

# 中央氣象局數值預報模式降水預測校驗 ：1999 與 2000 年

潘琦<sup>1</sup> 盧孟明<sup>2</sup>

<sup>1</sup>中央氣象局氣象資訊中心

<sup>2</sup>中央氣象局氣象科技研究中心

(中華民國九十年五月八日收稿；中華民國九十年八月二十九日定稿)

## 摘 要

本文以觀測資料校驗中央氣象局全球波譜作業模式 (GFST79 和 GFST120) 和有限區域作業模式 (LAFS) 在 1999 與 2000 年間對 12、24、36 與 48 小時預報之每 12 小時累積降水量，評估各模式的降水預測技術。校驗的地區乃是按觀測資料的完整性在東南與東北亞選取了 9 個區域，分別依年、季與月平均以及不同累積雨量類別統計校驗結果。

在年平均量方面，以每 12 小時累積降水量 0.1 公釐為判定標準時，大部分區域以 LAFS 的分數最低，並顯示 LAFS 有預報偏溼的特性。3 個模式在 9 區中以台灣地區 (B 區) 的預報得分相對較低，全球模式對南海附近 (A 區) 的降水預測也較不理想。另外，同一模式的降水預測能力有明顯的日夜差異，以夜間降水的預測準確率較差。平均而言，預報得分隨著預報時間的增加而降低，區域性差異則是變小。在季節變化方面，三組模式在七至九月的降水預報為整年中表現最差的部份，LAFS 尤其明顯。

按不同雨量類別累積量校驗後，發現不論是對大雨或小雨的預測，在東北亞的降水預測技術得分遠比東南亞高，也就是說三組模式對於日本與韓國地區的降水預測能力明顯地比南海與台灣地區準確。就全球模式而言，對於極小雨在南海與台灣地區會有預測偏溼的傾向。而有限區域模式對極小雨預測則於各區都偏溼。對於不同預報時間的比較，以 GFST120 的預測技術最穩定，四個時間的預報結果之間差別最小，然而大致上各模式都是以 24 小時的預測最為準確。值得注意的是，除了極小雨 (0.1 公釐) 以外，大致上 LAFS 在三個模式當中降水預報得分最高，以日本與韓國地區尤其明顯，表示區域模式對於中大雨預報具有較高參考價值。

若以月平均降水總量校驗不同的雨量類別，我們發現三個模式有共同的預報特性：小雨高估，偏差值大；大雨低估，偏差值較小。由於在溼季雨量的高低估現象明顯，相互抵銷平衡之下，以致在平均雨量方面的得分比其他季節為佳。冬季因大雨出現機會少，預報低估了中雨，因此降水總量預報以高估居多，比夏季差。若充分考慮有無降水以及降水大小類別的預測，結果顯示夏、秋季的得分實際上是比較低的，換言之模式降水預報準確度較低。

綜合而言，中央氣象局全球模式的降水預測技術與其他作業中心全球模式的成績大致相當，中央氣象局區域模式除小雨之外其他類別的降水預測技術可超越全球模式。本報告藉由校驗不同模式的預測結果，指出降水校驗的重要性與複雜性，同時顯示更精確仔細的長期預報校驗對於數值預報模式的選取與應用是絕對必要的。

**關鍵詞：**降水預測、模式校驗、預報得分

## 一、前言

降水預報是天氣預報作業的重點工作之一，數值預報模式產品又是極為重要的預報參考資料之一。降水預測內容包括了雨量大小與降水發生的時間和空間分布，難度甚高。因此，客觀地評估數值模式的降水預測技術，藉以顯示模式的優缺點，是校驗預報模式的一項重要工作。

關於模式降水預測校驗的方式，目前大都採用各項統計校驗指標（如 Ghelli and Lalaurette, 2000; McBride and Ebert, 2000）。由每日觀測資料，比對模式降水預測量，再透過一些統計方法，以了解模式降水預測的區域性特徵和隨季節變化的特性，並提供給模式發展人員作為改進模式的參考。在校驗工作中，如何使用觀測資料是非常重要的一個環節，模式輸出的網格通常不會恰好和測站位置相同，因此校驗結果的公平性往往會隨資料在空間上處理方式的不同而改變，這是值得注意的一個問題。Ghelli and Lalaurette（2000）以兩組不同的觀測資料校驗模式，一是在空間上分布密度較低且較不均勻的測站資料對比，他們以每個測站為中心點，將四週的模式預測值平均，再與觀測資料比較。另一組則是法國高密度降水觀測資料，其空間上的密度比模式的網格密度還大，因此是以模

式網格點為中心點來作校驗。結果顯示隨不同的觀測資料和處理過程中資料分類的方式，模式的預報得分會有明顯的差異。對於 ECMWF（European Center for Medium-Range Weather Forecasts）的 T213 作業模式而言，以高密度觀測資料做出的技術得分更能顯示模式的優點。

本文是以東亞地區一些測站的降水資料，校驗中央氣象局已上線作業的數值天氣預報模式的預報結果。第二節詳細說明資料的篩選方式和校驗方法，第三、四、五節分別討論年、季與月平均以及不同累積雨量類別統計的校驗結果，第六節為結論。

## 二、資料與分析方法

### （一）資料

本文以東亞地區地面測站降水觀測資料為標準，來校驗中央氣象局的數值天氣預報（NWP）模式的預報結果。校驗時間自 1999 年 1 月 1 日至 2000 年 12 月 31 日共兩年。測站觀測資料的來源是透過中央氣象局全球氣象資料即時傳輸系統（Global Telecommunication System）即時接收、解譯氣象電碼，所蒐集的每日逐時地面綜觀報。至於校驗的對象是在中央氣象局作業線上的 NWP 模式預測降水量。校

驗的模式包括了 T79 解析度之全球波譜預報(陳等, 1994) 和 T120 解析度之全球波譜預報(陳與劉, 1997) 等二種不同解析度的模式, 以及有限區域預報模式(葉等, 1994) 60 公里解析度的粗網格組, 各個模式分別有獨自的資料同化和客觀分析系統。文中分別以 GT79、GT120 和 CA 來標示以上三個模式的產品。

研究中直接採用數值模式每日二次(00UTC 和 12UTC) 輸出的預測 12 小時累積降水量, 所選取的預報領前時間(lead time)分別是 12、24、36 和 48 小時。在觀測資料方面, 我們直接根據預報模式每 12 小時累積降水量的輸出規則, 選擇了符合此一規則的測站。本研究為氣象局 NWP 作業線上降水預測校驗系統的第一項發展工作, 為了簡化工作的複雜度, 並未另外處理不符合模式輸出規則的觀測資料。在下一階段

的系統發展計畫, 將要解決這個問題, 以期充分利用 GTS 的地面觀測資料校驗模式。另一項篩選觀測資料的標準是以模式目前輸出存檔的格點型式(2.5° × 2.5°) 為根據, 將東亞地區劃分為網格, 挑選出每一網格內測站數大於 4 的測站作為校驗的觀測資料基礎。

經過一連串的篩選步驟後, 再將相鄰的網格合併, 最後以圖 1 所示的 9 塊區域進行校驗計算, 這些區域分別以 A 至 I 為代號, 同時各區域內的測站分布及數目亦清楚地標示在圖中。各區域內的測站數目為兩年平均的測站總數。每一區域內至少有 12 個符合選取標準的測站, 而測站分布大都集中在陸地上。對於廣大的大陸地區以及菲律賓、中南半島, 因 GTS 上的觀測資料降水編報較複雜, 不容易直接求得 12 小時累積降水量, 因而沒有採用。日後若能再

