

以多變量貝氏系集處理器配合球面調和函數波段拆解進行 展期極端低溫預報

朱心宇 張惠玲*

交通部中央氣象局

(中華民國 111 年 7 月 26 日收稿；中華民國 111 年 12 月 6 日定稿)

摘 要

透過統計後處理(statistical post-processing, SPP)方法將原始模式進行偏差修正，並降尺度到高解析格點或特定點位，可提供使用者更具參考價值的預報資訊；但偏差修正往往需要大量的訓練樣本以估算預報偏差。由於展期系集重預報(reforecast)需要大量計算資源，剛發展成熟之系集預報系統可能缺乏足夠之重(後)預報資料以進行偏差校正。本研究擬發展結合多變量貝氏系集處理器(Multi-Variate Bayesian Processor of Ensemble, BPE)與球面調和函數波段拆解(spherical harmonics decomposition, SHD)之先進 SPP 技術，目的在於探討以單一預報取代系集預報進行 SPP 之可行性，以解決上述系集預報系統缺乏大量訓練樣本，無法進行偏差校正之困境。

本研究將原始系集模式的控制模擬(Control Run)進行 SHD，以拆解後的不同波段作為 BPE 預報因子，進行 SPP 以得到臺灣測站點上之展期極端低溫機率預報。BPE 演算法利用預報因子及觀測值關聯結構產生的邊緣分布作為概似函數(likelihood)，並以觀測值的氣候分布作為先驗函數(prior)，再根據最新的預報資料予以結合成後驗函數(posterior，即為預報機率函數)；因其完全貝氏的架構，BPE 在重預報資料較少的狀況下，依然可以充分利用觀測氣候資料建立先驗函數，得到可信的後驗分布函數。

本研究以美國國家環境預測中心第 12 版全球系集預報(NCEP GEFS v12)之 (1)2 公尺最低溫系集平均作為單變量 BPE 預報因子 (實驗 SP_BPE)、(2)控制模擬(Control Run)之標準化 2 公尺最低溫拆解後的波段作為多變量 BPE 預報因子(實驗 MP_BPE_SH)，進行 8 至 14 天之日極端低溫預報以及第三、四週之週極端溫度預報校正。多項指標顯示：SP_BPE 與 MP_BPE_SH 兩種校正方式皆顯著提升原始模式的預報品質，但 SP_BPE 校正效果略優於 MP_BPE_SH，特別是在第三、四週之週極端低溫預報。上述結果應與預報因子的可預報度有關，因此未來擬利用可預報度較高之大尺度指標做為多變量 BPE 的預報因子，以提升第三、四週預報之校正成效。

關鍵字: 貝氏系集處理器、球面調和函數、極端溫度、展期預報、統計後處理

*通訊作者:張惠玲 博士 地址:臺北市中正區公園路 64 號 中央氣象局氣象科技研究中心，電子郵件:lingo@cwb.gov.tw

一、前言

近年來，農業及水資源單位對於精緻、客製化之展期天氣機率預報(或稱次季節、S2S，兩週以上之預報領先時間)之需求不斷增加(周等，2021)。然而，因為物理模型不完整、解析度及非線性過程造成預報不確定性的增加，造成原始預報皆有一定程度之系統性偏差。另外，由於展期全球模式之水平解析度多半在 25 公里×25 公里以上，不足以直接提供在地精緻化之機率預報。因此，統計偏差修正及降尺度(合稱統計後處理，Statistical Post-Processing, SPP)，成為模式預報產品加值必要的過程。由於氣候服務的蓬勃發展，過去十年內產學研界也積極將各種在國際上較先進及有效的 SPP 方法，導入對台灣重要的氣象變數上，包含林等人(2021)之貝氏模型平均(Bayesian Model Averaging, BMA)、類比統計後處理(Analog Post-Processing)及陳等人(2021)之系集核密度函數模式輸出統計方法(Ensemble Kernel Density Model Output Statistics, EKDMOS)。然而，多半 SPP 方法是針對傳統天氣時間尺度設計，在 S2S 時間尺度時，會面臨缺少建模所需的大量重預報(Reforecast)樣本及由於原始模式無預報技術導致過度擬合(Overfitting)的問題。朱等人(2021)及 Chu et al. (2020,2021)，參照 Krzysztofowicz & Evans (2008，後簡稱 KE08)，建立適用於台灣地

區均溫及高低溫之單變量貝氏系集處理器(Single-Predictor Bayesian Processor of Ensemble, Single-Predictor BPE)，其研究結果證實 BPE 確實有 KE08 理論架構內之優點，包含在有限樣本下建立可靠的 SPP 模型、預報變數(Predictors)無法對於目標變數(Predictand)有足夠可預報度時，透過貝氏性質客觀地將預報機率分佈趨近於氣候分佈，對於目前展期機率預報實務上需要大量重預報樣本及可預報度等難題提供務實的解決方法。因此，本研究進一步將單變量貝氏系集處理器，延伸至可容納多個預報因子之多變量系集處理器(Multi-Predictor BPE)，利用 NCEP-GEFSv12 Control Run 當作單一決定性預報，並進行濾波當作預報因子代替系集模式，評估以統計後處理方法濾除無法預報的雜訊，是否可以達成與系集平均相同的效果，以進一步節省產生統計後處理所需系集成員之大量運算資源及儲存空間。

二、系統概述及實驗設計

(一) 單變量與多變量貝氏系集處理器

假設有N個預報變數(predictor)，每個預報變數在時間維度上有 T 組資料，組成 $N \times T$ 之陣列 X_n ，預報領先時間 l 、預報初始時間為 t 、目標變數(predictand)為 w_{t+l} ，則多變量貝氏定理可表示為：

$$\Phi(w_{t+l}|X_1, X_2, \dots, X_n) = \frac{F(X_1, X_2, \dots, X_n | w_{t+l}) G(w_{t+l})}{K(X_1, X_2, \dots, X_n)} \quad \text{Eq.(1)}$$

Φ, F, G, K 皆為某一形式的機率密度函數 (Probability Density Function, PDF)。為了得到 $\phi(w_{t+l}|X_1, X_2, \dots, X_n)$ 的解析解 (analytical solution)，我們先簡化問題，從單變量貝氏系集處理器 ($N = 1$) 開始，Eq.(1) 可表示成：

$$\Phi(w_{t+l}|\bar{x}_t) = \frac{F(\bar{x}_t|w_{t+l})G(w_{t+l})}{K(\bar{x}_t)} \quad \text{Eq.(2)}$$

\bar{x}_t 可為決定性預報或者系集平均，而 w_{t+l} 為在預報目標時間 $t + l$ 的觀測值。左式為在預報出 \bar{x}_t 的條件下，觀測值為 w_{t+l} 的後驗機率(Posterior) Φ ；而在一個連續的定義域中可給定無限多個 w_{t+l} ，則 Φ 為一個累積機率分布函數(Cumulative Distribution Function, CDF)。為了求 Φ 的解析解，必須將問題拆解成右式三項：

- (一) 概似機率函數 $F(\bar{x}_t|w_{t+l})$: 在觀測到 w_{t+l} 的條件下，預報出 \bar{x}_t 的機率。
- (二) 先驗機率函數 $G(w_{t+l})$: 目標變數 w_{t+l} 先前發生之機率
- (三) 預測值機率函數 $K(\bar{x}_t)$: 預報變數 \bar{x}_t 先前發生之機率

由於 BPE 假設概似函數滿足高斯元模型(Meta-Gaussian Model)，必須先將變數進行常態分位數轉換(Normal Quantile Transform, NQT)，NQT 為將一個對於變數 x 經驗累積機率分布函數 (Empirical Cumulative Distribution Function, Empirical CDF) 最適配的有母數機率分布 $F(\cdot)$ ，以標準常態分位函數 Q^{-1} 映射到常態分布空間

的過程。對於經標準化後的預報變數 x'_k 及目標變數 w'_k :

$$V = Q^{-1}(G'(w'_k)), Z = Q^{-1}(K'(x'_k)) \quad \text{Eq.(3)}$$

使用 NQT 的優勢在於，對於任意有母數機率分布，皆可讓轉換後 (V, Z) 的關聯結構滿足：

$$F(Z|V = v) \sim N(av + b, \sigma^2) \quad \text{Eq.(4)}$$

讓機率分佈得以常態分佈的解析解表示。邊緣分布函數(Marginal Distribution Function) $F(Z|V = v)$ 之一階動差為 $av + b$ ，變異數為 σ^2 。參數 (a, b, σ^2) 則可利用最大概似法(Maximum Likelihood Estimation, MLE) 估計。估計出 (a, b, σ^2) 後即可代入後驗參數：

$$A = \frac{a}{a^2 + \sigma^2}, B = -\frac{ab}{a^2 + \sigma^2}, T^2 = \frac{\sigma^2}{a^2 + \sigma^2} \quad \text{Eq.(5)}$$

最後利用 (A, B, T^2) 計算出後驗(預報)累積機率 $\Phi(w|\bar{x}_t)$:

$$\Phi(w_{t+l}|\bar{x}_t) = Q\left(\frac{1}{T} [Q^{-1}(G'(w'_k))] - A Q^{-1}(K'(x'_k)) - B\right] \quad \text{Eq.(6)}$$

多變量 BPE 的部分，先假設 $N = 2$ ，且預報變數 (X_1, X_2) 之間獨立(independent)，則 Eq.(1) 可表示成：

$$\Phi(w_{t+l}|X_1, X_2) = \frac{F(X_1, X_2|w_{t+l})G(w_{t+l})}{K(X_1)K(X_2)} \quad \text{Eq.(7)}$$

由條件機率公式，右式 $l(X_1, X_2|w_{t+l})$ 可表示為：

$$F(X_1, X_2|w_{t+l}) = \frac{F(X_1, X_2, w_{t+l})}{G(w_{t+l})} = \frac{F(X_1|X_2, w_{t+l})Q(X_2, w_{t+l})}{G(w_{t+l})} \quad \text{Eq.(8-1)}$$

$$Q(X_2, w_{t+l}) = F(X_2|w_{t+l})G(w_{t+l}) \text{Eq.(8-2)}$$

由於 X_1, X_2 之獨立性，則 $F(X_1|X_2, w_{t+l}) = F(X_1|w_{t+l})$ ，將 Eq.(8-1)及 Eq.(8-2)代入 Eq.(7)，可推得：

$$\Phi(w_{t+l}|X_1, X_2) = \frac{F(X_1|w_{t+l})F(X_2|w_{t+l})G(w_{t+l})}{K(X_1)K(X_2)} \text{Eq.(9)}$$

