

# 應用貝氏模型平均法於2017年太平洋颱風路徑預報

陳昱璿<sup>1</sup> 馮智勇<sup>1</sup> 賈愛玫<sup>2</sup> 許乃寧<sup>2</sup>  
多采科技有限公司<sup>1</sup> 中央氣象局預報中心<sup>2</sup>

## 摘要

貝氏模型平均法(Bayesian Model Averaging, 以下簡稱 BMA), 可整合不同預報成員的預報並產製預報誤差的機率密度函數(Probability density function, 以下簡稱 PDF), 此方法是藉由訓練資料中各個成員的誤差分布情形, 以最大似然法則(Maximum Likelihood)尋找一組可在訓練資料中讓預報命中觀測機率最高的成員權重組合。

本研究以 2017 年太平洋颱風為案例, 將颱風預報拆為經度預報與緯度預報, 對 JMA、NCEP、EC、TWRP 四個預報成員的決定性預報進行誤差分布分析以及預報整合, 訓練資料保持以最接近整合時間的前 N 筆資料做動態訓練, 並且將等權重預報的表現作為路徑預報的比較基準, 分別檢視 BMA 在經度與緯度預報上的期望值預報表現、期望值路線的距離誤差、以及 BMA 在經度與緯度預報的機率預報表現情形。

關鍵字：貝氏模型平均 颱風路徑機率預報

## 一、前言

颱風路徑的誤差隨著近年各項技術的發展而不斷縮小, 其中包括了電腦技術的演進、數值模式的改良(Elsberry, 2014)、觀測資料的增加, 以及整合系集/多模式方式的方法演進(Krishnamurti et al., 2000; Williford et al., 2003; Elsberry et al., 2008; Dong et al. 2016)。除了路徑誤差之外, 靠著電腦運算能力與大量資料的統計分析, 颱風路徑的不確定性估計可藉由系集模式的散佈程度以及過去的誤差分布情形來推估(DeMaria, 2009; 蔡等人, 2011), Bayesian Model Averaging(以下簡稱 BMA)也是其中一種可用來進行系集/多模式整合以及估計不確定性的方式。

此方式已經廣泛的被應用在社會科學與醫學等領域中, 可以整合不同預報源的預測分佈, 在2005年時 Raftery 等人以 BMA 方法整合 Fifth-Generation Penn State/NCAR Mesoscale Model (MM5)系集模式在美國西南部地面氣溫與海平面氣壓, 以BMA法整合後的機率分布相較於直接以模式預報分散情形做為機率分布來的更加可靠, 並且以BMA的期望值預報作為決定性預報的成效, 也較系集平均的誤差來的更小。

張等人(2011)實作於國內測站地面溫度機率預報, 局內於2015年完成BMA溫度機率指引之發展(馮等人, 2015), 並且於2016年將BMA方法應用於颱風路徑預報, 完成BMA颱風路徑指引發展(馮等人, 2016), 本研究以2017年的颱風案例為例, 分析BMA方法用於2017年太平洋颱風之預報成效。

## 二、以BMA建置颱風路徑方式

貝氏模型平均法的算式如式子(1)所示, 其中  $y$  代表觀測的發生分布,  $f_k$  代表第k個成員的預報值,  $K$ 代表成員總數,  $y^T$  代表訓練期所使用的觀測資料。式子左側表示在已知各個成員的預報數值的條件下, 推估  $y$  的分布, 也就是發生在各個不同數值的機率大小。右側的  $p(f_k | y^T)$  代表在訓練期之中,  $f_k$  模式為k個成員中最佳預報的機率, 也可以把這個項視為該成員的權重。  $p(y | f_k)$  表示已知  $f_k$  的情況下觀測  $y$  的分布情形, 也可以說是誤差分布的估計情形。因此可以說這個式子要表達的是, 當我們有K個成員的預報時, 我們藉由訓練期的預報表現來給予不同成員權重, 並且把這權重對誤差分佈進行加權並且對各成員加總, 可以得到觀測分佈的PDF。

