

應用EKDMOS發展系集模式溫度機率預報

馮智勇¹、林佑蓉¹、劉家豪¹、蔡雅婷²、陳雲蘭³

多采科技¹ 中央氣象局資訊中心² 中央氣象局預報中心³

摘要

藉由已發展之模式輸出統計法(MOS)通用建模環境,本研究實作系集核密度模式輸出統計方法(EKDMOS)整合系集模式輸出層場資料,評估提供溫度機率預報產品的可行性。EKDMOS是以系集平均層場資料建置天氣要素迴歸模型,各系集成員以此模型估計天氣要素後,應用核密度估計方法(KDE)將各系集成員估計值整合成一個機率密度函數(PDF),最後透過微調預報分散性而得到預報PDF。

本研究使用歐洲中期天氣預報中心(ECMWF)全球模式2007~2008年系集資料對臺灣31個測站建置2月與6月的MOS策略測站溫度迴歸模型,再以2009年資料與PIT圖、CRD圖、MAE與偏估等指標校驗事後預報。結果顯示6月預報PDF較佳,2月預報PDF過於分散,並且第50%分位數之溫度預報有偏估現象,稍差於ECT決定性預報。為改善此現象,於此參考各延時的PIT圖形選擇適當分位數作為溫度預報值,結果顯示的確可降低MAE與偏估值至ECT決定性預報水準。

關鍵字：EKDMOS、系集模式、溫度迴歸模型、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系集模式(ensemble models)模擬結果整合技術,發展各項天氣要素可靠之機率預報指引。中央氣象局已於「強化災害性即時天氣預報」建置案發展完善的模式輸出統計法(Model Output Statistics, MOS)與完美預測(Perfect Prognosis, PP)降尺度策略通用迴歸模型建置環境,可針對測站點與網格點(2.5 km 解析度)建置中期天氣(7天)逐12小時及短期天氣(3天)逐3小時各項天氣要素迴歸模型,除可每日產製定量預報指引(guidance)供預報作業參考外,更可因應校驗需求進行離線重製預報指引[1],已具有相當完整且彈性的發展架構與流程,適合進一步發展機率預報產品。

目前美國國家氣象局(National Weather Service)隸屬之氣象發展實驗室(Meteorological Development Laboratory)針對單一或多種系集模式發展“系集核密度模式輸出統計法”(Ensemble Kernel Density Model Output Statistics, EKDMOS),並且實作於產製整合NCEP (National Centers for Environmental Prediction)與CMCE (Canadian Meteorological Centre's Ensemble Prediction System)兩種系集模式之地面溫度機率預報產品([2, 5]),校驗結果顯示使用EKDMOS對於溫度的估計上較優於決定性預報,並且認為使用兩系集預報系統能得到較單一系集模式為佳的可信度。

EKDMOS透過MOS、核密度估計(Kernel Density Estimation, KDE)與微調預報分散性(spread adjustment)等三個步驟整合系集模式預報結果而得機率預報產品。其中,MOS多變數迴歸模型係利用各預報點的觀測值與各種輸出層場資料之系集模式算術平均所建置;核密度估計方法用於整合各系集成員 MOS模型預報天氣要素值得到機率密度函數(Probability Density Function, PDF);最終將過於分散的系集成員預報天氣要素值依照系集模式特性向預報值中位數微幅調整。

EKDMOS的發展關鍵僅在於需具備系集模式輸出層場資料,非常適合氣象局採用並藉由既有通用迴歸模型建置環境整合作業化WRF系集模式,產製各項天氣要素的機率預報指引。本研究考量資料取得便利性,選擇採用可自歐洲中期天氣預報中心(European Centre for Medium-range Weather Forecast, ECMWF)官方網頁[3]下載的全球系集模式輸出層場資料進行技術發展,提供後續採用氣象局WRF系集模式時做參考。

以下先行說明Glahn等人[2]與Veenhuis[5]提出的EKDMOS機率預報法流程,再以2007年~2009年ECMWF全球系集模式資料與台灣測站溫度觀測進行模型建置與結果分析,最後綜整研究結論與建議,以為後續發展作業系統參考。

