

# 數值天氣預測之校驗方法綜評

方 力 僑

陳 慧 貞

中央氣象局  
資訊測站主任

中央氣象局  
約聘研究員

（收稿日期：78.1.18；定稿日期：78.3.7）

## 摘要

由於數值預報是利用有限的資料、不完全的物理及數值方法求得預報結果，故誤差必然存在，如何有效應用校驗方法修正其誤差，並且找出誤差來源是繼數值天氣預報系統建立後另一亟需且有待探討的課題。本文基於此需要，乃歸納整理各種有關的校驗方法，探討其特性及缺失，並且對於中央氣象局建立數值天氣預報校驗系統提供一些建議。

關鍵詞：校驗系統。

## 一、前言

就氣象上而言，校驗的終極目的在於增進天氣預報能力。早期，校驗工作是針對預報員所做的主觀天氣預報，透過校驗我們可以評估預報員的預報能力。然而，傳統的校驗方法乃是利用一些統計方法來求得某種指數（index）或分數（score）以表示預報得好或不好，有時預報員爲得較高的分數，竟遷就校驗方法的特性而非根據本意來做預報。例如：校驗方法採用均方根誤差（root-mean-square-error，RMSE），預報夜晚之最低溫，預報員認爲若當晚有雲則最低溫爲 $20^{\circ}\text{C}$ ，無雲則爲 $10^{\circ}\text{C}$ ，但爲了得較好的分數，而以平均值 $15^{\circ}\text{C}$ 來預報，儘管他確信絕不可能是 $15^{\circ}\text{C}$ 。像這種因統計方法的特性而有投機的機會，遂使得“最好”（分數最高）的預報，不見得是“最有價值”的預報（陳，1984）。因此在早期，校驗工作是否有存在的必要頗受當時人的爭議。

而今，天氣預報已步入數值天氣預報的時代。數值天氣預報是利用有限的資料、不完全的物理和數值方法所計算出來的，誤差必然存在，且其誤差之大小與性質又隨時空、氣流情況及模式特性而有

差異。要充分利用數值天氣預報產品，就必須先掌握模式的特性，以便修正其誤差，所以校驗、評估及相關之分析工作就變得十分重要了。何況，只要模式發展者不刻意設計，模式本身是不會“投機”的，故早期受爭議的癥結已不存在，因此校驗在數值天氣預報實施之後，即成爲一項不容置疑、忽視且刻不容緩的重要工作。

本文將於第二節詳細說明校驗的目的何在；對於數值預報模式產品的誤差來源將列於本文第三節；各種相關的校驗方法則於第四節中做一詳盡的敍述，此外更進一步探討各種方法的特性與缺失；最後並於末節做一總結與建議。

## 二、檢驗的目的

早期氣象上的校驗工作是針對預報員所做的天氣預報。其目的一方面在考核預報員的能力，檢定預報員的專長，做爲主管分配工作（譬如從事日常預報、特殊天氣預報或理論研究等）的參考依據，另一方面則可供預報員本身修正自己的預報偏差。然而由於前節所述之“投機”問題，校驗結果令人懷疑，使得校驗功能難以發揮。

近年來，許多國家實施數值天氣預報之後，校

驗對象已由對“人”轉為對“模式產品”，投機機會不再存在，且由於模式本身具有多種限制，使得預報誤差必然存在，因此校驗躍升為不可或缺的工作項目，甚至可說是數值天氣預報的一部分，當然，校驗的重要性也受到完全的肯定了。

至於對數值預報模式做校驗的目的何在，我們可以由兩方面來考慮：

- (一)計算預報模式的誤差傾向（例如系統出海時常有強度增強過快的情形，或是某系統經常移速太慢等），當預報員在使用模式產品做預報時，可憑之以為修正的參考依據。
- (二)分析誤差的特性，追溯誤差的來源（例如低壓系統發展太快，可能是積雲處理不恰當，造成降水太多，潛熱釋放太大；又如溫度誤差具有日夜變化，則可能是輻射參數化處理不完整等等），使模式發展人員在改進模式時，有個概略的指標。

總而言之，校驗的終極目的，就是在促使天氣預報結果更準確、更完美。

### 三、誤差來源

在尚未討論各種校驗方法之前，我們先來探討為何模式產品必有誤差存在。Haltiner 和 Williams ( 1975 ) 指出不適當的資料、客觀分析與初始化過程的誤差、積分過程所產生的截斷誤差 ( truncation error ) 、不正確的物理參數化及摩擦效應等都可造成數值預報的誤差。