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

在本研究中, 將二維空間的颱風位置分為經度預報以及緯度預報, 並且先對已知觀測情況下的觀測分佈型態, 也就是誤差型態進行分析, 緯度分析如圖1所示, 由左到右的圖框分別代表NCEP、EC、JMA、TWRP的誤差分布, 從上到下分別是預報時間12小時、24小時等逐12小時到96小時的誤差分布, 圖1與2的誤差分布資料都有先以平均加減3倍標準差做簡單的剔除離群值, 每張圖中央的數字是標準差, 右上角的P-Value 是以 Kolmogorov-Smirnov 進行檢定是否為常態的

P-Value值，檢定內容假設資料來自於以資料自身算出的平均與標準差所產生的常態分佈，因此當P-Value大於0.05代表通過檢定，誤差為常態，我們在圖中以綠色繪製誤差分布，若P-Value小於0.05則標示黃色，若小於0.001則標示為橘色，在圖1的緯度誤差中，僅有EC的12小時預報未通過檢定，而圖2的經度誤差中則是全數通過KS檢定。

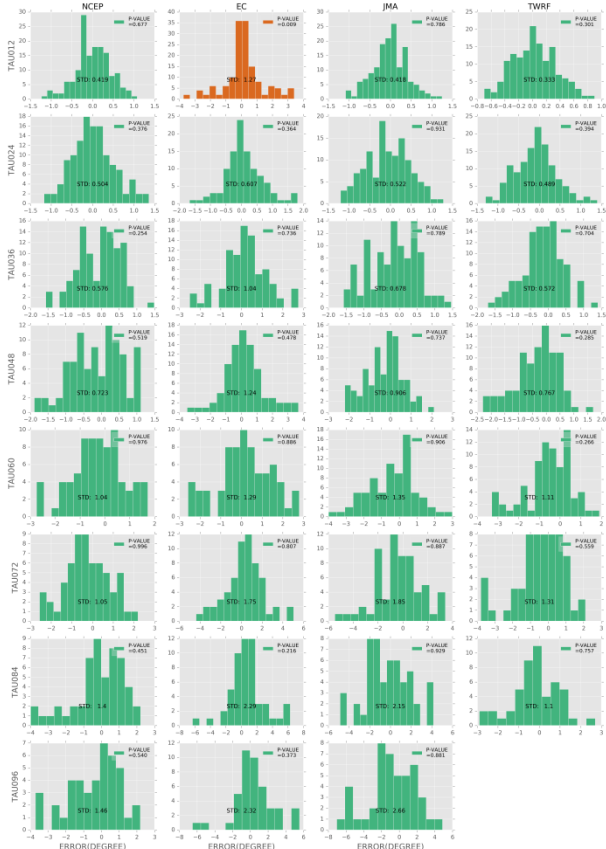


圖1 颶風路徑緯度預報誤差分布分析

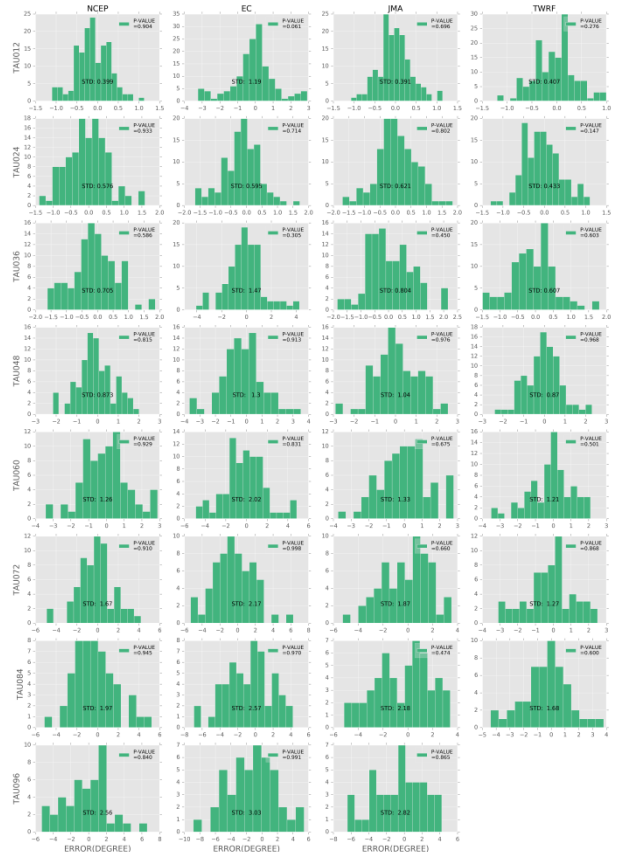


圖2 颶風路徑經度預報誤差分布分析

本研究參考Raftery等人(2005)以常態分佈誤差進行BMA訓練的方式進行，首先是估計收斂的目標設定，在此設定為最大化訓練期預報命中觀測的機率，並且將BMA算式簡化為式子(2)。

$$p(y | f_1, f_2, \dots, f_K) = \sum_{k=1}^K w_k g(y | f_k, \sigma_k) \quad (2)$$

式子(2)的右側將代表權重的條件機率  $p(f_k | y^T)$  以  $w_k$  表示，並且因為此處的誤差分布假設已設定為常態分佈，所以可以將  $p(y | f_k)$  改寫為  $g(y | f_k, \sigma_k)$ ，也就是以預報值作為平均值，並且包含一個未知標準差  $\sigma_k$  的常態分佈。收斂目標的表示如式子(3)，要將代表預報命中觀測機率地右式最大化(Maximum likelihood)，也就是找到一個適合的  $w_k$  與  $\sigma_k$ ，能夠將整個訓練期的命中率最大化。而此處取log是為了將機率的乘法轉為加法，簡化計算的複雜度(log-likelihood)。