二、EKDMOS機率預報法

(一)MOS策略測站溫度預報模型

MOS策略測站溫度迴歸模型建置時是以地面溫度觀測資料為預報元，選擇數值動力模式 τ 延時輸出層場為預報子資料，建置多變量線性迴歸模型如式(1)。整理本研究使用的ECMWF全球模式層場於下表1，其餘模型建置相關參數設定請參考[1]。

$$\hat{y} = a_0 + \sum_{i=1}^{10} a_i x_i \quad (1)$$

表1 建置MOS策略迴歸模型採用之輸出層場

模式	資料年份	輸出層場
ECMWF	2007~2009	HTUVQ@H100/925/850/700/500/300/250/200、T2M、10mu、10mv、MSPL

(二)核密度估計方法

核密度估計方法用於估計未知的PDF，為一種無母數方法(nonparametric method)，優點在於不需要數據的先驗知識，亦不針對數據的分布給予任何假設，而由數據本身找出特徵進而估計其PDF的方法。核密度估計之方程式如下式(2)所示，其中， x_o 為估計值， x_i 為系集成員預報值， n 為系集成員數目， h 為平滑參數， $K(*)$ 為核密度函數。常用之核密度函數包含有二次型(Quadratic)、三角型(Triangular)、四次型(Quartic)、高斯函數(Gaussian)。

$$\hat{f}(x_o) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x_o - x_i}{h}\right) \quad (2)$$

(三)預報分散性微調

預報分散性微調示意圖如圖1，左方的PDF表示為有最小預報值 F_{min} 之系集成員的PDF(標準差為 σ_{min})，其黃色區域表示為有最小預報值之系集成員且其值小於 F_{min} 之區域；而右方的PDF則為有最大預報值 F_{max} 之系集成員的PDF(標準差 σ_{max})，藍色區域則代表有最大預報值之系集成員且其值大於 F_{max} 之區域；灰色區域則為 F_{min} 與 F_{max} 之間的區域；因此當各個系集成員之間的預報PDF愈分散時，灰色區域越大，擴散調整的效果愈明顯，其擴散調整方式可以下式表示[2]。

$$x' = \frac{3(\sigma_{min} + \sigma_{max}) + sf(F_{max} - F_{min})}{3(\sigma_{min} + \sigma_{max}) + (F_{max} - F_{min})} \quad (3)$$

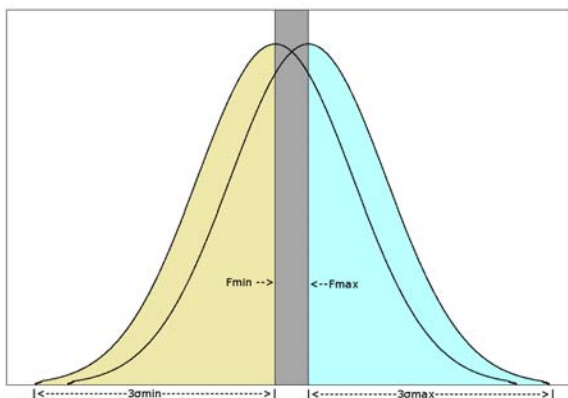


圖1 預報分散性微調示意圖， F_{min} 和 F_{max} 分別為系集模式預報之最小值與最大值， σ_{min} 和 σ_{max} 為其對應的標準差(摘自[2])。

式(3)中， sf (spread factor)為分散性調整經驗參數，Glahn等人[2]建議使用GEFS資料可採用 $sf = 0.5$ ，而Veenhuis [5]整合GEFS與CMCCE時採用 $sf = 0.4$ ，可知與系集模式相關。

