

預報成員選取對貝氏模型平均法預報的影響

張語軒、張庭槐、吳蕙如
中央氣象局

摘要

貝氏模型平均 (Bayesian model averaging, 簡稱BMA) 是利用統計方法將多組模式的預報及觀測資料進行後製處理, 建立符合真實天氣狀況的預測機率密度函數 (probability density function, 簡稱PDF)。

學者研究顯示, 利用BMA方法預報產生的系集產品, 在臺灣地區測站溫度之機率預報上能有效掌握其變化的不確定性, 建模過程採用分區域資料分別建立各區域BMA模型的預報能力明顯優於採用全區資料所建立的統一模型。本文將深入探討不同預報成員的選取對BMA預報的影響, 進而探討如何挑選預報成員可使BMA預報的表現最佳。

關鍵字: 貝氏模型平均、機率密度函數、系集預報

一、前言

我們永遠無法得知未來真正的天氣狀態, 天氣的演變充滿著許多不確定的因素, 要得到一個完全準確的決定性預報, 是不可能達成的事情, 因此我們期望能藉由系集預報的方法呈現數值模式預報之不確定性。然而, 如何透過預報的結果將不確定性表達出來, 是一個困難卻又重要的問題。若能從預報系統中建構出預測值的機率分佈, 是探究不確定性很好的一個方式, 透過此機率分佈可以很清楚地了解到未來觀測值所有可能發生的範圍及其機率, 進而能預測出更精確的氣象預報。

Hoeting等人 (Hoeting et al., 1999) 提出的貝氏模型平均 (Bayesian Model Averaging, 簡稱BMA) 法, 可應用在氣象預報上, 將多組模式的預報及觀測資料進行後製處理, 建立出各項天氣變數的預測機率分布。例如Reftery et al. (2005) 將BMA應用在美國西南部地面氣溫48小時預報的資料, 而在張等人 (2012) 也將此方法應用在臺灣地區的溫度機率預報上, 研究結果顯示BMA能有效掌握模式預報的不確定性。

在BMA模型下預測PDF的公式為

$$p(y | f_1, f_2, \dots, f_K) = \sum_{k=1}^K w_k g_k(y | f_k), \quad (1)$$

其中 y 為預測的天氣變數, 而 f_1, f_2, \dots, f_K 為 K 個預報成員, w_k 為第 k 個成員的後驗機率 (posteriori probability), 代表在 K 個預報成員中, 第 k 個成員預報最佳的機率, 其滿足

$\sum_{k=1}^K w_k = 1$ 。而 $g_k(y | f_k)$ 為條件PDF, 代表在第 k

個成員預報最佳且其預報值為 f_k 的條件下, y 的PDF。當預報變數為溫度時, $g_k(y | f_k)$ 假設為常態分布, 期望值為 $a_k + b_k f_k$, 標準差為 σ , 其中 a_k 和 b_k 是對預報值進行偏差調整時採用的參數。若用機率符號表示則是

$$y | f_k \sim N(a_k + b_k f_k, \sigma^2)。$$

關於BMA模型中參數的估計方式, a_k 和 b_k 使用最小平方法, 而 w_k 和 σ 則使用EM (expectation-maximization) 演算法, 參數估計的詳細方法可參考Reftery et al. (2005)。

由於BMA是對 f_1, f_2, \dots, f_K 共 K 個預報成員做後製處理, 因此預報成員的挑選會影響到BMA的預報能力, 而本文的主要目的是希望比較不同預報成員下BMA模型的預報能力表現。由於在張等人 (2012) 中提到分區BMA模型在臺灣地區的預報能力明顯優於全區BMA, 因此本文皆採用分區BMA模型來做比較。

二、預報模式介紹

本文使用的資料為臺灣地區16個綜觀測站的地面溫度觀測值及本局區域模式內插至觀測點上的第24小時預報值, 初始時間為每天00Z, 測站包含淡水、台北、竹子湖、基隆、花蓮、澎湖、台南、高雄、嘉義、台中、阿里山、玉山、新竹、蘭嶼、日月潭、台東, 採用資料的期間為2011年07月11日到2012年04月17日, 使用的區域模式預