$$l(w_1, w_2, \dots, w_K, \sigma_1^2, \sigma_2^2, \dots, \sigma_K^2) = \sum_t \log \left( \sum_{k=1}^K w_k g(y_t | f_k, \sigma_k) \right) \quad (3)$$

估計參數的方式在此使用Expectation Maximization algorithm (EM演算法)，將計算分為估計步驟(E)以及機率最大化步驟(M)並且不斷迭代，在E步驟中我們設定一個變數  $z_{kt}$  用以估計在每筆訓練資料t之中k個成員的表現，在同筆訓練資料t之中，得分依據每個成員在這筆訓練資料中命中觀測的機率值除以全部成員機率值的加總，因此總和會剛好為1，如式子(4)

所示，j代表迭代次數，第0步驟的權重設定為1/K，標準差為誤差值的標準差。

$$z_{kt}^{(j)} = \frac{w_k^{(j-1)} g(y_t | f_{kt}, \sigma_k^{(j-1)})}{\sum_{i=1}^K w_i^{(j-1)} g(y_t | f_{it}, \sigma_k^{(j-1)})} \quad (4)$$

在E步驟估計完畢  $z_{kt}$  的各成員各筆訓練資料的表現參考值之後，為了將機率最大化，因此將下一輪的權重採用將  $z_{kt}$  中該成員在全部訓練資料的得分進行平均，而標準差為了配合權重縮放，因此在標準差的算式前乘上對應的  $z_{kt}$  進行數值調整如式子(5)與(6)所示。

$$w_k^{(j)} = \frac{1}{n} \sum_t z_{kt}^{(j)} \quad (5)$$

$$\sigma_k^{2(j)} = \frac{\sum_t z_{kt}^{(j)} (y_t - f_{kt})^2}{\sum_t z_{kt}^{(j)}} \quad (6)$$

如此不斷地交替進行E步驟與M步驟的迭代，直到 log-likelihood 收斂為止，便完成了  $w_k$  與  $\sigma_k$  的估計。因為我們假設常態分佈的中心為預報值，因此BMA預報的期望值數值為預報值乘上權重的加總，在此以期望值預報值作為BMA的決定性預報。