(四)單一系集模式應用EKDMOS發展流程

參考Glahn等人[2]方法，應用EKDMOS方法在單一系集模式成員的流程如圖2所示，說明步驟如下：

1. 計算各系集成員層場資料之系集平均值
2. 以各層場系集平均值與測站觀測值建置MOS迴歸模型
3. 利用步驟2的MOS迴歸模型估計各系集成員之標準差與預報平均值
4. 以各系集成員於步驟3所得之標準差與預報平均值為採高斯分布核函數的參數，個別建置PDF如圖3中的紅線。
5. 透過式(2)整合各系集成員PDF為一，如圖3中的綠線。
6. 將整合PDF(綠線)以下式往系集預報平均之中位數 x_m 微調

$$x' = x_m + (x - x_m) \times \frac{3(\sigma_{min} + \sigma_{max}) + sf(F_{max} - F_{min})}{3(\sigma_{min} + \sigma_{max}) + (F_{max} - F_{min})} \quad (4)$$

7. 以步驟6微調後的曲線下方面積調整 x' 處的機率值

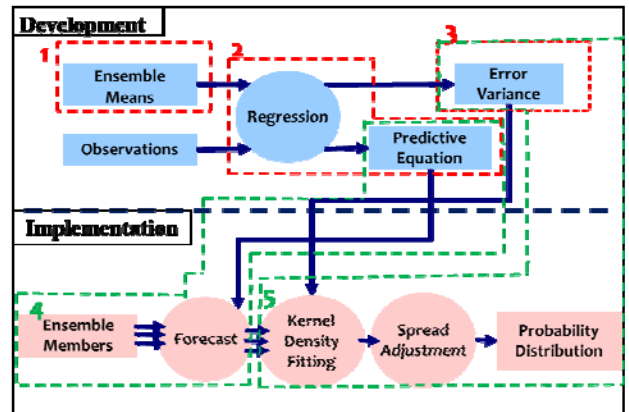


圖2 應用 EKDMOS整合單一系集模式成員流程

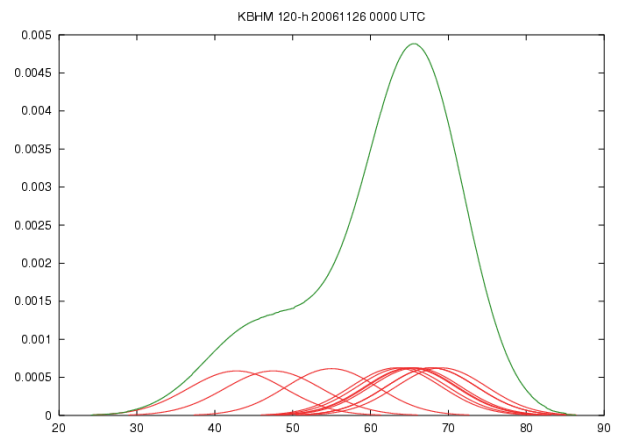


圖3 KDE整合示意圖。紅色線為各個系集成員之PDF，綠色線則為整合後的PDF(摘自[4])。

三、測站溫度EKDMOS機率預報實驗

本研究以2月與6月08時測站溫度機率預報為目標，取得2007至2009年1月至3月與5月與7月期間台灣31個人工站08時溫度觀測值，並且下載該段期間 ECMWF 模式00分析時間(UTC)之決定性預報(deterministic forecast)與21組系集成員(perturbed forecast)輸出層場資料，採用2007年與2008年資料建置 00、24、48、72、96與120等6組延時MOS策略迴歸模型，並以2009年資料進行事後預報。圖4為台北站利用2009年6月1日ECMWF預報場資料估計6月1日~6日08時溫度的EKDMOS機率預報產品，可供後續依實際觀測值進行校驗修正調整。

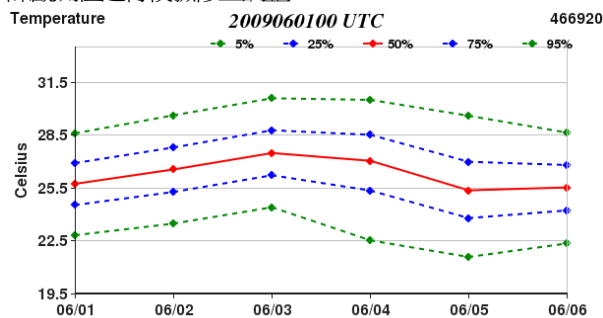


圖4 台北站2009年6月1日08時溫度EKDMOS機率預報產品，由上而下依序為95%、75%、50%、25%與5%累積機率於00、24、48、72、96與120延時之預報值連線