報資料包含2組非靜力預報系統（NFS，Non-hydrostatic Forecast System），分別命名為NFS_M00及NFS_M02、8組天氣研究與預報模式（WRF，Weather Research & Forecast model），共10組預報成員。其中8組WRF中有2組屬於中央氣象局WRF資料同化系統主作業成員，分別命名為WRF_M00及WRF_M02，另外6組為WEPS（WRF ensemble prediction system，區域系集預報系統）20組的預報中較具代表性的成員，依序命名為WEPS_E01~ WEPS_E06。WEPS是以WRF為基礎並搭配邊界擾動、初始場擾動和物理參數法擾動所建立之區域系集預報系統（李等人，2011）。

三、BMA模型比較

本文比較的準則，主要採用continuous ranked probability score（CRPS）與平均絕對誤差（mean absolute error，簡稱MAE）兩種校驗方法來做比較。CRPS（Hersbash，2000）是用來校驗預測CDF（cumulative distribution function），計算方法是使用布賴爾評分（Brier score）對所有門檻值積分，CRPS的公式為

$$CRPS = \int_{-\infty}^{\infty} [P(x) - O(x)]^2 dx, \quad (3)$$

其中 $P(x)$ 為預測CDF，而

$$O(x) = \begin{cases} 0, & x < x_o \\ 1, & x \geq x_o \end{cases},$$

x_o 為觀測值。CRPS可視為預測CDF準確性的得分，其值愈低代表預測CDF越準確。而MAE是用來校驗決定性預報（本文中採用BMA預報的中位數作為決定性預報）與觀測值之間的誤差，不同於CRPS校驗的對象為預測機率分布。MAE的公式為

$$\frac{1}{n} \sum_{i=1}^n |f_{di} - y_i|,$$

其中 f_{di} 為決定性預報而 y 為觀測值，其值越低越好，關於校驗方法可參考張等人（2012）更詳細的說明。

首先，我們利用全部10組預報成員資料來建立BMA預報模型（第一組BMA模型），並依照張等人（2012）的挑選方式進行評估後決定訓練期為25天。在此模型下預報能力的校驗結果如表一所示，此數據將在之後與其他組BMA模型比較。除了觀察BMA模型的預報能力，我們還可以觀察每一個預報成員在BMA模型中的重要性。在BMA模型參數中， w_k 代表的是第 k 個預報成員預報最佳的機率，亦是第 k 個預報成員在模型中的權

重，因此我們可藉由觀察 w_k 來分析每一個預報成員的重要性。表二為 w_k 的統計分析結果，表中列出每一個預報成員其 w_k 的第一四分位數（ Q_1 ）、中位數（ Q_2 ）、第三四分位數（ Q_3 ）及平均值（mean），藉由這些數字來了解 w_k 的分佈情形。若比較 w_k 的 Q_3 及平均值，可發現WEPS成員的權重明顯低於另外4組成員，表示WEPS成員在此BMA模型中的重要性相對較低。

	CRPS	MAE
第一組BMA模型	0.7221	0.9969

表一：第一組BMA模型的校驗結果。

第一組BMA模型預報成員	Q_1	Q_2	Q_3	mean
NFS_M00	0.0000	0.0000	0.2398	0.1440
NFS_M02	0.0000	0.0000	0.2340	0.1409
WRF_M00	0.0000	0.0000	0.1755	0.1283
WRF_M02	0.0000	0.0532	0.3679	0.2061
WEPS_E01	0.0000	0.0000	0.0012	0.0619
WEPS_E02	0.0000	0.0000	0.0094	0.0654
WEPS_E03	0.0000	0.0000	0.0000	0.0621
WEPS_E04	0.0000	0.0000	0.0001	0.0683
WEPS_E05	0.0000	0.0000	0.0000	0.0568
WEPS_E06	0.0000	0.0000	0.0390	0.0663

