

應用系集核密度模式輸出統計進行臺灣測站 1~14 天極端 低溫機率預報

陳昫靖 張惠玲*

中央氣象局氣象科技研究中心

(中華民國 111 年 7 月 26 日收稿；中華民國 111 年 12 月 14 日定稿)

摘 要

本研究採用美國國家環境預報中心(National Centers for Environmental Prediction, NCEP) 第 12 版全球系集預報系統(Global Ensemble Forecast System version 12, GEFS v12)之重預報資料，透過統計後處理技術-系集核密度模式輸出統計(Ensemble Kernel Density Model Output Statistics, EKDMOS)進行偏差校正與降尺度，目的在於得到臺灣地區特定測站點上具有良好預報品質與價值的 1-14 天極端低溫機率預報。

預報評估顯示: (1)原始系集預報有離散度不足的問題，且有明顯的預報偏差；相較之下，EKDMOS 有相當合理的系集離散度，並可移除絕大部分的預報偏差。(2)相較於模式原始系集預報，EKDMOS 能提供可信度更高且區辨能力更佳之機率預報。(3)EKDMOS 透過提升可信度與解析能力來提升原始模式的 BrSS。(4)更多的使用者可採用 EKDMOS 預報作為決策依據，而得到高於參考原始預報之經濟效益，特別是成本/損失比接近 1 與 0 的使用者有最顯著的經濟價值提升。(5)即便 EKDMOS 可以顯著地改善極端溫度機率預報品質，當預報時間拉長到 10 天以上而逐漸喪失預報訊息度或預報訊號時，EKDMOS 的預報分布仍無可避免地趨近於訓練樣本的氣候分布，這是數值天氣預報在氣象可預報度上的限制。

關鍵字：極端低溫機率預報、系集核密度模式輸出統計、可信度、區辨能力、預報訊息度

通訊作者：張惠玲博士，地址：10055 台北市中正區公園路 64 號

Email: lingo@cwbgov.tw，Tel: +886-2-2349-7967

一、前言

近年來，世界各國之氣象單位致力於氣象資訊的跨領域應用，如何將氣象資訊應用於各領域之決策，並提升經濟效益是氣候服務的重要目標。極端冷事件造成的災害與損失是許多產業關心的議題；以農業來說，低溫事件會對不同作物的不同生長期造成負面影響，進而影響產量，帶來經濟損失。為提前做好防災準備以降低損失，各產業對於展期(10-30 天)極端低溫預報的需求與日俱增。

在數值天氣預報過程中，由於數值方法、非線性方程、模式初始場與邊界條件的不確定性、次網格與參數化過程等數學與物理上的極限，不可避免地存在各種數值運算誤差、背景與觀測誤差，且誤差會隨著模式積分時間增加而放大，進而嚴重影響預報結果。因此，當預報拉長至 2 週以上，預報不確定性(forecast uncertainty)明顯增大，使用單一預報模式做為決策依據有著相當大的風險。為了提供預報之不確定性資訊，氣象學家發展出系集預報模式(Leith, 1974)，且系集預報漸漸取代單一模式做為決策依據。然而大部分的系集預報系統存在著系統性誤差，且離散度不足。

增加系集成員數量(Buizza and Palmer 1998)以及採用良好的資料同化方法(Buizza et al. 2005)都有助於改善模式的離散度，讓系集預報更能掌握各種可能的情境。考量計算資源以及作業時效，大多數預報作業中心之系集成員數有限，無法涵蓋預報過程中所有的不確定性來源，導致系集離散度不足，進而影響系集機率預報的可信度。此外，目前全球模式之網格解析度仍不夠精細，以美國國家環境預報中心(National Centers for Environmental Prediction, NCEP)目前的作業模式-第 12 版全球系集預報系統(Global Ensemble Forecast System version 12, GEFS v12)為例，網格解析度約為 25 km (Zhou et al., 2019)，無法解析臺灣地區複雜的地形(圖 1)，以至於難以直接針對特定地點或小區域提供具參考價值的預報資訊。另一種情況是使用者關心的近地面天氣要素並非數值模式的直接輸出場量，例如日最高溫、日最低溫、雲量、日射量、最大陣風等。上述原因導致使用者難以直接應用全球系集預報模式的展期預報輸出，必須透過統計後處理技術進行偏差修正、降尺度以及客製化預報產品。

常見的統計後處理技術包括模式輸出

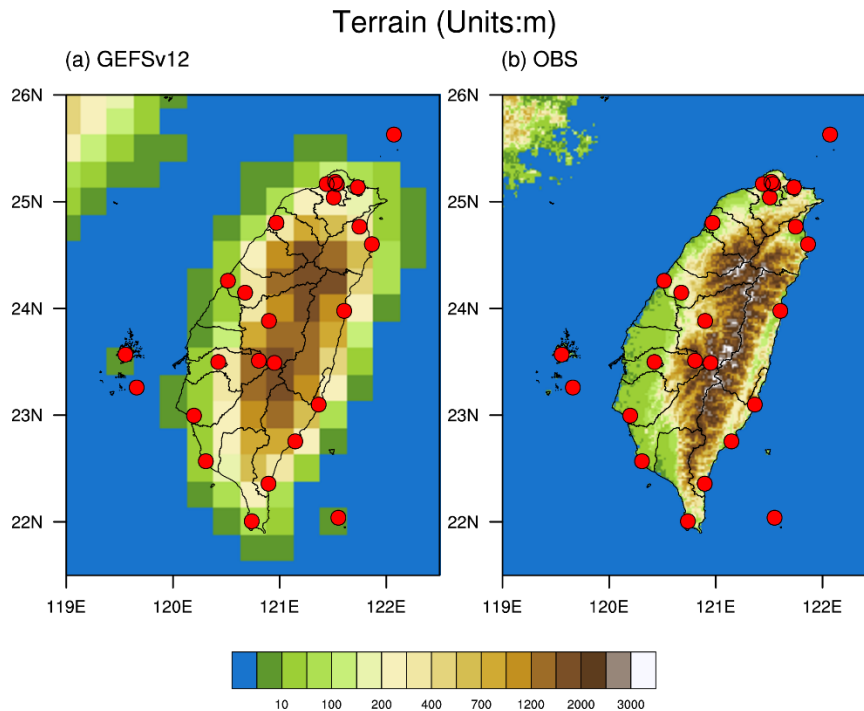


圖 1、本研究採用之測站分布圖，測站分別標示於(a)模式地形、(b)實際地形上。

統計 (Model Output Statistics, MOS; Glahn and Lowry 1972)、類神經網路技術 (Artificial Neural Network techniques, ANN; Yuan et al. 2007; Chu et al. 2007; 章等 2018)、線性迴歸方法 (Linear Regression method, LR; Yuan et al. 2008)、Decaying Average(Cui et al. 2012)、頻率配對校正 (Frequency Matching Method, FMM; Zhu and Luo 2015)、Quantile Mapping (Zhao et al. 2017) 等。其中 MOS 與線性迴歸方法一樣，是透過歷史長期訓練樣本資料，建立預報因子(predictor)與預測值(predictand)的關係式，可以同時包含降尺度與系統性誤差修正的功能(Glahn

and Lowry 1972)。MOS 主要是建立數值天氣預報(numerical weather prediction, NWP)與使用者關心之特定地區、特定氣象變數間的迴歸關係，特別是近地面的氣象參數。這是因為數值模式的解析度較粗，並且以參數化方式來近似邊界層的物理過程，導致模式對於使用者關心的近地面氣象參數無法提供具參考價值的預報結果。隨著系集預報的發展，隨之有以 MOS 為基礎發展出來的 Ensemble MOS 以及系集核密度模式輸出統計(Ensemble Kernel Density model output statistic, EKDMOS; Glahn et al. 2009)，並應用於短期溫度預報。Ensemble MOS 並非針對所有系集成員分

別建模，而是以動力模式的系集平均預報來建模，再將建好的統計模型套用到所有系集成員，這樣做的目的在於維持上游動力模式的系集離散度；若是不同的系集成員都分別建模，則所有成員的預報結果會非常接近觀測，而導致系集離散度嚴重不足。由於大部分系集預報系統有系集離散度不足的現象，因此即便 Ensemble MOS 採用系集平均預報來建模，再將建好的統計模型套用到所有系集成員，還是無可避免地出現系集離散度不足的問題。EKDMOS 則是進一步將 MOS 系集平均的誤差分佈，透過核密度估計(Kernel Density Estimation, KDE)套到 MOS 所有系集成員，以得到具有良好離散度的機率分布預報

陳等(2018)採 GEFS v11 之重預報資料(reforecast)，透過 EKDMOS 降尺度至臺灣測站，得到測站點上日最高、最低溫之預報分佈，再透過計算機率密度函數曲線下方之面積得到冬季日最低溫低於 10° C 之預報機率，研究結果顯示：透過 EKDMOS，可以明顯改善原始系集機率預報 14 天內的可信度與區辨能力，提供具參考價值的低溫事件機率預報。然而對於不熟悉機率預報的使用者來說，並不知道如