完整的數值天氣預報系統流程如圖一，事實上，幾乎每一個階段都有產生誤差的可能。若將模式產品經常重覆出現的誤差稱為系統性誤差，隨機出現者稱為非系統性誤差，我們可討論各階段出現誤差的原因及誤差屬性：

- (一)觀測：為非系統性之人為疏忽所造成的觀測隨機誤差。
- (二)資料接收與解碼：當通訊電路不穩定時，可能導致資料傳送錯誤，造成非系統性誤差。
- (三)資料檢定偵錯：此步驟之目的在管制資料品質，可消除部分前兩者所造成的誤差。然若理論應用得不恰當，亦可造成誤差。例如範圍檢定，若溫度大於  $x$  °C 則被剔除。當  $x$  值定義太低，則使得溫度場偏高，且可能影響溫度梯度。此種錯誤會重複發生，故可歸為系統性誤差。
- (四)資料客觀分析：客觀分析經水平內插過程比人為主觀分析圓滑，會將較淺天氣系統圓滑淺化甚而

抹殺之 ( Wash , 1984 )，屬系統性誤差。

(五)資料初始化：由於初始化是在調整質量場與風場，使之平衡，所以也會修勻資料 ( Rosmond , 1984 )，屬系統性誤差。

(六)數值預報模式：誤差來源有二，皆屬系統性誤差。

1. 模式物理參數化：次網格之物理過程（例如輻射過程、積雲對流之潛熱釋放及邊界層之摩擦效應等）模式無法直接處理，需參數化，故必有誤差產生。

2. 數值方法：將非線性偏微分方程化為計算機可解的差分方程，會有截斷誤差；此外，水平與垂直之解析度不足，非全球模式之側邊界條件及地形之模擬都會造成誤差。

第一項與第二項為非系統性誤差，不屬於校驗所感興趣的對象。至於檢定所造成的誤差，可經長期經驗之累積，修改檢定理論，使造成誤差的機會減至最低，故一般有關誤差之研究，亦不考慮此項。

### 四、檢驗方法

檢驗方法可分為兩大類，一為以統計方法計算出單一的指數或分數，來判定模式預報得“好”或是“不好”；另一則為對誤差做更深入的分析，以探討誤差的特性並追溯誤差的可能來源。

- (一)指數或分數

早期的校驗工作都是用這類方法，透過此類校驗方法，我們可以達到幾個目的：

I. 比較不同模式（或預報員）的預報能力。

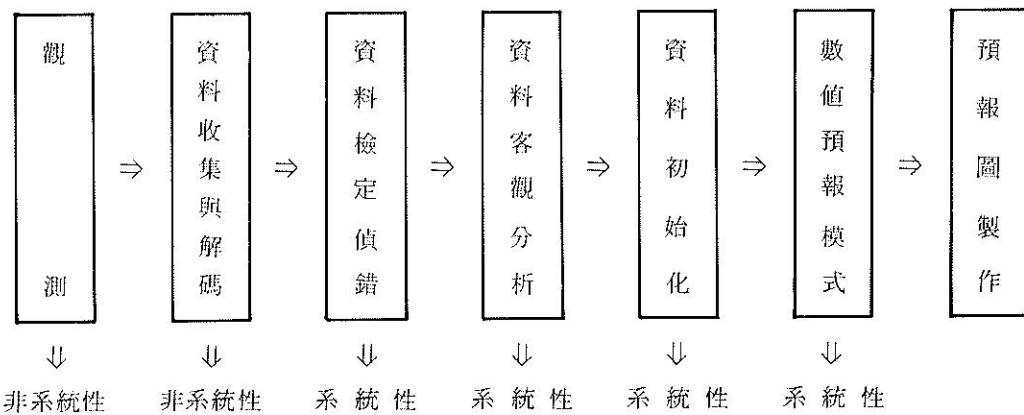
圖二為兩個物理參數化相異的模式與持續性預報之均方根誤差對預報時間的變化，由圖可比較兩組模式的預報能力。