為進一步了解EKDMOS機率預報的特性，參考Glahn 等人[2]設計實驗分成如表2所示之4組。其中，“ECT”代表採用ECMWF決定性預報層場資料所建置的溫度模型，“ECM”為以系集平均資料建置溫度模型，“KDE”同“ECM”但應用於各別系集成員得到溫度預報值後再使用核密度估計法取得預報PDF。

表2 測站溫度EKDMOS機率預報實驗設計

代號	MOS 模型 建置資料來源	MOS 模型 應用標的	機率分布型態
ECT	model output	model output	normal
ECM	Ensemble Means	Ensemble Means	normal
KDE	Ensemble Means	Ensemble Members	KDE
EKDMOS	Ensemble Means	Ensemble Members	KDE + spread-adjustment

利用各測站2009年2月1日~28日08時的4組實驗預報值時序與觀測值時序計算所得之平均絕對誤差(Mean Absolute Error, MAE)折線圖以及平均偏差值(BIAS)長條圖如圖5所示，橫軸為延時場，縱軸為溫度差值(°C)，以紅、藍、綠與黃色依序代表“ECT”、“ECM”、“KDE”與“EKDMOS”實驗組。其中，“KDE”與“EKDMOS”兩組實驗的預報值是採50%累積機率的估計值。由圖可知“ECT”的MAE最小，而整合系集成員得到“ECM”、“KDE”與“EKDMOS”的MAE皆相似，各組實驗的MAE均隨著延時增加而上升，並且自72延時“ECT”明顯優於其他3組；另一方面，所有

實驗之溫度預報結果皆為負偏差(低估)現象，除於24延時有最小偏差外，偏差值隨著延時增加而上升。

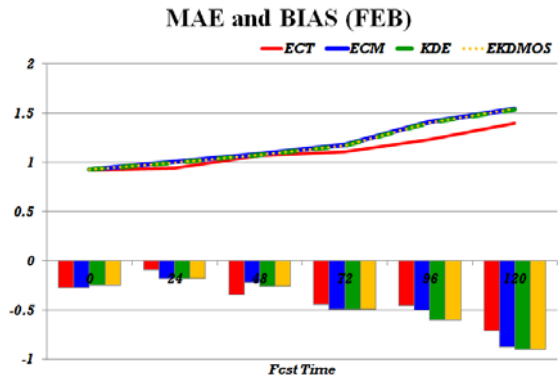


圖5 表2各組實驗在2009年2月08時之平均絕對誤差(折線)與平均偏差值(長條)隨預報延時00、24、48、72、96、120時的變化情況，紅、藍、綠與黃色依序代表“ECT”、“ECM”、“KDE”與“EKDMOS”實驗組。

圖6則為各組2009年6月1日~30日08時的誤差隨延時變化情況，可知各組實驗的MAE略優於2月預報結果並同樣隨延時的增加而上升，“ECT”的平均偏差值同樣於24延時場最小，延時72、96與120的平均偏差值幾乎皆為-0.2°C；“ECM”、“KDE”與“EKDMOS”的平均誤差值於00~72延時呈現負偏差且量值與“ECT”相當，但於96與120延時轉為正偏差且量值約為0.8°C~0.9°C，明顯增加許多。