如此一來， $F(X_1|w_{t+l}), F(X_2|w_{t+l})$ 等價於單變量 BPE 的概似函數，若 $F(\cdot)$ 符合高斯元模型(Meta-Gaussian Model)，則延伸到任意 $N = n$:

$$F(X_n|w_{t+l}) = \text{Norm}(a_n w_{t+l} + b_n, \sigma_n) = \frac{1}{\sqrt{2\pi}\sigma_n} \exp\left[-\frac{(X_n - a_n w_{t+l} - b_n)^2}{2\sigma_n^2}\right] \text{Eq.(10-1)}$$

$$\Phi(w_{t+l}|X_1, X_2, \dots, X_n) = \frac{\prod_1^n F(X_n|w_{t+l})G(w_{t+l})}{\prod_1^n K(X_n)} \text{Eq.(10-2)}$$

$\text{Norm}(\mu, \sigma)$ 為常態分佈機率密度函數 (Normal Probability Density Function)。將 Eq.(10-1) 代入 Eq.(10-2)，若 $G(w_{t+l})$ 及 $K(X_n)$ 也符合 Meta-Gaussian 假設，利用指數律，並且做一連串數學處理後，接著找出氣候資料最適配的累積機率分配函數 $G(w_{t+l})$ ，及預報變數最適配的累積機率分配函數 $K_n(x_n)$ 、預報變數與預報變數之間的共變異數矩陣 Σ 、預報變數與目標變數之間的共變異數向量 σ_{n0} 、預報變數的一階及二階動差向量 (μ_n, σ_n) 、氣候分布的一階及二階動差 (μ_0, σ_0) 來求得 ϕ 的累積分布函數 (Cumulative Distribution Function,

CDF) Φ 的解析解 (Krzysztofowicz and Maranzano, 2006)：

$$\Phi(w_{t+l}|X_1, X_2, \dots, X_n) = Q\left(\frac{1}{T} [Q^{-1}(G(w_{t+l})) - \sum_{n=1}^N c_n Q^{-1}(K_n(x_n)) - c_0]\right) \text{Eq.(11)}$$

其中， $Q(\cdot)$ 為常態累積機率分布函數 (Normal Cumulative Distribution Function, Normal CDF)，則 $Q^{-1}(\cdot)$ 為常態累積分位數， T 為 Q 的 Scaling 參數， (c_n, c_0) 為預報因子的權重及截距項。 (c_n, c_0, T_n) 可從 $(\Sigma, \sigma_{n0}, \mu_n, \sigma_n, \mu_0, \sigma_0)$ 推導出來。多變量 BPE 之演算流程如圖一。

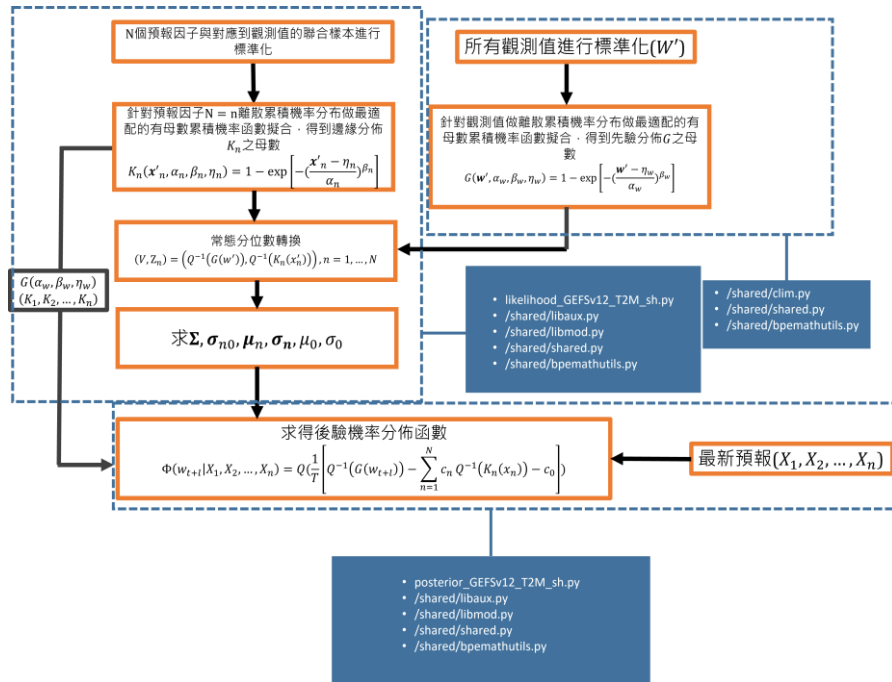
(二) 球面調和函數濾波方法

目前主流的全球系集模式仍為波譜模式，在積分時，先將控制方程映射到波譜空間離散化後，再於積分完成後映射回物理空間。在波譜空間上物理量皆以波動來描述，波數(n)與波長(λ)的關係滿足：

$$\lambda = \frac{2\pi a}{n} \text{Eq.(12)}$$

$$a = A \cos(\phi)$$

其中， A 為地球在赤道上之半徑， $A = 6371 \text{ km}$ ， a 為地球在緯度 ϕ 之半徑，定性上，波譜空間上波數較小(大)的波，其波長越大(小)，代表在映射回物理空間時，波數較小的波代表尺度較大的行星波動，例如羅士比波(Rossby Wave)、凱爾文波(Kelvin Wave)等，而波數較大的波代表尺度較小的大氣波動，例如從中尺度的中緯度準地



圖一、多變量貝氏系集處理器建模與預報流程

轉波動、到小尺度如對流激發出的重力波及地形波等等。在此研究中，我們參考Boer(2003)的結論，假設不同尺度的波動有不同的可預報時間尺度(Predictability Time Scale)，尺度較大的波的可預報時間尺度大於尺度較小的波。研究指出，波數大於 10 的波的可預報時間尺度快速遞減，n=10 的波可預報時間尺度約為 10-15 天，到了 n=20，其可預報時間尺度僅有 5 天左右。利用此概念，若將模式的某物理量U利用球面調和函數拆解(Spherical Harmonics Decomposition)：

$$U(\lambda, \mu) = \sum_{m=-M}^M [e^{i(m\lambda)} \sum_{n=|m|}^M X_n^m P_n^m(\mu)] \quad \text{Eq.(13)}$$

其中 P_n^m 為連帶勒讓德多項式(Associated-Legendre Polynomial)， λ 為經度， μ 為緯

度之函數， X_n^m 為波譜係數， m 為緯向波數， n 為經向波數。至於詳細波譜方法應用於模式積分及 X_n^m 之求法，可參考Machenhauer(1991)及郭與林(1991)，並非本研究主要探討的議題。

本研究應用 Eq.(13)中括號內展開項 $[\cdot]$ 定義為 ψ ，將波數 $m = w_1$ 到波數 $m = w_2$ 疊加得到一波段：

$$\psi_{w_1}^{w_2} = \sum_{i=w_1}^{w_2} \psi_i \quad \text{Eq.(14)}$$

則假設 Eq. (6)中的第 n 個預報變數：

$$X_n = \psi_{w_1}^{w_2} \quad \text{Eq.(15)}$$

並且讓 BPE 系統在建模時，依照不同預報領先時間給予不同波段不同權重 c_n ，在 Eq.(6)中達透過給予可預報度較低的波段較低的權重，達成與系集平均濾除不可預報之效果。在前人的統計後處理方法中，

也常用主成份分析(Principle Component Analysis, PCA), 或者簡單利用某物理量的區域平均做為大尺度指標來進行空間平滑及濾除不可預報的資訊, 預報展期時間尺度的目標變數。

(三) 實驗設計

1. 使用資料

本研究旨在於比較以下兩種不同統計後處理程序的預報表現：(1) 實驗組(MP_BPE_SH): 利用決定性預報之輸出, 透過球面調和函數濾波後的波段做為預報因子放入多變量 BPE 產出之機率預報；(2) 控制組(SP_BPE): 利用系集平均做為預報因子放入單變量 BPE 產出之機率預報。為了得到具有統計代表性且公平的結果, 本研究統一使用具有長達 20 年重預報之美國國家環境預測中心(National Centers for Environmental Prediction, NCEP)第 12 版全球系集預報系統(Global Ensemble Forecast System, GEFS) v12。實驗組(MP_BPE_SH)是以 NCEP GEFS v12 Control Run 做為「決定性預報」, 將兩公尺日最低溫(T_{\min})標準化後, 利用濾波程式將其分為下列五個波段:(1)Level 1(ψ_1^4)、(2)Level 2 (ψ_5^8)、(3)Level 3 (ψ_9^{12}) (4)Level 4(ψ_{13}^{16})、(5)Level 5 (ψ_{17}^{abv}), 參考 Wan et al.(2021) 之命名方式, 波段代表的空間尺度越大, 其 Level 數值越小。以此五個波段做為多變量 BPE 的預報因子。而控制組(SP_BPE)則是直接採