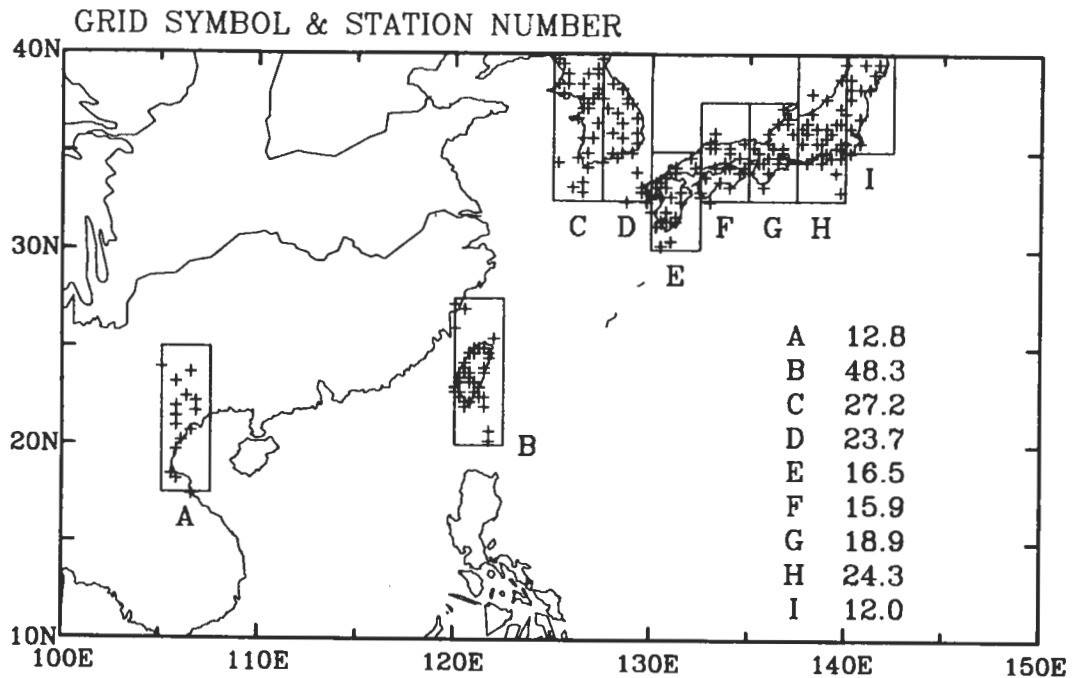


圖 1 A 至 I 九個校驗區域圖。+ 為測站位置, 圖右下數字為該校驗區域的總測站數。

將這些測站降水資料加以處理，即可擴大校驗範圍，做更詳盡的降水診斷工作。

## (二) 分析方法

在降水預報校驗上，我們採用列聯表 (Contingency Table)，也就是對於每一個網格點來校驗其降水事件預報正確與否。同時，也以不同的降水量閾值 (threshold) 作為降水類別命中與否的判別標準，本研究所採用的閾值分別為每 12 小時累積降水量 0.1、1、2、5、10、20 和 50 公釐等 7 個標準。在相同的時間點上，對於每一網格點進行降水校驗時，其結果必定屬於列聯表內四種類別中的一種。表一即為本文所採用的列聯表，其中 H (hit) 代表預測降水且實際上也發生降水的次數，M (miss) 表示實際發生降水卻未預測到的次數，F (false alarm) 表示預測降水但實際情況並未發生的次數，Z (correct no-rain forecast) 代表實際上沒有降水同時也預測不降水的次數。

根據列聯表 (表一) 中的次數統計，我們計算了下列幾種校驗模式的降水預報技術的指標。

1. 可預測率 (Probability of Detection, 參考 Wilks, 1995) 或陳等 (1991) 所稱的前估 (Pefigurance):

$$POD = \frac{H}{H + M}$$

表一 降水預報校驗之列聯表

觀測降水	預測降水	
	是	否
是	H	M
否	F	Z

表示所有的實際發生降水事件中，預測正確的比例。若 POD 愈接近於 1，則代表事先預測降水事件發生的能力愈高。但是由公式中可知，只要多預測降水，當真正發生降水時，自然就可提高預測正確的比例。因此 POD 並不能完全地代表預測技術的高低。

2. 偏離係數 (BIAS, 參考 Wilks, 1995):

$$BIAS = \frac{H + F}{H + M}$$

表示預測有降水和觀測到降水發生的相對頻率計算。BIAS 愈接近於 1，表示預測降水和實際發生的次數愈相近，亦即預測愈準確。它也可以表示成預測降水的區域和實際降水區域的比例。除了 POD 外，也需要一併探討 BIAS，才能具體了解降水預測的技術。

3. 公平得分 (Equitable Threat Score, 參考 Black, 1994):

$$ETS = \frac{H - C}{H - C + M + F}$$

其中 C 為隨機預測所預期的正確次數， $C = (H + M)(H + F) / T$ ，T 為總預報次數， $T = (H + M + F + Z)$ 。ETS 不因事件發生的頻率變化而有偏差，代表模式超出隨機的預測能力，為一比較公正的校驗指標，也是近年來各研究和作業單位在降水預報校驗上所廣泛應用的。

4. HKS (Hanssen-Kuipers Score, 參考 McBride and Ebert, 2000):

$$HKS = \frac{HZ - MF}{(F + Z)(H + M)} \\ = \frac{H}{H + M} + \frac{Z}{F + Z} - 1$$

公式中的第一項是預測降水發生的準確率

(POD)，第二項是預測沒有降水發生的準確率，第三項(-1)可視為使 HKS 的值介於 1 至 -1 之間標準化的步驟，使得 HKS 值愈接近於 1，表示預報能力愈好。由於 POD 只是單純地計算實際上有降水發生的預測率，不像 HKS 又另外加入計算無降水時的預報準確率，因此 HKS 是同時考慮有無降水兩方面的預報技術，能更完整地評估降水預報能力，此乃 HKS 的一大優點。在校驗過程中若能同時參考 POD，將能更適切地指出模式預測是在有降水發生的部份或是無降水發生時的預測技術較差。相對於 ETS 而言，我們認為 HKS 更能反映模式整體表現，對於預測校驗結果的分析探討甚有幫助。

校驗的降水預測按觀測時間分為 00UTC 和 12UTC 兩組，校驗每 12 小時的累積降水量。這裡 00UTC 是指觀測的降水量是前一日 12UTC 至當日 00UTC 之間的 12 小時累積降水量，而 12UTC 則是指觀測當日 00UTC 至 12UTC 間的 12 小時累積降水量。模式預測的累積降水時段和觀測是相同的，但是累積終止（模式輸出）的時間分別是 12、24、36 和 48 小時的預報。不論對於觀測或模式資料，都是取圖 1 網格內所有點的算術平均，來代表該區的降水情況。另外，分別校驗 00UTC 和 12UTC 可以判別模式的預測誤差是否在日夜有所不同。嚴格而論，算術平均並不是一個理想的校驗降水方式。特別是在測站較密的區域，若有相當局地性的降水，可能會導致對於觀測降水的高估。同樣地，若模式的格點密度遠高於測站，也會導致類似的觀測資料與模式產品在平均意義上的不一致性。如何避免這些不一致性，是我們未來的努力目標。

### 三、年平均降水預報得分之分析

首先我們以每 12 小時累積降水量 0.1 公釐為判定標準，針對不同的預報時間和結果，分區計算 1999 年 1 月 1 日到 2000 年 12 月 31 日整段時間的校驗指標，探討各模式降水預報的整體表現。圖 2 至圖 5 依序為 BIAS、POD、HKS 與 ETS 兩年平均的 12 小時累積降水預報技術得分，圖中包括了模式 12、24、36 與 48 小時預報平均，每一預報時間包含 00UTC 和 12UTC 兩組降水。此外每一圖也同時標示了 GT79、GT120 和 CA 三模式的結果，以方便比較。

在 BIAS 方面（圖 2），三個模式當中 CA 的降水預測偏離係數平均偏高，顯示中央氣象局的有限區域模式所預測的降水次數遠比實際發生的降水次數多，同時也多於全球模式的預報降水次數，並隨著預報時間的增加，偏離程度愈大。對於 GT79 和 GT120 而言，A 區的預報特性是明顯不同於其他各區的，除了其 BIAS 於各預報時間皆偏高外，還有明顯的日夜振盪，也就是說該模式對夜間降水預報偏溼的情形比日間降水要明顯。目前我們對於造成模式降水預報特徵的日夜差異的原因尚不清楚。整體來說，在 B 至 I 區 GT79 於各預報時間的 BIAS 大致為 0.9 至 1.2，GT120 則是 0.9 至 1.3，特別是 36 小時以後的降水預報明顯變差。

