

貝氏模型平均於溫度機率預報之應用

張語軒、張庭槐、吳蕙如
氣象資訊中心
中央氣象局

摘 要

系集預報的目的是希望盡可能掌握天氣發展的不確定性，尋找實際天氣狀況可能的分散範圍，若能計算出實際天氣的機率分佈，則能更精確地描述此不確定性。因此，本研究的目的在利用觀測資料透過統計方法後，建立類似實際天氣發生的機率分佈。

貝氏模型平均 (Bayesian Model Averaging, 簡稱BMA)，是根據多個模式的預報及測站實際觀測資料，透過統計方法建立出預測值的機率密度函數 (probability density function, 簡稱PDF)。經過校驗顯示BMA可以建立出較符合天氣真實狀況的預測PDF。

本文首先將介紹如何透過BMA方法來建立預測PDF，並利用台灣地區測站近地面溫度的觀測資料，來檢驗BMA方法的表現，最後說明如何將BMA應用在天氣預報上。

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。

利用多個模式的預報值來建構預測值的機率分佈，傳統的方法是Anderson在1996年提出使用的binned probability ensemble (簡稱BPE) 來進行，BPE方法的預測機率分佈說明如下：

若預報系統中有 K 個預報成員，將這些成員產出的 K 個預報值依照數值大小，在 $-\infty$ 到 $+\infty$ 的數線上從小到大依序排列可產生 $(K+1)$ 個區間，則觀測值會有相同的機率落在這 $(K+1)$ 個區間中，即落在每一個區間中的機率皆為 $1/(K+1)$ 。

但透過校驗我們會發現利用BPE得到的機率分佈往往與實際狀況差異很大，其主要原因是因為預報成員通常都會有離散程度不夠的情況。舉例說明，若有6個預報成員，依照BPE方法建構的機率分佈，則觀測值會有 $5/7$ (71.4%) 的機率會落在預報成員的最

大值與最小值之間，但在本篇文章使用的數據中 (數據將在後面做描述)，實際上卻只有約47.7%的觀測值落在預報成員的最大值與最小值之間，遠小於BPE假設的71.4%。

不論預報成員的離散程度是否足夠，如何透過預報資料的後處理，建構出一個有效的預測機率分佈，是本文的主要目的，而本文將採用Hoeting等人 (Hoeting et al., 1999) 提出的貝氏模型平均 (Bayesian Model Averaging, 簡稱BMA) 法進行研究，嘗試解決此問題。

BMA是被廣泛應用於經濟學、生物學、公共衛生學等領域的統計方法，近期也漸漸運用在氣象領域而有許多文獻發表。

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，其使用MM5 (The PSU / NCAR mesoscale model) 區域預報模式透過不同初始條件和側邊界條件產生的預報資料，經過校驗發現BMA建立出來的預測機率分佈比傳統方法表現更好，其應用在海平面氣壓也是類似的結果。此外，Sloughter et al. (2007、2010) 亦將BMA分別應用在降雨量及風速的預報上，表現都很不錯。而在統計自由軟體R，BMA已有套件可供使用 (Fraleley et al. 2007)，足見BMA的應用已越來越廣被接受。

本篇文章將會對BMA進行簡短的介紹，接著參考Reftery et al. (2005) 的做法，將BMA應用本局區域模式在台灣地區近地面溫度的48小時預報資料，建立溫度的預測機率分佈並進行校驗。

二、BMA法介紹

BMA模型是利用在得知預報成員的情況下，預測出對於觀測值的條件機率密度函數（conditional probability density function，簡稱 conditional PDF）。

令 y 為我們想要預測的天氣變數，而 f_1, f_2, \dots, f_K 為 K 個預報成員，則BMA模型下的預測PDF表示為：

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

其中 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。

BMA的預測PDF可視為許多PDF的加權平均，以圖一進行說明。圖中表示共有3個預報成員，每一條虛線分別是每一個預報成員對應到的 $g_k(y | f_k)$ 。若以每個成員的後驗機率 w_k 為權重進行加權平均，則得到圖中的粗實線，此粗實線即為BMA模型下的預測PDF，也就是公式(1)中的 $p(y | f_1, f_2, \dots, f_K)$ 。

若是溫度、氣壓等變數，其數據分佈相當對稱，會非常接近常態分佈（Normal distribution），因此可假設 $g_k(y | f_k)$ 為常態分佈，期望值為 $a_k + b_k f_k$ ，變異數為 σ^2 ，其中 a_k 和 b_k 是對預報值進行偏差調整時採用的參數。若用機率符號表示為