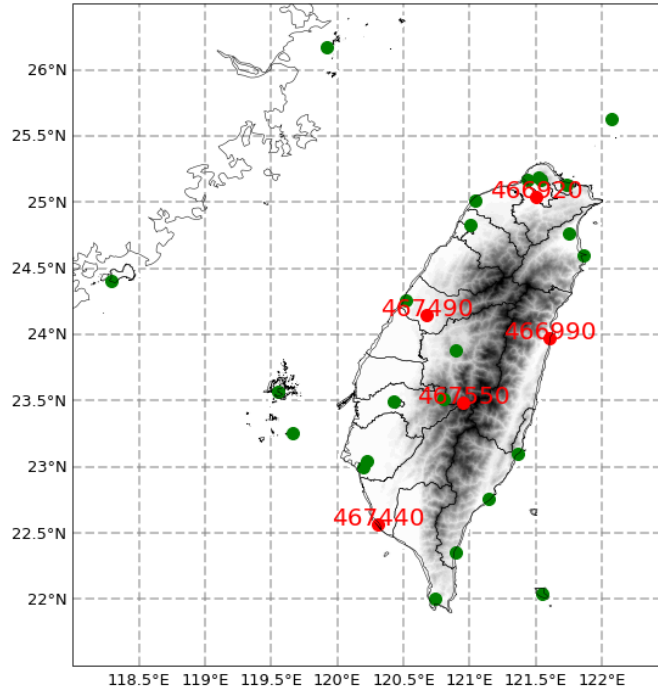
用 NCEP GEFS v12 系集平均預報, 將鄰近測站四個格點資料同時利用濕絕熱降溫率修正至海平面高度的溫度後, 進行雙線性內插到測站位置, 再利用濕絕熱降溫率修正至測站高度, 以此做為單變量 BPE 的預報因子。透過前述方法簡單修正因為雙線性內插時模式格點高度互不一致, 及模式格點高度及測站高度不一致造成的偏差, 後續對照之原始模式(RAW), 也進行了相同的修正。目標變數為台灣人工測站(圖二)冬季 8-14 天之日最低溫, 及第三週(15-21 天)、第四週(22-28 天)之週最低溫。由於控制組及實驗組來自同一系集預報系統, 可移除因為解析度、動力架構、參數化方法不同對結果產生的變因。接下來, 將過去 20 年之重預報資料分成 2000-2009 年與 2010-2019 年兩組, 每組分別做為訓練集, 再以另一組資料進行驗證(例:2000-2009 年為訓練集時, 2010-2019 年的資料即為驗證集)。 T_{\min} 的氣候分布則是以該測站過去 30 年觀測資料進行建構, 並且在建構時逐一去除驗證年, 以確保訓練資料及驗證資料完全獨立, 使用資料概述如表一。建模時採用分月建模, 即一年會有 12 組建模參數。每月的建模資料會納入前(後)半個月(當月-/+15 天)的資料, 以在季節性考量下合理增加訓練資料。

2. 資料前處理

首先，預報因子(x_n)及觀測值(w)進入 BPE 建模之前需要經過標準化：

$$x'_{nk} = \frac{x_{nk} - m_{nk}}{s_{nk}}, w'_k = \frac{w_k - m_k}{s_k} \quad \text{Eq.(16)}$$

m_{nk} 為變數 x_n 在一年之中第 k 天之氣候平均， s_{nk} 為標準差。要特別說明的是，在 MP_BPE_SH 實驗中， $\psi_{w_1}^{w_2}$ 已經在濾波前採用全球模式之氣候場事先進行標準



圖二、校驗所使用測站點位置，紅色點位為過去研究常用之代表測站。

表一、實驗資料及設定概述

實驗名稱	控制組 (SP_BPE)	實驗組 (MP_BPE_SH)
原始資料	NCEP GEFSv12 輸出兩公尺日最低溫取系集平均後雙線性內插至測站點上。若目標變數為週最低溫，再透過日最低溫找出七天最小值作為週最低溫。	NCEP GEFSv12 輸出兩公尺日最低溫，利用模式本身氣候分佈標準化後進行濾波，再雙線性內插至測站點上。若目標變數為週最低溫，則透過各波段日最小值找出七天最小值作為預報變數。
時間長度	2000-2019	2000-2019
預報領先時間	第 8 到 14 天、第三週(第 15-21 天)、第四週(第 22-28 天)	
預報變數	5 個系集成員之系集平均 (第 8 到 14 天) 11 個系集成員之系集平均(第三週、第四週)	Level 1 (ψ_1^4) Level 2 (ψ_5^8) Level 3 (ψ_9^{12}) Level 4 (ψ_{13}^{16}) Level 5 (ψ_{17}^{abv})
空間解析度	0.5° × 0.5°	
目標變數	日最低溫(第 8 到 14 天) 週最低溫(第三週、第四週)	

化，濾波後不再進行標準化。而觀測值 w_k 則是利用對應測站第 k 天的氣候平均值(m_k)及標準差(s_k)進行標準化。實驗 SP_BPE 中的 x_k ，同樣也利用對應測站第 k 天的氣候平均值及標準差進行標準化。

3.套配函數分析、預報因子選用分析

由於資源與時間的考量，且溫度之分佈特性近似於常態分佈，本研究採用韋伯分布(Weibull Distribution)：

$$F(x) = \frac{\beta}{\alpha} \left(\frac{x-\eta}{\alpha}\right)^{\beta-1} \exp\left[-\left(\frac{x-\eta}{\alpha}\right)^\beta\right] \text{Eq.}(17)$$

來套配所有預報變數及目標變數的經驗機率分佈，得到 Eq.(6) 中的 $K_n(x_n)$ 與 $G(w_{t+l})$ 。由於韋伯分布具有相當大的彈性，例如在 $\alpha = \sqrt{2}\sigma, \beta = 2$ ，可用於近似常態分佈，又可以在資料分佈特性為偏態時(skewed)，合理擬合資料，也符合極端值定理(Extreme-Value Theory) 的 Type III 分佈，可用於擬合具有極端特性的資料，所以為原 BPE 研究(KE08)及本研究首先選用之分佈。至於其他分佈是否更適配於極端溫度機率分佈，以及在預報其他變數時需用哪一種分佈更為適合，有待後續研究分析。

接下來，我們利用訊息度得分 (Informativeness Score, IS, Krzysztofowicz, 1992，後簡稱 K92)：

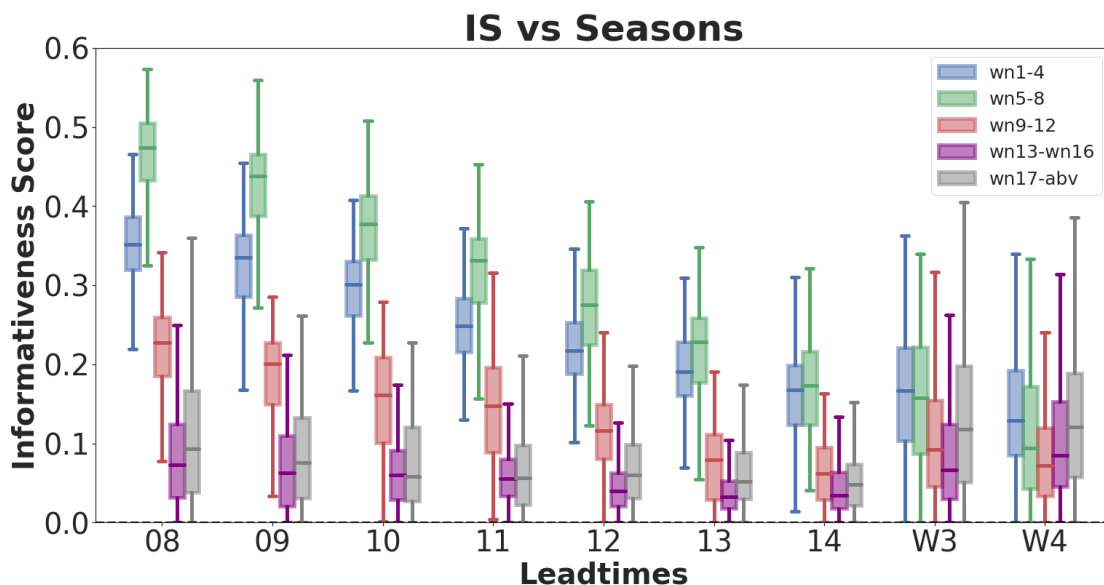
$$IS = \left(1 + \left(\frac{\sigma}{as}\right)^2\right)^{-\frac{1}{2}} \text{Eq.}(18)$$

分析各個波段訊息-雜訊比值 (Signal to Noise ratio)，做為預報因子選擇的準則。IS 介於 0-1 之間，IS=0 代表預報因子不具有訊息度 (Informativeness)，而 IS=1 時為完美預報因子。而 Eq.(18) 所需各個預報變數的 (a, σ) ，在多變量 BPE 中可用下列公式求得：

$$a = \frac{\sigma_{n0}}{\sigma_0^2}, \sigma^2 = \sigma_n^2 - a\sigma_{n0}, S = \sigma_0 \text{ Eq.}(19)$$

圖三為台灣 29 測站，11-3 月所有建模參數計算出的 IS 隨預報領先時間變化之盒狀圖。在 95%信心水準下進行 Welch's t 檢定後發現，Level 2 在第二週時最具預報訊息度，但隨著時間拉長而遞減；而到了第三與第四週 Level 1 之 IS 大於 Level 2 (以中位數做為最佳估計)，顯示大尺度特徵在展期時間尺度的主導性。比較令人意外的是，小尺度的特徵 Level 5，在第三及第四週主導地位比 Level 3 與 Level 4 高，顯示小尺度局地氣候特徵，在展期週尺度預報上提供的訊息不亞於大尺度環流，在高度時空不連續的變數，例如雨量機率預報，選擇預報因子上或許可提供重要的參考。

K92 指出，可先利用 IS 作為預報因子，評估加入另一變數後是否顯著增加貢獻或改善預報表現，以類似向前選擇法 (Forward Selection) 選擇預報因子。本研究則是直接讓多變量 BPE 依照計算出來之 $(\sigma_{n0}, \sigma_n^2, \Sigma)$ 自行決定 $\psi_{w_1}^{w_2}$ 權重，最後將權重



圖三、不同波段之 Informativeness Score 隨領先時間變化之盒狀圖，顏色分別代表： ψ_1^4 (藍)、 ψ_5^8 (綠)、 ψ_9^{12} (紅)、 ψ_{13}^6 (紫)、 ψ_{17}^{abv} (灰)之 IS。

代入 Eq.(11)，產出 $\Phi(w_{t+1}|X_1, X_2, \dots, X_n)$ 。至於加入向前選擇法是否可改進預報表現，留待後續研究做評估。在第三章預報校驗評比部分，原始模式(RAW)的 CDF 是利用最新預報的系集平均及標準差，套配常態分佈而產生；而氣候分布的 CDF 則是拿測站過去 30 年歷史資料做排序得到累積機率後，套配韋伯分佈而產生。

三、校驗結果與個案探討

為了得到具有統計代表性的校驗結果，本研究透過交叉驗證方式增加驗證樣本，以證明 BPE 校正成效，同時使用多種校驗指標來評比 MP_BPE_SH 及 SP_BPE 之機率預報表現，以避免某些校驗指標是預報系統透過「對沖」(hedging, Wilks, 2011)來提升某些技術指標之得分；許多常

見的技術指標，包含 Accuracy、Probability of Detection、False Alarm Ratio 和 Threat Score 皆可以透過 hedging 來提升得分，例如透過過度預報(Over-Forecasting)可提升 Probability of Detection (POD)。

