

應用貝氏模型平均法發展臺灣測站氣溫機率預報指引

馮智勇¹ 劉家豪¹ 許乃寧² 賈愛玲² 黃嘉美²
多采科技有限公司¹ 中央氣象局氣象預報中心²

摘要

貝氏模型平均法(Bayesian Model Averaging, BMA)採用混合模型概念整合多種模擬結果，避免選擇相信單一模擬結果所造成的誤差，中央氣象局於「強化災害性即時天氣預報」建置案中發展利用BMA整合多模式之模式統計輸出法(Model Output Statistics, MOS)策略溫度定量預報指引為機率預報指引的方式，並且採逐站估計BMA方程式中之各模式權重係數參數，以符合臺灣各區測站特性不同的應用差異。

本研究為納入各模式對於測站的預報變異度，首先分析Expectation-Maximization演算法(EM演算法)計算式代表意義後，設計調整各測站BMA模型建置A、B與C方案。其中，A方案為對照組，採逐站建立BMA模型方式，但假設各模式的預報變異度均相同；B方案同樣採逐站建立BMA模型方式，並且允許不同模式間存在預報變異度；而C方案則是建置一組用於所有測站的BMA模型，但同樣允許不同模式間存在預報變異度。

採臺灣測站2014年4月15日至5月15日每日08時之ECA1模式降尺度預報、MOS策略之NCEP、JMA、ECA1與WRF模式預報以及動態MOS策略之ECA1模式預報等六種定量預報資料進行BMA模型事後預報分析。結果顯示，納入各模式預報變異度的B方案與C方案，可改善大部分測站的平均誤差(ME)、平均絕對誤差(MAE)、均方根誤差(RMSE)與相關係數表現。進一步比較後認為C方案除了收斂速度較快之外，所得BMA模型中的各模式權重係數隨時間變動幅度較小，可以客觀反應各模式對於臺灣測站預報能力，故建議採用。

關鍵字：貝氏模型平均、EM演算法、機率預報

一、前言

真實大氣環境包含許多不確定性，而數值模式模擬也包含參數化物理模式、使用參數、初始條件與邊界條件等等不確定性，因此各國氣象單位積極研究整合不同數值模式(以下稱多模式)或系集模式(ensemble models)模擬結果技術，以期發展各項天氣要素可靠之機率預報指引。

貝氏模型平均法(Bayesian Model Averaging, BMA)衍伸自貝氏定理，已被廣泛應用在經濟學、生物學等領域。Barnard [3]將模型混合的概念應用在飛機旅客的資料上；Roberts [10]提出由兩個模型以及後驗機率加權平均得到所需的模型，此為BMA雛型概念；Leamer [7]指出此方法可以處理模型選擇上之不確定性問題；Hoeting等人[6]拓展BMA雛型概念，提出將所有可能的模型，以其後驗(posterior)機率加權平均後，得到一個混合模型，再進行推估與預測，避免選擇相信單一模型所致誤差。Raftery 等人[9]直接採迴歸方式求得模式後驗機率而應用於估計美國西南部地面氣溫與海平面氣壓後，大幅簡化模型建置與推求後驗機率的流程，使得BMA得以廣為應用。

用於測站地面溫度[3]、雨量[8, 11]、風速[12]與風向[13]的機率預報。

中央氣象局已於「強化災害性即時天氣預報」建置案採模式統計輸出法(Model Output Statistics, MOS)統計降尺度策略建置ECA1、JMA、NCEP與WRF等數值動力模式之各項天氣要素迴歸模型，每日定時產製具相當可靠度之中期天氣(7天)逐12小時及短期天氣(3天)逐3小時定量預報指引(guidance)，提供測站點與網格點(2.5 km 解析度)預報作業參考[1]，也完成應用於發展臺灣測站BMA溫度機率預報產品的可行性評估[2]。

由於發展BMA所需資料可同時採用如ECA1、WRF等多模式以及系集模式各個成員的模擬結果，必須進一步考量不同模式對於測站的預報變異度差異，以期建置更具使用彈性的溫度機率預報產製作業流程。本研究以氣象局現有之MOS策略、動態MOS策略與模式直接降尺度等多種地面溫度定量預報資料為基礎，藉由調整Expectation-Maximization演算法(EM演算法)計算式並進行事後預報的方式，評估適合的參數估計方案。以下先行說明Raftery等人[9]提出的BMA模型參數估計方式，再行設計考量模式預報變異度的參數估計方案，並以2014年4月至