何以機率預報作出最佳決策。Richardson (2000) 提出經濟價值分析，除了評估模式潛在經濟價值外，也有助於使用者使用機率預報得到最佳決策。Chang et al. (2015) 指出預報系統的潛在經濟價值主要取決於預報的區辨能力，而具有不同成本/損失比的使用者可以透過選擇自己的最佳機率門檻，以決定是否採取防災行動，而得到最大的經濟效益。綜合上述，模式可以透過 EKDMOS 提升對極端冷事件之區辨能力，進而提升預報的經濟價值，讓下游使用者透過參考 EKDMOS 機率預報做為決策依據，得到更大的經濟效益。

簡言之，由於原始模式預報具有系統性偏差，無法提供使用者可信的預報，由於解析度過粗，無法滿足使用者對於特定點位的氣象預報需求，因此本研究的重點是對原始模式預報進行偏差修正與降尺度。由於極端溫度預報具有高度不確定性，特別是 10 天以上的展期極端溫度預報，比較恰當的呈現方式是機率型預報而非單一決定性預報；因此本研究是針對 1-14 天極端低溫機率預報，採用 EKDMOS 方法進行偏差修正與降尺度。

2020 年美國 NCEP GEFS v12 正式上線，相較於 v11 除了預報表現明顯提升外

(Zhou et al., 2019), 也提供了完整的歷史重預報資料(Guan et al., 2020)。長期的重預報資料有助於建立更穩定的統計後處理模型, 以利於對作業模式的即時預報進行偏差修正, 得到可信的機率預報。因此本研究主要使用 NCEP GEFSv12 歷史重預報資料, 透過 EKDMOS 進行偏差校正與降尺度, 目的在於得到臺灣地區特定測站點上具有良好預報品質與預報價值的 1-14 天極端低溫機率預報。第二章將簡介模式以及觀測資料, 第三章為 EKDMOS 方法介紹, 第四章是預報品質評估結果, 第五章則是經濟價值評估結果, 最後一章是結論與展望。

二、資料說明

(一) 預報資料

本研究使用美國國家環境預報中心(National Centers for Environmental Prediction, NCEP)第 12 版全球系集預報系統(Global Ensemble Forecast System version 12, GEFS v12)重預報(reforecast)資料, 透過統計後處理技術進行臺灣地區 1-14 天極端低溫機率預報。資料年份為 2000-2019 年, 採用每日 00 UTC 的預報資料; 資料水平解析度於前 10 天為 0.25 度,

之後為 0.5 度, 初始日期為星期三的預報有一個 control run 以及 10 個成員並預報至 35 天, 其餘日期皆為一個 control run 加上 4 個成員, 並只預報至 16 天。在此要特別說明, 由於本研究只分析到預報第 14 天, 因此會納入所有預報資料, 另外限於頻寬與儲存空間, 氣象局目前接收 NCEP GEFS v12 即時預報時, 雖然在預報第 1 至 10 天有更高解析度的資料, 但只儲存 0.5 度的資料。為配合即時作業所能取得之預報解析度, 因此本研究將重預報第 1 至 10 天資料處理成 0.5 度之解析度, 以利於進行預報之即時校正作業。此外, 由於全球模式預報資料初始時間為 00UTC 每 6 小時輸出一筆最低溫資料, 由於 UTC 與台灣當地時間有 8 小時之偏差, 再將 6 小時最低溫預報資料處理成日最低溫資料時, 以與台灣時間 00-24 重疊時間最多為原則。舉例來說, 預報初始時間為 2001010100 UTC lead time 為 18 小時之預報資料對應台灣時間為 2001010120, 而 lead time 為 24 小時對應時間為 2001010202。依此類推預報 18-36 小時對應台灣時間為 2001010120-20010220, 而預報 24-42 小時對應台灣時間為 20010202-20010302。由於 2001010100UTC 預報 24-42 小時與台灣時

間 2001010200-2001010300 重疊時間最高，因此取 lead time 為 24、30、36、42 四筆預報最低溫中預報值最低那筆資料作為第一天預報，後續預報天數依此類推。

(二) 觀測資料

MOS 建模以及後續預報評估所需之觀測資料來自中央氣象局 25 個人工測站之對應模式預報目標日期之觀測日最低溫，測站位置分布如圖 1。

(三) 建模及驗證

本研究聚焦在極端冷事件好發之冬季(12 至 2 月)，在建模階段，不同測站、不同預報時間與不同月份會分開建模。為了避免不同月份間有預報不連續性的現象，特定月份之建模除了使用該月份資料外，同時納入該月份前後 15 天之預報及觀測資料進行建模。

在驗證階段，本研究使用留一交叉驗證 (Leave-one-out cross-validation, LOOCV)方式來進行預報評估，以 2000 年 1 月的預報為例，其訓練資料來自 2001 至 2019 年所有 12 月 16 日至隔年 1 月 15 日之預報及觀測資料進行建模，再將 2000 年 1 月之重預報資料代入建好的模式輸出統計(Model output statistics, MOS)模型中，即

可得到偏差校正後之 2000 年 1 月預報。2001 年則是使用 2000 年以及 2002 至 2019 年同月份之資料建模，依此類推至所有年份。本研究最後呈現的預報評估結果，是所有年份分別作為驗證樣本後，再將所有驗證樣本合併進行預報校驗的結果。

三、預報方法與校驗評估

(一) 系集核密度模式輸出統計(Ensemble Kernel Density Model Output Statistics, EKDMOS)

EKDMOS 預報流程如下，並以台北測站預報第 8 天為例：

1. 將每個系集成員離測站最近的四個模式格點之日最低溫預報，透過距離平方反比的方法內插至測站點上，得到各成員在測站點上的預報。
2. 在測站點上，以日最低溫預報之系集平均作為預報因子，與測站觀測之日最低溫建立迴歸關係式，此即本研究採用之單預報因子 MOS；將驗證樣本的預報值代入迴歸方程式，即可得到偏差修正後的預報值。
3. 如圖 2 所示，直方圖為 MOS 系集平均預報誤差分布，其分布近似於常態分布，因此可以透過計算預報誤差之平均

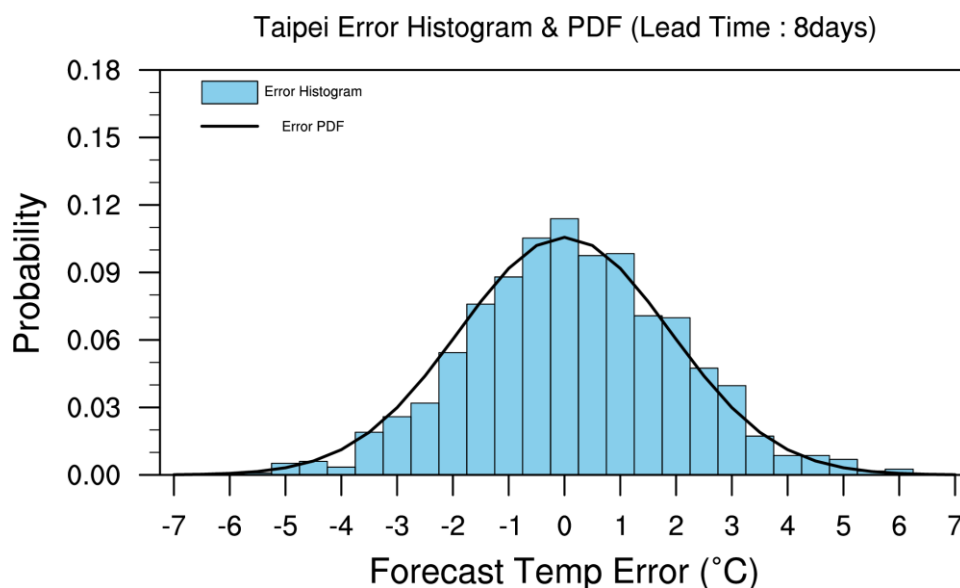


圖 2、臺北測站第 8 天日最低溫預報之誤差直方圖(藍色)與套配後的常態分布機率密度函數(黑色曲線)，橫軸為誤差值(°C)，縱軸為機率。

表 1、使用資料來源、時段整理

觀測資料	25 人工站(圖 1)	
模式資料	使用模式	NCEP GEFSv12
	解析度	0.5*0.5 度
	成員數	每週三 1 control run + 10 members 非週三 1 control run + 4 members
	預報因子	最靠近測站之 4 個模式格點內插至測站點的日最低溫預報
建模及驗證 (模式+觀測)	(1)各測站於不同月份、不同預報天數分別建模 (2)採留一交叉驗證	
	訓練樣本	(1) 2000 年預報使用 2001-2019 年之資料建模，2001 年預報使用 2000 年及 2002-2019 年資料建模，依此類推。 (2) 為減緩跨月導致之預報不連續，建模時除了當月的資料外，也會納入該月前後 15 天之資料建模。(例如 2 月建模之訓練樣本會納入 01/17 - 03/15 之資料)
	驗證樣本	每年 12、1、2 月(冬季)
極端門檻 (觀測)	不同測站根據歷史觀測找出各自的極端溫度門檻	
	年份	2000-2019 年
	月份	12、1、2 (冬季)
	極端門檻	將所有冬季樣本的日最低溫觀測排序(由低至高)後，取第 5、第 10 以及第 15 百分位做為不同極端程度之冷事件門檻。

與標準差來套配出誤差機率密度函數 (Probability Density Function, PDF)，如圖中的黑色曲線所示。

4. 將各系集成員(表 1)內插至測站點上的預報，代入系集平均建立的迴歸方程式中，得到 Ensemble MOS (EMOS)。
5. 最後將誤差的 PDF 作為 EMOS 所有成員之「kernel」，得到各個成員之預報 PDF。由於長期統計下來，每個成員之預報能力相當，因此本研究將所有成員之預報 PDF 等權合成單一 PDF，此即為 EKDMOS 之預報 PDF，如圖 3 之黑線所示。