(一) Rank Histogram/PIT Histogram

Rank Histogram(RH)為系集預報常用之校驗方法，以校驗樣本中的觀測值，是否均勻落在系集預報的範圍內，來檢驗系集預報的離散度是否合理反映預報的不確定性。若為連續型的機率預報，則採用 Probability Integral Transform Histogram (PITH)進行分析。一個離散度合理的機率型預報需要滿足所有觀測值均勻落在所有秩內，又稱秩均勻(Rank Uniformity)。舉例來說，假設有 1200 個校驗樣本，而 NCEP-GEFSv12 每個校驗樣本有 11 個成員(包含

Control Run)，則 11 個成員排序後會有 12 個秩，分別是小於數值最小的系集成員、介於第 1 個與第 2 個成員間、介於第 2 個與第 3 個成員間...、到大於數值最大的系集成員。理想狀況下，1200 個觀測值均勻落在系集成員的所有秩內的話，每個秩皆會有 100 個觀測樣本。將 12 個秩的樣本數除以總樣本數後，就可以依據這 12 個數值繪製直方圖，即為 Rank Histogram。若繪製出來的直方圖偏離秩均勻，則可依據直方圖的偏態了解系集預報系統的問題，主要有以下四種狀況：(1)離散度過大 (Over-dispersion) (2)離散度不足 (Under-dispersion) (3)過度預報 (Over-forecasting) (4) 不足預報(Under-forecasting)。

第 8-14 天之日最低溫機率預報之 RH 及 PITH[圖四(a)、(b)、(c)]顯示 RAW 有離散度不足之情形，SP_BPE 接近秩均勻，而 MP_BPE_SH 有稍微預報不足的情形。第三、第四週之週低溫機率預報[圖四(d)、(e)]顯示 RAW 有過度預報，SP_BPE 仍是非常接近秩均勻，而 MP_BPE_SH 不足預報的情形與第 8-14 天相比，明顯許多。

(二) Continuously Ranked Probability Skill Score

連續型機率預報校驗常見的技術指標之一為 Continuously Ranked Probability Skill Score(CRPSS)，透過 Continuously

Ranked Probability Score (CRPS; Hersbach, 2000)：

$$CRPS \equiv \int_{-\infty}^{\infty} [F(x) - F_o(x)]^2 dx \quad \text{Eq.(20)}$$

計算預報系統之累積機率密度函數 $F(y)$ 與觀測值的「機率密度函數」 $F_o(x)$ 在定義域的方差後進行積分。 $F_o(x)$ 為單位階躍函數(unit step function)，定義如下：

$$F_o(y) = \begin{cases} 0, & y < \text{Ground Truth} \\ 1, & y \geq \text{Ground Truth} \end{cases} \quad \text{Eq.(21)}$$

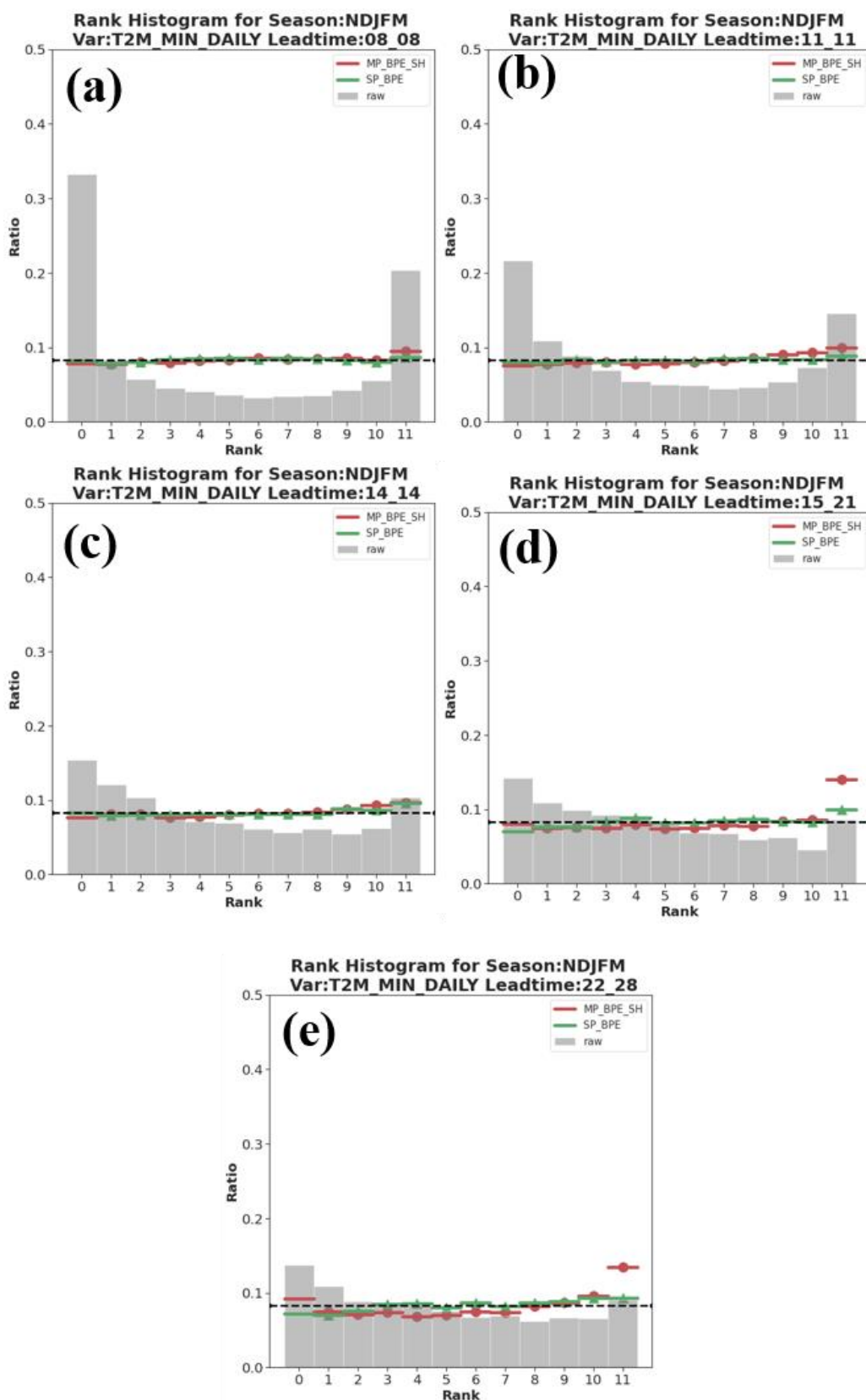
機率預報誤差 CRPS 與決定性預報的均方差概念類似，其值域為 $0 < CRPS \leq \infty$ ，0 為完美預報。

利用機率預報系統之 CRPS 長期平均，並以用氣候分佈做機率預報計算出來之 CRPS 當作基準(Baseline)，則可得的到 CRPSS：

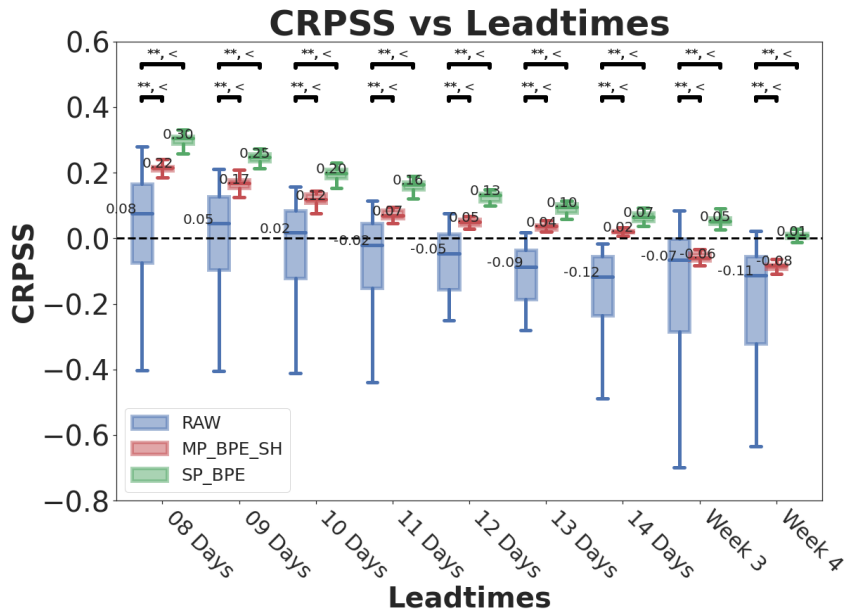
$$CRPSS = 1 - \frac{\overline{CRPS}_{forecast}}{\overline{CRPS}_{climate}} \quad \text{Eq.(22)}$$

其值域為 $-\infty < CRPSS \leq 1$ ；當 $CRPSS < 0$ 時，代表相對於氣候分布沒有預報技術，當所有校驗樣本皆為完美預報時 ($\overline{CRPS}_{forecast} = 0$)，則 $CRPSS = 1$ 。

圖五為分別計算各測站 CRPSS 後，綜合校驗期內 29 個測站的預報校驗結果所繪製出 CRPSS 隨預報領先時間變化的盒狀圖。不論是原始模式、SP_BPE、MP_BPE_SH，CRPSS 皆隨預報領先時間增加而遞減。校正效果方面，雖然 SP_BPE 之 CRPSS 較 MP_BPE_SH 來的高，但是兩者相對於氣候值皆具有預報技術，且較



圖四、RAW 實驗(灰色)、MP_BPE_SH(紅色線)、SP_BPE(綠色線)，在領先時間(a)第 8 天(b)第 11 天(c)第 14 天，預報之日最低溫及(d)第三週 (e)第四週預報之週最低溫之 Rank Histogram。虛線為秩均勻時各個秩的占比。



圖五、11-3 月 29 個測站第 8-14 天之日最低溫，及第三週、第四週之週最低溫機率預報 CRPSS 盒狀圖。藍色盒狀圖為 RAW，紅色盒狀圖為 MP_BPE_SH 實驗，綠色為 SP_BPE 實驗。數字為 CRPSS 分布之中位數。*及**分別代表在 95%及 99%信心水準下，通過 Welch' s t 檢定。<及>代表檢定之兩組實驗中位數之相對大小。

