

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、劉家豪、許乃寧、賈愛玫、黃嘉美



多采科技有限公司
中央氣象局預報中心

2015.09.15

簡報大綱

一、緣起與目的

二、BMA機率預報模型建置方案

三、測站溫度機率預報方案評估

四、結語與建議

緣起

- 各國氣象單位積極整合**多模式或系集模式**以發展天氣要素**機率預報指引**
 - ▣ 大氣環境 & 數值模式的不確定性
 - ▣ 維運系集模式所費不貲
- 中央氣象局具**多種數值動力模式決定性預報、系集模式與相關指引**，適合發展**機率預報產品**
 - ▣ JMA, ECA1, NCEP, WRF & NFS 等多種模式輸出
 - ▣ WEPS系集模式輸出
 - ▣ MOS, DMOS, PP 等降尺度策略定量預報指引
 - ▣ ...

目的

- 中央氣象局已於 103年應用貝氏模型平均法(BMA)發展**台灣測站溫度機率預報雛形**
 - 使用 JMA, ECA1, NCEP, WRF & NFS 等模式MOS策略指引
 - 實作 Raftery 等人(2005)參數估計方案
 - 採用 UW-MM5 系集成員資料，並假設各成員預報變異度相同
- 建置**考量模式預報變異度**的測站溫度機率預報模型，以具**彈性利用各種預報資料**的擴充性
 - 多模式預報
 - 系集模式預報
 - 統計降尺度預報
 - ...



BMA機率預報模型建置方案

貝氏定理與貝氏模型平均

- 貝氏模型平均(BMA)採混合模型概念，以各個模型推估觀測值的**條件機率密度函數**(conditional PDF)進行**權重整合**

$$\begin{aligned} p(y|D) &= \frac{p(D|y)p(y)}{p(D)} = \frac{p(y,D)}{p(D)} \\ &= \sum_{k=1}^K \frac{p(y, M_k, D)}{p(D)} = \sum_{k=1}^K \frac{p(M_k, D)}{p(D)} \frac{p(y, M_k, D)}{p(M_k, D)} \\ &= \sum_{k=1}^K \frac{p(D|M_k)p(M_k)}{p(D)} \frac{p(y, M_k, D)}{p(M_k, D)} \\ &= \sum_{k=1}^K \frac{p(D|M_k)p(M_k)}{p(D)} \frac{p(M_k, D|y)p(y)}{p(M_k, D)} \\ &= \sum_{k=1}^K p(M_k|D) p(y|M_k, D) \end{aligned}$$

- y : 待估計的物理量
- D : 已知資料
- M_k : 第 k 個模型

後驗機率

條件機率

BMA 機率預報模型

□ Raftery 等人(2005) 應用於 溫度 與 海平面氣壓

- 當 y 為溫度或海平面氣壓時，認為其 條件機率 可表示為一平均值為 $a_k + b_k f_k$ 而 標準差為 σ 之 常態分布
 - f_k 為「具預報訊息」的第 k 個系集成員預報值

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

後驗機率

條件機率

□ y : 待估計物理量

□ D : 已知資料

□ M_k : 第 k 成員

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

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

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

BMA 機率預報模型參數估計

- 利用預報日前 m 天資料求解參數
 - 一筆資料為 $\{y_{st}, f_{1st}, f_{2st}, \dots, f_{Kst}\}$
 - a_k & b_k : 採簡單線性迴歸
 - ω_k & σ : EM 演算法, 最大化概似函數

$$l(w_1, w_2, \dots, w_k, \sigma^2) = \sum_{s,t} \log \left(\sum_{k=1}^K w_k g_k(y_{st} | f_{kst}) \right)$$