(二) 極端低溫門檻計算方法

本研究依照以下步驟計算出觀測氣候第 5、10 與 15 百分位對應之溫度值，做為不同程度之極端低溫門檻。

1. 將各測站 2000 至 2019 年冬季(12 月至隔年 2 月)之所有日最低溫觀測資料由低到高進行排序。
2. 計算步驟 1 所有觀測資料之第 5、10 與 15 百分位溫度值，做為不同程度之極端低溫門檻。

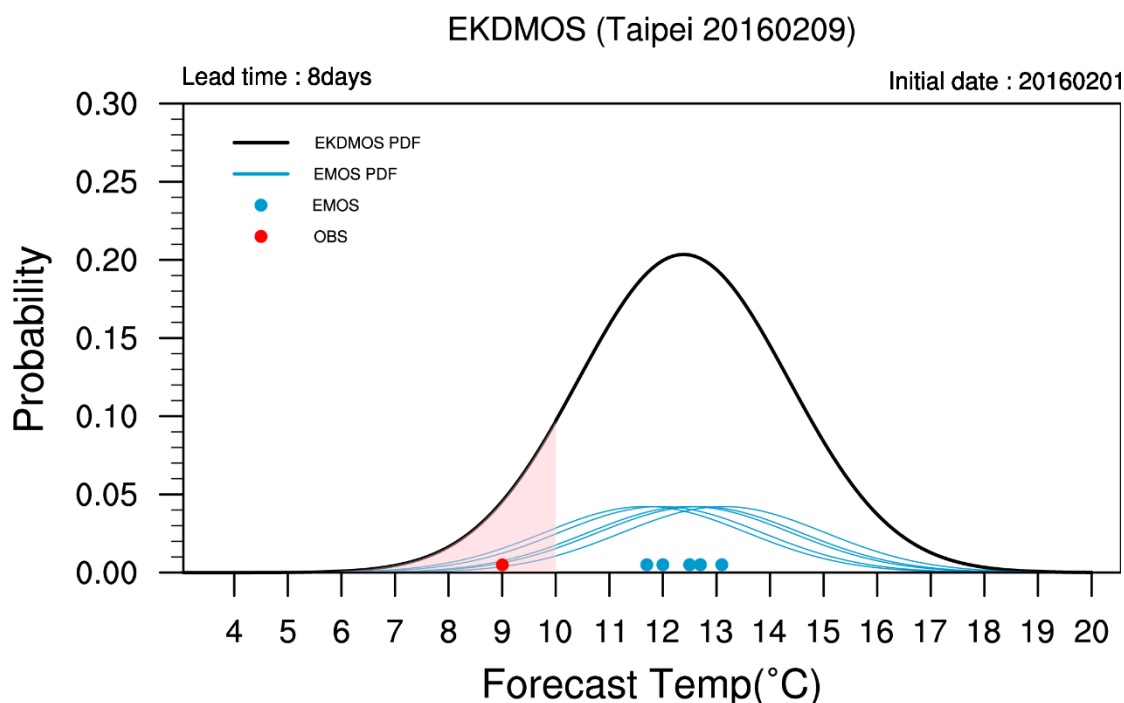


圖 3、EKDMOS 示意圖，橫軸為溫度值(°C)，縱軸為機率。藍點為 EMOS 所有系集成員的預報值，紅點為觀測值，藍色曲線為 EMOS 各成員套上誤差 PDF 後，再乘以相等權重(權重為成員數的倒數)之結果，黑色曲線為所有 EMOS 成員之預報 PDF 等權合成後的預報 PDF，即為 EKDMOS 的預報 PDF。

(三) 機率預報產製方法

本研究針對模式內插至測站的預報 (RAW)、Ensemble MOS(EMOS) 以及 EKDMOS 的機率預報進行評比, 機率預報的產製方法如下: EKDMOS 產製預報機率的方法如圖 3 所示, 計算 PDF 中溫度低於極端事件門檻的面積(粉紅陰影部分), 即可得到極端冷事件之機率預報。至於 RAW 以及 EMOS 則計數有幾個成員預報的日最低溫低於極端事件門檻, 再除以總成員數, 即可得到預報機率。

以上關於預報與觀測資料、建模、驗證階段以及觀測氣候採用之資料整理如表 1。

(四) 校驗方法

本文將透過以下方法校驗系集預報離散度之合理性以及機率預報之可信度、區辨能力與預報技術。

1. PIT(Probability integral transform) 與 RH(Rank Histogram)

針對離散型系集預報 (RAW 及 EMOS), 本研究採用 rank histogram (RH; Anderson 1996)校驗其系集離散度是否合理反映預報的不確定性。而針對連續型系集分布預報 (EKDMOS) 則採用 PIT(Gneiting et al.,2005)。RH 之校驗原理如下所述: 觀測與 K 個系集成員進行排序, 觀測可落在 K+1 個區間; 具有良好品質的系集預報系統, 每次預報所對應的觀測落在各區間的機率會相當。因此當樣本數足夠多時, 所有觀測樣本落在各區間的數量會差不多, 繪製出來的 PIT 會呈現均勻分布如圖 4a 所示。圖 4b 至圖 4e 則是呈現離散度不足、過度離散、低報(under-forecast) 以及高報 (over-forecast) 的分布情形 (Statistical Methods in the Atmospheric Sciences, figure 8.26; Wilks 2011)。

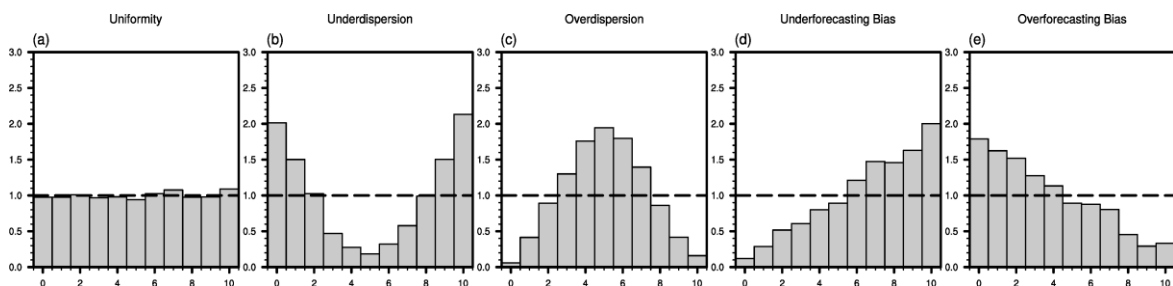


圖 4、Rank Histogram 以及 Probability Integral Transform(PIT)示意圖。(a)-(c)分別代表系集離散度良好、不足(under-dispersion)與過度離散(over-dispersion); (d)與(E)代表系集預報低報(under-forecasting)及高報(over-forecasting)。

2. Reliability Diagram

Reliability Diagram 主要用來校驗機率預報是否可信(Hsu and Murphy, 1986; Hamill, 1997), 若信賴曲線貼近對角線, 代表預報機率等於實際觀測發生頻率, 即為完美可信(perfectly reliable); 若信賴曲線落在對角線下方, 代表過度預報; 落在對角線上方代表低報。

通常 reliability diagram 會搭配 sharpness diagram 一同分析; sharpness diagram 顯示落在各預報機率區間之樣本數, 可以用來判斷一個預報系統能否提供具有信心的預報。若呈現 U 型分布, 代表模式預報機率較常接近 0 或 1, 顯示模式有足夠的信心判斷事件是否發生。若大多數樣本都接近氣候發生頻率, 代表模式沒有足夠信心判斷事件是否會發生, 只好根據氣候頻率進行較為保守的預報。

3. Relative Operating Characteristics (ROC)

ROC 曲線為給定各種不同機率門檻, 並計算不同機率門檻下之 hit rate 與 false alarm rate, 其中

$$\text{Hit rate} = \text{hit} / (\text{hit} + \text{miss}) \quad (1)$$

$$\text{False alarm rate} = \text{false alarm} / (\text{false alarm} + \text{correct negative}) \quad (2)$$

最後將不同機率門檻之(hit rate, false alarm

rate)相連, 即可繪出 ROC 曲線。曲線下面積(Area under the Curve of ROC, AUC)為評估預報是否具有區辨能力的指標, 當 AUC 等於 1 時為完美區辨能力, AUC 大於 0.7 時代表預報具有區辨能力, 當 AUC 低於 0.5 時代表預報表現不如隨機預報(Buizza et al., 1999)。

4. Brier Skill Score (BSS)

Brier Skill Score 定義如下:

$$\text{BSS} = (\text{BS}_{cli} - \text{BS}_{fc}) / (\text{BS}_{cli} - 0.) \quad (3)$$

BSS 大於 0 代表具有預報技術。其中 BS_{fc} 為預報的 Brier Score(BS), 而 BS_{cli} 代表氣候預報(以氣候發生頻率作為預報機率)的 Brier Score。機率預報誤差 Brier Score 定義如下:

$$\text{BS} = \sum_{i=1}^N (P_i - Y_i)^2 / N \quad (4)$$

其中 N 為樣本數, P 為預報機率, Y 為對應的觀測指標; 當事件發生時, Y 等於 1, 事件未發生, 則 Y 等於 0。根據(Mason, 2004), BSS 可拆解成

$$\text{BSS} = (\text{Reliability} - \text{Resolution}) / \text{Uncertainty} \quad (5)$$