表二：第一組BMA模型中各預報成員的 w_k 參數統計數據。表中包含了第一四分位數（ Q_1 ）、中位數（ Q_2 ）、第三四分位數（ Q_3 ）及平均值（mean）。

根據上述的分析，接著我們試圖將WEPS從預報成員中刪除，使用剩下的4組預報成員建立BMA模型（第二組BMA模型），第二組BMA模型所採用的訓練期同樣為25天。第二組BMA模型的校驗結果如表三所示，將表三與表一比較，顯示出兩組BMA的校驗結果非常接近，差距微小，表示兩組BMA模型的預報能力差異不大。表四為第二組BMA模型中 w_k 的統計分析結果，相對於第一組BMA，第二組BMA模型中每一個成員的 w_k 彼此之間較接近，表示在第二組BMA模型中每一個預報成員都具有相當的重要性。

	CRPS	MAE
第二組BMA模型	0.7235	1.0036

表三：第二組BMA模型的校驗結果。

第二組				
-----	--	--	--	--

BMA模型 預報成員	Q ₁	Q ₂	Q ₃	mean
NFS_M00	0.0000	0.0280	0.3643	0.2013
NFS_M02	0.0000	0.0417	0.3799	0.2108
WRF_M00	0.0000	0.1611	0.4997	0.2754
WRF_M02	0.0000	0.2366	0.5465	0.3125

表四：第二組BMA模型中各預報成員的 w_k 參數統計數據。

分析結果顯示，儘管第二組BMA模型比第一組少了6組預報成員，但其預報能力與第一組不相上下。這樣的現象，主要歸因於這6組WEPS預報成員在第一組BMA模型中的權重 (w_k) 都相當低，因此把WEPS除去對BMA預報能力的影響不大。至於為何WEPS的權重會顯得較低，我們可藉由簡單的統計分析來尋找可能原因。

首先我們觀察每一個預報成員各自的預報能力。表五為各個預報成員經過偏差調整後的預報值與觀測值之MAE，偏差調整的方式與BMA相同，使用 a_k 和 b_k 作線性調整。比較MAE後我們可以發現WEPS的預報誤差均高於其他4組預報成員，這應是導致WEPS在第一組BMA模型中權重較低的可能原因之一。

由於BMA模型中權重參數 w_k 反映出的是該預報成員相對於其他成員的重要性，因此 w_k 不僅隱含著該成員的預報能力，還有該成員與其他成員之間的相關性。表六為將所有預報成員經過偏差調整後，利用其與觀測值之間的誤差所計算出來的相關係數矩陣。表中我們可以觀察到WRF_M01與WEPS成員間彼此的相關係數都較高，因此在已經擁用WRF_M01預報資訊的情況下，WEPS預報資訊的重要性就會降低，這也是造成WEPS權重較低的可能原因。

在表五中顯示NFS_M00、NFS_M02、WRF_M00這三組預報模式的準確度差異不大，但在表二中WRF_M00的權重低於NFS_M00、NFS_M02，因為在第一組BMA模型中，預報成員包含了WEPS，而WEPS與WRF_M00的相關性很高，因此造成了WRF_M00的相對重要性降低而具有較低的權重。但若比較第二組BMA模型的權重，表四中WRF_M00卻高於NFS_M00、NFS_M02，因為在第二組BMA模型刪去了WEPS預報成員，使得WRF_M00相對重要性增加。相反地，由於在第二組BMA模型下，四組成員中NFS_M00、NFS_M02彼此之間的相關性最高，使這兩者各別的重要性降低而具有相對較低的權重。不論在第一組或第二組BMA模型中，WRF_M02皆具有最高的權重，這是因為WRF_M02有最高的準確度（見表五），且與其他成員的相關性較低（見表六）。