5月的事後預報結果進行分析，最後綜整研究結論與建議，以為後續發展作業系統參考。

二、BMA機率預報模型與建置方式

假設測站溫度觀測值為 y ，而 K 種模式的預報值為 M_1, M_2, \dots, M_K 並且已知資料 D 的情況下，BMA 模型描述物理量發生的機率為：

$$p(y|D) = \sum_{k=1}^K p(M_k|D) p(y|M_k, D) \quad (1)$$

其中， $p(y|M_k, D)$ 為各模式利用已知資料 D 估計 y 發生之條件機率密度函數，而 $p(M_k|D)$ 則為各模式的後驗機率，代表模式擬合 D 資料的優劣程度。由於當模式數量大於 1 時，必須滿足 $\sum_{k=1}^K p(M_k|D) = 1$ 的條件，因此後驗機率可視為權重值。

Raftery 等人[9]假設 f_1, f_2, \dots, f_K 分別代表 K 種系集模式成員經過系統性誤差調整後的『最佳』預報值，而 $g_k(y|f_k)$ 為在已知 f_k 時的 y 條件機率密度函數，則 BMA 模型描述溫度發生機率為：

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

其中， ω_k 為第 k 個成員的後驗機率，代表第 k 個模式在訓練期中成為最佳預報的機率，並滿足條件 $\sum_{k=1}^K \omega_k = 1$ ； $g_k(y|f_k)$ 表示在給定預報值 f_k 條件下 y 的 PDF。因此 BMA 模型的預報 PDF 可視為 K 個模式預報 PDF 的加權平均，意即以每個模式的後驗機率 ω_k 為權重進行加權平均。當物理量為溫度時，模式預報與實際觀測數據分布會非常接近常態分布(normal distribution)，因此可假設 $g_k(y|f_k)$ 為常態分布，其平均值為 $a_k + b_k f_k$ ，標準差為 σ ，可表示為：

$$y|f_k \sim N(a_k + b_k f_k, \sigma^2) \quad (3)$$

因此，BMA 模型描述物理量發生機率的期望值為：

$$E(y|f_1, f_2, \dots, f_K) = \sum_{k=1}^K \omega_k (a_k + b_k f_k) \quad (4)$$

BMA 模型中之未知參數 a_k 、 b_k 、 ω_k 及 σ ，需透過訓練期資料進行估計。其中， a_k 與 b_k 可使用簡單線性迴歸(simple regression)逐站進行估算，代表各模式系統性誤差校正參數； ω_k 及 σ 則透過最大化概似函數(maximum likelihood)的過程進行估計。假設預報誤差在時間(t)與空間(s)上不相關，對數概似函數可寫為 $\sum_{s,t} \log(\sum_{k=1}^K \omega_k g_k(y_{st}|f_{kst}))$ ，則可透過 EM 演算法求解 ω_k 及 σ 。

EM 演算法分為 E 步驟(Expectation)與 M 步驟(Maximization)找出最大化概似函數的最佳化參數。首先假設一個潛在變數 z_{kst} ，當 $z_{kst} = 1$ 時，代表第 k 個模式在 s 測站及 t 時間時有為最佳預報值，反之則 $z_{kst} = 0$ ；意即 $\{z_{1st}, z_{2st}, \dots, z_{Kst}\}$ 中只有一個為 1，其餘皆為 0。求解時先給定待估參數 ω_k 及 σ 初始猜測值，於第 j 次迭代的 E 步驟中計算：

$$\hat{z}_{kst}^{(j)} = \frac{\omega_k^{j-1} g(y_{st}|f_{kst}, \sigma^{(j-1)})}{\sum_{i=1}^K \omega_i^{j-1} g(y_{st}|f_{ist}, \sigma^{(j-1)})} \quad (5)$$

式中， $g(y_{st}|f_{kst}, \sigma^{(j-1)})$ 代表期望值為 $a_k + b_k f_{kst}$ ，標準差為 $\sigma^{(j-1)}$ 的常態分布函數。計算 $\hat{z}_{kst}^{(j)}$ 後，再於 M 步驟中計算 ω_k 及 σ 為：

$$\omega_k^{(j)} = \frac{1}{n} \sum_{s,t} \hat{z}_{kst}^{(j)} \quad (6)$$

$$\sigma^{2(j)} = \frac{1}{n} \sum_{s,t} \sum_{k=1}^K \hat{z}_{kst}^{(j)} (y_{st} - f_{kst})^2 \quad (7)$$