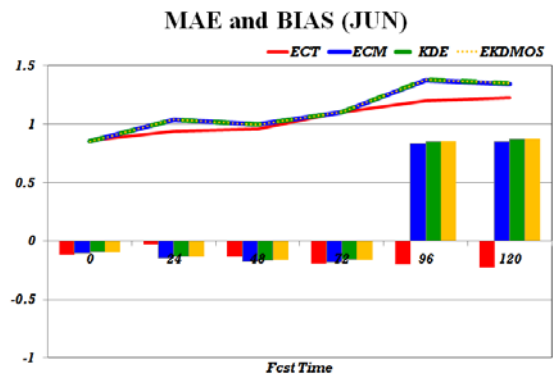


圖6 同圖5，但為表2各組實驗在2009年6月08時之平均絕對誤差(折線)與平均偏差值(長條)隨預報延時變化情況。

為進一步了解各組實驗預報特性，以2009年2月1日~28日08時的各組實驗預報與觀測資料繪製00、24、48、96與120延時的(Probability Integral Transform, PIT)圖如圖7所示，“ECT”與“ECM”兩組實驗於00~72延時呈現「鐘形」分布(72延時與48結果相似)，意即實際觀測值於0.9和0.1的兩側出現的頻率相對於其它區間來的小，代表溫度預報 PDF過於分散，使得觀測值集中出現在機率預報區間為0.4至0.8之間。此外，“KDE”與“EKDMOS”兩組實驗的PIT圖幾乎完全相同，此乃因系集成員溫度估計值差異不大，使得微調預報分散性的結果無由於PIT圖中呈現差異；各延時均呈現「鐘形」分布的情況，觀測值出現頻率最高的區間由00延時的0.6隨著延時增加而逐漸往橫軸右方移動，代表兩組實驗的預報值逐漸低估，與圖5所見吻合。