在以下的實驗中，我們以NCEP、EC、JMA、TWRF 整合預報時間12小時到84小時的模式經度與緯度預報，84小時以上則是以NCEP、EC、JMA三個成員進行，並且都以每次預報的前40筆資料作為訓練資料進行動態訓練，原因是我們認為各個模式的表現優劣會隨著時間有所變化，因此取較為近期的資料做訓練資料，並且就文獻與先前的實驗結果而言，大約30筆資料開始權重會趨於穩定。

### 三、 結果與討論

圖3是BMA的期望值預報與等權重平均預報比較的結果，我們將2017年的資料逐筆測試，並且分別算出經度預報誤差、緯度預報誤差、路徑誤差三個項目的平均進行比較，最上面為經度預報誤差，在預報時間較長(60小時以上)的部分開始出現表現劣於等權重平均的情形，預報時間較短則無，緯度預報則改善幅度相對較好，兩者皆有改善幅度隨著預報時間增加而越趨不穩定的趨勢，路徑誤差則是只有84小時與108小時的預報有平均誤差大於等權重平均的情形。

圖4是BMA的期望值預報誤差與等權重預報誤差結果比較圖，每張圖的右下標明預報時間長度，左上則是以T檢定比較兩者差異性是否顯著，從T檢定的結果來看只有12小時的預報有顯著性的改善，其餘的則無。在24小時到48小時可以看到資料點有稍微向右下分佈的情形，60與72小時則是有幾個點改進較為明顯，可能是造成平均誤差略勝的原因。