其中， n 為訓練資料數。反覆迭代計算式(5)~式(7)的過程將使概似函數值逐漸增加，直到相鄰兩次迭代的對數概似函數值的相對誤差小於 $1E-10$ 時(或迭代次數達 10,000 次)，參數 ω_k 及 σ 收斂完成求解[5]。

2.1 各測站逐一建置BMA模型

由式(2)與式(4)可知 BMA 模型是將不同模式估計值各以 a_k 與 b_k 參數完成系統性誤差調整後，再利用係數 ω_k 予以加權平均標準差為 σ 的常態分布而得。因此，若於參數 ω_k 與 σ 估計時不進行空間(即測站)累加(式(8)與式(9))，則可因應台灣測站變化差異顯著而個別建置測站溫度 BMA 模型[2]。於此稱式(5)、式(8)與式(9)為參數估計 A 方案，作為對照組。

$$\omega_k^{(j)} = \frac{1}{n} \sum_t \hat{z}_{kst}^{(j)} \quad (8)$$

$$\sigma^{2(j)} = \frac{1}{n} \sum_t \sum_{k=1}^K \hat{z}_{kst}^{(j)} (y_{st} - f_{kst})^2 \quad (9)$$

觀察式(9)可知 A 方案並未考量各種模式對於測站的預報變異度，當移除模式累積迴圈並加入最小值限制如式(10)，並調整潛在變數 z_{kst} 估計式如式(11)，可得由式(11)、式(8)與式(10)構成的參數估計 B 方案。

$$\sigma_k^{2(j)} = \text{Max} \left[\frac{1}{n} \sum_t \hat{z}_{kst}^{(j)} (y_{st} - f_{kst})^2, \sigma_k^{2(0)} \times 50\% \right] \quad (10)$$

$$\hat{z}_{kst}^{(j)} = \frac{\omega_k^{j-1} g(y_{st}|f_{kst}, \sigma_k^{(j-1)})}{\sum_{i=1}^K \omega_i^{j-1} g(y_{st}|f_{ist}, \sigma_k^{(j-1)})} \quad (11)$$

2.2 所有測站一起建置BMA模型

雖然 B 方案可藉由逐站估計參數 ω_k 及 σ_k 而更能反應台灣測站變化差異，但亦因收斂速度較慢而可能形成即時上線作業的時效問題。因此，再行提出於 Raftery 等人[9] 參數估計方式中考量模式預報變異度的參數估計 C 方案，其計算式包含式(11)、式(6)與式(12)，亦即除了參數 a_k 與 b_k 由各測站單獨進行估計外，所有測站均有相同的模式權重係數與變異度。

$$\sigma_k^{2(j)} = \frac{1}{n} \sum_{s,t} \hat{z}_{kst}^{(j)} (y_{st} - f_{kst})^2 \quad (12)$$

三、參數估計方案評估

3.1 使用資料

評估參數估計方案是以 2014 年 4 月 15 日至 5 月 15 日每日 08 時的台北(466920)、台中(467490)、台南

(467410)、高雄(467440)、台東(467660)、宜蘭(467080)、嘉義(467480)、日月潭(467650)、阿里山(467530)、玉山(467550)、淡水(466900)、基隆(466940)、花蓮(466990)與恆春(467590)等14個測站觀測值為事後預報比較依據。建置BMA模型時設定訓練期為50天，採用包含ECA1模式降尺度預報、MOS策略之NCEP、JMA、ECA1與WRF模式預報以及動態MOS策略(Dynamic MOS, DMOS)之ECA1模式預報等六種模式00Z初始時的000延時定量預報資料。

以事後預報2014年4月15日08時溫度為例，將使用2月24日至4月14日共50組資料(若有任一模式缺漏，則捨棄該日資料，往前取至補足)進行BMA模型參數估計後，再以六種模式於4月15日00Z初始時的000延時資料進行預報；同理，4月16日事後預報則採用2月25日至4月15日的資料進行訓練，並以4月16日的資料進行預報。

3.2 評估方式

A、B與C方案的評估方式將以事後預報所得之各測站估計值與觀測值序列計算的平均誤差(Mean Error)、平均絕對誤差(Mean Absolute Error)、均方根誤差(Root Mean Square Error)與相關係數等統計指標為主，另外亦採用各模式權重係數表現穩定性與參數收斂速度進行參考。