其中 Reliability 為一負向指標, 值越接近 0 代表預報機率愈可信。Resolution 用於評估機率預報所對應之觀測頻率與氣候發生頻率的距離, Resolution 越大代表對應之觀

測頻率有別於氣候發生頻率。Uncertainty 則代表氣候不確定性。本研究同時分析經統計後處理的機率預報技術提升(BSS 提高)，主要來自 Reliability 或 Resolution 的貢獻，而氣候不確定性只和氣候發生頻率有關，並不會在校正過程中有所改變。

四、預報校驗結果

為了增加樣本以提高預報校驗結果的統計代表性，此章節所有校驗是合併 25 個測站的結果。要特別說明的是，在預報評估部分所採用的極端溫度門檻是以特定的觀測氣候百分位(例如第 5、10、15 百分位)來決定，而非以特定溫度值(例如 8、10、12°C) 作為極端溫度門檻；這是因為對山區和平地測站或者對北部和南部測站而言，特定溫度值可能是不同極端程度的溫度，例如就南部平地的測站而言，12°C 的低溫是相當極端的，但對玉山測站而言，卻是尋常的低溫。因此，合併 25 個測站在相同觀測氣候百分位(而非在相同溫度值)的結果進行預報評估，是比較合理的做法。

(一) 系集預報品質

PIT 以及 RH(圖 5) 評估結果顯示:(1) 不論預報領先時間長短，模式原始預報都呈現暖偏差(RH 呈現 L 型)，且有離散度不

足的情況；(2)EMOS 移除部分暖偏差，但同時也增加了冷偏差樣本，因此 RH 呈現 U 型分布，代表 EMOS 沒有解決離散度不足的問題；(3)EKDMOS 呈現相當均勻平坦的 RH 分布，代表其系集離散度能合理的反映預報的不確定性(預報誤差)。這是因為 EKDMOS 是根據 MOS 過去長期的預報誤差分布情形，套配出預報誤差的 PDF，因此其調整後的離散度能合理的表達預報之不確定性。

(二) 機率預報可信度

我們分析預報日最低溫小於觀測氣候第 5、10 和 15 百分位的 reliability diagram 與相對應的 sharpness diagram；由於三個百分位的分析結果相似，限於篇幅，只呈現第 10 百分位的結果(圖 6)。因為本研究主要針對極端事件作分析，因此預報機率等於 0 或接近 0 的樣本佔大多數，繪圖時預報高機率的樣本太少，繪製 Sharpness Diagram 時幾乎貼近底線導致難以分析，因此除了原始樣本之外，另外呈現取 log 後的樣本值，便於進行後續分析。結果顯示：(1) 任何極端程度的冷事件，RAW 以及 EMOS 隨著預報時間拉長，信賴曲線逐漸偏離對角線，代表可信度隨著預報時間

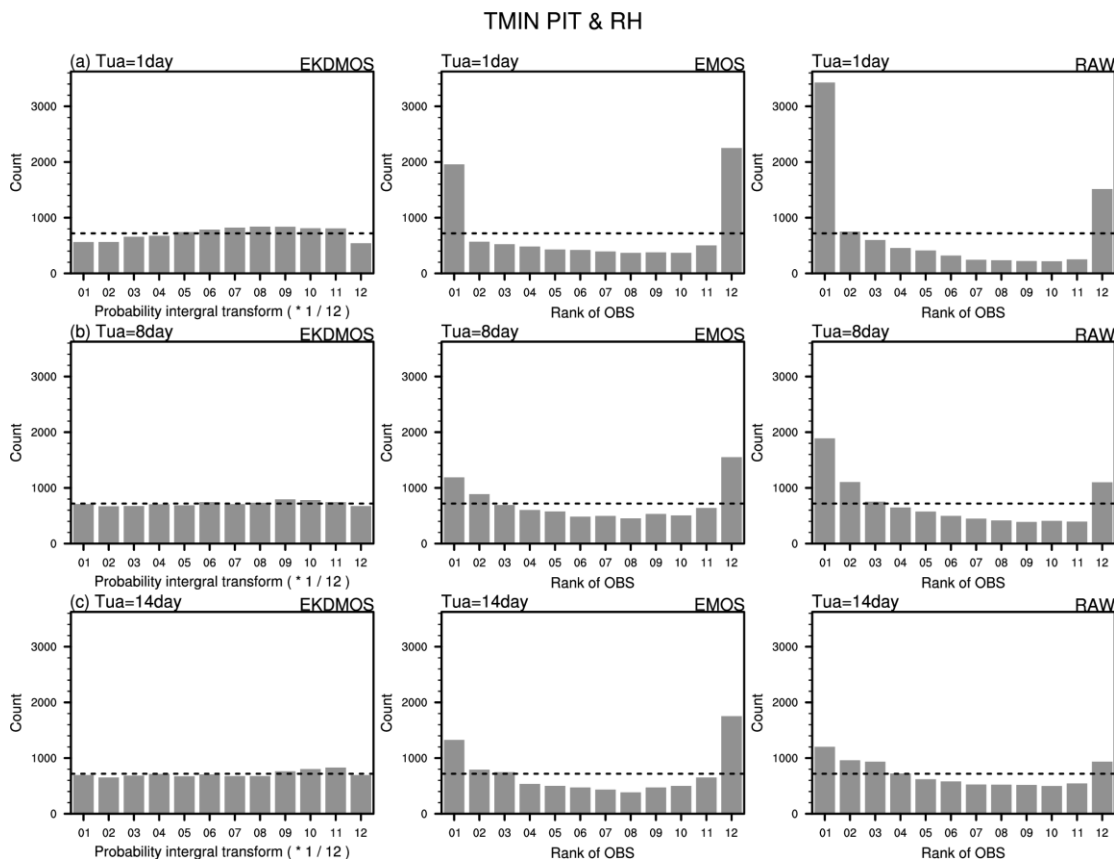


圖 5、EKDMOS(左欄)、EMOS(中欄)與 RAW(右欄)之 PIT/RH。離散型機率預報(RAW 與 EMOS)採用 RH，而連續型機率預報(EKDMOS)採用 PIT 來評估系集離散度的合理性。(a)至(c)分別為預報第 1、8 與 14 天之結果。

拉長而降低；但 EKDMOS 於所有百分位以及預報時間的信賴曲線貼近對角線，接近完美可信。(2) 同 Glahn et al. (2009)的結果，sharpness diagram 顯示 EKDMOS 預報機率隨著預報時間拉長，逐漸接近氣候發生頻率(低預報機率樣本數逐漸增加)；這結果反映模式原始預報隨著預報時間拉長，預報能力 (skill) 或預報訊息度 (informativeness) 逐漸降低。

(三) 預報區辨能力

預報日最低溫小於觀測氣候第 10 百分位的 ROC 圖 7 顯示：在所有不同預報天數，相對於 RAW 以及 EMOS，EKDMOS 的 AUC 都有明顯的提升，代表 EKDMOS 可以提升 RAW 與 EMOS 對極端冷事件之區辨能力，預報第 5 以及第 15 百分位也顯示出相應的結果 (未附圖)。EMOS 離散度不足的現象從圖 3 可見一斑；離散度不足的系集預報系統很難預報到極端事件，透過 EKDMOS 調整系集離散度，使其能合理反映預報不確定性後，才能抓到極端