若參考其他數值模式，McBride and Ebert (2000) 以 1997 年 9 月至 1998 年 8 月的資料，校驗並比較七個模式的 24 小時累積降水量預報，整年平均的結果顯示 JMA (Japanese Meteorological Agency, Global spectral model)、LAPS (Bureau of Meteorology, Australia, Nested

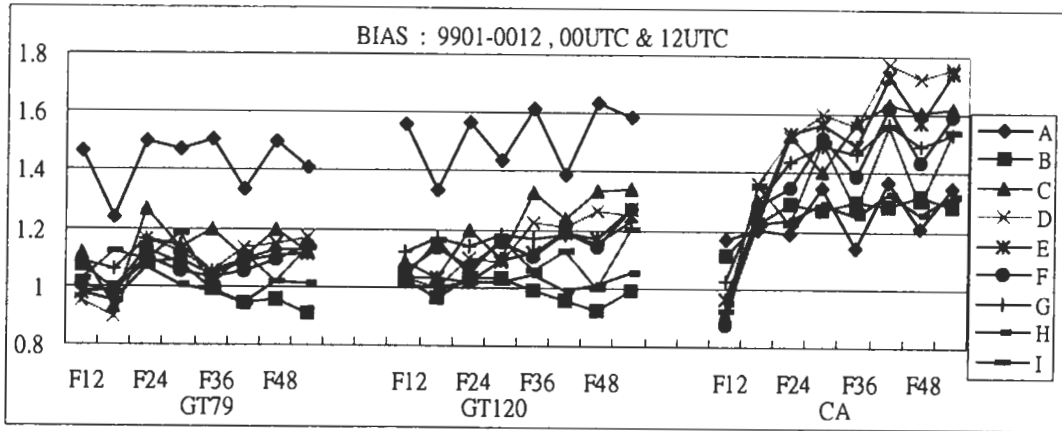


圖 2 GT79、GT120 與 CA 三模式 12 小時累積降水量預報的 BIAS 校驗，為 1999 年 1 月至 2000 年 12 月時間平均，並分別對 A 至 I 九個區域及 12、24、36 和 48 小時預報，每一個預報時間又依序分 00UTC 和 12UTC 個別計算。

regional gridpoint model)、AVN (National Centers for Environmental Prediction, United States, Global spectral model) 高估了 50% 的降水次數，即 BIAS 大於 1.5。GASP (Bureau of Meteorology, Australia, Global spectral model) 是高估了 100%，即 BIAS 平均大於 2.0。UKGC (Meteorological Office, United Kingdom, Global gridpoint model)、ECMWF (European Centre for Medium-Range Weather Forecasts, United Kingdom, Global spectral model)、DWD (Deutscher Wetterdienst, Germany, Global spectral model) 的 BIAS 則平均為 0.9，低估降水發生次數。雖然他們分析的時間與本報告並不相同，但是大體而言中央氣象局全球模式（不論 GT79 或 GT120）BIAS 校驗結果還不太差，平均來看又以 GT79 稍好，特別是 36 小時之後的預測模式之間的差別比較明顯。至於中央氣象局有限區域模式（CA）在每 12 小時累積降水量 0.1 公釐的 BIAS 校驗表現是明顯地比全球模式差。另外需要注意的一點，上述七個模式的校驗得分 (McBride and Ebert, 2000) 是針對

24 小時累積降水量和整個澳洲陸地，包括約六千個測站校驗而得的結果，這和本文中以三個模式 12 小時累積降水量，並且僅對東亞部份地區做校驗之情形不同。

在 POD 方面 (圖 3)，其值愈接近 1，表示降水預測愈準確。正如 BIAS 的校驗結果，A 區的表現明顯不同於其它各區。若對於 GT79 和 GT120 兩個模式，A 區的 POD 是均居於其它區域之冠，並有相當清楚的日夜振盪現象。相反地，在 CA 方面，A 區的平均 POD 反而出現最低值。由圖中顯示對於 B 至 I 區大都以 CA 的 POD 為最高，分布於 0.85 至 1.0 之間，唯獨 00UTC 的 12 小時預報 POD 偏低，降水預測較不準確，這個特性正好和 BIAS 的結果是完全不同的。由此可知，和 GT79、GT120 相比，CA 預測有降水的次數有偏多的傾向，導致其 POD 雖然接近 1，預測到降水發生的機率較高，但是實際上是因遠遠高估降水次數 (BIAS 偏高) 所產生的效果，並不是真正有比較好的預報技術。這種模式預報誤差特性，亦可由 POD 和 BIAS 的曲

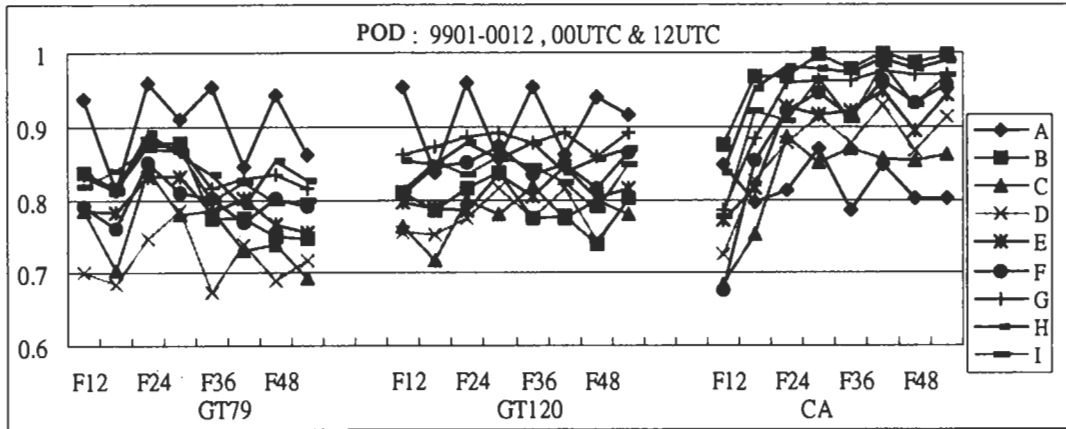


圖 3 同圖 2，但為 POD 校驗。

線隨預報時間增長的變化推敲得知。同時 A 區在 GT79 和 GT120 中其 POD 為最高分的特徵，同樣地亦是由於模式高估降水次數所致。若單單比較全球模式，不論是 00UTC 或 12UTC，區域 D、E、F 與 G 在各預報時間以 GT120 的 POD 為高，其餘各區或各預報時間則是互有優劣。

正如分析方法中所述，HKS 是同時評估有與沒有發生降水兩方面預報技術的校驗指標。圖 4 是各模式的 HKS 分布，三模式的 HKS 皆隨預報時間的增加而降低，表示預報時間愈長，

模式降水預報技術的能力愈差，這和一般數值預報模式的預報趨勢是類似的 (McBride and Ebert, 2000)。GT79 和 GT120 的 HKS 校驗結果是大致相同：A 與 B 兩區的技术得分明顯地比其他區域低 (GT120 略高於 GT79)，約 0.1 至 0.2，E 區為平均表現最好的區域，36 小時以前的預報得分大致為 0.5 以上。McBride and Ebert (2000) 計算七個數值預報模式 (如前述) 的 HKS，校驗區域為整個澳洲時，HKS 大都高於 0.5，若分為南、北區域，分數分布範圍則是 0.21 至 0.67。由此可見，同一模式的降水預測能力