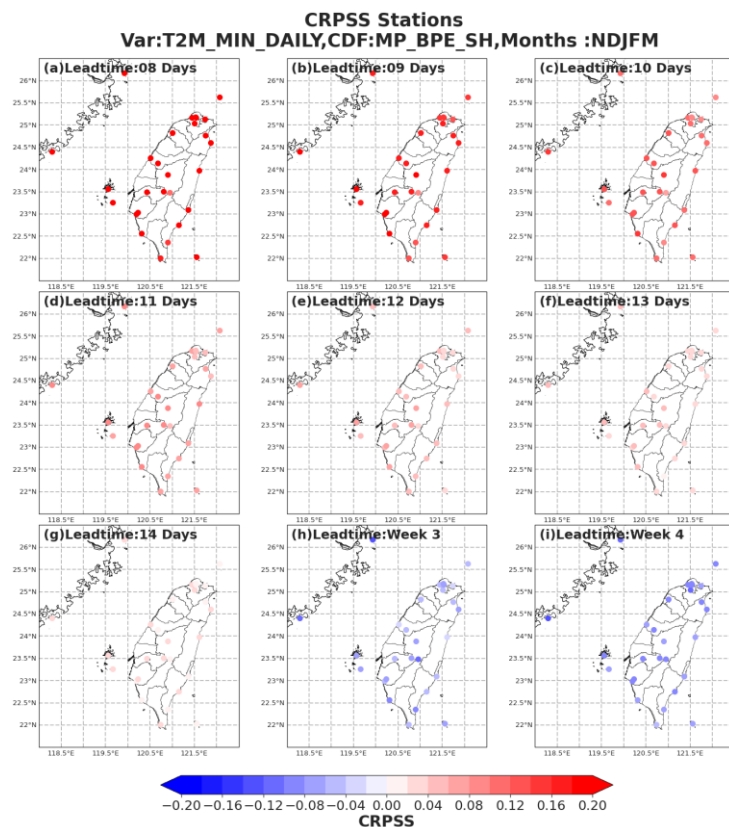
較 RAW 的 CRPSS 高的假設，在 95%信心水準下在各個 Leadtime 皆通過 Welch's t 檢定。各測站 MP_BPE_SH、SP_BPE、RAW 之 CRPSS 空間分佈隨預報時間變化圖(圖六到八)顯示，MP_BPE_SH 實驗在預報 14 天內，大多數測站的 CRPSS 皆優於氣候值(圖六)，而 SP_BPE 甚至到第三、四週多數測站的 CRPSS 皆優於氣候值(圖七)。原始模式預報在第 8-11 天於中南部測站的 CRPSS 較差，北部測站則是略優於氣候值(圖八)；第 11 天到第三、四週，全台(包含離島)測站都沒有預報技術。台灣中南部為農漁業重鎮，透過 BPE 校正原始系集預報，可有效提升極端溫度預報技術。由此可知，應用統計後處理方法於展期時間尺

度上的模式加值及強化農漁溫害預警，是極具防災效益的研究課題。

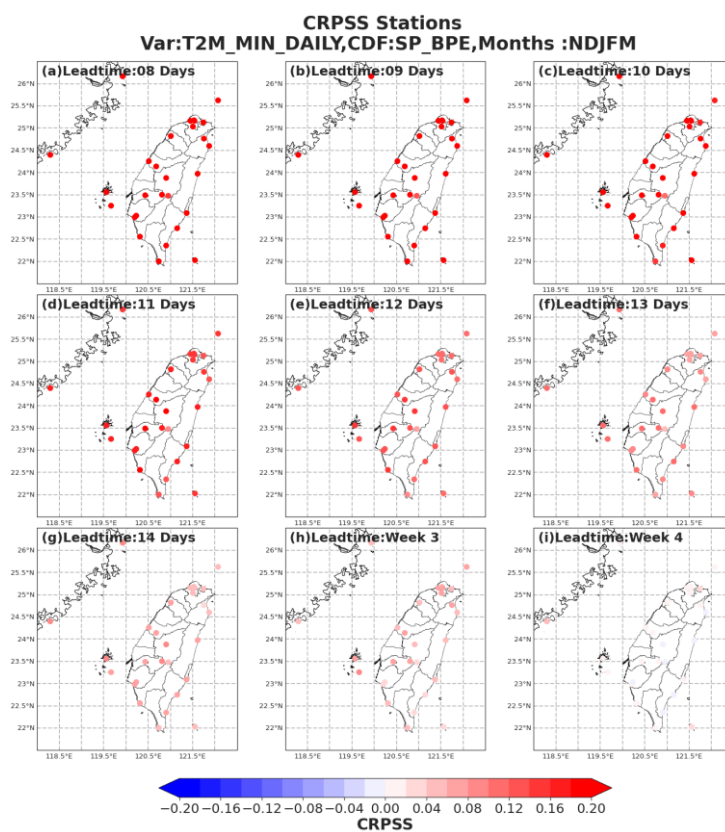
(三) Reliability Diagram 與 Brier Skill Score (BSS)

除了整體機率函數之預報誤差 (CRPS)，作業實務上，針對某(極端)溫度門檻進行機率預報校驗也相當重要，因此本研究同時使用 Brier Score (BS) 來評估特定門檻機率預報誤差。BS 為針對預報二元事件(Dichotomous Event) 設計之機率預報校驗指標，可解讀為機率預報在特定門檻之誤差，BS 定義如下：

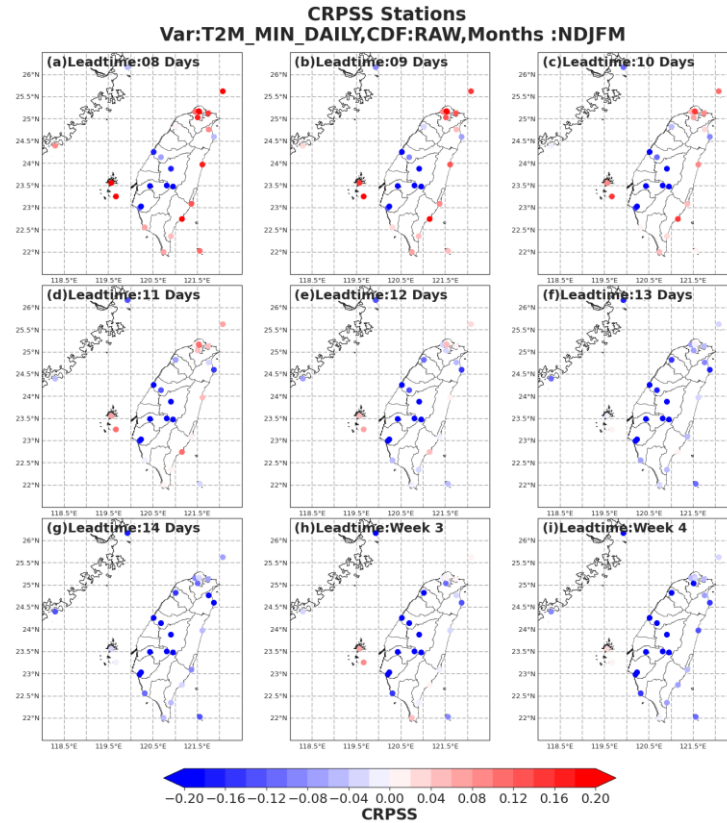
$$\text{Brier Score} \equiv \frac{1}{N} \sum_{k=1}^N (y_k - o_k)^2 \quad \text{Eq. (23)}$$



圖六、台灣人工測站 MP_BPE_SH 實驗 CRPSS 空間散布圖，由(a)-(g)分別為第 8-14 天之日最低溫，(h)第三及(i)第四週之週最低溫之 CRPSS。



圖七、同圖六，但為 SP_BPE 實驗。



圖八、同圖六，但為 RAW。

其中 y_k 為第 k 個預報之機率， o_k 為當次事件是否發生，若有發生，則 $o_k = 1$ ，反之則 $o_k = 0$ 。舉例來說，若對於 $T_{min} \leq 10^\circ\text{C}$ 的機率預報，第 $k=3$ 之預報機率為 0.6，而實際上溫度確實小於 10°C 時，此次預報在計算 BS 的數值為 $(0.6 - 1)^2$ 。根據 Murphy (1973)，BS 可分解成三項：可靠度 (Reliability)、解析度(Resolution)和不確定性 (Uncertainty)：

$$Brier\ Score = \frac{1}{n} \left[\sum_{i=1}^n N_i (y_i - \bar{o}_i)^2 - \sum_{i=1}^n N_i (\bar{o}_i - \bar{o})^2 \right] + \bar{o}(1 - \bar{o})$$

(Reliability) (Resolution) (Uncertainty)

Eq.(24)