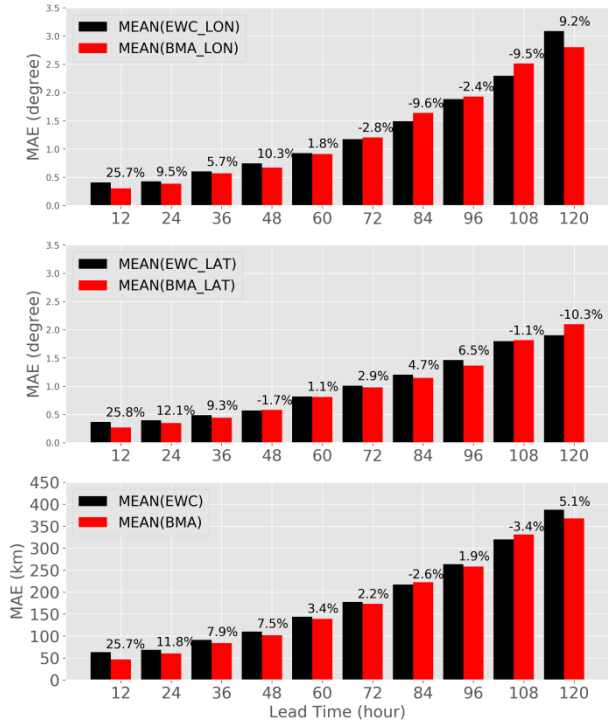


圖3 BMA期望值預報與等權重平均的平均誤差比較

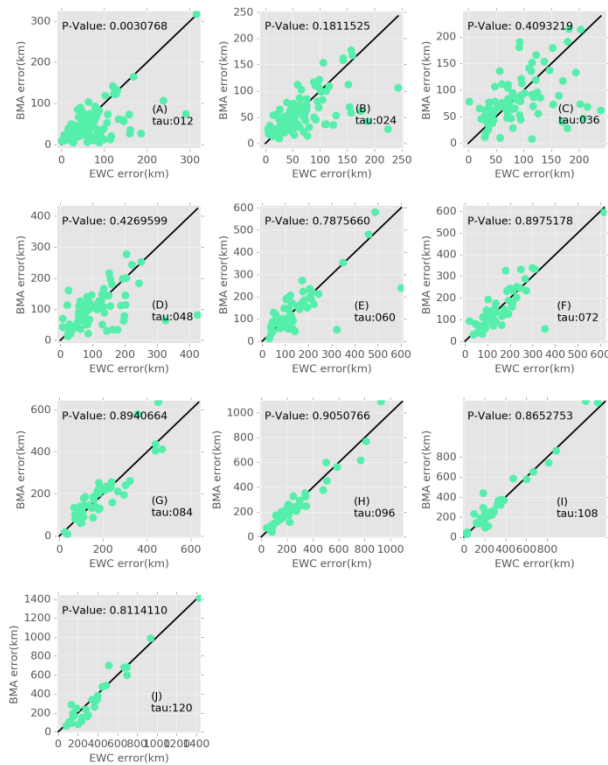


圖4 BMA期望值預報與等權重平均的誤差分布比較

圖5與圖6是BMA的機率預報檢定結果，我們採用 probability integral transform (PIT) Histogram 以及 continuous ranked probability score (CRPS) 兩個指標來做校驗，PIT的做法是將累積機率分佈(Cumulative

Distribution Function, CDF)拆為N等分，看每一個區間所對應到的觀測發生頻率為多少，這裡將CDF拆為10等份，並且將總資料量標準化為10，因此每一個區間的理想發生相對頻率應該為1，若是發生相對頻率高於1則代表低估該區間發生率，反之則是高估。

CRPS是將布賴爾評分(Brier Score)對所有門檻值積分並且將所有的案例取平均得到量化數據，當機率報的分布越分散，或者集中在離觀測發生的位置較遠，則CRPS值會隨之上升，反之當機率預報的機率分布越集中並且也越準確時，其值越低，代表機率預報的表現越好，如式子(7)所示。(Hersbach, 2000)

$$CRPS = \frac{1}{ncase} \sum_{acase} \int_{-\infty}^{\infty} [P(x) - O(x)]^2 dx, \quad O(x) = \begin{cases} 0, & x < obs \\ 1, & x \geq obs \end{cases} \quad (7)$$

圖5與6每張圖上方的標題除了預報時間以外還有CRPS的值，圖中橫向虛線是標準化後的理想觀測頻率發生值，從圖5中可以看到預報時間較短機率預報不論在PIT分布上還是CRPS數值上表現都較佳，在預報時間24到60小時皆低估了CDF在0.7區間以上的發生機率，也就是說低估了颱風路徑偏北的機率，在預報時間更大的預報中也有類似情形，但是因為整體而言穩定度下降導致低估偏北的特徵相對不明顯。

圖6中CRPS的數值相較圖5來的小，代表經度預報的機率誤差表現較緯度誤差來的好，而預報時間48到120皆有出現程度不等的CDF 0.7區間以上低估情形，表示低估了颱風路徑偏東的機率，可能代表BMA預報的某些成員在2017年的預報中出現數次颱風路徑比預報情形更偏東北的情形。