	MAE
NFS_M00	1.094438
NFS_M02	1.095907
WRF_M00	1.098542
WRF_M02	1.057189
WEPS_E01	1.147145
WEPS_E02	1.142319
WEPS_E03	1.161016
WEPS_E04	1.143561
WEPS_E05	1.163553
WEPS_E06	1.117129

表五：各個成員預報值（經過偏差調整後）的MAE。

	NFS M00	NFS M02	WRF M00	WRF M02	WEPS E01	WEPS E02	WEPS E03	WEPS E04	WEPS E05	WEPS E06
NFS_M00	1.0000	0.9419	0.7221	0.7357	0.7171	0.7246	0.7133	0.7135	0.6995	0.7205
NFS_M02	0.9419	1.0000	0.7193	0.7308	0.7181	0.7252	0.7134	0.7151	0.6993	0.7234
WRF_M00	0.7221	0.7193	1.0000	0.8294	0.8729	0.8559	0.8765	0.8591	0.8636	0.8621
WRF_M02	0.7357	0.7308	0.8294	1.0000	0.7604	0.7556	0.7703	0.7605	0.7578	0.7554
WEPS_E01	0.7171	0.7181	0.8729	0.7604	1.0000	0.8861	0.8848	0.8824	0.8695	0.8694
WEPS_E02	0.7246	0.7252	0.8559	0.7556	0.8861	1.0000	0.8709	0.8940	0.8507	0.8719
WEPS_E03	0.7133	0.7134	0.8765	0.7703	0.8848	0.8709	1.0000	0.8819	0.8622	0.8756
WEPS_E04	0.7135	0.7151	0.8591	0.7605	0.8824	0.8940	0.8819	1.0000	0.8595	0.8887
WEPS_E05	0.6995	0.6993	0.8636	0.7578	0.8695	0.8507	0.8622	0.8595	1.0000	0.8672
WEPS_E06	0.7205	0.7234	0.8621	0.7554	0.8694	0.8719	0.8756	0.8887	0.8672	1.0000

表六：各個成員預報誤差（將偏差調整後的預報值與觀測值相減）的相關係數矩陣。

四、模擬實驗

第三章中關於預報成員對BMA模型的影響，皆是利用歷史資料所進行的推測，由於歷史資料有限，而且真實資料往往會受到外在因素(例如模式改版)影響到觀察結果。因此在這一章中，我們希望藉由數值模擬來證實第三章中的推測。

為了單純探討預報成員的相關性和準確度對BMA模型的影響，因此在數值模擬中我們不考慮溫度的季節變換以及預報模式系統性偏差等等的問題。在不失一般性下，我們假設觀測值與各成員預報值皆來自期望值為0及標準差為1的標準常態分布(Standardized Normal Distribution)，而彼此之間的關係可以用多變量常態分布(multivariate normal distribution)來描述：

$$\begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \end{bmatrix} \sim N_4 \left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho_{01} & \rho_{02} & \rho_{03} \\ \rho_{01} & 1 & \rho_{12} & \rho_{13} \\ \rho_{02} & \rho_{12} & 1 & \rho_{23} \\ \rho_{03} & \rho_{13} & \rho_{23} & 1 \end{bmatrix} \right),$$

其中 y_0 代表觀測值， y_1 、 y_2 、 y_3 分別代表3組預報成員的預報值，而 ρ_{ij} 代表的是 y_i 與 y_j 之間的相關係數。我們可以利用 ρ_{ij} 值的調整來設計預報成員的準確度以及預報成員之間的相關性。

接續將針對四個主題做不同的數值模擬實驗，每一組實驗都會透過設計好的多變量常態分布生成 10^5 筆數據，代表同一地點未來 10^5 天的觀測值與成員預報值，再以25天為訓練期建立每一天的BMA預報模型。

(一) 預報成員準確度對BMA模型權重的影響

在第一組實驗中，我們設計的觀測值與預報成員之間的關係為：

$$\begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \end{bmatrix} \sim N_4 \left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0.8 & 0.7 & 0.6 \\ 0.8 & 1 & 0.9 & 0.9 \\ 0.7 & 0.9 & 1 & 0.9 \\ 0.6 & 0.9 & 0.9 & 1 \end{bmatrix} \right),$$