3.3 方案比較結果

圖1~圖4所示為以A方案(黑線)、B方案(粉紅十字線)與C方案(綠空心圓線)計算各測站(橫軸)BMA模型期望值與觀測值序列的四個統計指標。比較B方案與A方案可知，兩者誤差行為類似，考量模式預報變異度的B方案有助於改善部分測站BMA模型的表現，但改善幅度並不顯著且統計指標互有增減，亦有如花蓮站各項指標均略為下降的情況。另一方面，C方案則是明顯改善除阿里山與玉山外大多數測站的平均絕對誤差、均方根誤差與相關係數，尤以日月潭站的改善幅度最大。

圖5為日月潭站於2014年4月15日至5月15日間的觀測值與三種方案預報值比較，可知大部分時間三種方案預報結果相近，但C方案在4月20日至22日以及4月26日至5月3日兩個區間明顯勝出許多，使得統計指標改善幅度顯著。圖6與圖7分別為阿里山與玉山站的觀測值與三種方案預報值比較，明顯可見A、B方案預報結果與走勢非常相近，並且較C方案預報結果接近觀測值，因此在平均絕對誤差、均方根誤差與相關係數上表現較C方案為佳。

圖8與圖9所示分別為日月潭站採A與B方案進行事後預報時所估計的模式權重係數，橫軸代表時間(依序為4月15日、16日...)，而圖10則為採C方案的結果(所有測站均採用此組係數)，可知採C方案有助於獲得穩定模式權重係數。

四、結論與建議

4.1 結論

為納入各模式對於測站的預報變異度，本研究藉由調整EM演算法計算式提出建置BMA模型的A、B與C方案。透過比較14個測站進行2014年4月15日至5月15日間的事後預報結果統計指標可知，B方案雖有助於改善採A方案的BMA模型預報成效，但幅度有限；C方案則可明顯改善大多數測站的BMA模型預報成效，尤以日月潭站改善幅度最為顯著，並且BMA模型中之各模式權重係數表現穩定，以MOS策略WRF模式預報、動態MOS策略ECA1模式預報與ECA1模式降尺度預報的權重最大，可客觀反應各模式對於臺灣測站預報能力。

4.2 建議

為了可彈性整合多模式預報資料或者是系集模式資料為機率預報產品，本研究建議於BMA模型參數估計時納入各模式的預報變異度，用以協助篩選解釋能力較佳者，並且以C方案改善效果最為顯著且一致。

對於如阿里山與玉山等性質較為特殊的測站，建議進一步對於各模式進行校驗，以確認適當的BMA模型建置方式。

五、參考文獻

- 1.交通部中央氣象局，2013:「102年度強化災害性即時天氣預報—資料處理與統計降尺度建置案」委託研究計畫成果報告。
- 2.馮智勇、劉家豪、林佑蓉、蔡雅婷、陳雲蘭, 2014: “應用BMA發展多模式MOS策略溫度機率預報”，103年天氣分析與預報研討會。
- 3.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。
4. Barnard, G. A., 1963: “New methods of quality control”. J. Roy. Stat. Soc., 126A, 255-258.
5. Dempster, A. P., N. M. Laird, and D. B. Rubin, 1977: “Maximum likelihood from incomplete data via the EM algorithm”. J. Roy. Stat. Soc., 39B, 1-39
6. 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
7. Leamer, E. E., 1978: Specification Searches. Wiley, 370 pp
8. Liu, J., and Z. Xie, 2014: “BMA Probabilistic Quantitative Precipitation Forecasting over the Huaihe Basin Using TIGGE Multi-model Ensemble Forecasts”. Mon. Wea. Rev., 142, 1542-1555.
9. 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.

10. Roberts, H. V., 1965: "Probabilistic prediction". *J. Amer. Stat. Assoc.*, 60, 50-62.
11. 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.
12. Sloughter, J. M., T. Gneiting, and A. E. Raftery, 2010: "Probabilistic wind speed forecasting using ensembles and Bayesian model averaging". *J. Amer. Stat. Assoc.*, 105, 25–35.
13. Sloughter, J.M., T. Gneiting, and A.E. Raftery, 2013: "Probabilistic wind vector forecasting using ensembles and Bayesian model averaging". *Mon. Wea. Rev.*, 141, 2107-2119.