Expectation

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

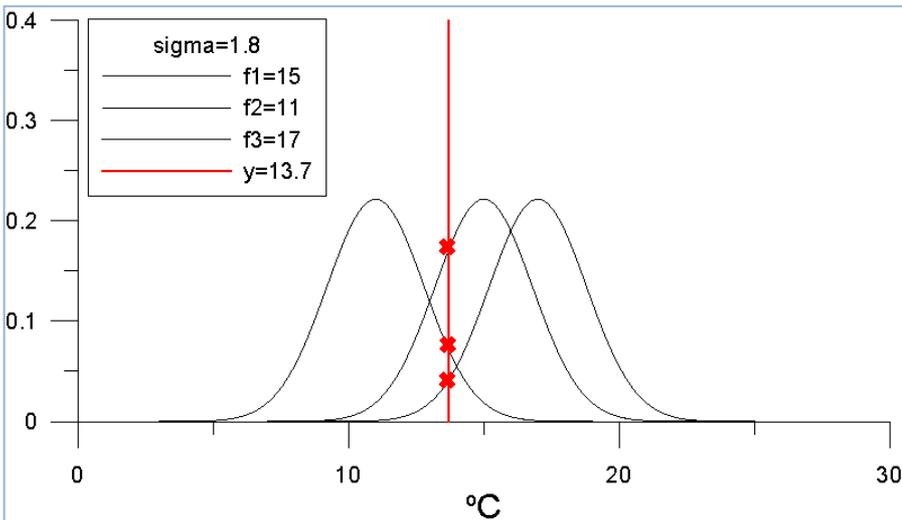
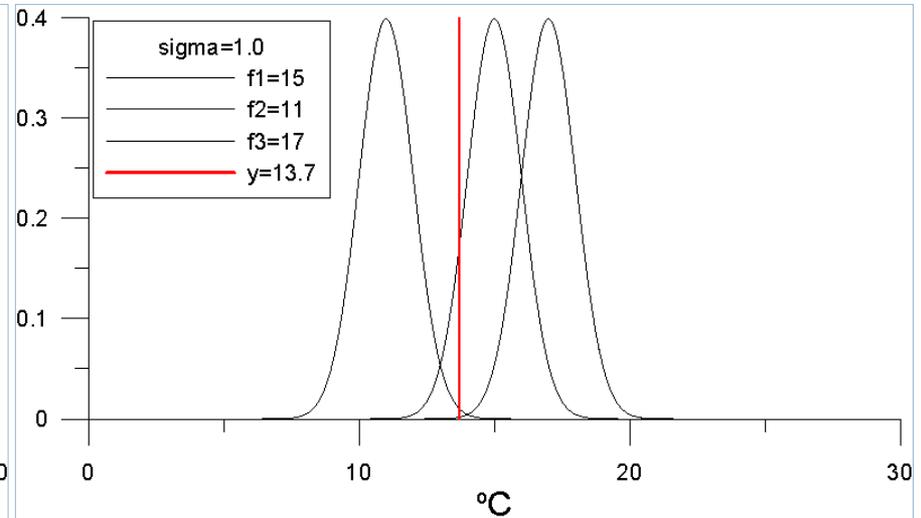
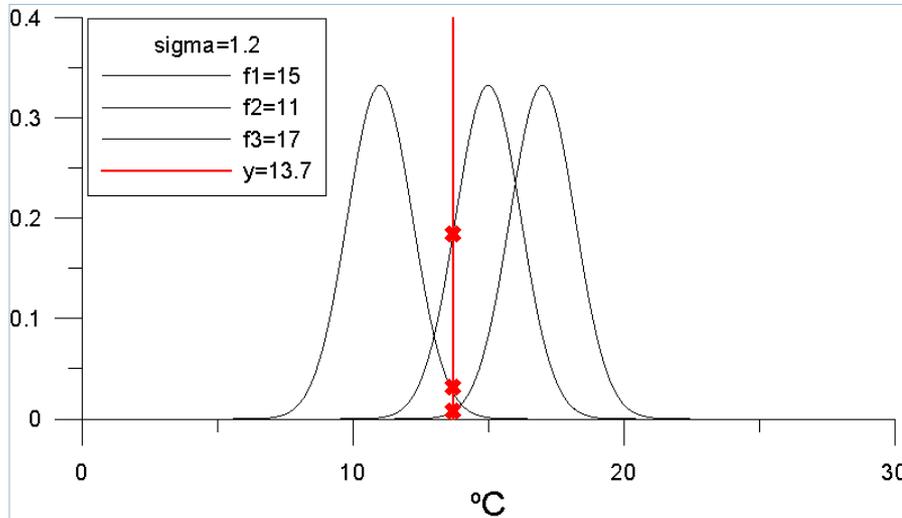
收斂時每筆資料(s,t), 只有1個 $z_{ist}=1$, 代表最佳估計; 其餘為零。

Maximization

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

EM演算法迭代示意

假設A測站一筆資料 $\{y_{st}=13.7, f_{1st}=11, f_{2st}=15, f_{3st}=17\}$



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

$$w_k^{(j)} = \frac{1}{n_{s,t}} \sum \hat{z}_{kst}^{(j)}$$

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

BMA機率預報模型建置方案

□ 逐測站建置

- A 方案(對照組)：各模式預報標準差相同 (103年度)

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

$$w_{s,k}^{(j)} = \frac{1}{n_t} \sum \hat{z}_{kst}^{(j)}$$

$$\sigma_s^{2(j)} = \frac{1}{n_t} \sum \sum_{k=1}^K \hat{z}_{kst}^{(j)} (y_{st} - f_{kst})^2$$

- B 方案：各模式預報標準差不同

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

$$w_{s,k}^{(j)} = \frac{1}{n_t} \sum \hat{z}_{kst}^{(j)}$$

$$\sigma_{s,k}^{2(j)} = \text{Max} \left\{ \frac{1}{n_t} \sum \hat{z}_{kst}^{(j)} (y_{st} - f_{kst})^2, \frac{\sigma_s^{2(0)}}{2} \right\}$$