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

而預測PDF的期望值則為

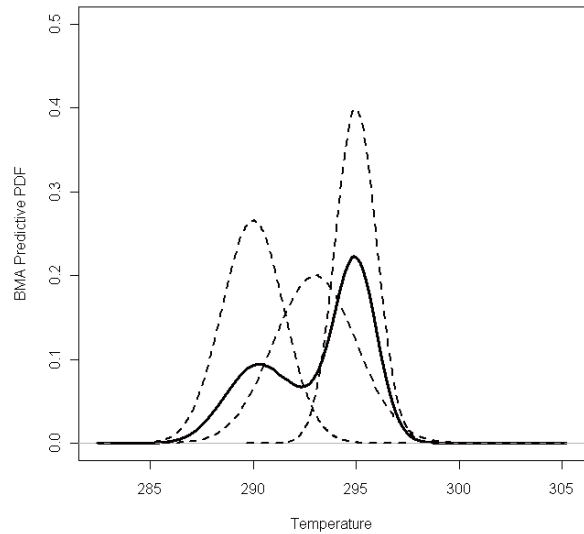
$$E(y | f_1, f_2, \dots, f_K) = \sum_{k=1}^K w_k (a_k + b_k f_k), \quad (2)$$

期望值可視為BMA方法下的決定性預報。

而在此常態分佈的假設下，BMA模型有 a_k 、 b_k 、 w_k ($k=1, \dots, K$) 及 σ^2 這些參數，我們將預報發佈的前 N 天視為訓練期，利用訓練期的數據來估計參數。訓練期是可變的，本文參考Reftery et al. (2005) 的設定，使用25天的資料進行訓練。參數估計的方法用到一些統計方法如最小平方方法（least squares method）、最大似估計法（maximum likelihood estimation），詳細估計步驟可參考Reftery et al. 的文章。

由於本文探討的是溫度預報，因此採用常態分佈假設下的BMA。若是針對雨量或是風速這一類非對稱的數據，可改用gamma分佈來做逼近，而關於

gamma分佈假設下的BMA模型，可參考Sloughter et al. (2007、2010) 的文章。



圖一：BMA方法下的預測PDF示意圖。圖中的虛線代表每一個模式對應到的條件PDF，將三條虛線作加權平均後，得到的實線即為觀測值的預測PDF。

三、實際分析與校驗

(一) 資料來源

本文使用的資料為台灣地區16個測站的地面溫度觀測值及區域模式內插至觀測點上48小時預報值，使用的測站包含淡水、台北、竹子湖、基隆、松山、花蓮、澎湖、台南、高雄、嘉義、台中、阿里山、新竹、蘭嶼、日月潭、台東，資料的時間為2010年10月20日到2011年6月5日，使用的預報模式為2組NFS（Non-hydrostatic Forecast System）、1組WRF（Weather Research & Forecast model），每組模式都分別包含解析度5公里及15公里預報資訊，共有6個模式預報資料，將這些格點資料內插至測站後即是預報的成員。

(二) 校驗方法

依Anderson (1996) 的研究，BPE方法下觀測值的預測機率分佈檢驗，可藉由rank histogram (RH) 來觀察此機率分佈是否準確。若有 K 個預報成員，將其預報值在 $-\infty$ 到 $+\infty$ 的數線上從小排到大可產生 $(K+1)$ 個區間，因為BPE假設觀測值有相同的機率落在這 $(K+1)$ 個區間中，因此校驗的方法就是去計算觀測值是否平均落在這 $(K+1)$ 個區間中。而RH就是藉由計算觀測值落在這 $(K+1)$ 個區間的比例所畫出的直方圖，若觀測值有相同的機率落在 $(K+1)$ 個區間中，則得到的RH圖將會是很平坦的。

針對BPE下的預測機率分佈可以很簡單的採用RH來作校驗，但若是針對BMA模型下的預測PDF，Gneiting et al. (2005) 建議採用probability integral transform (PIT)直方圖進行校驗。

設 $p(x)$ 代表BMA模型下的預測PDF，而 $P(x)$ 為其對應的累積分佈函數(cumulative distribution function，簡稱CDF)，可表示成 $P(x) = \int_{-\infty}^x p(y)dy$ 。將所有觀測值 x_o 對應的 $P(x_o)$ 畫成直方圖，即為PIT直方圖。若我們的預測PDF很準確，則對應到的PIT直方圖會是很平坦的。其機率理論是因為對於任意隨機變數 x ，所對應到的 $P(x)$ 會呈現均勻分佈(uniform distribution)。