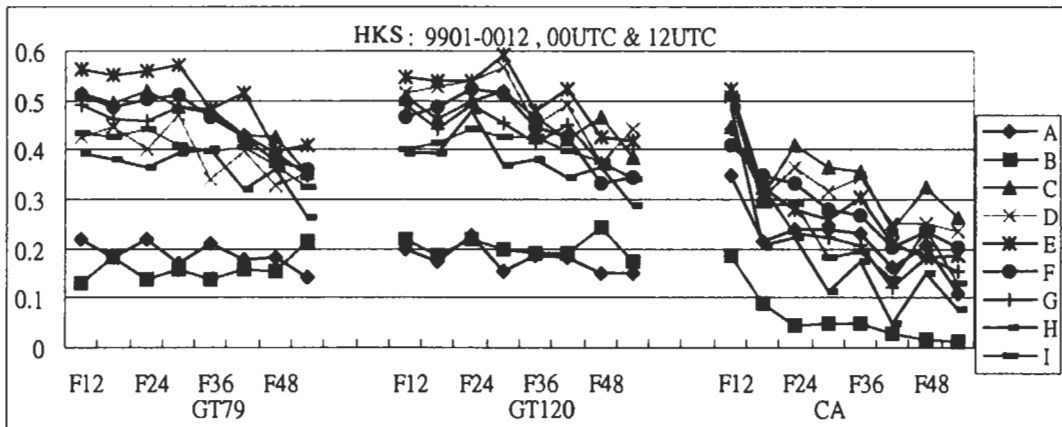


圖 4 同圖 2，但為 HKS 校驗。

可以有很大的區域性差異。對照圖 3 POD 分布則發現雖然 A 區的 HKS 偏低，但是 POD 卻為最高分，這表示 A 區在有發生降水時的預測準確率最好，但表現在 HKS 的低分應是由沒有降水現象時不佳的預報所導致的。

若仔細比較 GT79 和 GT120 對每 12 小時累積降水量 0.1 公釐之各預報時間的 HKS，我們發現在 B 和 D 兩區不論是哪一个預報時間，GT120 的技術得分都比較高，顯然 GT120 在這兩個地區降水預報技術的能力是優於 GT79 的，至於其他區域或預報時間，二模式的預報得分則是互有高下。再來看看 CA，很明顯地自 00UTC 的 12 小時預報以後，CA 的 HKS 隨著預報時間的增加而迅速地下降，整體而言中央氣象局區域模式的 HKS 是低於中央氣象局的全球模式。若參考圖 3 CA 的 POD 就可得知，CA 平均預測偏溼，也就是說預測會降水的次數偏高，導致表面上 POD 最接近 1，預報分數高，但是在沒有降水時預測一定是比較不準，才會造成 HKS 偏低，這種特徵在 B 區最為明顯。值得注意的是在 A 區除了 12UTC 的 48 小時預測外，其餘

預報時間的 HKS 都以 CA 分數最高，其 POD 又以 CA 的分數最低，顯示 CA 在 A 區沒有明顯偏溼的預報特性，同時預報得分也比全球模式高。也由此再次證明 HKS 同時配合 POD，更容易顯示模式的預報特性。

圖 5 是三個模式對每 12 小時累積降水量 0.1 公釐的 ETS 得分分布，顯示出的模式特性和前面圖 4 HKS 所得的結論相同。GT79 和 GT120 的 ETS 範圍約在 0.08 至 0.42，而在 McBride and Ebert (2000) 中提出的 ETS 約在 0.15 至 0.5，表示本局全球模式降水預測能力尚可，但不可忽略的是圖 5 中的得分有很大的地區性差異。例如，區域 A 與 B 的分數就遠低於 E 與 F 地區，不過預報時間愈長 E 與 F 地區的得分隨之降低，區域性差異因而減小。從以上校驗結果得知，若不區分累積降水多寡的類別，從兩年的平均結果看來，三個模式當中區域模式因預報有降水的次數過多導致結果最差，GT79 和 GT120 的預報技術相當，但在東南亞的預報結果比東北亞明顯較差。在東南亞的 A 與 B 二區，GT120 的預報結果略優於 GT79，顯示模式的解析度提

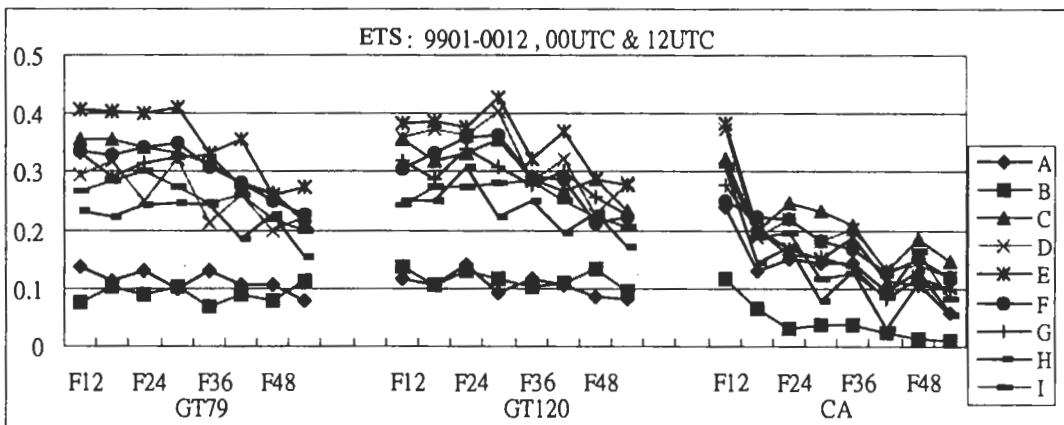


圖 5 同圖 2，但為 ETS 校驗。



高對於改進緯度較低區域的降水預報似有幫助。然而，不論是 GT120 或是區域模式的 60km 解析度，均不能掌握中小尺度的天氣系統，這有可能是導致東南亞預報結果比東北亞明顯較差的一個重要原因。

#### 四、降水預報得分之季節變化

為要顯示模式的降水預測技術會隨著季節變化而改變，我們利用 1999 和 2000 兩年資料，仍以每 12 小時累積降水量 0.1 公釐為判定標準，探討三組模式 HKS 得分的季節性差異。我們先計算了月平均 HKS，發現各模式的預報技術的確隨著不同月份而起伏變化（圖略）。然後我們根據月平均 HKS 的特徵，自一月起，每三個月分為一組，分別得到季平均 HKS 分布。

圖 6 是 GT79、GT120 與 CA 所得的 HKS，先做四組預報（12、24、36 和 48 小時）的平均之後，再取季平均的結果，時間是自 1999 年 1 月至 2000 年 12 月，每三個月為一組，每組又分為 00UTC 和 12UTC 兩個時間。結果顯示三組降水預報的 HKS 得分有明顯的季節性變化，皆以 JAS（七至九月）平均分數最低。GT79 和 GT120 除了 A 與 B 兩區外，其餘各區冬季的平均預報得分表現不錯。各區的得分差異性相當大，在夏季特別明顯，但是 CA 的區域性差異較小，唯平均 HKS 得分偏低，其中又以 JAS 平均分數偏低，JAS 的 12UTC 則最低。由此可看出模式在七至九月的降水預測為整年中表現最不好的季節，CA 各區的得分大都在 0.15 以下，分數偏低。此外，GT79 和 GT120 各季的 HKS 幾乎都是以 A 與 B 兩區為最低，整年平均後也是這兩區的分數最低，這個結果和圖 4 是相同

的。對於台灣而言，降水預測可說是夏秋兩季預報作業的重點工作，但是由這一部份的校驗結果看來，HKS 分數不盡理想，是值得注意的問題，有必要針對夏秋季節來改善模式的降水預測。

如前一節所討論的，對於三組模式降水預測來說，基本上預測的準確度會隨著預報時間的增長而下降。圖 7 是計算 GT79、GT120 與 CA 個別於四個預報時間的 HKS 標準差，同樣地為分區季平均分布。圖中很清楚地發現 GT79 和 GT120 的 HKS 標準差比較小，平均在 0.3 以下，而 CA 各區的 HKS 標準差值偏大，分布較不平均，同時又以冬、春季的預報差異最明顯。由此可知，中央氣象局全球預報模式在各個預報時間的降水預測較為相似，而有限區域模式除了各預報時間的預測差異性偏大之外，各區的預報分數差別也偏高，降水預測準確度的穩定性是比較低的。這些結果和我們前面討論過的年平均特性（圖 4）是一致的，也就是說除了 12 小時降水預報的 HKS 分數相對地較高外，隨預報時間的增加，CA 的 HKS 是呈現急劇地下降。