II. 顯示模式預報準確度隨時間的變化。

由圖二可看出模式產品的均方根誤差隨預報時間的增長而增加，即表預報準確度隨預報時間而降低。

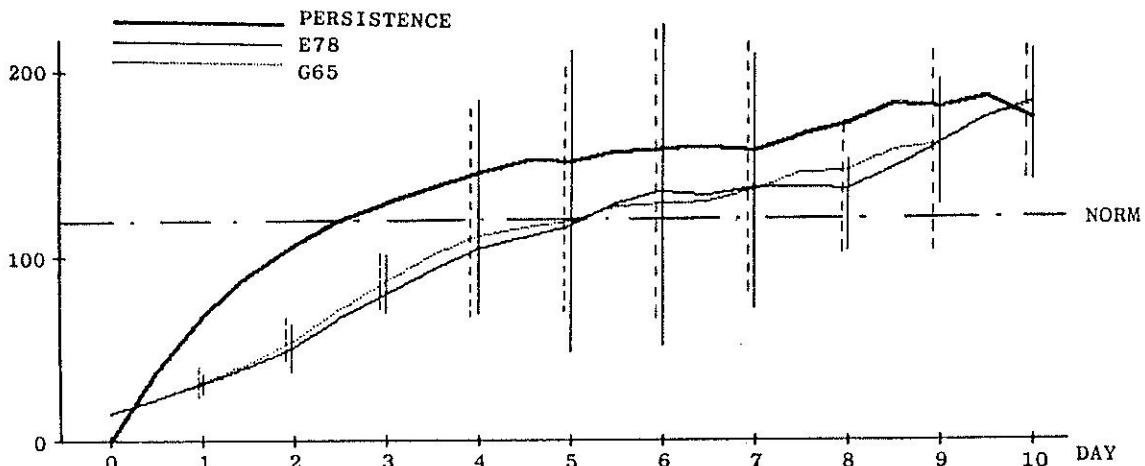
III. 顯示預報能力之逐年變化。

圖三為美國國家氣象中心 ( National Meteorological Center , NMC ) 預報 36 小時北美地區 500 mb 環流的技術得分、誤差及 S1 分數，由 1955 年到 1980 年的逐年變化，其間預報方法由主觀預報、正壓模式、準地轉模式、半球模式而至加入風場資料。由圖顯見每一階段的改變都提高了預



圖一 完整的數值天氣預報系統流程圖及其誤差性質。

Fig. 1 Flow chart of the numerical weather prediction system and its nature of error.



圖二 持續性預報（粗實線）與兩種不同模式（細實線與點線）預報歐洲區域高度場（經垂直平均）之均方根誤差隨預報時間的變化（Hollingsworth 等，1980）。

Fig. 2 Time evolution of vertical average RMSE of geopotential height (m) for European area for persistence and for two different models (Hollingsworth et al, 1980).

報能力。

### 三、了解模式預報能力隨季節的變化。

圖四為日本氣象廳區域模式預報24小時海平面氣壓與500 mb高度場和實際觀測值的相關係數（月平均）。由圖可知此模式預報於冬季結果較佳，夏季則較差，顯然此模式較能掌握冬季天氣系統之變化。