TMIN10 Reliability Daigram

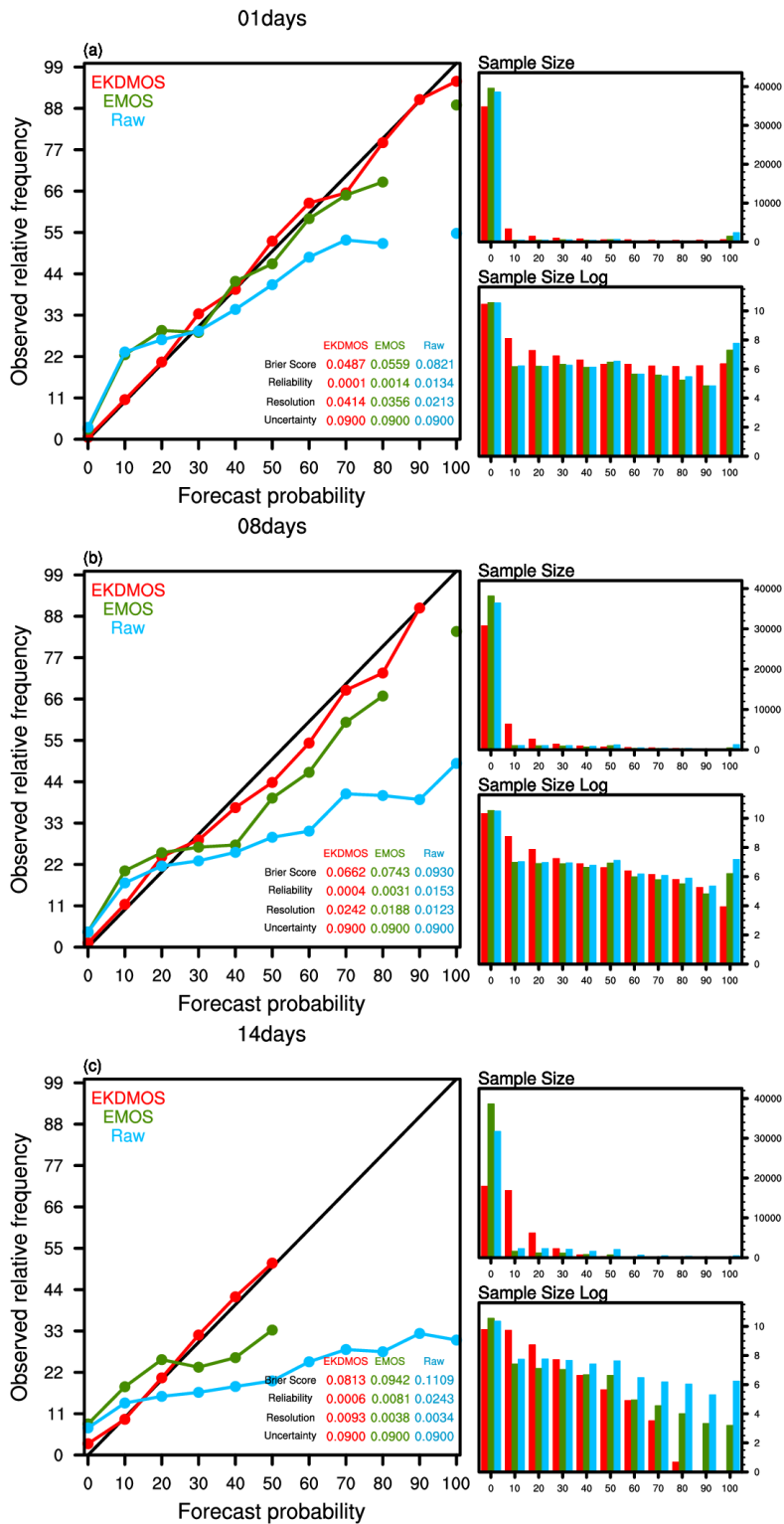


圖 6、EKDMOS(紅色)、EMOS(綠色)與 RAW(藍色)預報日最低溫低於第 10 百分位之機率預報 Reliability Diagram 以及 Sharpness Diagram；Sharpness Diagram 採用兩種方式來呈現不同預報機率區間的樣本分佈，分別是採用原始值(上)與對數值(下)之結果。(a)至(c)分別為預報第 1、8 與 14 天之結果。

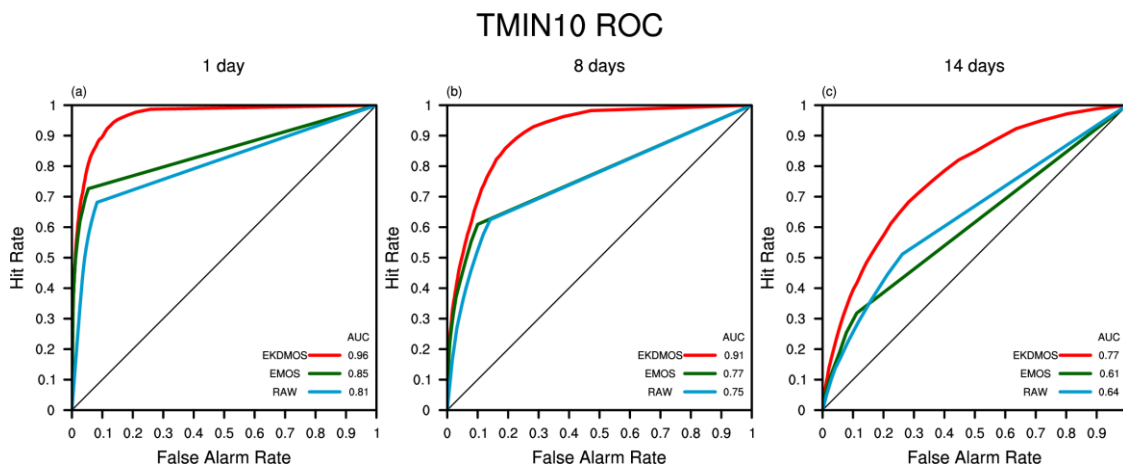


圖 7、預報 (a)第 1、(b)第 8 與(c)第 14 日最低溫低於觀測氣候第 10 百分位機率預報之 ROC。紅、綠與藍色曲線分別為預報 EKDMOS、EMOS 與 RAW 之結果。

事件，並有可信的機率預報。這從圖 6 的 sharpness diagram 也可以看出：相較於 RAW 和 EMOS，EKDMOS 預報機率為 0 的次數降低了，被調整為低機率樣本。總言之，EMOS 及 RAW 預報極端事件的機率大多為 0，無從分辨極端事件與非極端事件，但 EKDMOS 則可以透過預報機率的調整進而提高預報區辨能力。

不同極端程度之冷事件於所有預報天數之 AUC(圖 8)顯示：EKDMOS 可以提升 RAW 與 EMOS 的區辨能力；且隨著事件越極端以及預報天數越長，區辨能力提升的幅度越高。此外，圖中也顯示出 EMOS 與 RAW 在預報 12 天候 ROC 面積有交叉的情況，其原因可能為線性迴歸方程一般是以最小平方法(最小化預報誤差的平方和)來尋找預報因子(predictor)與預測值

(predictand)間之最佳函數匹配。因此，當預報時間拉長，預報與觀測間相關性越來越低時，透過迴歸方程修正後的預報就會越貼近訓練期的觀測氣候值。由於 EMOS 是所有成員都套入迴歸方程式，因此在 lead time 較長時會出現所有成員都往氣候值貼近的現象，而造成系集離散度不足，這也代表 EMOS 系集相對於 RAW 更難預報出極端事件，隨著 lead time 增加時，當 spread 降低之負貢獻大於系統性誤差之正貢獻時，EMOS 之 ROC 面積反而會低於 RAW。

(四) Brier Skill Score

Brier Skill Score (BSS)主要用以評估 EKDMOS 相較於 EMOS 及 RAW 預報表現改進的程度；不同極端程度以及不同預

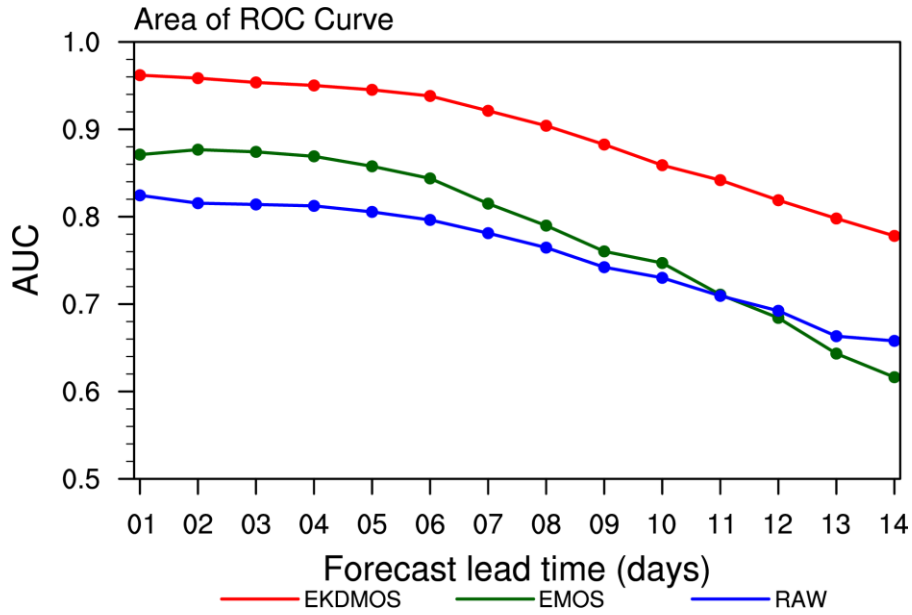


圖 8、不同方法於不同預報天數預報日最低溫低於觀測氣候第 10 百分位機率預報之 ROC 曲線下面積。橫軸為預報天數，縱軸為 ROC 曲線下面積，紅線 EKDMOS，綠線為 EMOS，藍線為 RAW 之結果。