#### 五、各種累積雨量之預報得分校驗

前面兩節討論的僅是針對模式預測有無降水這個部份所做的校驗，即判斷閾值（threshold）為每 12 小時累積降水量 0.1 公釐。然而台灣一年四季各有截然不同的降水型態，降水量的多寡直接影響到人民生活作息，因此近年來中央氣象局除了做降雨機率預報外，也因應特殊天氣情況，額外加發豪雨特報（日雨量大於 130 公釐）或大雨特報（日雨量小於 130 公釐但大

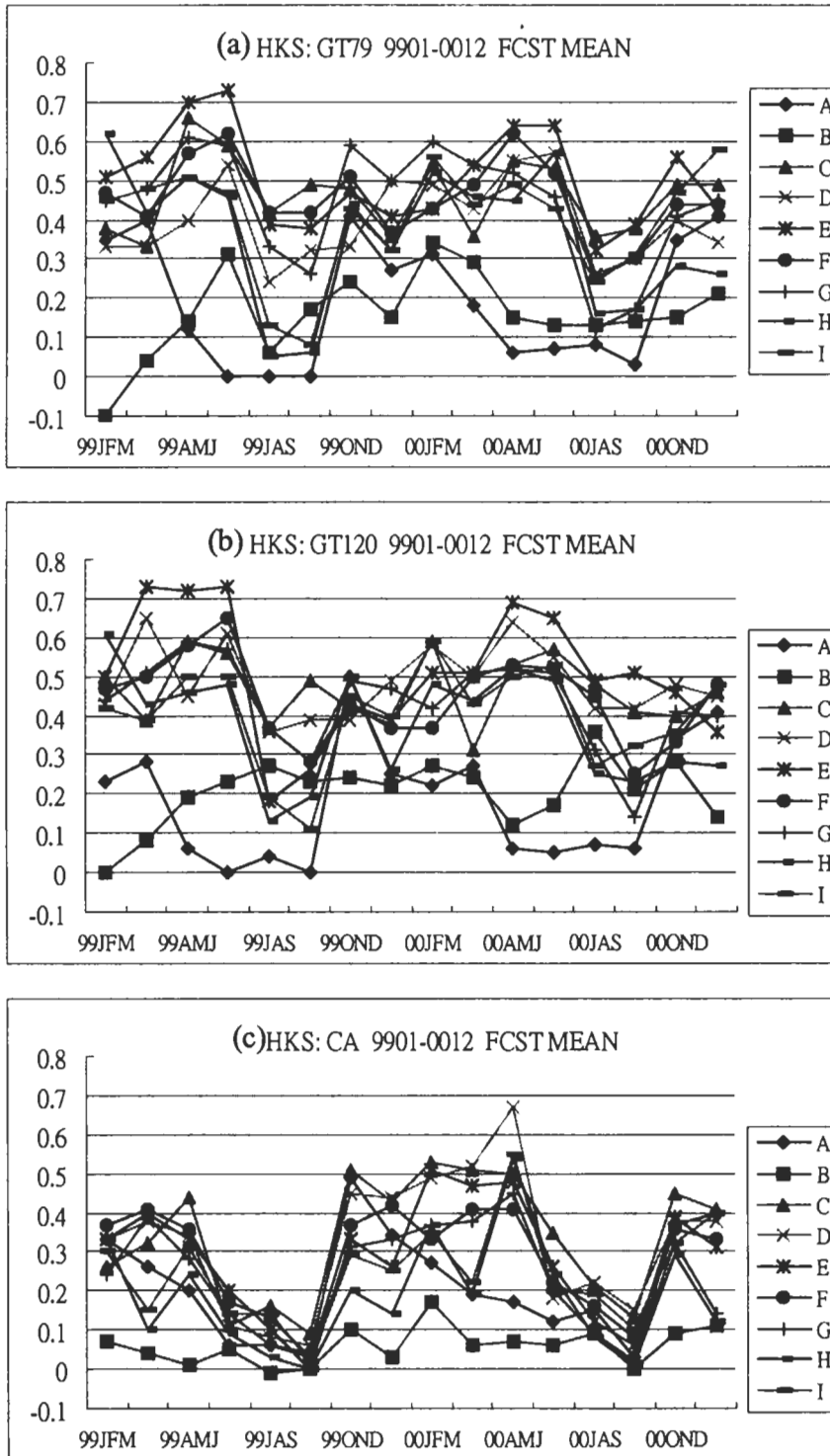


圖 6 (a) GT79 (b) GT120 (c) CA 三模式 12 小時累積降水量預報的季平均 HKS 校驗，結果為四個預報時間（12、24、36 和 48 小時預報）的平均。每一個季平均又分 00UTC 和 12UTC 個別計算。