在此設計中，3組預報成員彼此之間的相關係數皆為0.9，代表相關性都相同，但與觀測值之間的相關係數分別為0.8、0.7、0.6，代表 y_1 、 y_2 、 y_3 這3組預報成員的準確度是由高到低。

而這組實驗中預報成員的MAE如表七所示。表中顯示這3組預報成員實際的準確度也是由高到低，而在這樣的條件下，我們觀察它們在BMA模型中的權重。表八為成員權重的統計分析，表八顯示成員的權重大小排列與表七中的準確度完全一致。因此證明了在預報成員相關性相同的條件下，BMA模型中的權重可以確實反映出預報成員的準確度。

	MAE
y_1	0.5042
y_2	0.6195
y_3	0.7155

表七：第一組實驗中各預報成員的MAE。

	Q ₁	Q ₂	Q ₃	mean
y_1	0.9283	1.0000	1.0000	0.9175
y_2	0.0000	0.0000	0.0680	0.0816
y_3	0.0000	0.0000	0.0000	0.0010

表八：第一組實驗中各預報成員的 w_k 參數統計數據。

(二) 預報成員相關性對BMA模型權重的影響

在第二組實驗中，我們設計的觀測值與預報成員之間的關係為：

$$\begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \end{bmatrix} \sim N_4 \left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0.7 & 0.7 & 0.7 \\ 0.7 & 1 & 0.9 & 0.6 \\ 0.7 & 0.9 & 1 & 0.6 \\ 0.7 & 0.6 & 0.6 & 1 \end{bmatrix} \right),$$

這3組預報成員的準確度都相同，而 y_1 與 y_2 之間有很強的相關性， y_3 則與其他兩組的相關性較低。

這組實驗中各成員的準確度幾乎沒有差異，如表九所示。儘管3組成員的準確度沒有差異，但在表十中，我們可以看到 y_3 的權重較高，表示 y_3 在BMA模型中的相對重要性高於其他兩者，而 y_1 與 y_2 因為彼此之間的相關性較高而降低了本身的相對重要性。此實驗證明儘管預報成員各自的準確度相同，彼此之間的相關性還是會影響成員本身的相對重要性。

	MAE
y_1	0.6178
y_2	0.6188
y_3	0.6169

表九：第二組實驗中各預報成員的MAE。

	Q ₁	Q ₂	Q ₃	mean
y ₁	0.0000	0.2389	0.4876	0.2846
y ₂	0.0000	0.2213	0.4813	0.2769
y ₃	0.2535	0.4387	0.6167	0.4385

表十：第二組實驗中各預報成員的 w_k 參數統計數據。

(三) 預報成員相關性對BMA預報能力的影響

關於預報成員相關性對BMA預報能力的影響，我們設計了兩組實驗來比較BMA模型的預報能力。第三組實驗中，觀測值與預報成員之間的關係設定為：

$$\begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \end{bmatrix} \sim N_4 \left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0.6 & 0.6 & 0.6 \\ 0.6 & 1 & 0.9 & 0.9 \\ 0.6 & 0.9 & 1 & 0.9 \\ 0.6 & 0.9 & 0.9 & 1 \end{bmatrix} \right),$$

而第四組實驗中，觀測值與預報成員之間的關係設定為：

$$\begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \end{bmatrix} \sim N_4 \left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0.6 & 0.6 & 0.6 \\ 0.6 & 1 & 0.7 & 0.7 \\ 0.6 & 0.7 & 1 & 0.7 \\ 0.6 & 0.7 & 0.7 & 1 \end{bmatrix} \right),$$