雖然PIT與RH兩個直方圖的計算方式不同，但它們的機率概念其實是一樣的，對於圖形的解釋方法也相似，因此兩者也可直接進行比較。

除了透過直方圖來做圖形法的校驗外，我們也使用continuous ranked probability score (CPRS) 以及平均絕對誤差(mean absolute error，簡稱MAE) 進行數值上的校驗。

CRPS(Hersbash, 2000)是將布賴爾評分法(Brier score) 對所有門檻值積分，公式為

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

其中 $P(x)$ 為預測CDF，而

$$O(x) = \begin{cases} 0, & x < x_o \\ 1, & x \geq x_o \end{cases}, \quad x_o \text{ 為觀測值。}$$

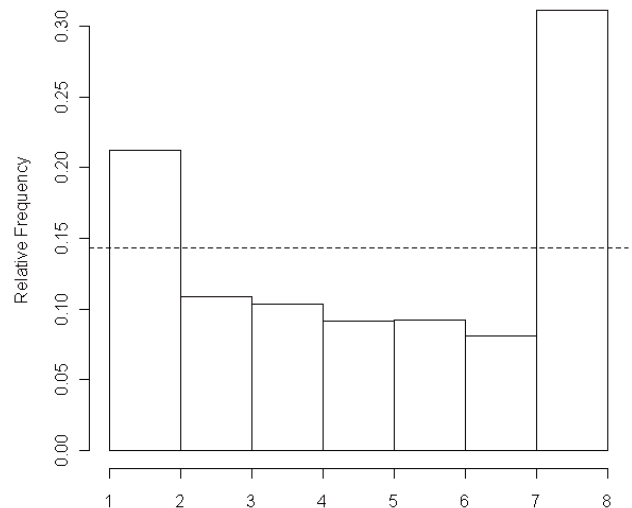
CRPS可視為預測PDF準確性的得分，其值愈低代表預測機率分佈越準確。

MAE是計算決定性預報與觀測值的差距，其值也是越低越好。對於傳統的作法，決定性預報為預報成員的平均值，而BMA的決定性預報為其預測機率分佈的期望值(見公式2)。

(三)校驗結果

利用上述的台灣地區地面溫度資料，所畫出的RH如圖二所示，圖中的虛線代表在完美的情況下直方圖的高度。圖二可發現此RH圖很不平坦，因此觀測值並非均勻的落在由預報成員構成的7個區間中，代表BPE方法下的預測機率分佈與現實狀況差距很大。而直方圖的左側及右側特別突出(呈現U型)，這代表著大部份的觀測值，都小於預報成員的最小值，或是大於預報成員的最大值，只有47.7%的觀測值介在預報成員的最大值與最小值之間，不到半數，這反應出預報成員存在著離散程度不夠的現象。

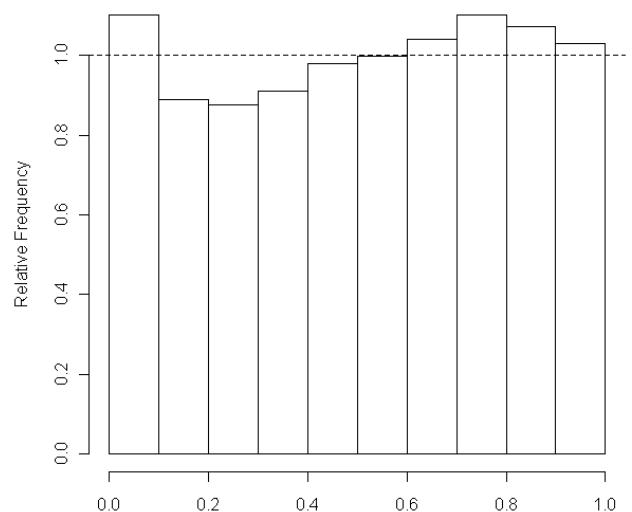
Rank Histogram



圖二：rank histogram (RH) 圖，橫軸代表將預報成員之預報值排序後的7個區間，每一個塊形的面積代表觀測值落在該區間的頻率。虛線代表在完美情況下，直方圖的高度。

接續校驗BMA方法下建立出來的預測機率分佈與真實狀況的差距。相同的資料在經過BMA方法建立出預測PDF後，將所有觀測值對應到的CDF畫出PIT直方圖如圖三，圖中的虛線也代表完美情況下的高度。此直方圖並非完全的平坦，代表BMA方法下的預測機率分佈與真實狀況依然存在差距。圖中最左側的塊形高於虛線，表示有超過10%的觀測值，小於我們原本預測的10%百分位數。而偏左側低於虛線的