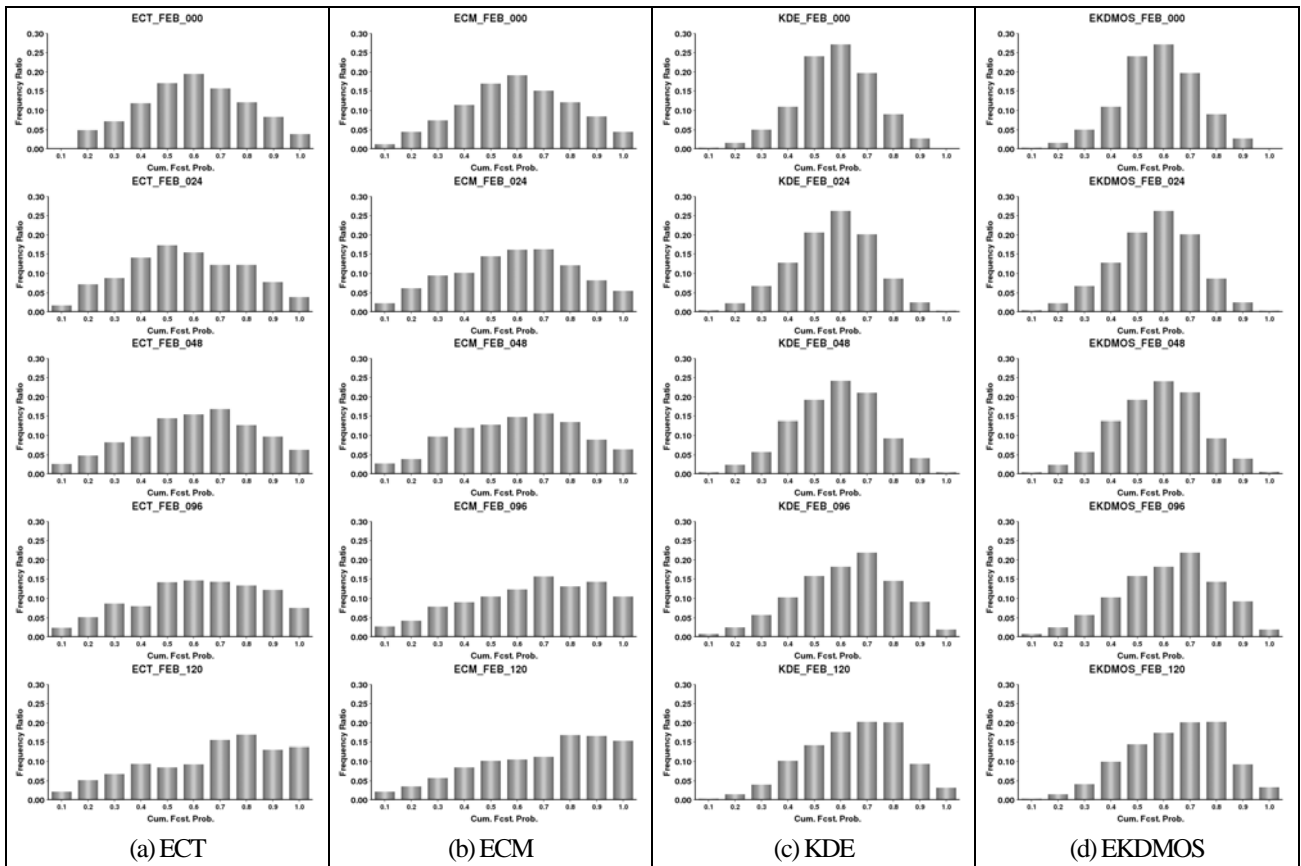


圖7 2009年2月08時“ECT”、“ECM”、“KDE”與“EKDMOS”實驗組溫度預報結果於各延時的PIT圖。由上而下依序為00、24、48、96與120延時

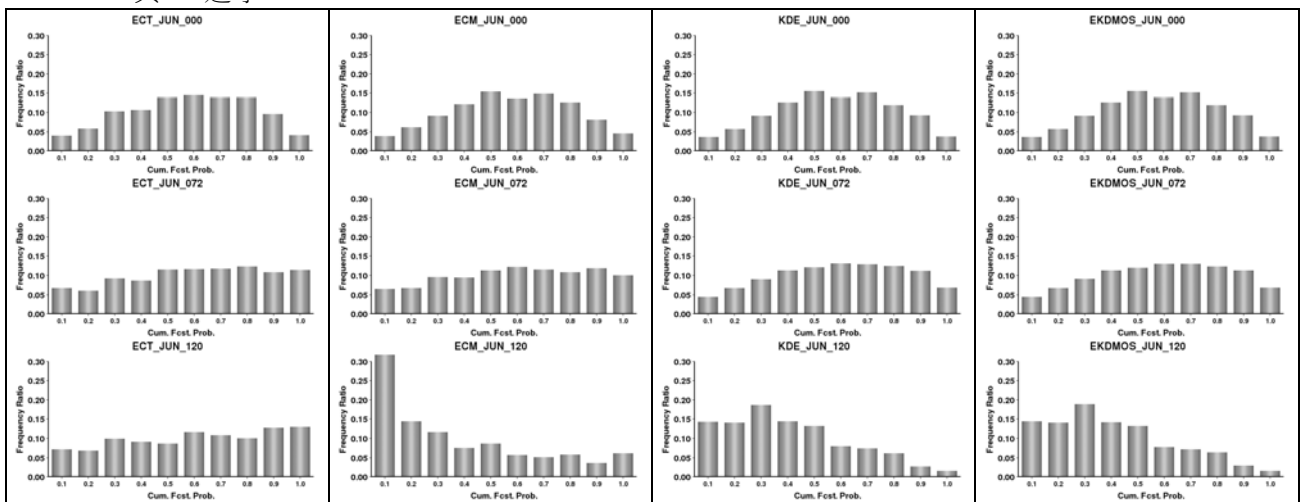


圖8 2009年6月08時“ECT”、“ECM”、“KDE”與“EKDMOS”實驗組溫度預報結果於各延時的PIT圖。由上而下依序為00、72與120延時

圖8則為以2009年2月1日~28日08時的各組實驗預報與觀測資料繪製00、72與120延時的PIT圖，由於延時24與48時之PIT相似於00延時，而96則相似於72或120時，因此僅顯示3個預報延時的圖形。由圖可知，00延時各組均類似「鐘形」分布並且較2月PIT圖為平緩，“KDE”與“EKDMOS”結果近乎相同。此外，觀測值出現於“ECT”各機率預報區間的頻率隨著預報延時增加約略相同，代表預報PDF良好；反觀其餘3組預報PDF則由00延時的過於分散，轉變為120延時

的分散不足而致預報高估的情況(與圖6吻合)，尤其以“ECM”於預報區間0.1處的觀測值出現機率高達0.3為最。相較而言，“KDE”與“EKDMOS”兩組可略為改善溫度預報PDF不夠分散的現象，但其極大頻率值出現的位置仍偏向較小的機率預報區間。