□ 所有測站一起建置

- C 方案：各模式預報標準差不同

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

$$w_k^{(j)} = \frac{1}{n_{s,t}} \sum \hat{z}_{kst}^{(j)}$$

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



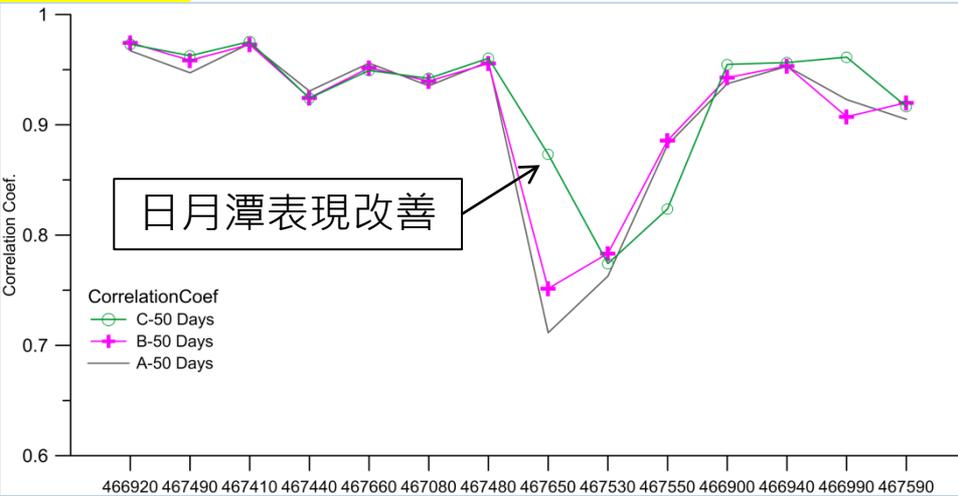
測站溫度機率預報方案評估

使用資料

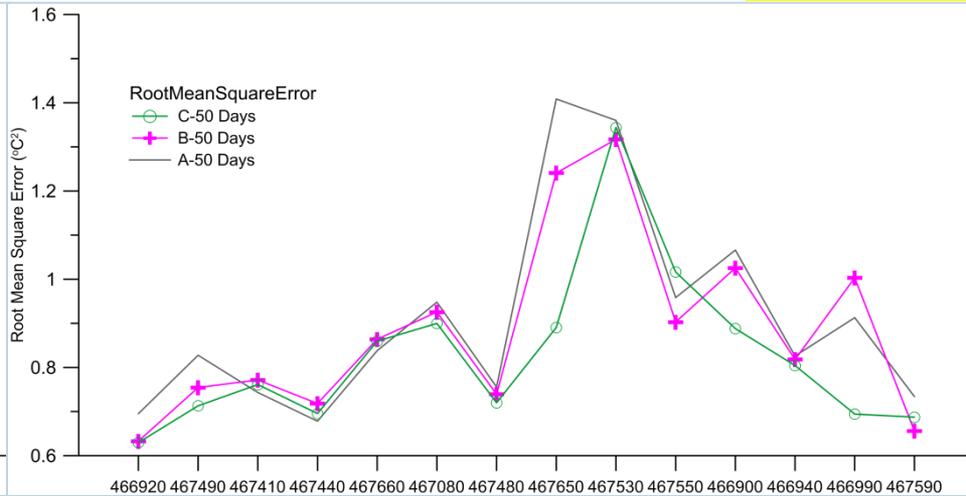
- 2014.04.15~05.15 測站資料
 - 測站觀測：08 時溫度
 - 台北(466920)、台中(467490)、台南(467410)、高雄(467440)、台東(467660)、宜蘭(467080)、嘉義(467480)、日月潭(467650)、阿里山(467530)、玉山(467550)、淡水(466900)、基隆(466940)、花蓮(466990)與恆春(467590)
 - 預報：
 - 模式直接輸出：ECA1
 - MOS 預報指引：NCEP、JMA、ECA1與WRF
 - DMOS 預報指引：ECA1
- 訓練期採用 50 天