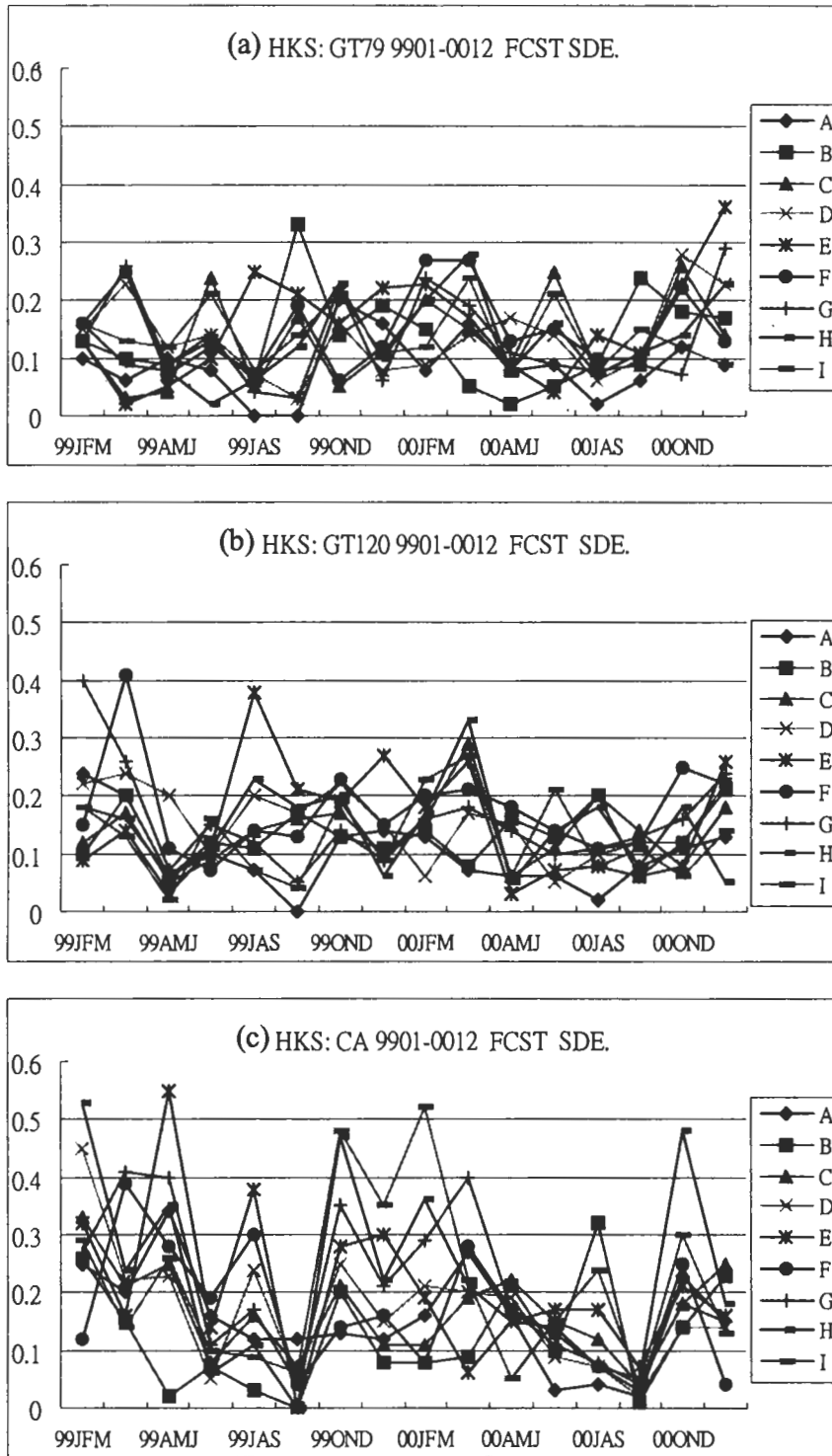


圖7 同圖6，但為HKS標準差。

於 50 公釐) 等相關資訊, 提供民眾參考。有鑑於此, 我們針對不同累積降水量的預報, 進行 POD 和 HKS 方面的校驗, 將判斷標準分為七類, 分別為 12 小時累積降水量 0.1 公釐、1 公釐、2 公釐、5 公釐、10 公釐、20 公釐與 50 公釐等, 校驗的區域仍是圖 1 所示 A 至 I 九區。

1999 年 1 月至 2000 年 12 月整段時間平均的校驗結果顯示 A 與 B 兩區具有相同的預報特徵, C 至 I 區又呈現另外一種分布, 因此特別挑選 B 區 12UTC 和 D 區 12UTC 為代表, 說明並比較各模式降水預報的區域特性。圖 8 為 GT79、GT120 和 CA 於 B 區 12UTC 的 HKS 校驗, 圖中四條曲線分別為 12、24、36 和 48 小時預報。圖 9 則和圖 8 類似, 但為 POD 方面的分析。由圖 8 比較可知, 三模式預報得分在此區是偏低的, 皆在 0.4 以下, 特別是在預報 12 小時累積降水量 5 公釐以下的部份, 顯示了中央氣象局模式對於極小雨預報得分偏低的共同特性, 但是其中又以 CA 最為明顯。再參考圖 9 POD 分布, 基本上三組模式預報曲線皆隨累積雨量的增加而下降, 表示降水預報準確度也降低了, 這個結果是和 McBride and Ebert (2000) 的分析相同的, 為一般數值預報模式的普遍特性。另外, 三組模式預報各自在判斷標準為 0.1 公釐的部份, 其 HKS 是比 1 公釐或 5 公釐部份要低, 但是 POD 反而是各部份中最高分的, 表示模式在極小雨部份預報具有偏溼的特性, 這種特性也以 CA 特別明顯, 其 HKS 為三模式中最低分小於 0.1 以下, 其 POD 卻為最高分, 接近 1.0。對於不同的預報時間比較, 以 GT79 的四個預報結果差異最大, 平均是以 24 小時預報得分最高。相較於 GT79, GT120 的四個預報結果的差別很小, 顯示該模式的預報穩定性較高。CA 則

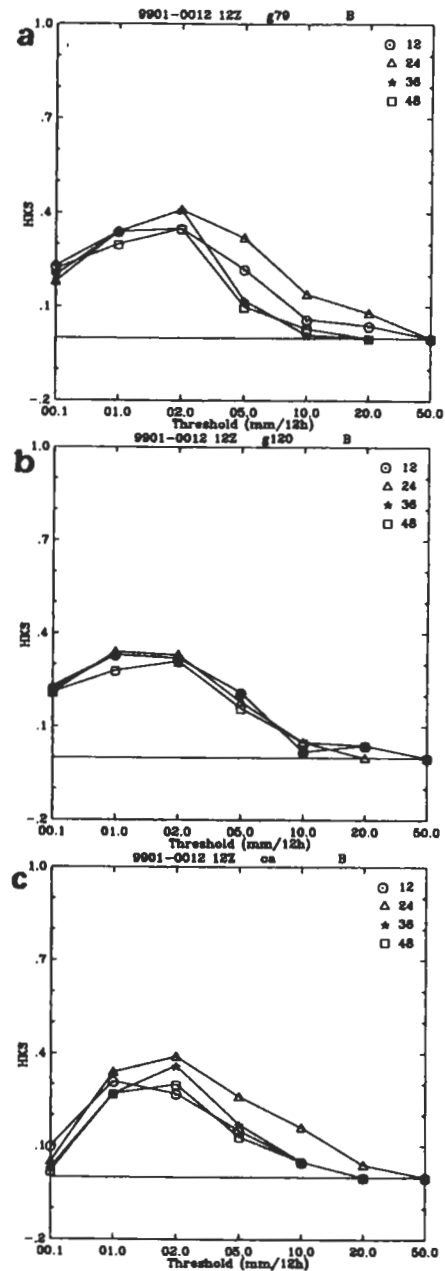


圖 8 B 區各類 12 小時累積降水量的 HKS, 為 1999 年 1 月至 2000 年 12 月 12UTC 平均, 圖中四條曲線各代表 12、24、36 和 48 小時模式預報。(a) GT79, (b) GT120, (c) CA。