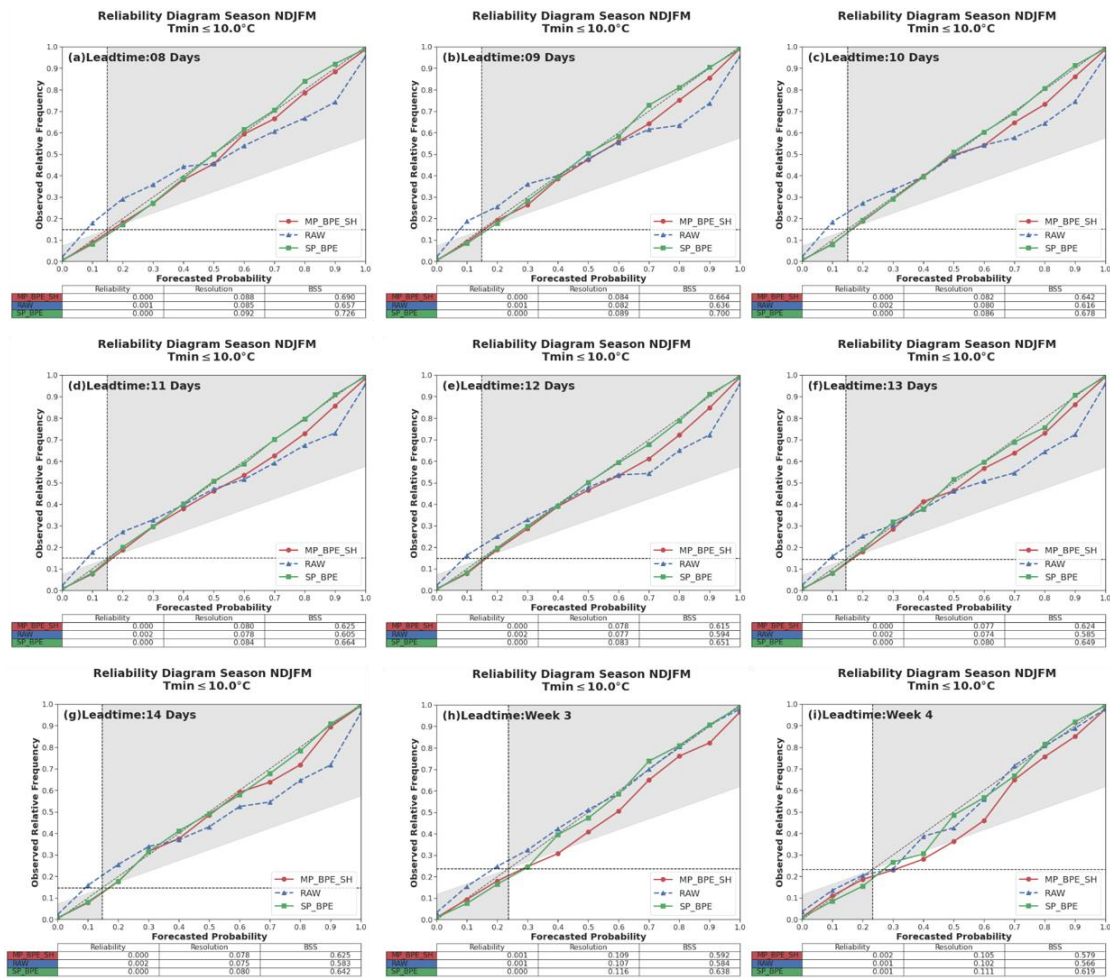
BS 之值域介於 0-1，值越低代表機率預報誤差愈小。

本研究同時採用 BSS 評估預報系統在某個(極端)溫度門檻之機率預報技術。BSS 為評估預報之 BS 相對於另一個基準預報(Baseline)之表現，數學式定義為：

$$BSS = \frac{Resolution-Reliability}{Uncertainty} \quad Eq.(25)$$

BSS 值域介於 $(-\infty, 1]$ ，若 $BSS=1$ 為完美預報， $BSS<0$ 則相對於基準無預報技術。

圖九為 $T_{min} \leq 10^\circ\text{C}$ 之 Reliability Diagram；信賴曲線(reliability curve)越貼近對角線，則機率預報越可靠；在對角線上方代表觀測頻率大於預報機率 $P(observed|forecasted) < P(observed)$ ，為不足預報 (under-forecasting)，在對角線下方則為過度預報



圖九、(a)-(g)為 11-3 月台灣人工測站第 8-14 天之日最低溫 $\leq 10^\circ\text{C}$ 、(h)第三及(i)第四週之週最低溫 $\leq 10^\circ\text{C}$ 之 Reliability Diagram。藍色虛線為原始模式之機率預報，紅色實線為 MP_BPE_SH 之 Reliability Curve，綠色實線為 SP_BPE。

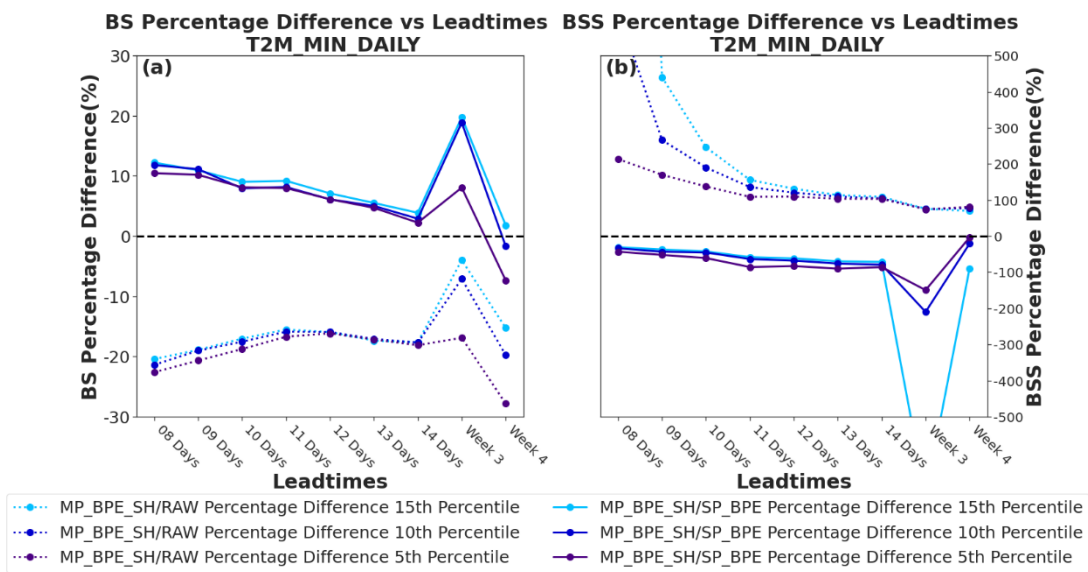
(over-forecasting)。因此，由圖九可知 RAW 在低預報機率樣本 ($P(T_{min} \leq 10^\circ\text{C}) \leq 0.5$) 有低報現象，在高預報機率樣本 ($P(T_{min} \leq 10^\circ\text{C}) > 0.5$) 有過度預報。經 BPE 修正後，不論 MP_BPE_SH 或 SP_BPE，在第 8-14 天[圖九(a)-(g)]皆可明顯改善可信度(信賴曲線較 RAW 貼近對角線)。SP_BPE 實驗在 $P(T_{min} \leq 10^\circ\text{C})$ 可靠度方面表現略優於 MP_BPE_SH，但無論如何，在 BPE 校正之後之最低溫機率預報

對於此門檻有幾乎完美的可靠度。而在預報第三、四週[圖九(h)-(i)]，SP_BPE 可改善可靠度，但是 MP_BPE_SH 則改善有限。此外，本研究也分析 MP_BPE_SH、SP_BPE， $T_{min} \leq 10^\circ\text{C}$ 之 ROC 曲線及 AUC，SP_BPE 預報第 8-14 天之日最低溫及第三、四週之週最低溫，AUC 都在 0.95 以上，而 MP_BPE_SH 之 AUC 都在 0.91 以上，代表對於 $T_{min} \leq 10^\circ\text{C}$ 具有區辨能力。

雖然固定溫度門檻之機率預報在預報實務上較為常用，然而依據 WMO 極端事件之定義，極端低溫須以各測站氣候分佈的百分位而定。本研究極端事件機率門檻採用較嚴格的定義，為各測站過去歷史觀測紀錄在目標預報日第 5、10 及 15 百分位為該測站極端低溫之門檻。

MP_BPE_SH 相較於 RAW，及 MP_BPE_SH 相較於 SP_BPE 實驗之 BS 及 BSS 百分比差異如圖十所示。百分比差異定義如下：

$$\text{percentage difference} \equiv \frac{A - B}{|B|} \times 100(\%) \quad \text{Eq.(26)}$$



圖十、(a): 11-3 月台灣人工測站第 8-14 天之日最低溫，及第三四週之週最低溫機率預報第 15、10、5 百分位之 BS 百分比差異。(b):11-3 月 29 個測站第 8-14 天之日最低溫，及第三四週之週最低溫機率預報第 15、10、5 百分位之 BSS 百分比差異。虛線為 MP_BPE_SH 相較於 RAW 之百分比差異，實線為 MP_BPE_SH 相較於 SP_BPE 百分比差異。

MP_BPE_SH 的 BS, BSS 當作 Eq.(26) 中之 A，而 RAW 及 SP_BPE 分別當作 Baseline(B)。由於 BS 越小越好，所以當 Eq.(26) 小於 0，則 MP_BPE_SH 相較於 Baseline 有進步，而大於 0 則反之。BS 評估結果顯示：MP_BPE_SH 相較於 RAW，約改善 10-20%，而 SP_BPE 相較於 MP_BPE_SH 的進步幅度在 10% 以內 [圖十(a)]。由於 BSS 越大越好，所以當 Eq.(26) 大於 0，則 MP_BPE_SH 較好，而小於 0 則反之，BSS 評估結果顯示：

MP_BPE_SH 相較於 RAW，約改善 100-200%，而 SP_BPE 相較於 MP_BPE_SH 的進步幅度在 100% 以內 [圖十(b)]。

(四) Extended Calibration Score

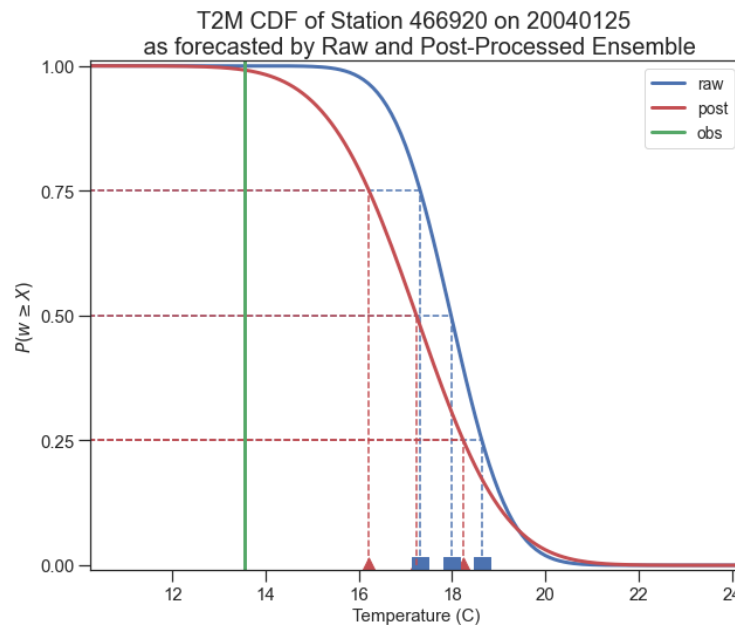
Reliability Diagram 及 BS 可以針對某一個類別的事件及百分位門檻進行校正效果分析，但是無法看出整體機率分布之校正效果。此外，在評估不同預報系統的表現時，門檻訂定的資料選擇、定義方式不

一樣也會造成評比上標準不一致的問題(例:某測站溫度第 85 百分位的門檻值要取自 6-9 月所有溫度樣本,還是不同太陽日分別求第 85 百分位門檻),計算出不同的 Brier Score。因此,本研究延伸 Krzysztofowicz & Sigrest (1999) 提出的 Calibration Score (CS),提出 Extended Calibration Score(ECS):