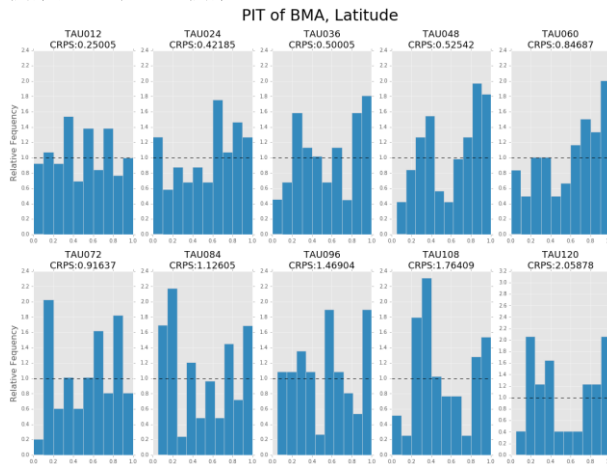


圖5 BMA整合緯度預報之機率預報表現情形

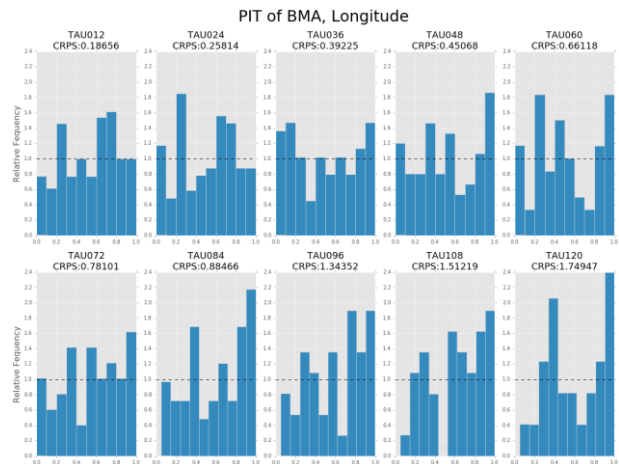


圖6 BMA整合經度預報之機率預報表現情形

## 四、 結論

BMA可以整合訓練資料中的誤差分布，藉由機率最大化的演算目標，給予不同成員權重將多成員預報整理為更為可靠的機率預報，並且可同時提供路徑預報以及不確定性估計資訊。

以2017年的實驗結果而言，颱風路徑的決定性預報上，相對於整合成員的等權重平均整體有所改善，但是在預報時間較長的部分穩定度並不高，平均誤差在某些預報時間會出現大於等權重平均的情形。

以BMA方法而言，在誤差分布已經通過檢定的情形下，若是預報表現不夠理想，有可能是訓練期的資料性質無法有效的反應出當下要整合的新預報所在的時空環境，未來將朝向調整訓練期的選擇方式著手。

## 五、 參考文獻

1. 張語軒、張庭槐與吳蕙如，2011: “貝式模型平均於溫度機率預報之應用”。氣象學報，49(1)，19-38。
2. 馮智勇、劉家豪、許乃寧、賈愛玫、黃嘉美，2015: “應用貝氏模型平均法發展台灣測站氣溫機率預報指引”，104年天氣分析與預報研討會。
3. 馮智勇、陳昱璿、黃嘉美、賈愛玫，2016: “應用貝氏模型平均法發展颱風路徑機率預報指引”，105年天氣分析與預報研討會。
4. 蔡孝忠、呂國臣、許乃寧、賈愛玫，2011: 蒙地卡羅法在颱風侵襲機率估計的應用。大氣科學期刊，39(3)，269-288
5. DeMaria, M., J. A. Knaff, R. Knabb, C. Lauer, C. R. Sampson, and R. T. DeMaria, 2009: A new method for estimating tropical cyclone wind speed probabilities. Wea. Forecasting, 24, 1573–1591
6. Dong, L., and F. Zhang, 2016: OBEST: An observation-based ensemble subsetting technique for tropical cyclone track prediction. Wea.

Forecasting, 31, 57–70, doi:10.1175/  
WAF-D-15-0056.1.

7. Elsberry, R. L., J. R. Hughes, and M. A. Boothe, 2008: Weighted position and motion vector consensus of tropical cyclone track prediction in the western North Pacific. *Mon. Wea. Rev.*, 136, 2478–2487.
8. Elsberry, R. L., 2014: Advances in research and forecasting of tropical cyclones from 1963–2013. *Asia-Pac. J. Atmos. Sci.*, 50, 3–16.
9. Krishnamurti, T. N., C. M. Kishtawal, Z. Zhang, T. LaRow, D. Bachiochi, E. Williford, S. Gadgil., and S. Surendran, 2000: Multimodel ensemble forecasts for weather and seasonal climate. *J. Climate*, 13, 4196-4216.
10. 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.
11. Williford, C. E., T. N. Krishnamurti, R. Correa Torres, S. Cocke, Z. Christidis, and T. S. Vijaya Kumar, 2003: Real-time multi-model super-ensemble forecasts of Atlantic tropical systems of 1999. *Mon. Wea. Rev.*, 131, 1878-1894.