BMA模型建置方案統計指標

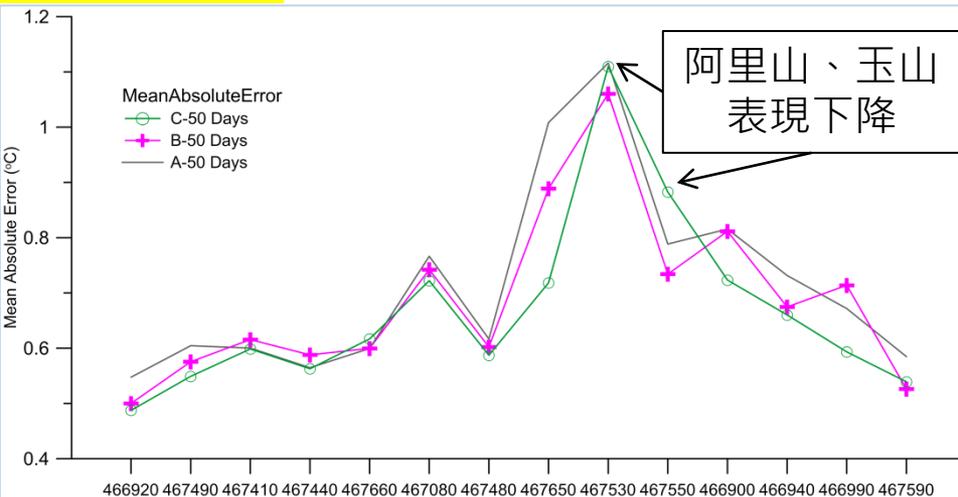
相關係數



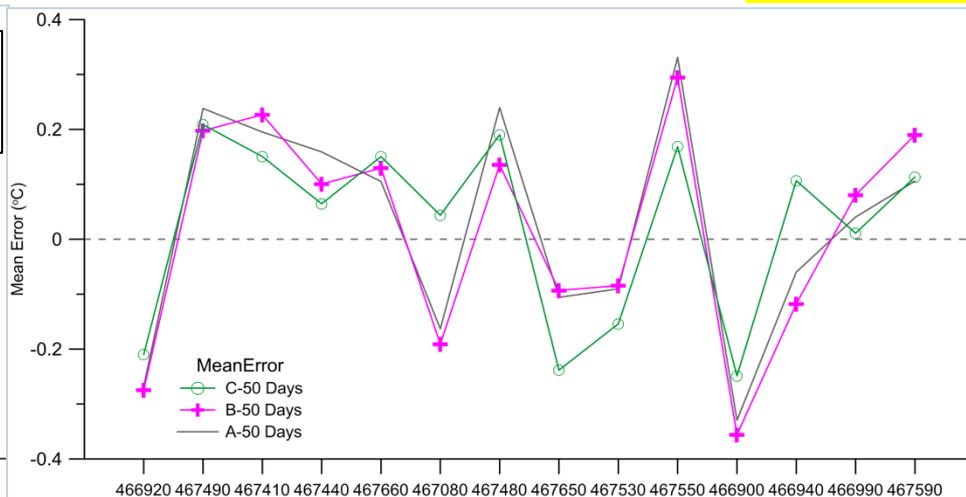
