. *Geophys Res Lett*, 28(2): 327-330.
- 郭其蕴. 1994: 东亚冬季风的变化与中国气温异常的关系. *应用气象学报*, 5(2): 218-225.
- Li S, Hendon H H, Alves O, et al. 2012. How Predictable is the Indian Ocean Dipole? *Mon Wea Rev*, 140: 3867-3884.
- 李晓燕, 翟盘茂. 2000. ENSO 事件指数与指标研究. *气象学报*, 58(1): 102-109.
- Luo J J, Masson S, Behera S, et al. 2007. Experimental forecasts of the Indian Ocean dipole using a coupled OAGCM. *J Clim*, 20: 2178-90.
- Jin E K, Kinter J L, Wang B, et al. 2008. Current status of ENSO prediction skill in coupled ocean-atmosphere models. *Clim Dyn*, 31 (6): 647-664.
- Kang D, Lee M I, Kim J I D, et al. 2014. Prediction of the Arctic Oscillation in boreal winter by dynamical seasonal forecasting systems. *Geophys Res Lett*, 41: 3577-3585.
- Klein S A, Soden B J, Lau N C. 1999. Remote sea surface temperature variations during ENSO: Evidence for a tropical atmospheric bridge. *J Clim*, 12: 917 - 932.
- Kumar A, Chen M Y, Xue Y. 2015. An analysis of the temporal evolution of ENSO prediction skill in the context of the equatorial Pacific ocean observing system. *Mon Wea Rev*, 143: 3204-3213.
- Ren H L, Jin F F. 2011. Niño indices for two types of ENSO. *Geophys Res Lett*, 38, L04704.
- Riddle E E, Butler H A, Furtado J C, et al. 2013. CFSv2 ensemble prediction of the wintertime Arctic Oscillation. *Clim Dyn*, 41(3): 1099-1116.
- Saji N H, Goswami B N, Vinayachandran P N, et al. 1999. A dipole mode in the tropical Indian Ocean. *Nature*, 401: 360-363.
- Scaife A A, et al. 2014. Skillful long-range prediction of European and North American winters. *Geophys Res. Lett*, 41: 2514-2519.
- Thompson D W J, Wallace J M. 2000. Annular modes in the extratropical circulation. Part I : month-to-month variability. *J Clim*, 13: 1000-1016.
- Wang B, Li J Y, Kang I S, et al. 2009. Advance and prospectus of seasonal prediction: assessment of the APCC/CliPAS 14-model ensemble retrospective seasonal prediction (1980-2004). *Clim Dyn*, 33: 93-117.
- Wang B, Xiang B, Lee J Y. 2013. Subtropical high predictability establishes a promising way for monsoon and tropical storm predictions. *Proceedings of the National Academy of Sciences* 110(8): 2718-2722.
- WMO. 1999. Abridged final report with resolutions and recommendations. Secretariat of the WMO, Geneva, Switzerland, No.896: 154-157.
- WMO. 2006. Standardised Verification System (SVS) for Long-Range Forecasts (LRF): New attachment II-8 to the manual on the GDPFS (WMO-No. 485), Volume I.
- 张庆云, 陶诗言, 陈烈庭. 2003: 东亚夏季风指数的年际变化与东亚大气环流. *气象学报*, 61(4): 559-568.
- Zuo J Q, Li W J, Sun C H, et al. 2013. Impact of the North Atlantic seasurface temperature tripole on the East Asian summer monsoon. *Adv Atmos Sci*, 30(4): 1173-1186.
- 施能, 鲁建军, 朱乾根, 1996: 东亚冬、夏季风百年强度指数及其气候变化, *南京气象学院学报*, 19:168-177.
- 朱艳峰, 2008, 一个适用于描述中国大陆冬季气温变化的东亚冬季风指数, *气象学报*, 66(5), 781-788.
- Wang. B. and Z. Fan, 1999: Choice of South Asian Summer Monsoon Indices. *Bull. Amer. Meteor. Sci.*, 80: 629-638.