六、附圖

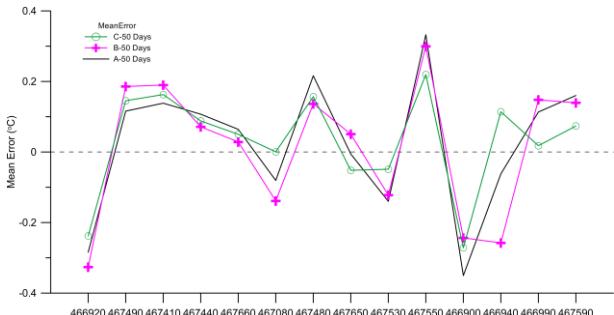


圖1 BMA模型建置A方案(黑線)、B方案(粉紅十字線)與C方案(綠空心圓線)於台北(466920)、台中(467490)、台南(467410)、高雄(467440)、台東(467660)、宜蘭(467080)、嘉義(467480)、日月潭(467650)、阿里山(467530)、玉山(467550)、淡水(466900)、基隆(466940)、花蓮(466990)與恆春(467590)站事後預報平均誤差比較。

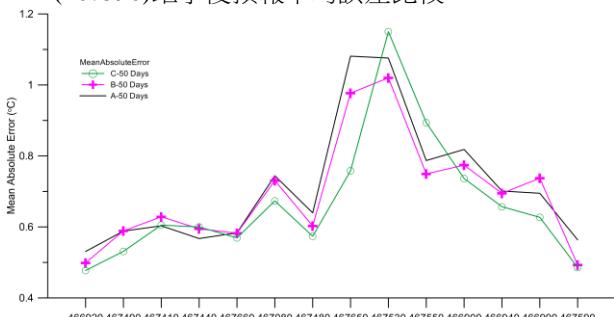


圖2 BMA模型建置A方案(黑線)、B方案(粉紅十字線)與C方案(綠空心圓線)於各測站事後預報平均絕對誤差比較。

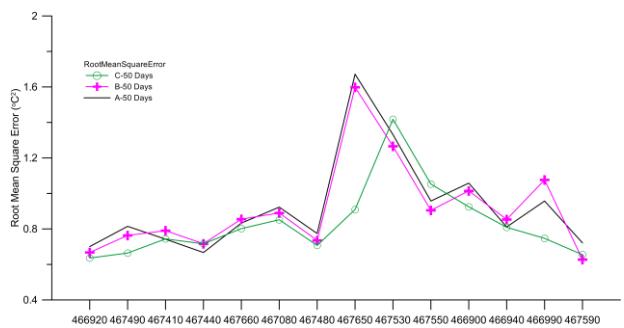


圖3 BMA模型建置A方案(黑線)、B方案(粉紅十字線)與C方案(綠空心圓線)於各測站事後預報均方根誤差比較。

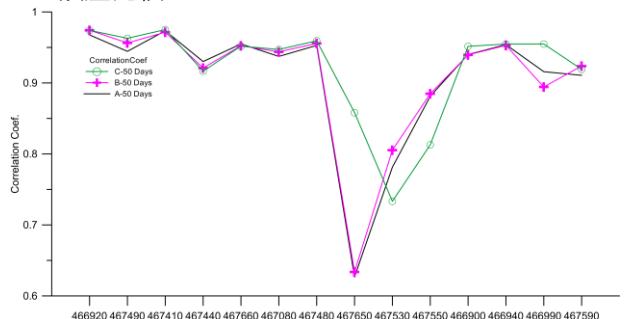


圖4 BMA模型建置A方案(黑線)、B方案(粉紅十字線)與C方案(綠空心圓線)於各測站事後預報相關係數比較。

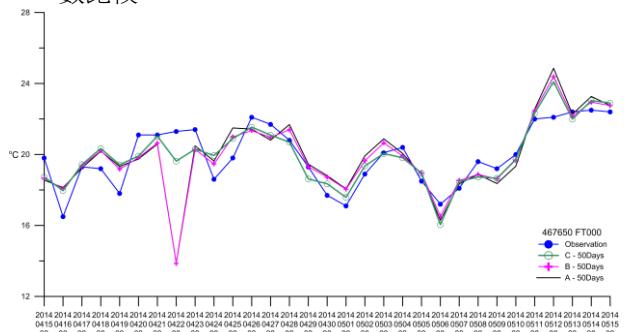