方均根誤差



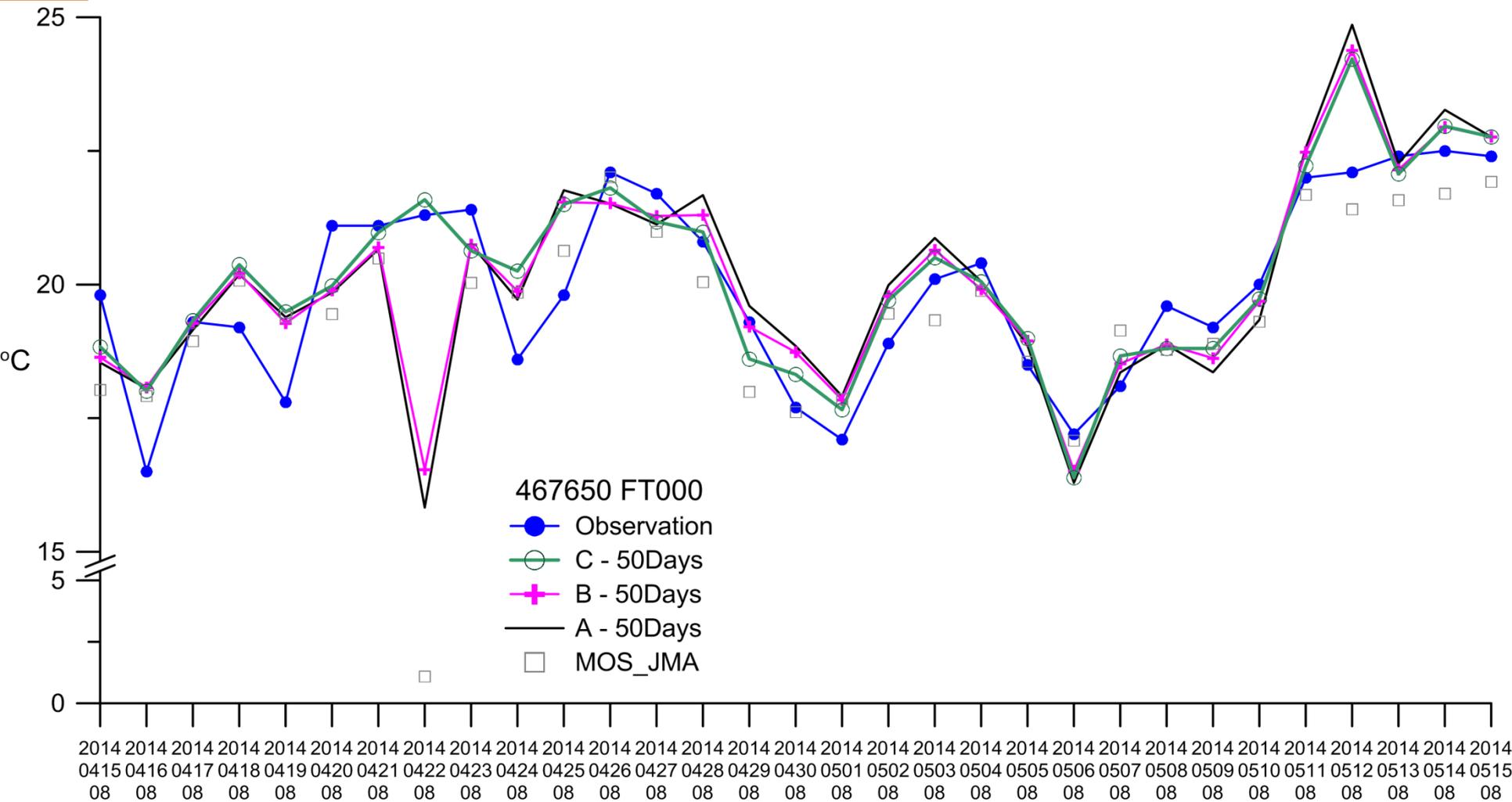
平均絕對誤差



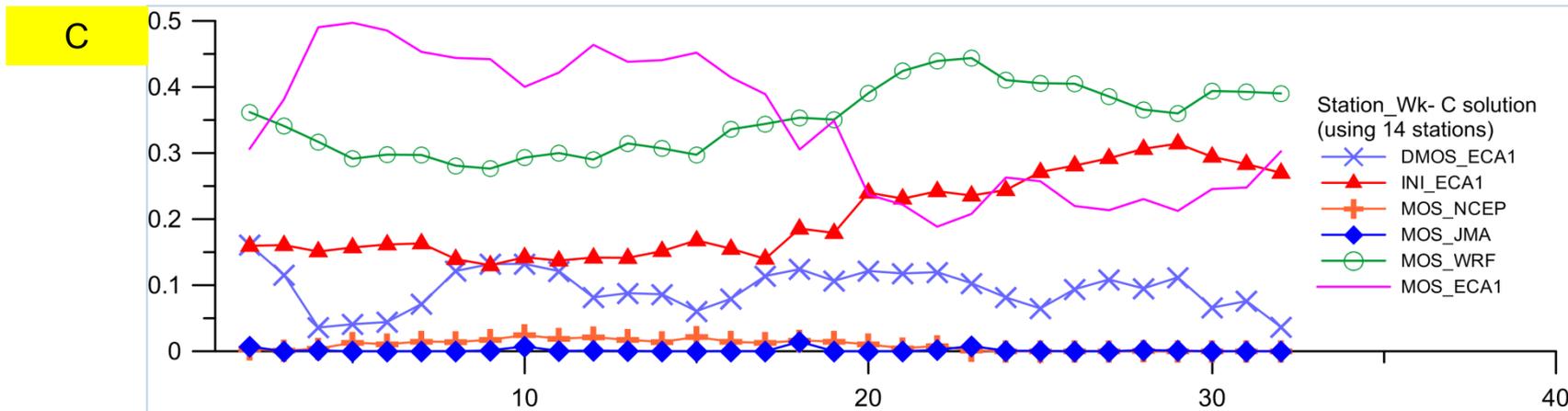