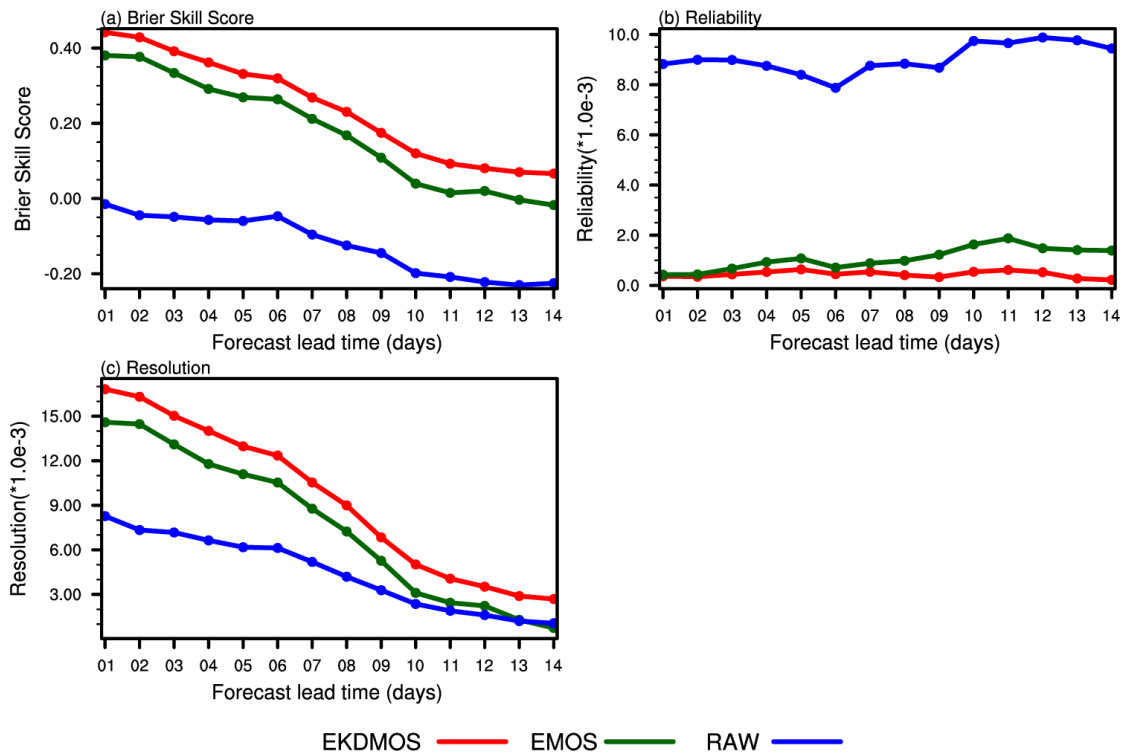


圖 9、同圖 8，但分別為 Brier Skill Score、Reliability 以及 Resolution 之結果。

報天數的 BSS(圖 9a)顯示：透過迴歸校正系統性偏差後(EMOS)，BSS 有明顯的提升，再經由 EKDMOS 調整系集離散度後，BSS 又再進一步提升，使原本無預報技術的 RAW 與 EMOS(主要是較極端以及預報時間較長的狀況)都提升至有預報技術。Brier Score 分解後的 Reliability 以及 Resolution 分析結果(圖 9b 和 c)顯示：不論 Reliability 或 Resolution 皆以 EKDMOS 表現最佳。相較於 EMOS，EKDMOS 之 Reliability 於較不極端且預報時間較長的情況改進程度較大；同樣地，Resolution 也在較不極端的事件中有較明顯的改進。此

結果也反映在 EKDMOS 相對於 EMOS 的 BSS 中，亦即在較不極端的事件中有較高的技術得分(圖未附)。這和 AUC 的結果正好相反，AUC 於較極端的事件中改進較為明顯。這是因為 Reliability 及 Resolution 是以「預報」作為條件機率的條件依據(conditioned on forecast)，但在極端事件預報中，由於事件發生的次數少，因此不論 EKDMOS、EMOS 以及 RAW 預報機率等於 0 或者接近 0 且事件沒有發生的情形非常多。這導致真正需要關注之高預報機率的權重相對很小，因此使用這些校驗得分去評估預報表現時，EKDMOS 改進的幅度

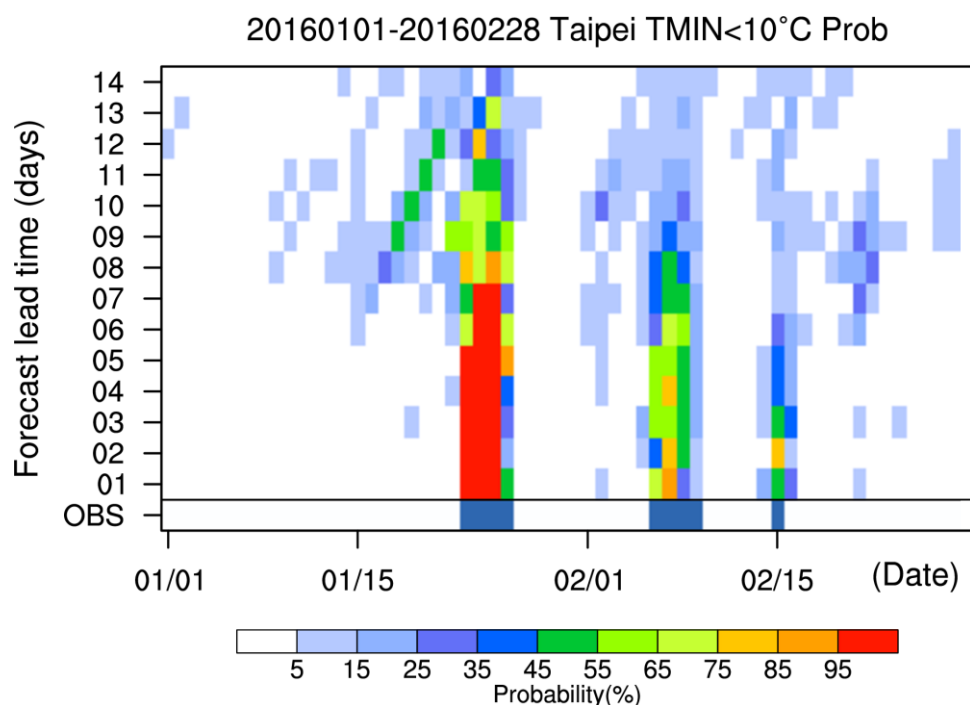


圖 10、2016 年 1 月 1 日至 2 月 28 日，EKDMOS 預報臺北測站日最低溫低於 10°C 之機率。橫軸為觀測日期，縱軸為預報領先時間，最下面一列為觀測結果，深藍色代表臺北測站實際觀測到日最低溫低於 10°C。

較小。而 ROC 是以「觀測」作為條件機率的條件依據(conditioned on observation)，因此不會有高預報機率權重被稀釋的問題，更能凸顯出 EKDMOS 調整離散度後，明顯改善 EMOS 以及 RAW 因離散度不足導致常常出現預報機率為 0 的問題。

接著探討 EKDMOS 之 BSS 隨著預報天數增加而下降的原因；EKDMOS 在所有預報時間(1~14 天)皆接近完美可信 [reliability ~0；圖 9a]，但 AUC(圖 8)及 resolution(圖 9b)則隨著預報時間有明顯下降的趨勢。EKDMOS 對於實際極端低溫個案(日最低溫 $< 10^{\circ}\text{C}$)之預報(圖 10)中，同樣可看出上述特性：預報時間較短時，EKDMOS 較能預報出高機率或接近 0 的機率，且預報結果大多能良好地反映事件是否發生；但隨著預報時間拉長，預報機率趨近於極端事件的氣候發生頻率(統計 2000-2019 年間，12 月至隔年 1 月，台北測站低於 10°C 之氣候頻率約為 4.9%)，亦即 EKDMOS 當預報時間拉長時，對於極端事件不具有區辨能力。

上述分析結果顯示：EKDMOS 可以良好地修正預報偏差，得到可信的預報機率；同時可提升模式原始預報的區辨能力，但當預報時間拉長，原始預報的可預報度顯

著降低，甚至完全不具有預報資訊(informativeness 或 signal)時，任何統計後處理方法皆無法提升預報資訊，這是數值天氣預報在可預報度上的科學限制。對於事件之區辨能力除了透過離散度的調整外，還取決於預報因子的好壞，也就是要提升模式本身的表現或者找出更適合的預報因子，例如預報天數拉長可以選擇使用與測站溫度高相關之大尺度環流指標來當預報因子，而不是鄰近格點的溫度場。

五、經濟價值分析

本研究除了評比 EKDMOS、EMOS 與 RAW 對於極端冷事件的機率預報品質外，同時評估使用者以預報資訊做為決策依據所能得到的經濟價值(Chang et al.2015)。預報系統所能提供給使用者的經濟價值定義如下(Richardson,2000)：

$$EV = (E_c - E_f) / (E_c - E_p) \quad (6)$$

其中 E_c 、 E_f 以及 E_p 分別是以事件氣候發生頻率、預報系統資訊以及完美預報系統資訊做為決策依據的費用期望值。完美的預報系統是指預報系統對於天氣事件的發生與否，永遠提供正確的預報；因此， E_p 是 3 個費用期望值中最小的一個。根據式(6)

的定義，EV 代表的涵義為：相對於採用氣候資訊做決策，如果完美的預報系統可以讓使用者省下 x 元，那麼經濟價值是 EV 的預報系統，將可為使用者省下 x*EV 的費用。EV 的值域從負無限到 1；EV 等於 1 代表預報系統為完美預報；EV 介於 0 到 1 之間代表參考預報系統之費用期望值低於參考氣候發生頻率，亦即參考預報系統做決策相較於參考氣候頻率有更大的經濟價值；EV 小於 0 代表使用者參考氣候頻率做決策反而有較高的經濟效益。

在經濟價值分析中，我們假定一個決策者是否採取防災行動，完全根據預報資訊而定(亦即，預報有天氣事件會發生，決策者就會採取防災行動，反之則不採取行動)。因此，根據預報系統過去長期的預報表現，我們可以借助 2x2 的列聯表(表 2)求出此預報系統所能提供的經濟價值。表中列出 4 種可能狀況的發生頻率(h、m、f 和 c)，4 個頻率的加總是 1。此外，表中也列