圖5 日月潭站2014年4月15日至5月15日間BMA模型建置A方案(黑線)、B方案(粉紅十字線)與C方案(綠空心圓線)預報值與觀測值(藍實心圓線)比較。

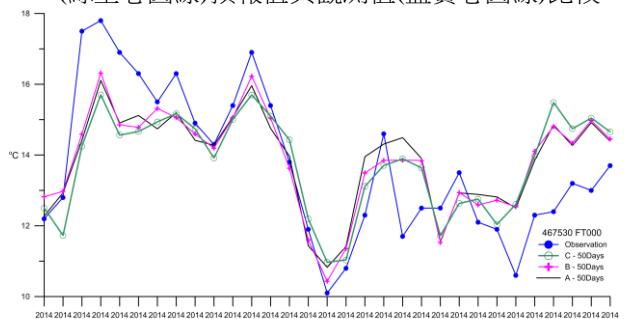


圖6 阿里山站2014年4月15日至5月15日間BMA模型建置A方案(黑線)、B方案(粉紅十字線)與C方案(綠空心圓線)預報值與觀測值(藍實心圓線)比較。

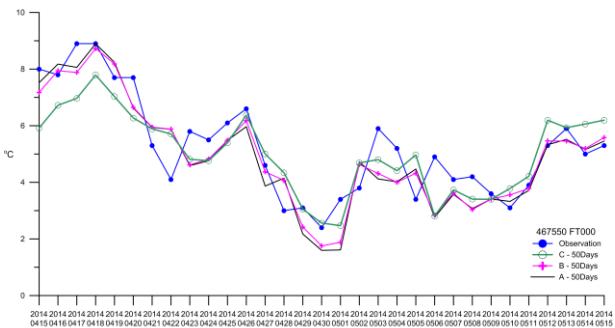


圖7 玉山站2014年4月15日至5月15日間BMA模型建置A方案(黑線)、B方案(粉紅十字線)與C方案(綠空心圓線)預報值與觀測值(藍實心圓線)比較。

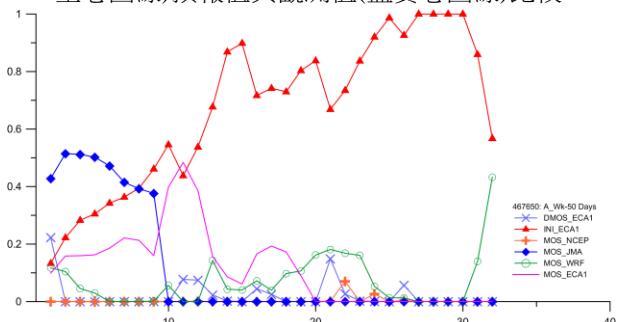


圖8 日月潭站採A方案建置的BMA模型中，各模式權重係數於2014年4月15日至5月15日間的變化情形。

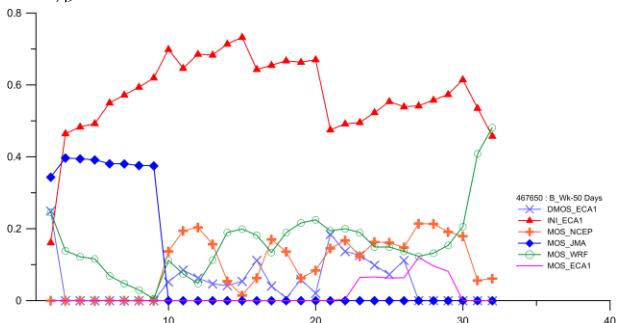


圖9 日月潭站採B方案建置的BMA模型中，各模式權重係數於2014年4月15日至5月15日間的變化情形。

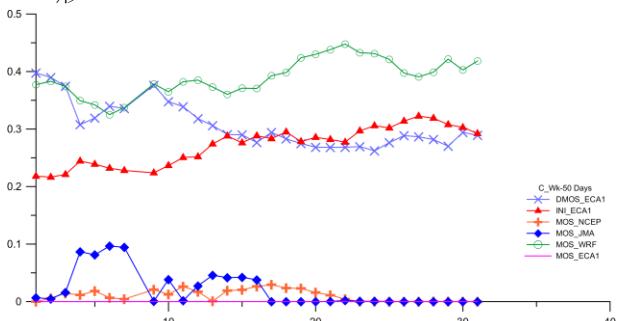


圖10 採C方案建置的BMA模型中，各模式權重係數於2014年4月15日至5月15日間的變化情形。