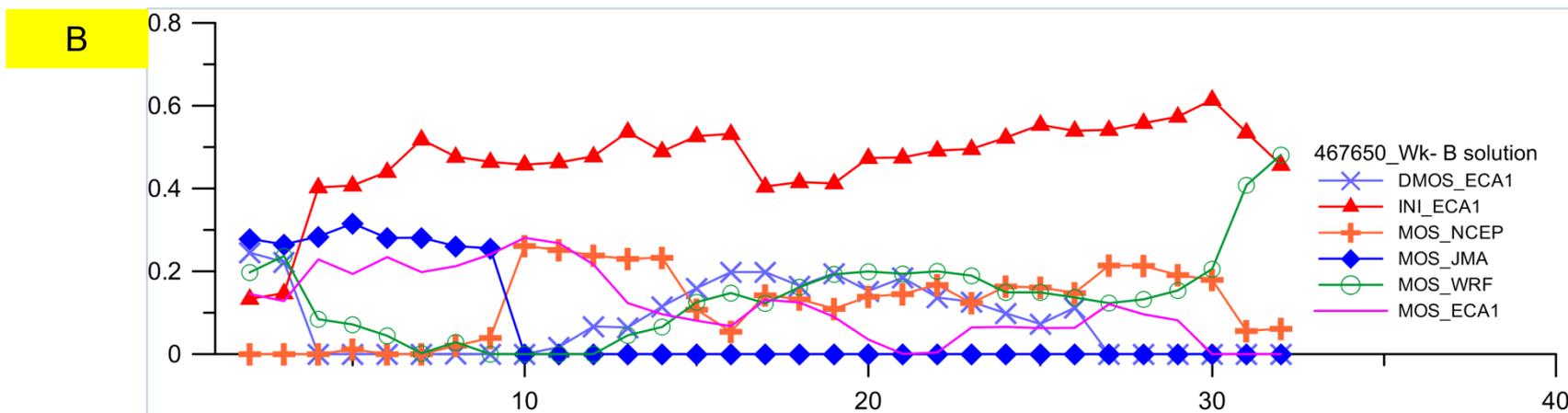
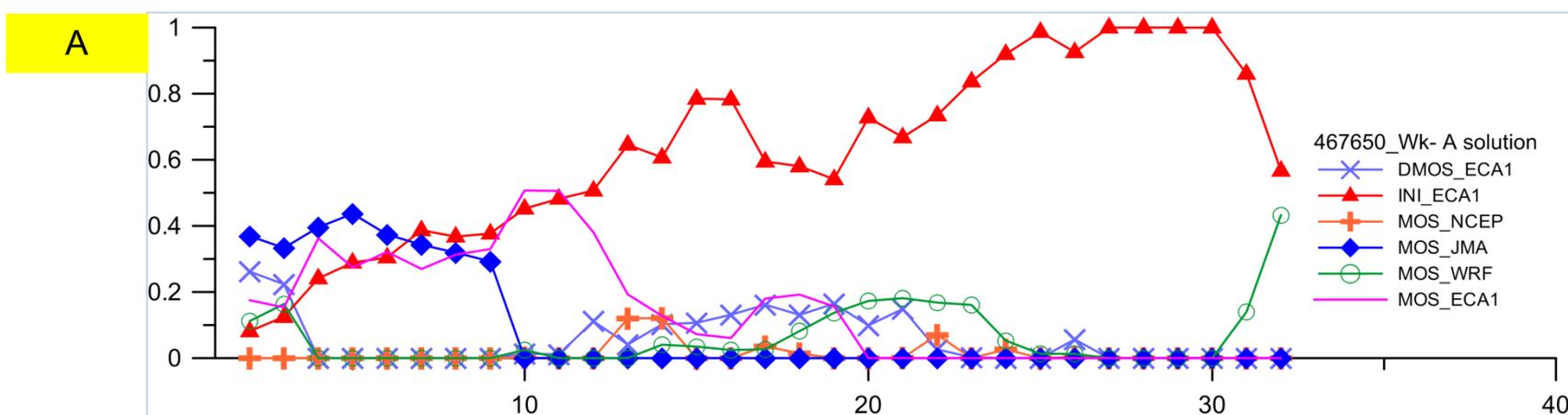
平均誤差