Probability Integral Transform



圖三：probability integral transform (PIT) 直方圖，橫軸代表預測CDF值，每一個塊形面積代表觀測值的CDF落在該區間的頻率。虛線代表在完美情況下，直方圖的高度。

部分，表示低於40%的觀測值介在我們預測的10%到50%百分位數之間。但如果拿圖三與圖二比較，會發現圖三相較而言比圖二更平坦，代表BMA方法下得到的預測機率分佈與真實狀況比較接近。

上述方法僅能透過圖形進行機率分佈優劣程度的判斷，並無法了解預報得分，因此還需透過CRPS來比較BMA與BPE的預測PDF得分。CRPS計算的結果列在表一，表中BMA的CRPS比較低，代表BMA方法建立的預測PDF比起BPE有更高的得分。

	BPE	BMA
CRPS	1.5225	1.1179

表一：BEP 和 BMA 機率模型兩者預測 PDF 的 continuous ranked probability score (CPRS)。

以預報成員平均值作決定性預報和以BMA期望值作決定性預報所計算出的MAE列在表二中。表中BMA期望值有較低的MAE，代表BMA期望值比預報成員平均值為更準確的決定性預報。

	預報成員平均值	BMA 期望值
MAE	1.9111	1.5736

表二：預報成員平均值與BMA期望值的平均絕對誤差 (MAE)。

此外，在BMA期望值的公式中(式2)，由於 a_k 與 b_k 是透過最小平方方法計算，其可代入最小平方方法的公式再對式2作化簡，最後得到

$$E(y | f_1, f_2, \dots, f_K) = \bar{O} + \sum_{i=1}^K w_i b_i (f_i - \bar{f}_i),$$

其中 \bar{O} 為訓練期中的觀測值平均， \bar{f}_i 為訓練期中第 k 個預報成員的平均，此公式與Krishnamurtiet al. (2000)中的預報有相同的公式，只差在兩者的參數估計方式不同。

四、BMA在氣象預報上的應用

由於透過BMA，可以建立出一個完整的預測PDF，因此在氣象預報上的應用有非常大的彈性。除了前面提到BMA可產出決定性預報之外，也可透過BMA的預測PDF得到90%的預測區間(任意的百分比皆可)，或是90%的上界(或下界)。若我們對每一條PDF，取5%到95%的百分位數作為90%的預測區間，則這分溫度資料中，觀測值有約88.6%的比例確實落在此區間中，非常接近理想中的90%。也可藉由觀察

預測區間的長度，來了解預報的不確定性，預測區間長度越長則不確定性就越高。另外，我們也可設定任意的門檻值，來計算溫度高於(或低於)此門檻值的機率，當作極端天氣發生的機率。