在這兩組實驗的設定中，每個預報成員的準確度都相同，但第三組實驗中預報成員彼此之間的相關性都很高，而第四組實驗中預報成員彼此之間的相關性較低。

表十一顯示這兩組實驗中BMA模型的預報能力。雖然兩組實驗中，預報成員的預報能力都相同，但表中顯示第四組實驗下BMA模型的預報能力略優於第三組，因此預報成員之間的相關性會影響到BMA的預報能力，相關性越高則預報能力越差。另外我們再比較這兩組實驗下BMA模型的權重參數 w_k ，如表十二及表十三所示。由於在任一組實驗中，每個成員都具有相同的設定，因此估計出來的 w_k 都很接近1/3，但若比較兩組實驗 w_k 的變異程度，表中顯示第三組實驗的變異較大（可比較Q₁與Q₃的差距），這是由於第三組實驗中，成員間彼此的相關性太強，導致使用EM演算法估計 w_k 參數時較難達到收斂，因此估計出來參數的變異程度較大，也使得建立BMA模型時需要花更多的時間估計參數。

	CRPS	MAE
第三組實驗	0.4697	0.6584
第四組實驗	0.4555	0.6372

表十一：第三組實驗與第四組實驗中BMA模型的校驗結果。

	Q ₁	Q ₂	Q ₃	mean
y ₁	0.0000	0.1964	0.6383	0.3353
y ₂	0.0000	0.1861	0.6236	0.3289
y ₃	0.0000	0.2007	0.6369	0.3358

表十二：第三組實驗中各預報成員的 w_k 參數統計數據。

	Q ₁	Q ₂	Q ₃	mean
y ₁	0.0016	0.2871	0.5524	0.3308
y ₂	0.0005	0.2852	0.5547	0.3296
y ₃	0.0043	0.3036	0.5640	0.3396

表十三：第四組實驗中各預報成員的 w_k 參數統計數據。

(四) 預報成員多寡對BMA預報能力的影響

關於預報成員多寡對BMA預報能力的影響，我們同樣設計了兩組實驗來做對照，第五組實驗的設計為：

$$\begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \end{bmatrix} \sim N_4 \left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0.6 & 0.6 & 0.6 \\ 0.6 & 1 & 0.7 & 0.7 \\ 0.6 & 0.7 & 1 & 0.7 \\ 0.6 & 0.7 & 0.7 & 1 \end{bmatrix} \right),$$

而第六組實驗的設計為：

$$\begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \end{bmatrix} \sim N_8 \left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0.6 & 0.6 & 0.6 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.6 & 1 & 0.7 & 0.7 & 0.9 & 0.9 & 0.7 & 0.7 \\ 0.6 & 0.7 & 1 & 0.7 & 0.7 & 0.7 & 0.9 & 0.9 \\ 0.6 & 0.7 & 0.7 & 1 & 0.7 & 0.7 & 0.7 & 0.7 \\ 0.5 & 0.9 & 0.7 & 0.7 & 1 & 0.9 & 0.7 & 0.7 \\ 0.5 & 0.9 & 0.7 & 0.7 & 0.9 & 1 & 0.7 & 0.7 \\ 0.5 & 0.7 & 0.9 & 0.7 & 0.7 & 0.7 & 1 & 0.9 \\ 0.5 & 0.7 & 0.9 & 0.7 & 0.7 & 0.7 & 0.9 & 1 \end{bmatrix} \right),$$

這兩組實驗設計中，都有設定相同的 y_1 、 y_2 、 y_3 成員，與觀測值之間的相關係數皆為0.6，彼此之間的相關性皆為0.7，不同的地方在於第六組實驗多了 y_4 、 y_5 、 y_6 、 y_7 這4組成員，其中 y_4 與 y_5 準確度較低（相關係數0.5），但彼此之間與 y_1 有很強的相關性（相關係數0.9）。同樣地， y_6 與 y_7 準確度也較