日月潭BMA模型時間序列

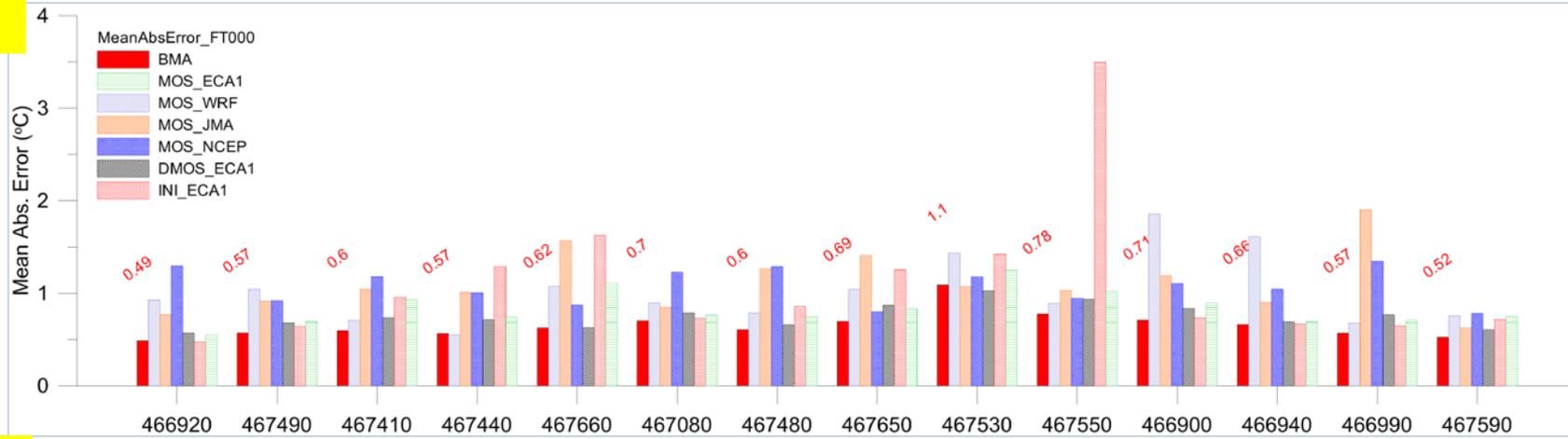


日月潭BMA模型權重係數時間序列

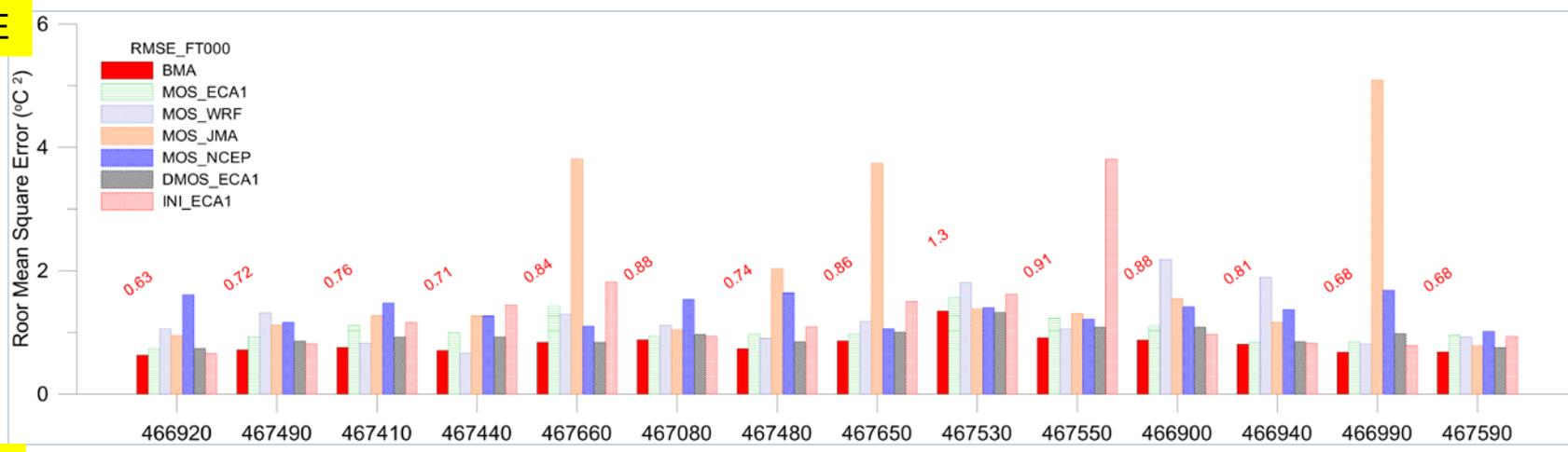


與原始預報模式比較

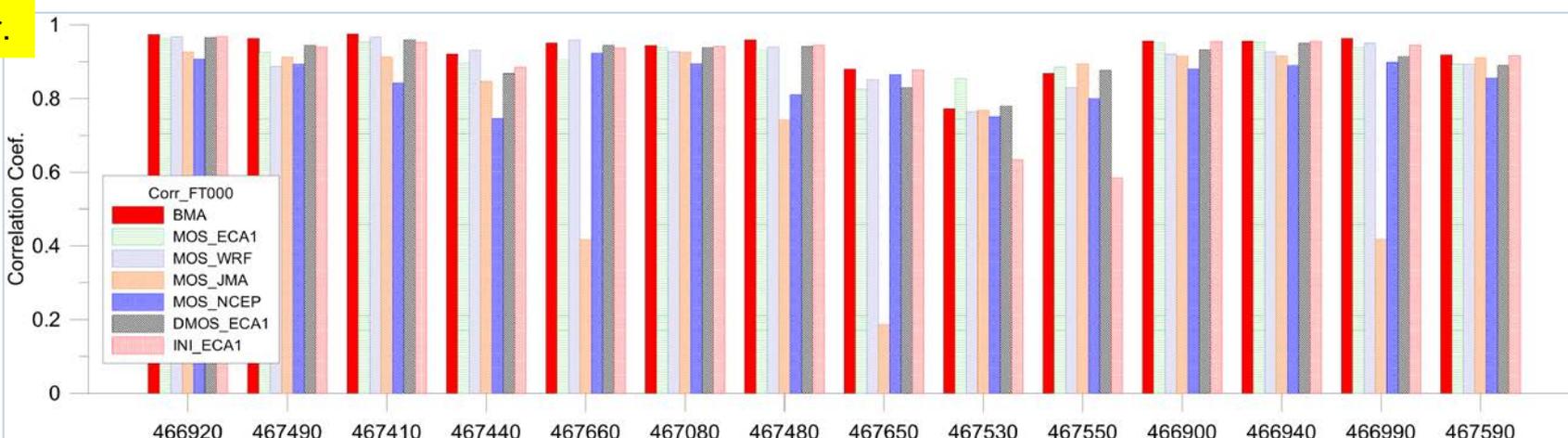
MAE



RMSE



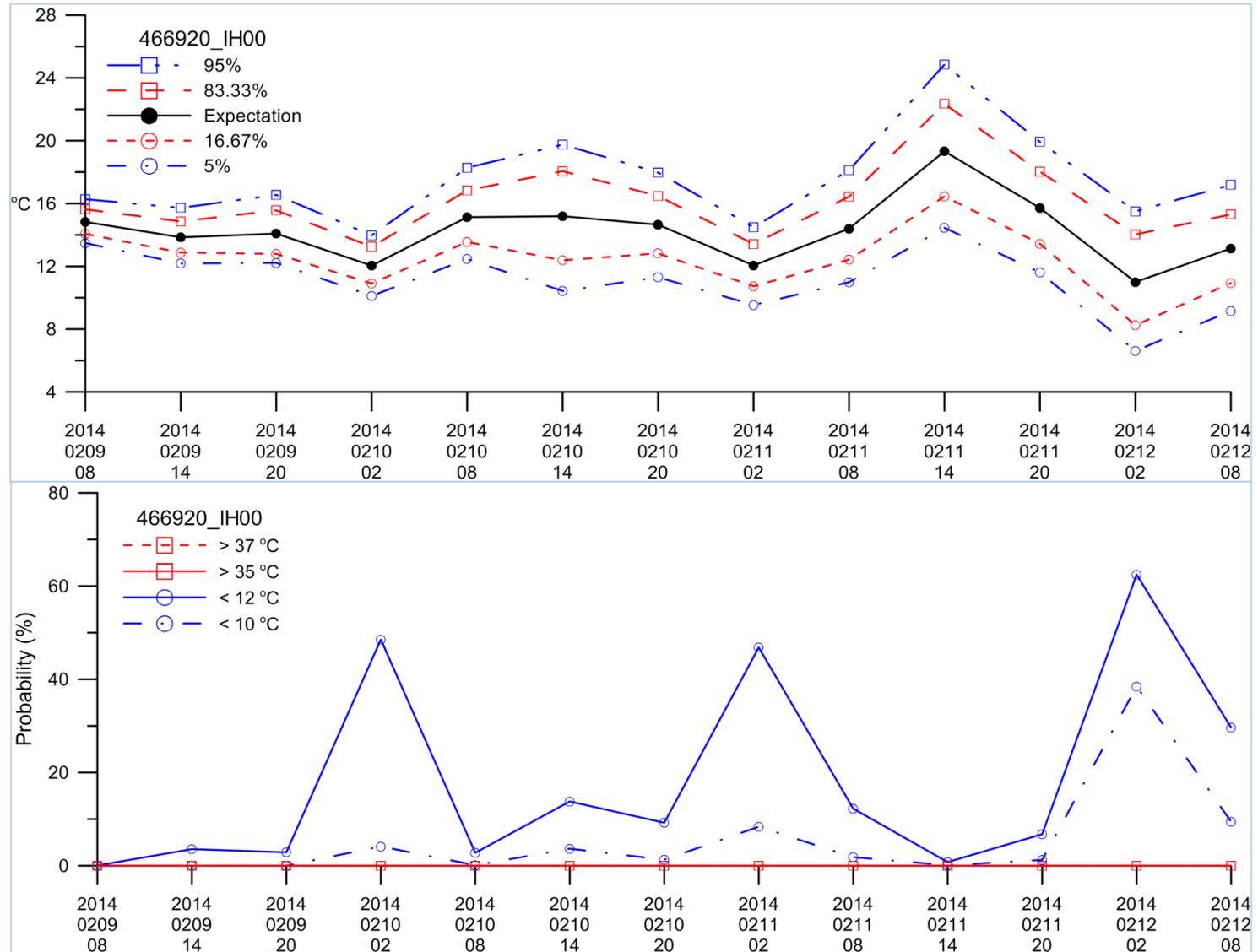
Corr.



結語與建議

- 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，**考量模式預報變異度有助於改善BMA模型成效**。
 - ▣ 以所有測站一起建置BMA模型的C方案成效最顯著
 - ▣ 另行建置性質差異較大的測站(如阿里山、玉山)
- 建議發展測站雨量BMA機率預報模型

台灣測站氣溫機率預報產品



簡報結束 敬請指教

www.manysplendid.com