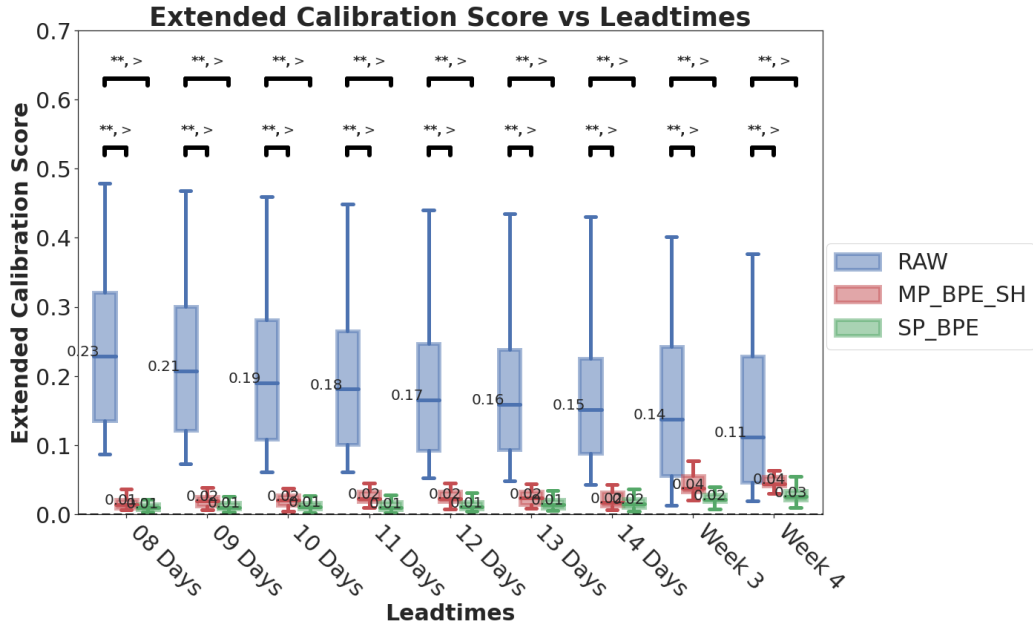
$$ECS = \sqrt{\frac{1}{10} [\sum_{i=1}^{10} (r_{0.1i} - 0.1i)^2]} \quad \text{Eq.(27)}$$

$r_{0.1i}$ 表示在所有樣本 N 中,與觀測值超越 Φ 第 0.1i 百分位對應到的溫度量值數量 $n_{0.1i}$ 的比值,本研究之 ECS 採用的百分位為 [0.1,1.0] 區間中 0.1 的倍數。以圖十一舉例說明,當樣本中某一目標時間之溫度觀測值為綠色垂直線時,由於觀測值皆低於

三個範例百分位(0.25,0.5,0.75)的預報值,則 $n_{0.25}, n_{0.5}, n_{0.75}$ 皆不累加。從圖中很明顯從超越機率函數曲線看出,RAW 的機率函數(藍色曲線)高估了當天的日均溫,導致觀測值落在預報 CDF 極右側(即原始模式認為當日均溫為觀測值之機率为 0)。經過 BPE 修正之後的 CDF(紅線)較為合理,有稍微修正誤差,但仍為高估。反之若觀測值皆高於三個範例百分位,則 $n_{0.25}, n_{0.5}, n_{0.75}$ 皆累加 1。ECS 與 BS 中 Reliability 項相似,長期統計下來,當 $r_{0.1i}$ 之量值越接近 0.1i,代表 Φ 在此超越百分位的預報機率为可信,即為越良好校正(well-calibrated)。將各百分位之 $r_{0.1i}$ 取 RMSE,即反應出整個 Φ 良好校正的程度。由 Eq.(27)可看出,機率为分布的 ECS 值越接近



圖十一、超越機率为示意圖。藍色線(raw)為原始模式預報 20040125 日均溫超越機率为,紅色線(post)為經過 BPE 修正的日均溫超越機率为。在 x 軸上三角形為 post 的超越機率为 75、50、25 百分位對應到的溫度值,而正方形為 raw 的超越機率为 75、50、25 百分位對應到的溫度值。



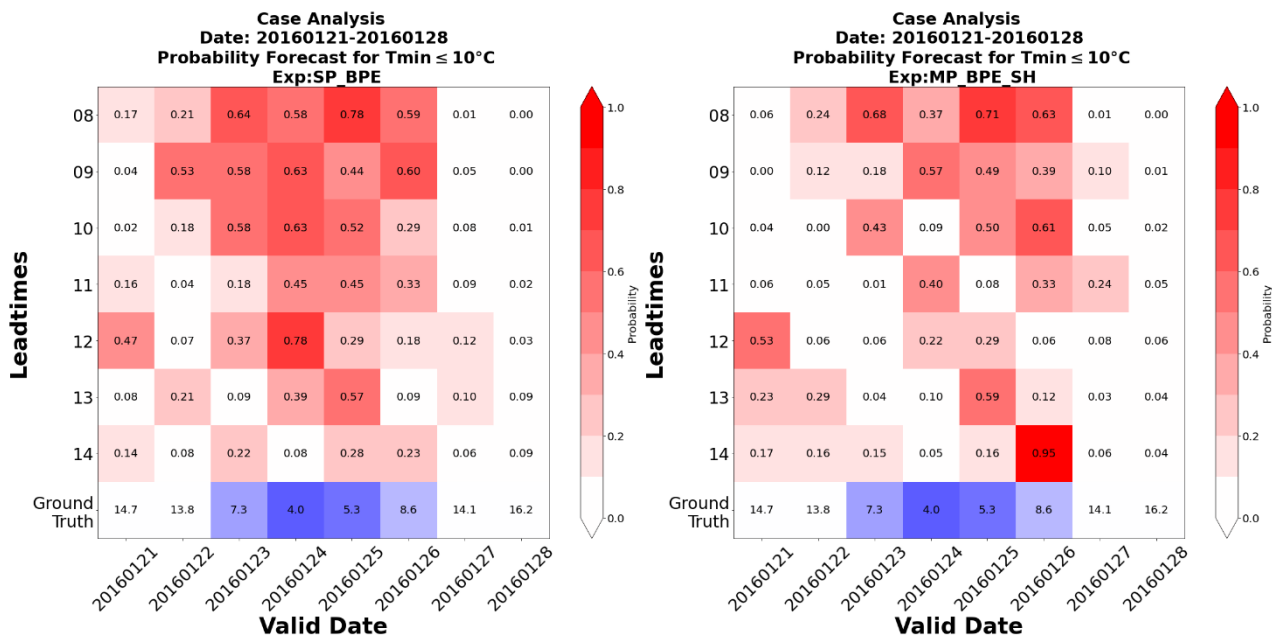
圖十二、11-3 月台灣人工測站第 8-14 天之日最低溫，及第三四週之週最低溫機率預報之 Extended Calibration Score 盒狀圖，分別為 RAW(藍色)、MP_BPE_SH(紅)、SP_BPE(綠色)。*及**分別代表在 95%及 99%信心水準下，通過 Welch’ st 檢定。<及>代表檢定之兩組實驗中位數之相對大小。

0，代表越接近良好校正。圖十二顯示 MP_BPE_SH 及 SP_BPE 經 BPE 校正後，其相對於 RAW 的 ECS 有大幅度的進步，皆在 99%的信心水準下通過 Welch’s t-test，並且接近於完美校正 (perfect-calibrated; ECS ~ 0)，惟 SP_BPE 的 ECS 仍略優於 MP_BPE_SH。

(五) 個案探討-2016 年超級寒流

2016 年 1 月下旬，北極震盪強負相位造成冷空氣大舉南下，讓台灣許多平地測站創下破紀錄的低溫。因此，本研究以 2016 年 1 月 21 至 27 日超級寒流為個案，展示 BPE 校正後的預報對極端低溫事件的掌握能力。圖十三為 BPE 機率預報在台

北測站的個案校驗結果，左圖和右圖分別是 MP_BPE_SH 及 SP_BPE 在 8 至 14 天前預報日最低溫<10°C的機率，最下方列為測站的日最低溫觀測值。預報領先時間較短(8-10 天)時，BPE 對於低溫事件(< 10 °C)有較高的預報機率，而對於非低溫事件(>10°C)有接近於 0 的機率，顯示 BPE 校正後的預報具有良好的區辨能力(亦即，預報結果能夠區辨事件的發生與否);但隨著預報時間拉長(11 天以上)，預報機率逐漸喪失區辨能力。這是數值天氣預報在可預報度上的科學限制；但當預報時間拉長，原始預報的可預報度明顯降低，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



圖十三、格點圖第 1 -9 列:領先時間 8-14 天預報台北測站 2016 年 1 月 21 至 27 日日間最低溫低於 10 度機率,不同顏色對應不同量值的預報機率。左圖為 SP_BPE 實驗,右圖為 MP_BPE_SH 實驗。最下方一列為溫度觀測值,上色代表觀測溫度低於 10°C。

資訊。因此,10 天以上的預報,比較恰當的呈現方式是以「週」為單位來呈現,例如週均溫、週累積雨量、週最高溫或最低溫等,而不再以「日」為單位來呈現(White et al.2017)。這是因為預報能力(forecast skill)與天氣系統的空間尺度高度相關;當預報時間拉長,只剩較大尺度天氣系統具有可預報度(Boer 2003)。時間上的平均(例如週均溫)或累積(如週累積雨量)反映的就是較大尺度的特徵。

四、總結

本研究的重要結論如下:

1. 對於可預報度較低的展期天氣及極端溫度事件,本研究透過具有嚴謹數學理論的 BPE 方法對模式報進行偏差修正及

降尺度,大幅改善提升原始模式預報的預報技術與可信度。

2. 利用球面調和函數對全球模式的預報輸出進行波段拆解,並將其與台灣測站最低溫進行訊息度分析,結果顯示:訊息度最高的波段會隨領先時間改變;在預報領先時間 8~14 天佔主導地位的波段 ψ_5^8 ,其 IS 隨著領先時間逐步遞減,到了第三、四週,代表最大尺度的 ψ_1^4 變成訊息度最高的波段,而代表較小尺度的 ψ_{17}^{abv} ,則轉換為訊息度次高的波段。
3. 在第 8-14 天之日最低溫機率預報方面,以單一預報濾波後的波段作為多變量 BPE 預報因子之技術,雖然未達到與以系集平均做為預報因子相當之校正成效。然而,我們證明了利用濾波方法搭