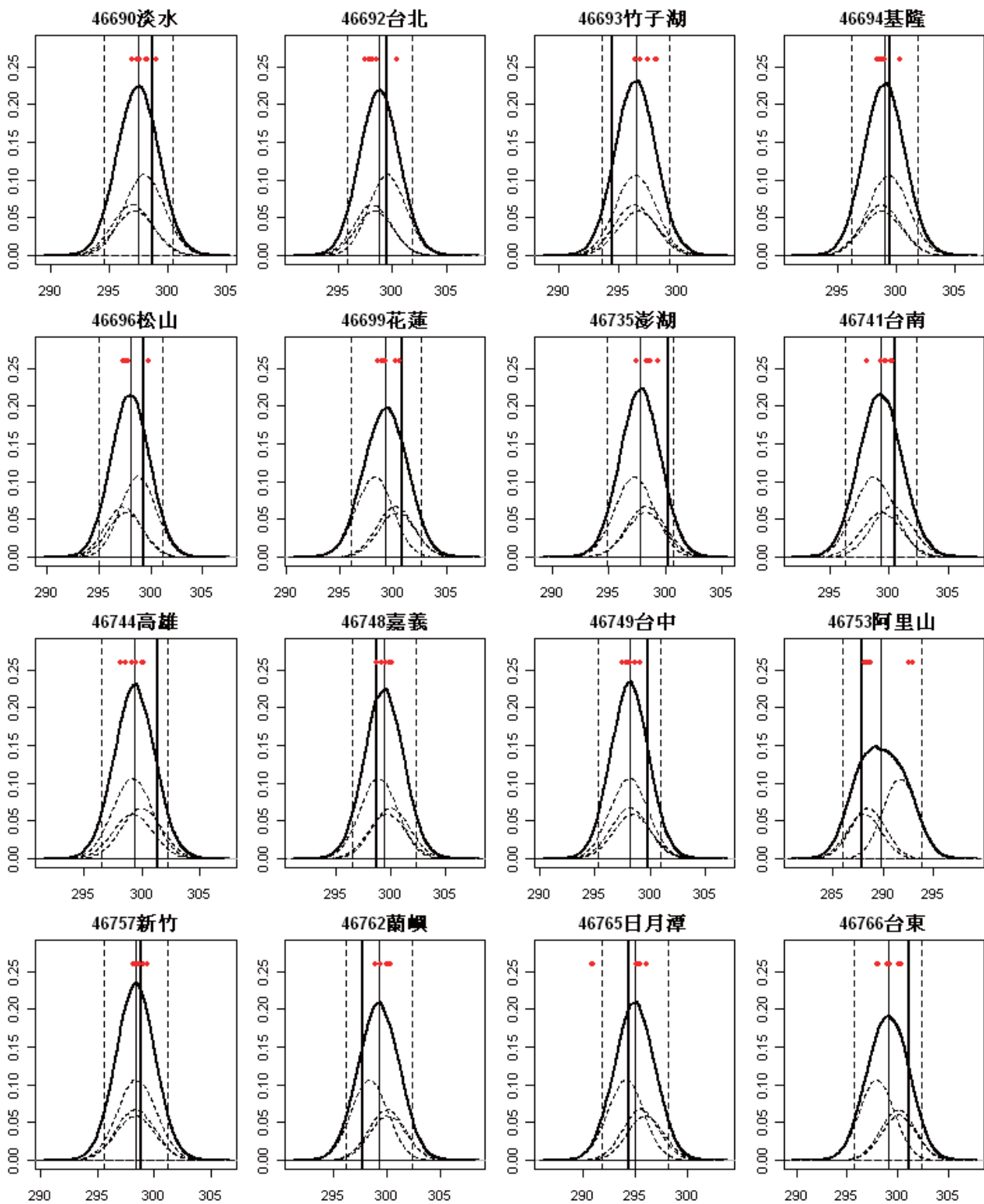
圖四為2011年6月4日00UTC時各測站的預測PDF。圖中實曲線代表BMA方法下的預測PDF，虛曲線為每一個預報成員的PDF乘上權重，垂直線的部份，細實線為期望值，虛線為90%預測區間，粗實線為觀測值，上方的圓點則分別代表每一個預報成員的預報值。在這16張圖中，可以發現有10張圖的預報值全部都大於或小於觀測值，佔了一半以上，這反應了圖二中rank histogram呈現U型的結果。儘管我們的預報成員分散範圍無法確實包含觀測值，但利用BMA方法建立出來的90%預測區間，在這個例子中全都涵蓋了觀測值。由此可見，BMA方法能有效的掌握數值預報的不確定性。

在BMA模型中，每一個預報成員都會對應到一個後驗機率參數 w_k ，此參數代表第 k 個預報成員預報最佳的機率，也是BMA的預測PDF中 $g_k(y | f_k)$ 所對應到的權重。若 w_k 越低，則代表第 k 個預報成員在BMA模型的建立中重要性越低，反之 w_k 越高則越重要。因此，我們可將 w_k 視為第 k 個預報成員相對於其他成員的重要性指標。

若將這份資料的BMA分析中，每一個成員的 w_k 平均值計算出來，結果如表三所示，表中也計算出每一個成員經過偏差調整後的預報值跟觀測值之間的MAE。由表可看出， w_k 從大到小的順序，與MAE從小到大的順序完全一致，表示我們估計出的 w_k 能反應出成員的準確性。再來從表中發現，雖然NFS M00 15km與NFS M02 15km這兩個預報成員的MAE相差不多，但 w_k 卻差異蠻大，我們可以從預報成員的相關係數(correlation coefficient)來解釋這樣的現象。

	w_k	MAE
NFS M00 5km	0.1323	1.8537
NFS M00 15km	0.0028	2.3634
NFS M02 5km	0.0871	1.8711
NFS M02 15 km	0.0005	2.3675
WRF M00 5 km	0.7692	1.5933
WRF M00 15 km	0.0081	2.0747