總體而言，將ECMWF之全球模式系集資料以EKDMOS方法應用在臺灣地區31個觀測站進行預報時，2月溫度的預報PDF過於分散，但6月的預報PDF則分散性不錯，而CDF上第50%分位數之溫度預報有

低估溫度的現象，其MAE相較於決定性預報(“ECT”)稍差。

由於EKDMOS方法的優勢在於可同時提供機率預報PDF與定量預報值，上述結果顯示採用50%累積機率的估計值為定量預報的結果無法保有ECMWF模式決定性預報的準確性。以2009年6月溫度預報為例，由圖6與圖8可看出“EKDMOS”的50%累積機率預報結果在00~72延時低估08時溫度，代表觀測值多落在“EKDMOS”累積機率50%以上區間，而96與120延時呈現高估則代表觀測值多落在累積機率50%以下區間。因此，本研究於此嘗試於“EKDMOS”預報延時48、72、96與120時的定量預報分別改用55%、55%、35%與30%累積機率的估計值，重新繪製MAE與平均偏差如圖9中“TEST”(綠色)所示，明顯降低“EKDMOS”定量預報的平均偏差。

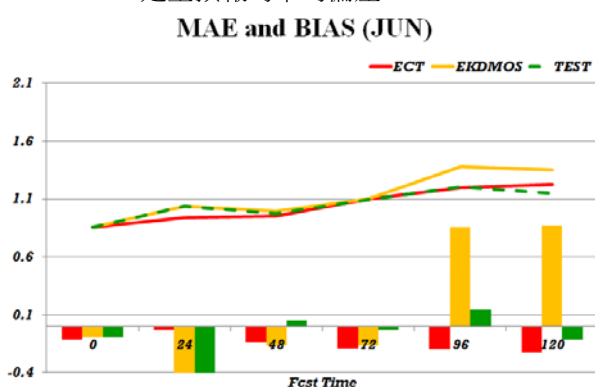


圖9“ECT”(紅色)、“EKDMOS”(黃色)與調整後之“TEST”實驗組2009年6月08時之平均絕對誤差(折線)與平均偏差值(長條)隨預報延時00、24、48、72、96、120時的變化情況。

四、結論與建議

本研究主要在建構EKDMOS模型並評估其預報之可行性，由應用ECMWF全球模式系集資料對臺灣地區31個測站進行2009年2月與6月之溫度預報之校驗結果顯示，適當的應用機率預報產品，可以得到相對於決定性預報而言有更好的預報結果，並且可有效修正其溫度偏差。因此，當EKDMOS方法用於CWB WRF系集模式時(或其它較高解析度之系集預報模式)，可以依循相同的方式獲得機率預報產品。

而EKDMOS之預報溫度校驗結果皆初步分析臺灣地區31個測站之平均值，並無針對各別測站做較深入的探討。雖然分析結果顯示出機率預報產品可提供較為準確的預報，但臺灣地區地形起伏大，平原及高山測站有其不同的溫度型態，由平均值之校驗結果無法得知各測站的預報結果，故建議未來可針對各別「測站」找出適當的機率預報產品；同理，對於「季節」或「系集模式系統」以及不同「預報延時」也建議分別選擇適當的CDF分位數來做預報，以得到最佳的預報結果。

由於ECMWF之各系集成員對於台灣測站08時溫度預報結果差異不大，因此無由評估使用Glahn等人[2]或Veenhuis[5]建議之預報分散性微調方法成效，建議未來若利用氣象局WRF系集模式進行實作時，可進行測試予以評估。

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. Glahn, H.R., M. Peroutka, J. Wiedenfeld, J. Wagner, G. Zylstra, B. Schuknecht, and B. Jackson, 2009: “MOS Uncertainty Estimates in an Ensemble Framework”. *Mon. Wea. Rev.*, 137, 246–268.
3. <http://apps.ecmwf.int/datasets/data/tigge/>
4. http://www.dtcenter.org/events/workshops11/nuopc_2011/presentations/EKDMOS_NUOPC_John_Wagner.pdf
5. Veenhuis, B., 2013: “Spread Calibration of Ensemble MOS Forecasts”. *Mon. Wea. Rev.*, 141, 2467–2482