配多變量 BPE，仍然可以只使用決定性模式(Deterministic Model) 在展期時間尺度產出具可信度與預報技術的機率預報。結合 BPE 相對不需太多重預報資料即有良好校正成效的優勢，若因為運算資源有限而無法產製長期的系集重預報時，本研究發展之統計後處理方法可作為替代方案之一。

4. 在第三、四週之週最低溫機率預報方面，以單一模式濾波後的波段作為多變量 BPE 預報因子之技術未如預期，推測以最低溫濾波後的波段當作預報因子可能不是預報展期週尺度最低溫的最佳策略。未來擬嘗試使用代表大尺度環境特徵的 500 百帕高度場或其他對台灣週極端溫度有預報訊息度之變數進行濾波，當作多變量 BPE 之預報因子，以期改善校正成效。

致謝

本研究感謝中央氣象局科技中心劉邦彥技士提供濾波程式及濾波方法技術指導、美國國家海洋暨大氣總署全球系統實驗室(NOAA-GSL)資深科學家 Dr. Zoltan Toth 在 BPE 研發過程中給予寶貴建議，以及中央氣象局提供計算資源。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

- Boer, G.J., 2003: Predictability as a function of scale, *Atmosphere-Ocean*, 41:3, 203-215, DOI: 10.3137/ao.410302.
- Chang, H.-L., Yang, S.-C., Yuan, H., Lin, P.-L., & Liou, Y.-C., 2015: Analysis of the relative operating characteristic and economic value using the LAPS ensemble prediction system in Taiwan. *Monthly Weather Review*, 143(5), 1833-1848.
- Chou, S. C., H. L. Chang, K. J. Chen, and J. S. Hong, 2021: Forecast Evaluation and Economic Value Analysis of the Probabilistic Forecasts of 100-meter Wind Speed from WRF Ensemble Prediction System over Taiwan Area. 2021 Asia Oceania Geosciences Society 18th Annual Meeting.
- Chu, H.-Y., Chang, H.-L., Chou S.-C., Lo T.-W., Chen Y.-J., 2020: Week 3 Temperature Probabilistic Forecasts using Bayesian Processor of Ensemble. 109 年天氣分析與預報研討會，論文彙編，中央氣象局，臺灣，臺北，A3-19
- Chu, H.-Y., Chang, H.-L., Toth, Z., 2021: Extended-Range Temperature Probabilistic Forecasts Using the Bayesian Processor of Ensemble Over Taiwan. 2021 Asia Oceania

- Geosciences Society 18th Annual Meeting, AS30-A013
- Hersbach, H., 2000: Decomposition of the continuous ranked probability score for ensemble prediction systems. *Weather and Forecasting*, 15(5), 559-570.
- Krzysztofowicz, R., 1992: Bayesian correlation score: A utilitarian measure of forecast skill. *Monthly Weather Review*, 120(1), 208-220.
- Krzysztofowicz, R. & Maranzano, C., 2006: Bayesian Processor of Output for Probabilistic Quantitative Precipitation Forecasting. AGU Spring Meeting Abstracts.
- Krzysztofowicz, R., & Evans, W. B., 2008: Probabilistic Forecasts from the National Digital Forecast Database. *Weather and Forecasting*, 23(2), 270-289. doi:10.1175/2007waf2007029.1
- Krzysztofowicz, R., & Sigrest, A. A., 1999: Calibration of probabilistic quantitative precipitation forecasts. *Weather and Forecasting*, 14(3), 427-442.
- Machenhauer, B., 1991: Spectral Methods. Paper presented at the Seminar on Numerical Methods in Atmospheric Models, 9-13 September 1991, Shinfield Park, Reading. <https://www.ecmwf.int/node/10901>
- Murphy, A. H. (1973). A New Vector Partition of the Probability Score, *Journal of Applied Meteorology and Climatology*, 12(4), 595-600. Retrieved Mar 2, 2022, from https://journals.ametsoc.org/view/journals/apme/12/4/15200450_1973_012_0595_anvpot_2_0_co_2.xml
- Wang, Y., Zhang, X., & Toth, Z., 2018: Application of the Bayesian Processor of Ensemble to the Combination and Calibration of Ensemble Forecasts. Paper presented at the International Conference On Signal And Information Processing, Networking And Computers.
- Wan, Z.Y., Dodov, B.A., Lessig, C., Dijkstra, H.A., & Sapsis, T.P. (2021). A data-driven framework for the stochastic reconstruction of small-scale features with application to climate data sets. *J. Comput. Phys.*, 442, 110484.
- Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences* (Vol. 100). Academic press.
- White, C.J., Carlsen, H., Robertson, A.W., Klein, R.J., Lazo, J.K., Kumar, A., Vitart, F., Coughlan de Perez, E., Ray, A.J., Murray, V., Bharwani, S., MacLeod, D., James, R., Fleming, L., Morse, A.P., Eggen, B., Graham, R., Kjellström, E., Becker, E., Pegion, K.V., Holbrook, N.J.,

- McEvoy, D., Depledge, M., Perkins-Kirkpatrick, S., Brown, T.J., Street, R., Jones, L., Remenyi, T.A., Hodgson-Johnston, I., Buontempo, C., Lamb, R., Meinke, H., Arheimer, B. and Zebiak, S.E. (2017), Potential applications of subseasonal-to-seasonal (S2S) predictions. *Met. Apps*, 24: 315-325. <https://doi.org/10.1002/met.1654>
- 朱心宇、張惠玲、陳昀靖、羅存文、周柿均，2021: 貝氏系集處理器應用於展期極端溫度機率預報，110 年天氣分析與預報研討會，論文彙編，中央氣象局，臺灣，臺北，A3-12
- 林涵芳、陳苡甄、馮智勇、羅存文、張惠玲，2021: 貝氏模型平均法與多變量類比法應用於臺灣測站第 3 週均溫預報之校驗評估，110 年天氣分析與預報研討會，論文彙編，中央氣象局，臺灣，臺北，A2-61
- 周筱倩、張惠玲、馮智勇、陳奕翰，2021: 高低溫機率預報決策支援系統之發展與建置，110 年天氣分析與預報研討會，論文彙編，中央氣象局，臺灣，臺北，A4-29
- 郭鴻基、林沛練，1991: 發展中的氣象局全球預報系統，80 年天氣分析與預報研討會，論文彙編，中央氣象局，臺灣，臺北，20
- 陳昀靖、張惠玲、陳孟詩、羅存文、周柿均、朱心宇，2021: 台灣測站第三第四週極端低溫預報，110 年天氣分析與預報研討會，論文彙編，中央氣象局，臺灣，臺北，A3-15

附錄

(一) Welch's t-test

Welch's t-test 又稱為 Unpaired t-test，為 Student's t-test 的變形，使用時機為比較之兩組樣本大小不一樣，及樣本抽樣來源的母群體有不同的變異數(Variance)時。其檢定目的與 Student's t-test 相同，皆是用來檢定兩組樣本之母群體是否有相同的平均值。Welch's t-test 之 t statistic 為：

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}}$$

其中， \bar{X}_i 為樣本 i 的平均， s_i 為標準誤差(Standard Error)， N_i 為樣本大小。而 Student's t-test 之 t statistic 為：

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_p \sqrt{\frac{1}{N_1} + \frac{1}{N_2}}}$$

s_p 為合併方差(pooled variance)，類似於將不同樣本的變異數做加權平均。由此可以看出，Welch's t-test 由於少了合併方差這個動作，所以會對變異數更為敏感。

Applying Multi-Variate Bayesian Processor of Ensemble and Spherical Harmonics Waveband Decomposition to Extended-Range Extreme Low Temperature Probabilistic Forecasts

Hsin-Yu, Chu, Hui-Ling, Chang

Central Weather Bureau

(manuscript received 26 July 2022 ; in final form 6 December 2022)

ABSTRACT

To produce elaborate, location specific extended-range probabilistic forecasts, it is required to bias-correct and downscale the raw model. If statistical methods are used, then the two processes can be collectively phrased as “statistical post-processing (SPP)”. To infer a representative estimate of the raw model’s systematic bias, large reforecast sets have to be produced. However, creating reforecast for a recently operational, state-of-the art ensemble that prolongs to extended-range forecasting is very computationally extensive. In practical operations, the size of reforecast are often limited, and far smaller than observations. In this research, we aim to evaluate if the forecast performance of filtering non-linear noise using ensemble mean, is achievable by wave decomposition of a single model with the best initial condition and post-processing the wavebands using state-of-art SPP method. We hope this will reduce the reliance on both ensemble and large reforecast in SPP, providing a solution to aforementioned issues.

The control run from raw ensemble model are decomposed to individual wavebands using spherical harmonics, and the wavebands are used as predictors. These predictors are then feed into a state-of-the-art, multi-variate Bayesian Processor of Ensemble (BPE), to produce post-processed extreme minimum temperature probabilistic forecasts located on specific meteorological stations in Taiwan. As a full Bayesian Method, BPE utilizes the copula of the predictors-predictand pair to generate a marginal distribution as the likelihood, and the climatological distribution constructed by the observational data as prior. The predictive distribution, or the posterior, are generated using fusion of likelihood and prior once receiving the latest run from the raw model. Due to its Bayesian structure, the advantage of BPE is the

capability to derive well-calibrated posterior under limited reforecast data, from using a larger observation dataset to construct an informative prior.

This research uses (1) The ensemble mean of 2-meter minimum temperature (Experiment SP_BPE), and (2) The wavebands decomposed from the standardized 2-meter minimum temperature of the control run (Experiment MP_BPE_SH) of NCEP GEFSv12 as the predictor(s) to calibrate daily minimum temperature for lead times of 8-14 days and weekly minimum temperature for lead times of week 3 and 4. We use various performance benchmarks to compare the different aspects of probabilistic forecast and found that both SP_BPE and MP_BPE_SH could improve the quality of probability forecasts over raw ensemble model. It is worth noting the performance of SP_BPE is better than MP_BPE_SH, especially in the lead times of week 3 to week 4, but we think that the cause should be related to the selection of predictors, instead of a deficiency in the methodology. In the future, alternative variables can be decomposed, or use large scale indices as predictors to improve the calibration of week 3 and 4.

Keywords : Bayesian Processor of Ensemble, Spherical Harmonics, Extreme Temperature, Extended Range Forecasts, Statistical Post-Processing. doi: 10.53106/025400022022125003002