表 2、防災決策與對應之成本損失

		是否採取防災行動	
		是	否
事件是否發生	是	Hit (h) 所需花費：C + Lu	Miss (M) 所需花費：Lp + Lu
	否	False Alarm (f) 所需花費：C	Correct Rejection (c) 所需花費：0

出使用者在不同狀況下的費用期望值，其中 C 是防災行動的成本，L 是天氣事件造成的所有損失，包括採取防災行動後可避免的損失 Lp 和無法避免的損失 Lu。Ec、Ef 以及 Ep 之計算方式如下：

$$E_c = \text{Min}[\bar{o}(L_p + L_u), C + \bar{o}L_u] \quad (7)$$

其中 \bar{o} 為事件之氣候發生頻率。

$$E_f = h(C + L_u) + fC + m(L_p + L_u) \quad (8)$$

$$E_p = \bar{o}(C + L_u) \quad (9)$$

將 E_c 、 E_f 與 E_p 代入式(6)後，經由簡單的推導可以得到下式：

$$EV = \frac{\text{min}[\bar{o}, r] - FARr(1 - \bar{o}) + HR\bar{o}(1 - r) - \bar{o}}{\text{min}[\bar{o}, r] - \bar{o}} \quad (10)$$

其中 HR 以及 FAR 定義如式(1)及式(2)，r 定義為 C 與 Lp 的比值($r=C/L_p$)。r 大於 1 代表防災成本高於可避免的損失，決策者在此情況下通常不採取防災行動。r 等於 0

代表防災行動不需要成本，使用者只要持續採取行動即可，不需參考預報決定是否行動。由此可知 r 介於 0 到 1 之間的使用者才是預報系統的潛在用戶。經濟價值分析結果(圖 11)顯示：對於各種不同程度的極端低溫事件與所有預報時間，EKDMOS 比 EMOS 或 RAW 可提供更多的使用者(r 值更廣)更高的經濟價值。Chang et al. (2015)指出：EV 與 ROC 具有密切關聯，提高機率預報之區辨能力為提高經濟價值的關鍵；本研究再次驗證此結論 --

EKDMOS 透過提升 EMOS/RAW 的區辨能力，以提升預報的經濟價值。

EKDMOS、EMOS 與 RAW 在不同極端程度以及不同預報時間下之最大經濟價值分析(圖 12)顯示：(1)三者所能提供任何使用者的經濟價值，幾乎都隨著預報時間的拉長而降低；(2)幾乎對於所有使用者和所有預報時間，EKDMOS 所能提供的經濟價值都高於 EMOS 及 RAW，其中 r 接近 0 與 1 的使用者採用 EKDMOS 做為決策依據，可以得到經濟價值有最明顯的提升。

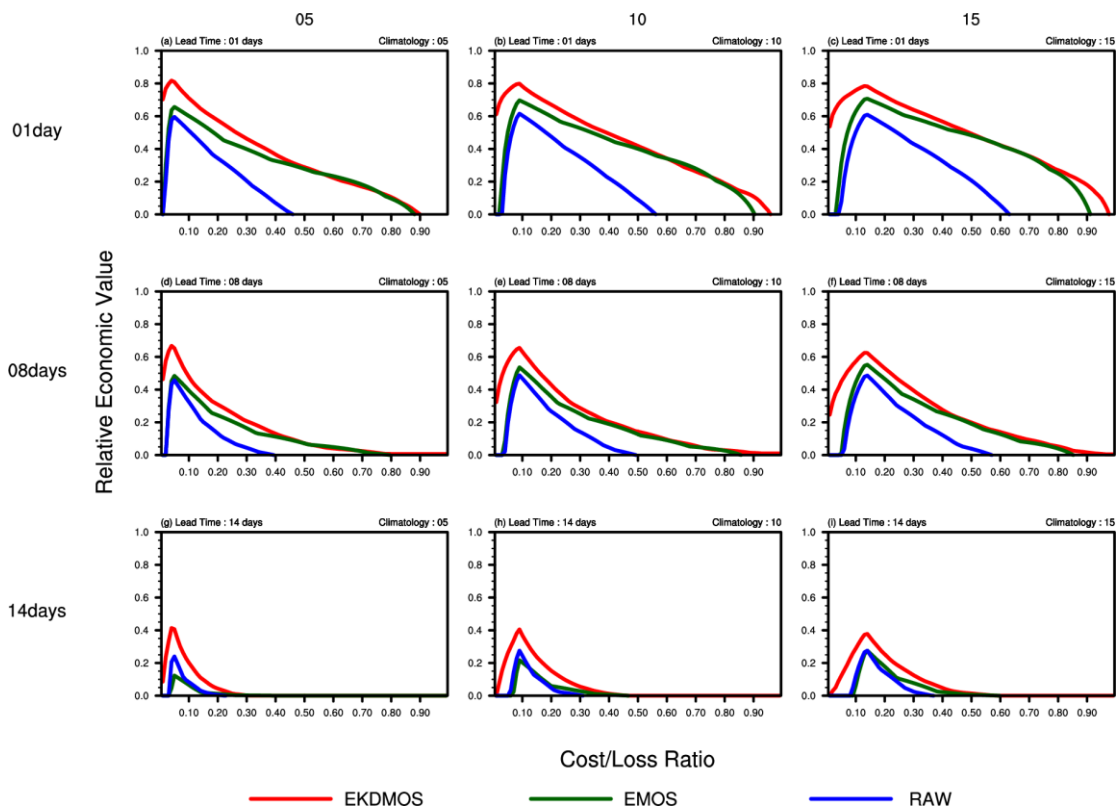


圖 11、EKDMOS(紅色曲線)、EMOS(綠線)與原始預報(藍線)之於不同預報時間及不同極端程度之最大經濟價值分佈圖，縱軸為經濟價值，橫軸為 r 值(r =防災成本/防災後可避免的損失)。第 1、2 與 3 列分別為預報第 1、8 與 15 天的結果，左、中與右欄分別為日最低溫小於觀測氣候第 5、10 與 15 百分位的結果。

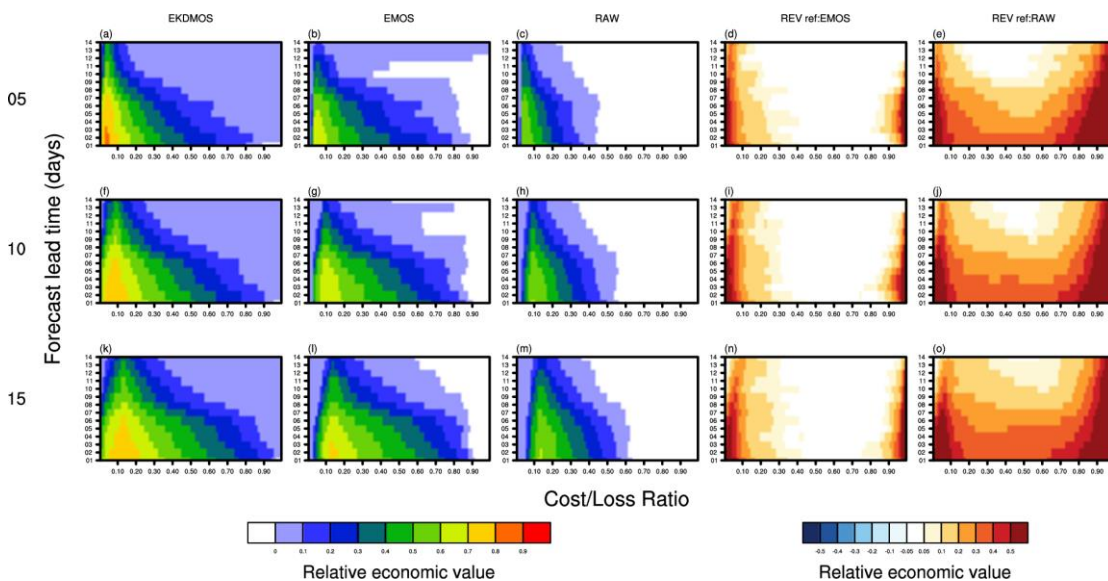


圖 12、不同極端程度以及不同預報時間下之最大經濟價值(不同顏色代表不同量值的相對經濟價值)，橫軸為 r 值，縱軸為預報天數。左到右共 5 欄，第 1、2 與 3 欄分別是 EKDMOS、EMOS 與 RAW 相對於觀測氣候的最大經濟價值，第 4 和 5 欄分別是 EKDMOS 相對於 EMOS 和 RAW 的最大經濟價值。上、中與下列分別是日最低溫小於觀測氣候第 5、10 與 15 百分位的結果。

(3) 對於所有使用者和所有預報時間，EKDMOS 之最大經濟價值幾乎都大於 0，代表在 14 天內，幾乎所有使用者皆能依據 EKDMOS 預報做決策而得到經濟效益。

(4) 在所有使用者中， r 值等於事件氣候發生頻率的使用者能得到最大的經濟價值，同 Richardson (2000) 的結論。

六、結論與展望

本研究聚焦在極端低溫事件，透過統計後處理技術 EKDMOS 對 NCEP GEFS v12 進行預報偏差修正與降尺度，結果顯示可以得到臺灣地區特定測站點上具有良好預報品質與預報價值的 1-14 天極端低