表三：每一個預報成員，在BMA模型中後驗機率 w_k 的平均值，以及經過偏差調整後的MAE。



圖四：2011年6月4日00 UTC，台灣地區各測站地面溫度預報的預測PDF。實曲線代表BMA方法下的預測PDF，而虛曲線則為每一個預報成員的PDF乘上權重。垂直線的部分，細實線為期望值，虛線為90%預測區間，粗實線為觀測值。上方的圓點則分別代表每一個預報成員的預報值。

相關係數如表四所示，表中NFS M00 15km與NFS M02 15km有相當高的相關係數（0.9933），表示在已知預報成員NFS M00 15km預報值的條件下，增加另一個成員NFS M02 15km的預報值所帶來的額外

資訊較少，因此NFS M02 15km的重要性就會降低許多。類似的情況也發生在NFS M00 5km與NFS M02 5km身上，其相關係數為0.9913。

另外，表三的 w_k 中，三個解析度為15公里的預報成員對應到的數值很小（皆小於0.01），其他三個解析度為5公里的預報成員則較大，這反應了圖四中，每一張圖都只看得到三條虛曲線（虛曲線為每一個預報成員的PDF乘上 w_k ）。

	NFS M00 5km	NFS M00 15km	NFS M02 5km	NFS M02 15 km	WRF M00 5 km	WRF M00 15 km
NFS M00 5km	1.0000	-	-	-	-	-
NFS M00 15km	0.9340	1.0000	-	-	-	-
NFS M02 5km	0.9913	0.9293	1.0000	-	-	-
NFS M02 15 km	0.9306	0.9933	0.9360	1.0000	-	-
WRF M00 5 km	0.9377	0.8634	0.9357	0.8639	1.0000	-
WRF M00 15 km	0.9175	0.9381	0.9158	0.9366	0.9406	1.0000

表四：預報成員之間的相關係數。

五、結論

本文的目的是希望在預報系統中，透過統計方法的後處理，來建立出有效的預測機率分佈，達到更精確的氣象預報。研究過程採用BMA方法嘗試來達成此目的，並利用台灣地區測站近地面溫度的觀測資料，進行實際的分析校驗。

從資料的分析中，我們發現利用BMA比起預報的傳統方法，可得到更準確的預測機率分佈，與真實的觀測值誤差較小。若要採用決定性的預報，BMA的期望值顯然比預報成員的平均值更準確。

研究中，BMA機率模型建立的設定，均是參考Raftery et al. (2005) 的作法，但這篇文章探討的是美國西南部地區，因此用在台灣地區似乎有必要作一些調整。原本使用25天做為訓練期，但若改採用20天的訓練期，則發現建立出的預測機率分佈有較低的CRPS與MAE，其 $CRPS = 1.1155$ ， $MAE = 1.5696$ ，因此對於台灣地區的最佳的訓練期為何？有待更進一步的研究。

此外，我們也嘗試分區域建立BMA機率模型，利用單一測站的資料來估計模型參數，結果發現這種方式下建立出的預測機率分佈具有更低的CRPS與MAE，其 $CRPS = 0.9636$ ， $MAE = 1.3318$ 。如何針對

台灣地區建立出最合適的BMA機率模型亦需更多的研究來探討。

六、參考文獻

- Anderson, J. L., 1996: A method for producing and evaluating probabilistic forecasts from ensemble model integrations. *J. Climate*, 9, 1518–1530.
- Fraley, C., A. E. Raftery, T. Gneiting and J. M. Sloughter, ensembleBMA: An R Package for Probabilistic Forecasting using Ensembles and Bayesian Model Averaging, Technical Report No. 516R, Department of Statistics, University of Washington, 2007 (revised 2010).
- Gneiting, T., A. E. Raftery, A. H. Westveld, and T. Goldman, 2005: Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. *Mon. Wea. Rev.*, 133, 1098–1118.
- 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.]
- Krishnamurti, T. N., Z. Zhang, T. E. LaRow, D. R. Bachiochi, C. E. Willifor, S. Gadgil, and S. Surendran, 2000: Multi-model ensemble forecasts for weather and seasonal climate. *J. Climate*, 13, 4196–4216.
- 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.
- Sloughter, J. M., A. E. Raftery, T. Gneiting, and C. Fraley, 2007: Probabilistic quantitative precipitation forecasting using Bayesian model averaging. *Mon. Wea. Rev.*, 135, 3209–3220.
- , T. Gneiting, and A. E. Raftery, 2010: Probabilistic wind speed forecasting using ensembles and Bayesian model averaging. *J. Amer. Stat. Assoc.*, 105, 25–35.