是除了 0.1 公釐部份外, 其餘都是以 24 小時預報結果為最好。

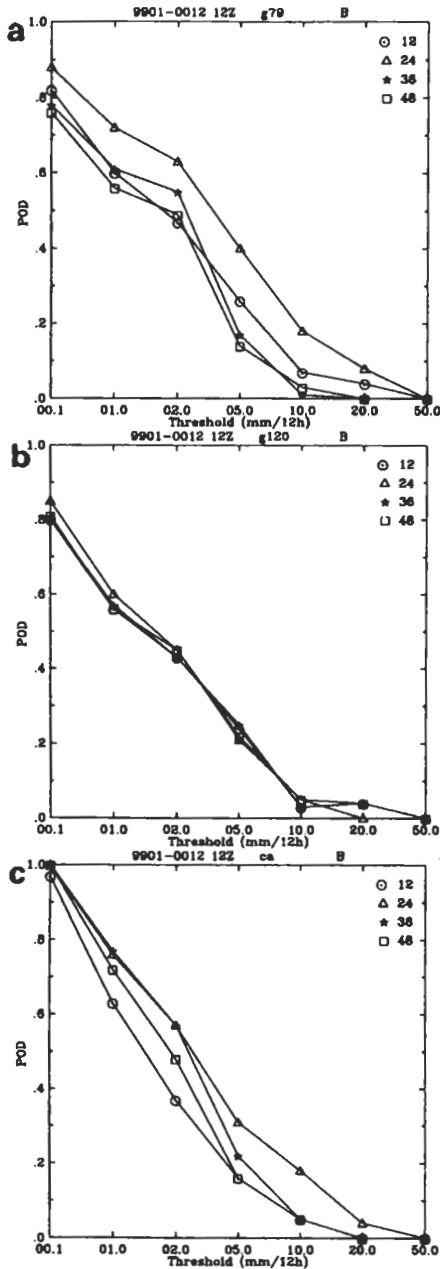


圖 9 同圖 8，但為 POD。

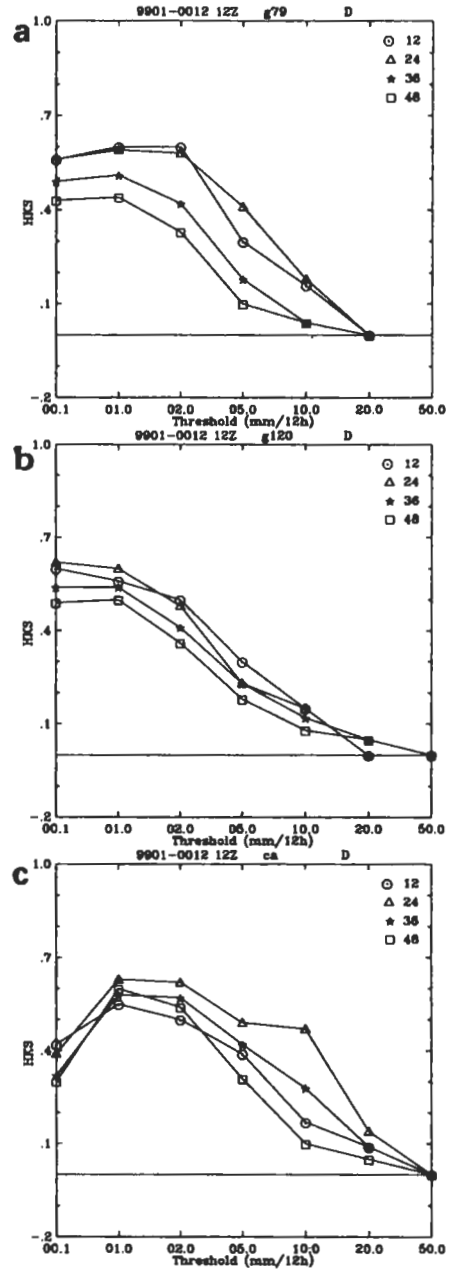


圖 10 同圖 8，但為 D 區。

圖 10 和圖 11 分別與圖 8 和圖 9 相同的，只是分析的區域為 D 區。由圖 10 和圖 11 可知，GT79 和 GT120 都是隨預報累積雨量的增加其 HKS 和 POD 逐漸降低。和 A 與 B 區特徵相似

的是 GT79 也比 GT120 於各個預報時間的預測差異性大。此外，二模式 12 與 24 小時預報得分是比 36 與 48 小時的表現好。而 CA 除了在 0.1 公釐部份外，其 HKS 和 POD 都是逐步下降的。

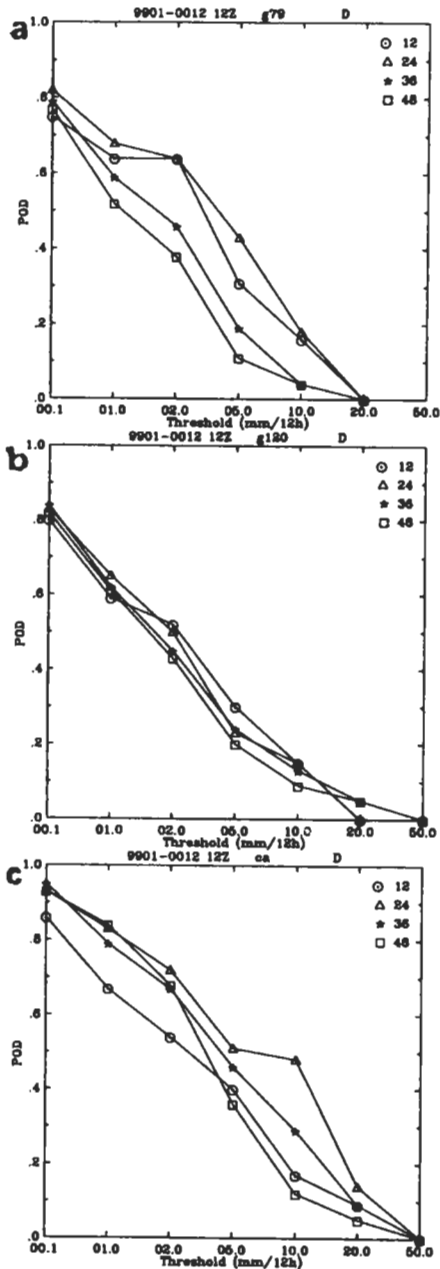


圖 11 同圖 10，但為 POD。

同樣地 CA 在極小雨部份 (0.1 公釐) 仍有偏溼的預報特性，這是 GT79 和 GT120 在 C 至 I 區中所沒有的預報特徵。對於累積雨量超過 2.0 公釐的部份，CA 的得分則是比 GT79 和 GT120 高，並以 24 小時預報結果為最好。

綜合上述結果，中央氣象局數值預報模式不論是大雨或小雨的預報，C 至 I 區的降水預報技術得分都遠比 A 與 B 兩區高，也就是說對於日本與韓國地區的降水預報能力是明顯地比南海與台灣地區準確。就中央氣象局全球模式來說，當預報極小雨時，對於南海與台灣地區會有預報偏溼的傾向。而中央氣象局的有限區域模式在預報極小雨時，則於各區都存在著預報偏溼的特徵。大體而言，各模式都是以 24 小時的預報最為準確。此外，除了在極小雨 (0.1 公釐) 部份，其餘累積降水量的預報得分普遍以 CA 最高，日本與韓國地區尤其明顯。這些結果顯示出分類校驗降水的必要性。以區域模式為例，雖然該模式有高估小雨的系統性誤差，但是在預報上比較關切的中大雨預測方面它的表現可以超越全球模式，達到對區域模式的需求功能。

我們也特別針對 CA 在極小雨預報時，各區有降水預報偏溼的問題，做進一步的分析。方式是當觀測無雨或微量 (12 小時累積降水量小於 0.1 公釐) 時，則計算預報降水量和觀測的差值，再做整段分析時間的平均，同時又分區 (A ~ I)、分時段 (00UTC 和 12UTC) 比較。結果顯示並沒有在任一個特別的區域或時間出現明顯地預報偏差，如同前述平均預報的累積降水量是偏溼的：12 小時預報差值為 0.68 公釐、24 小時預報差值為 0.80 公釐、36 小時預報差值為 0.9 公釐、及 48 小時預報差值為 0.99 公釐。四個預報時間的降水預測偏差值都相差無幾。

另外，又分析了月平均各類 12 小時累積降水量的預報偏差 (圖略)，基本上三模式是有共同的預報特性：小雨部份 (12 小時累積降水量小於 2 公釐) 預報高估，偏差值大；大雨部份 (12 小時累積降水量大於 20 公釐) 預報低估，

偏差值較小。小雨的高估特徵是有季節性的變化，在溼季或夏季比較明顯，而大雨的低估特性只有些許的季節性變化。若是以月平均降水總量校驗，由於夏季或溼季的小雨高估、大雨低估現象明顯，雨量高低估相互抵銷平衡之下，反而是夏季或溼季的整體預報結果比較好。冬季小雨仍有高估情形，但低估的卻是中雨部份，因此整體預報呈現的是高估居多，並比夏季差。雖然以月平均總量校驗是以夏、秋季預報表現比較好，但這是由於夏季或溼季的小雨高估、大雨低估相互抵銷所造成的假象；實際上，我們由前面章節的討論可知，就降水預報技術的能力（HKS）來看，夏、秋季的得分比較低，顯示模式降水預報準確度較低。

## 六、結 論

本文以 1999 至 2000 年這段時間，根據列聯表計算 POD、BIAS、ETS 與 HKS 四種指標，校驗中央氣象局全球作業模式和有限區域作業模式的降水預報產品，評估比較各模式的降水預報技術，進而提供模式研發和預報人員參考。校驗區域如圖 1 所示分為 9 個區域，分別以 A 至 I 為代號。

在年平均方面，若不區分降水類別僅以每 12 小時累積降水量 0.1 公釐為判定標準時，除了 A 區（南海地區）之外，全球模式（GT79 和 GT120）的 BIAS 得分尚佳，平均來看又以 GT79 稍好，特別是 36 小時之後的預報。有限區域模式（CA）的表現則是明顯地比全球模式差。POD 方面，卻是以 CA 的得分最高，但是並不代表 CA 就真正有比較好的預報技術；這是由於 CA 有預報降水次數偏多的傾向，導致其 POD 雖然接近 1，預測到降水發生的機率較高，實際上卻是因

遠遠高估降水次數（BIAS 偏高）所產生的效果。我們也可由 HKS 和 ETS 的分數看出端倪，這兩項的校驗結果都是以 CA 的分數最低，顯示 CA 有預報偏溼的特性。GT79 和 GT120 的預報技術和其他作業中心數值預報模式大致相當。