溫機率預報。EKDMOS 主要是透過長期的歷史訓練資料找出模式預報與測站觀測間的關係，進行系集成員的偏差修正，再透過 kernel density estimation (KDE) 產生系集機率分布預報，改進原始模式與 EMOS 離散度不足的問題，因而顯著提升模式對極端冷事件機率預報之可信度以及區辨能力，同時讓更多的使用者(更廣的成本/損失比)可以透過參考 EKDMOS 預報進行決策而得到更高的經濟效益。

Murphy (1977) 指出：若是採用完全可信(perfectly reliable)的機率預報，使用者得到最大經濟價值之最佳機率門檻恰好就等於自身的成本/損失比(r)。本論文證明

EKDMOS 能提供接近完美可信之機率預報，使用者只要以預報機率是否達到 r 值作為採取防災行動之依據，即可得到最大的經濟效益。

在未來工作部分，除應用 EKDMOS 在夏季日最高溫外，也考慮延伸應用在其他氣象變數，包括高度和壓力場的校正，因其誤差分布符合常態分佈假設。此外，氣象局科學技術發展的最終目的在於提供更高品質的氣象預報資訊，並幫助使用者運用氣象預報資訊做出更好的決策，得到更高的經濟效益。因此，本研究的後續工作規劃還包括以下兩項：

- (1) 將 EKDMOS 的應用從測站點擴展到高解析格點：本論文是透過 EKDMOS 進行偏差修正，並降尺度到測站點，後續將進一步降尺度至高解析(1km 解析度)格點，以滿足下游農業單位在特定作物專區的極端溫度預報需求
- (2) 建置極端溫度決策支援系統：亦即結合 EKDMOS 極端溫度機率預報與經濟價值分析，找出最佳的防災行動門檻(也就是得到最大經濟效益的最佳機率門檻)，提供使用者是否採取防災行動的決策建議。

總言之，本研究的技術發展是氣象局推動氣候服務過程中極為重要的前端部份。整個氣候服務規劃的通盤性解決方案，先是將預報資料(原始模式預報結果)轉為預報資訊(偏差修正與降尺度後的預報)，再將預報資訊連結到決策建議；期能在改善預報品質外，進一步提升預報價值。

致謝

感謝科技部自然司之研究經費補助，本研究是在專題研究計畫補助下完成(MOST 110-2111-M-008-021)。同時感謝兩位審查者提供寶貴的建議，使論文內容更為嚴謹且豐富。

參考文獻

- 陳昀靖、陳孟詩、張惠玲、羅存文，2018：臺灣地區測站之第二週日最高/最低溫機率預報的發展與評估，107 年天氣分析與預報研討會論文全文彙編，A4-17
- 章鶴群、陳冠儒、劉豫臻、張惠玲、洪景山、楊舒芝，2018：臺灣地區 WRF 颱風系集降雨機率預報之評估、校正與經濟價值分析 - 第二部分：校正。大氣科學，第 46 期，107 - 127。

- Anderson, J. L., 1996: A method for producing and evaluating probabilistic forecasts from ensemble model integrations. *J. Climate*, **9**, 1518–1530.
- Buizza, R., and T. N. Palmer, 1998: Impact of ensemble size on ensemble prediction. *Mon. Wea. Rev.*, **126**(9), 2503–2518.
- _____, M. Miller, and T. N. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System. *Q. J. R. Meteorol. Soc.*, **125**, 2887–2908.
- _____, P. L. Houtekamer, G. Pellerin, Z. Toth, Y. Zhu, and M. Wei, 2005: A comparison of the ECMWF, MSC, and NCEP global ensemble prediction systems. *Mon. Wea. Rev.*, **133**(5), 1076–1097.
- Chang, H. L., S.-C. Yang, H. Yuan, P. L. Lin and Y. C. Liou, 2015: Analysis of relative operating characteristic and economic value using the LAPS ensemble prediction system in Taiwan area. *Mon. Wea. Rev.*, **143**, 1833–1848.
- Chu, P.-S., Zhao X., C.-T. Lee, and M.-M. Lu, 2007: Climate prediction of tropical cyclone activity in the vicinity of Taiwan using the multivariate least absolute deviation regression method. *Terr. Atmos. Ocean. Sci.*, **18**, 805–825.
- Cui, B., Z. Toth, Y. Zhu, and D. Hou, 2012: Bias correction for global ensemble forecast. *Wea. Forecasting*, **27**, 649–667.
- Glahn, H. R., and D. A. Lowry, 1972: The use of Model Output Statistics (MOS) in objective weather forecasting. *J. Appl. Meteorol.* **11**: 1203–1211.
- _____, Peroutka, M., Weidenfeld, J., Wagner, J., Zylstra, G. and Schuknecht, B. (2009). MOS uncertainty estimates in an ensemble framework. *Monthly Weather Review*, **137**, 246–268.
- 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.

- Guan, H., Y. Zhu, E. Sinsky, B. Fu, X. Zhou, W. Li, X. Xue, D. Hou, B. Cui, J. Peng, 2020: The NCEP GEFS-v12 Reforecasts to Support Subseasonal and Hydro-meteorological Applications. *Climate Prediction S&T Digest, 44rd NOAA Climate Diagnostics and Prediction Workshop special issue*, NOAA's National Weather Service
- Hamill, T. M., 1997: Reliability diagrams for multicategory probabilistic forecasts. *Wea. Forecasting*, **12**, 736–741, doi:10.1175/1520-0434(1997)012<0736:RDFMPF>2.0.CO;2.
- Hsu, W.-R., and A. H. Murphy, 1986: The attributes diagram: A geometrical framework for assessing the quality of probability forecasts. *Int. J. Forecasting*, **2**, 285–293, doi:10.1016/0169-2070(86)90048-8.
- Leith, C. E., 1974: Theoretical skill of Monte Carlo forecasts. *Mon. Wea. Rev.*, **102**, 409–418.
- Mason SJ., 2004: On Using “Climatology” as a Reference Strategy in the Brier and Ranked Probability Skill Scores. *Monthly Weather Review*, **132**, 1891–1895.
- Murphy, A. H., 1977: The value of climatological, categorical and probabilistic forecasts in the cost–loss ratio situation. *Mon. Wea. Rev.*, **105**, 803–816,
- Richardson, D. S., 2000: Skill and relative economic value of the ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **126**, 649–667, doi:10.1002/qj.49712656313
- Wilks, D. S., 2011: *Statistical Methods in the Atmospheric Sciences*. 3rd ed. Academic Press, 704 pp.
- Yuan, H., X. Gao, S. L. Mullen, S. Sorooshian, J. Du, and H. H. Juang, 2007: Calibration of probabilistic quantitative precipitation forecasts with an artificial neural network. *Wea. Forecasting*, **22**, 1287–1303.
- _____, J. A. McGinley, P. J. Schultz, C. J. Anderson, and C. Lu, 2008: Short-range precipitation forecasts from time-lagged multimodel ensembles

- during the HMT-West-2006 campaign. *J. Hydrometeor.*, **9**, 477–491.
- Zhao, T., J. C. Bennett, Q. J. Wang, A. Schepen, A. W. Wood, D. E. Robertson, and M.-H. Ramos, 2017: How suitable is quantile mapping for postprocessing GCM precipitation forecasts? *J. Climate*, **30**, 3185–3196.
- Zhou, X., Y. Zhu, B. Fu, D. Hou, J. Peng, Y. Luo and W. Li, 2019: The development of next NCEP Global Ensemble Forecast System. *Climate Prediction S&T Digest*, 43rd NOAA Climate Diagnostics and Prediction Workshop special issue, NOAA's National Weather Service, 159-163, doi:10.25923/ar2c-v522.
- Zhu, Y., and Y. Luo, 2015: Precipitation calibration based on the frequency-matching method. *Wea. Forecasting*, **30**, 1109–1124.

1-14-day Probabilistic Cold Extreme Forecasts over Taiwan Using the Ensemble Kernel Density Model Output Statistics (EKDMOS)

Yun-Jing Chen and Hui-Ling Chang

Research and Development Center, Central Weather Bureau, Taipei, Taiwan

(manuscript received 26 July 2022 ; in final form 14 December 2022)

Abstract

This study applies the Ensemble Kernel Density Model Output Statistics (EKDMOS) in the 20-year reforecasts of the National Centers for Environmental Prediction (NCEP) Global Ensemble Forecast System version 12 (GEFS v12) to perform bias correction and downscaling. The purpose is to produce 1-14-day probabilistic cold extreme forecasts with good quality and value at specific stations over Taiwan.

Forecast evaluation shows that the raw ensemble forecast is under-dispersive with an obvious bias. In contrast, the EKDMOS has a much more reasonable ensemble spread with most of the bias removed. The EKDMOS has more reliable and higher discrimination than the raw ensemble forecast. The EKDMOS increases the Brier skill score (BrSS) of the raw ensemble forecast relative to the sample frequency by both improving the reliability and resolution. Users with a much wider spectrum of cost/loss ratio can obtain more benefit from the EKDMOS as compared to the raw ensemble forecast. Furthermore, users with cost-loss ratio close to 1 and 0 can benefit more than other users. Even with improved forecast quality and value, the forecast distribution of the EKDMOS unavoidably approaches the climatology of the training sample when the forecast informativeness is lost beyond 10 days.

Keywords : probabilistic cold extreme forecast, Ensemble Kernel Density Model Output Statistics (EKDMOS), reliability, discrimination, forecast informativeness. doi:

10.53106/025400022022125003003