低，但彼此之間與 y_2 有很強的相關性。這樣設計的目的，是爲了產生3群預報成員，每一群皆以不同的模式爲基礎，因此不同群的預報成員彼此之間相關性較低（相關係數0.7），而同一群內成員彼此相關性高（相關係數0.9），但同一群中只有一個成員的準確度較高（相關係數0.6），其他成員的準確度較低（相關係數0.5）。對照實際數據，同一群成員的關係與 WRF_M00、WEPS之間的關係類似，而不同群的關係就像是NFS、WRF之間的關係。

而這兩組實驗中，BMA模型的校驗結果如表十四所示。表中可以觀察到，雖然第六組實驗具有較多成員的預報資訊，但BMA模型的預報能力反而較差，這樣的現象主要是因爲過度配適(overfitting)所造成。當我們在建立任何的統計預測模型時，若加入過多不重要的參數，反而會使得模型過度配適導致預測能力降低。所以，在建立BMA模型時，並非預報成員越多越好。

爲了避免形成過度配適的現象，在實際預報作業中，我們不建議加入WEPS爲BMA模型的預報成員。雖然在我們第三章實際數據的校驗中，尚未看到過度配適的證據出現，但這並不表示沒有過度配適的可能。

	CRPS	MAE
第五組實驗	0.4555	0.6372
第六組實驗	0.4614	0.6457

表十四：第五組實驗與第六組實驗中BMA模型的校驗結果。

五、結論

在第三章的比較中，我們發現第一組（10組預報成員）與第二組（4組預報成員）BMA模型的預報表現差異不大。但由於BMA方法在計算模型參數 w_k 時，使用EM演算法需要花較多的時間，若加入越多的預報成員會使得計算成本大幅增加，例如第一組BMA模型建模所花的時間大約是第二組的三倍。另外，還有模型因參數過多而導致過度配適的潛在可能。因此在預報能力差異不大的情況下，我們會偏好使用成員較少的第二組BMA預報模型來進行預報作業。

而關於BMA模型中的預報成員挑選方法，可以根據模型中的參數 w_k 來做挑選， w_k 代表的是第 k 個預報成員在所有成員中預報最佳的機率，也是第 k 個預報成員相對於其他成員的重要性指標。而本文利用觀察歷史資料和數值模擬兩種方法，證明影響 w_k 值大小的包含該成員的預報準確度以及與其他成員之間的相關性，而且成員彼此之間的相關性若太強，甚至還可能會降低了BMA模型的預報能力。因此在

建立BMA模型時，預報成員的挑選不僅要考量成員的預報能力，還需要考量成員彼此之間的相關性。

在本文的測試中，由於WRF_M00、WRF_M02以及WEPS都是屬於相同的模式，彼此之間有很強的相關性，而WRF_M00與WEPS預報是採用更多相同的條件，若將全部的預報成員加入BMA模型中不會有更佳的預報能力反而增加了計算成本，因此從中取一些較具代表性的即可。然而，若能加入與原有預報成員相關性低且具有相當預報能力的新成員，或許能提升BMA模型的預報能力。因此，未來希望能加入與NFS或WRF差異性較大的模式來做測試，期待BMA模型會有更好的預報表現。

六、參考文獻

張語軒、張庭槐、吳蕙如，2012：貝氏模型平均應用於臺灣地區溫度機率預報。氣象學報（已接受）。

李志昕、洪景山，2011：區域系集預報系統之校驗與評估。建國百年天氣分析與地震測報研討會，122頁。

Hersbach, H., 2000: Decomposition of the continuous ranked probability score for ensemble prediction systems. *Wea. Forecasting*, 15, 559–570.

Hoeting, J. A., D. M. Madigan, A. E. Raftery, and C. T. Volinsky, 1999: Bayesian model averaging: A tutorial (with discussion). *Stat. Sci.*, 14, 382–401. [A corrected version is available online at www.stat.washington.edu/www/research/online/hoeting1999.pdf.]

Raftery, A. E., T. Gneiting, F. Balabdaoui, and M. Polakowski, 2005: Using Bayesian model averaging to calibrate forecast ensembles. *Mon. Wea. Rev.*, 133, 1155–1174.