另外，模式的降水預報能力有很大的區域性差異。譬如 A 區，在 GT79 和 GT120 中其 BIAS 與 POD 於各預報時間皆偏高，並有清楚的日夜差異，在 CA 則有相反的結果。而 HKS 與 ETS 卻又以 CA 分數較高，顯示了 CA 在 A 區並沒有明顯的偏溼預報特性，同時預報能力也比全球模式高。整體而言，三模式在 B 區（台灣地區）的預報得分大致比其他區域差，全球模式對 A 區的降水預報能力也不佳。基本上，除了 A、B 二區之外各模式的預報得分是會隨著預報時間的增加而降低，區域性的差異則是變小。

在季節性差異方面，三組模式降水預報的 HKS 皆以 JAS（七至九月）平均分數最低，表示中央氣象局數值模式在七至九月的降水預報為整年中表現最不好的月份，尤其是 CA 各區的得分大都 0.15 以下，分數偏低。此外，GT79 和 GT120 各季的 HKS 幾乎是以 A 與 B 兩區為最低，因此整年平均後也是以這兩區的分數最低。對於 B 區（即台灣地區），HKS 分數不盡理想，這是值得我們注意的問題，必須針對夏秋季節來改善模式的降水預報。

針對不同類別的降水累積量校驗，A 與 B 兩區具有相同的預報特徵，C 至 I 區又是另外一種特性。整體比較發現不論是大雨或小雨的預報，中央氣象局數值預報模式在 C 至 I 區的降水預報技術得分都遠比 A 與 B 兩區高，也就是說對於日本與韓國地區的降水預報能力是明顯地比南海與台灣地區準確。就全球模式來說，

當預報極小雨時，對於南海與台灣地區會有預報偏溼的傾向。而有限區域模式在預報極小雨時，則於各區都存在著預報偏溼的特徵。對於不同的預報時間比較，以 GT120 的四個預報結果的差別最小，顯示模式的預報準確率較為穩定，但是大致上各模式都是以其 24 小時的預報最為準確，為何以致如此則是一個需要繼續追蹤的問題。此外，除了在極小雨（0.1 公釐）部份，其餘累積降水量的預報得分普遍是以 CA 最高，日本與韓國地區尤其明顯。由於中雨和大雨在預報上的重要性平均而言較小雨要高，因此區域模式雖然因預報小雨的次數過多導致降水預報的平均準確率稍差，但在中大雨方面的校驗結果證明其預報還是甚具參考價值的。

若以月平均各類 12 小時累積降水量總量校驗，三模式是有共同的預報特性：小雨預報高估，偏差值大；大雨預報低估，偏差值較小。由於夏季或溼季的小雨高估、大雨低估現象明顯，雨量高低估相互抵銷平衡之下，反而是夏季或溼季的整體預報結果比較好。冬季小雨仍有高估情形，但大雨出現機會少，低估的卻是中雨部份，因此整體預報呈現的是高估居多，並比夏季差。降水總量和預報技術得分的校驗結果是不同的，即使以月平均總量校驗是夏、秋季預報表現比較好，但這是由於夏季或溼季的小雨高估、大雨低估相互抵銷所造成的假象。實際上，就降水預報技術的能力（HKS）來看，夏、秋季的得分是比較低分的，模式降水預報準確度降低。

在校驗的過程中，對於觀測資料是採取直接應用原始接收資料，不加以任何計算處理，因此能夠用來校驗的測站數量有限。日後若能加強處理測站資料，例如統一計算各站 6 小時或 12 小時的累積降水量，就可擴大或增加校驗

區域，將能更靈活、客觀地評估模式的降水預報能力。另外，從本研究當中可以看出模式在緯度較低的區域降水預測技術比緯度較高的區域差，這也是普遍存在於其他作業中心數值預報模式的問題。但是因校驗的區域有限，本研究的結果並不能肯定熱帶降雨預測的區域性差異。未來準備利用中央氣象局接收的 GMS 衛星資料以及其他的全球衛星資料來加強診斷熱帶降水誤差。最後，本研究中所使用的模式格點資料均為插到  $2.5^{\circ} \times 2.5^{\circ}$  網格點上的資料，未來應該要用更忠實於模式的細格點進行校驗，較為合理。本研究藉由比較不同模式降水預報的結果來呈現降水預報校驗工作的複雜性與重要性，顯示唯有長期細心校驗模式預報結果，才有可能適當評量模式的預報能力，作為日後改進預報之參考。

## 誌 謝

本文是在中央氣象局「氣象業務全面電腦化第三期計畫」下完成。特別感謝中央氣象局氣象資訊中心秦照萍小姐在觀測降水資料檢定與使用方面的協助，以及潘華陸博士和葉天降博士的建設性討論。

## 參考文獻

- 陳泰然，謝信良，陳來發，陳清得，1991：台灣地區現階段豪（大）雨預報能力。大氣科學，19，177-188。
- 陳建河，劉其聖，馮欽賜，汪鳳如，蕭志惠，1994：中央氣象局第二代全球波譜模式之介紹。天氣分析與預報研討會論文彙編，中央氣象局，1-10。
- ，劉其聖，1997：中央氣象局全球模式之更新---T120 作業模式。天氣分析與預報研討會論文彙編，中央氣象局，279 頁。



葉天降，林淑卿，李尚武，陳雯美，張偉正，彭順台，1994：中央氣象局第二代有限區域預報系統之建置與現況。天氣分析與預報研討會論文集編，中央氣象局，11-20。

Black, T. L., 1994: The new NMC mesoscale Eta model: Description and forecast examples. *Weather and Forecasting*, **9**, 265-278.

Ghelli, A. and F. Lalaurette, 2000: Verifying precipitation forecasts using upscaled observations. *ECMWF Newsletter Number*

87 – Spring 2000. 9-17.

McBride, J. L., and E. E. Ebert, 2000: Verification of Quantitative Precipitation Forecasts from Operational Numerical Weather Prediction Models over Australia. *Wea. Forecasting*, **7**, 103-121.

Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences. An Introduction.* Academic Press, 477 pp.

## **Rainfall Evaluation of the CWB Forecast Model : Year 1999 and 2000**

**Chi Pan      Mong-Ming Lu**

**Central Weather Bureau**

**( Manuscript received 8 May 2001 ; in final form 29 August 2001 )**

### **ABSTRACT**

The 12-, 24-, 36- and 48-hour forecasted 12-hour accumulated precipitation during 1999 and 2000, produced by the Global Forecast System using T79 and T120 (GFST79 and GFST120) and the Limited Area Forecast System (LAFS) at CWB, are verified against the synoptic station data. Nine areas in Southeast and Northeast Asia are selected for the evaluation. The results are summarized in forms of annual, seasonal and categorized rainfall averages. In terms of annual average, we find the skill of LAFS is the lowest among three models. The LAFS has clear wet bias. The forecast skill in lower latitudes is worse than that in high latitudes. The skill in Taiwan surrounding area is the lowest among 9 areas. In addition, we find that the prediction skill oscillates diurnally. The forecast skill of the night precipitation is worse than the skill of the day. Considering the results predicted by the same model, the skill differences in regional contrasts and overall scores decrease as the forecast lead-time increases. In terms of the seasonal average, we find that the skill during summer season from July to September is the lowest among all seasons.

Regardless of the categories of accumulated rainfall amount, the prediction in Southeast Asia is worse than that in Northeast Asia. Both global models have wet bias in the small rain ( $< 0.1\text{mm}/12\text{hr}$ ) category in Southeast Asia, while the LAFS has wet bias in all categories. The performance of the LAFS is superior to the global models, particularly in Northeast Asia, except in the small rain category. The skill differences among the predictions with different forecast lead-time are smaller in GFST120 compared with other models. The 24-hour forecast is in average better than the forecasts with other lead-times. On the basis of monthly average, all models have the tendency of overestimate the small rain and underestimate the large rain ( $> 50\text{mm}/12\text{hr}$ ). We find that particular caution needs to be exercised if some months in the wet season showing better scores than in the dry season because the superiority may be a false impression due to the compensation of different types of errors in different rainfall categories. Generally speaking, the rainfall forecast skills of CWB global models are comparable to the global models at other operational centers.

**Key Words: Rainfall forecast, Model evaluation, Prediction skill**