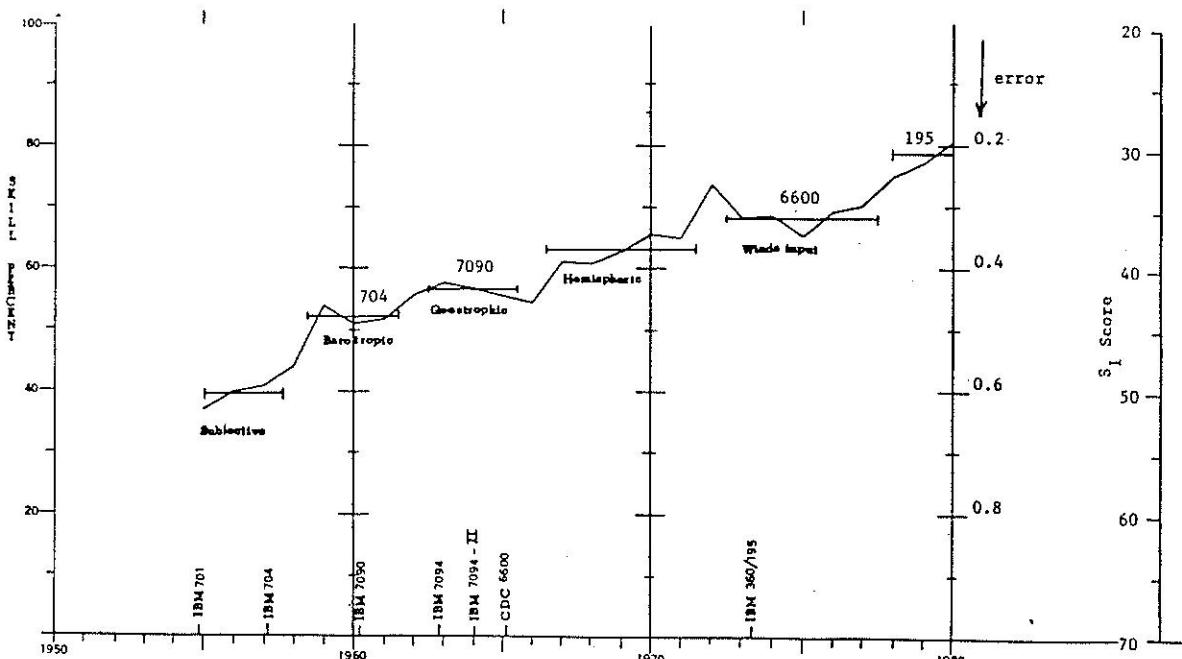
以下我們將詳細討論各種此類校驗方法的內

容、特性、適用範圍與缺點。

### 1.列聯表（contingency table）

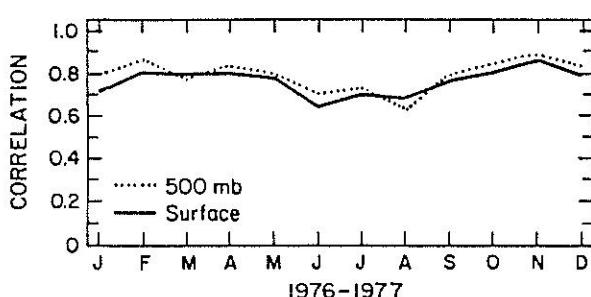
將預報結果和實際觀測所得製成列聯表表示之。此種表示方法雖不直接包括校驗意義，但經由此表可計算出許多校驗指數與分數。適用於範疇式（categorical）預報，即下雨不下雨，或大雨、中雨、小雨等之預報。表一即為列聯表的例子。

此法對於數值天氣預報系統之產品僅適用於降



圖三 1955 年至 1980 年美國 NMC 預報 36 小時北美地區 500 mb 環流之技術得分，平均誤差及 S1 分數 ( Rosmond, 1984 )。

Fig. 3 1955-1980 skill percent error and s1 score for the NMC 36h, 500mb circulation predictions over North America. (Rosmond, 1984).



圖四 日本區域模式24小時預報與觀測值之相關係數(月平均)。實線為海平面氣壓，虛線為500 mb高度場(Anthes, 1983)。

Fig. 4 Annual variation of the monthly mean correlation coefficient between the observed and LFM 24 h forecast (Anthes, 1983).

水一项。

2. 正確百分比 ( correct percentage , CP ) 、  
偏差率 ( bias , B ) 、預兆分數 ( threat score  
, TS ) 、前估 ( pre figuration , PF ) 、未命  
中率 ( no hit rate , NH ) 、後符 ( post  
agreement , PA ) 、錯誤警報率 ( false al-  
arm , FA ) 及技術得分 ( heiclke skill score  
, HSS )

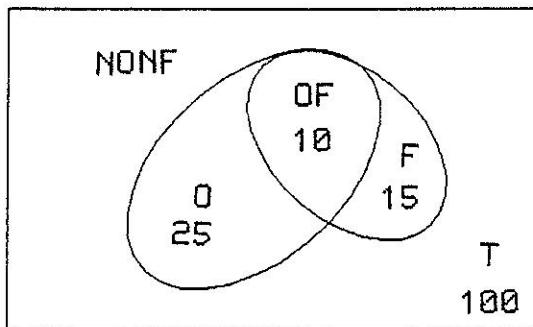
這幾種分數或指數都可由列聯式求得，故亦適用於範疇式預報，此外面積（以測站數或網格點數表示之）預報也可用其中某些方法來校驗。圖五為一面積預報之示意圖。

下面我們將以表一及圖五的例子來說明各種校驗方法。

(1) 正確百分比 ( CP ) =  $\frac{\text{預報正確次數 (面積)}}{\text{預報總次數 (總面積)}}$

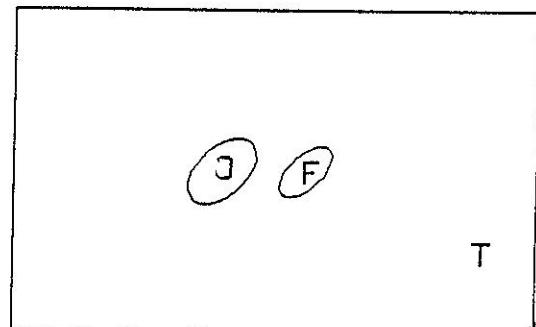
表一 範疇式預報——降水預報之列聯表  
Table 1 Contingency table for a categorical forecast - precipitation

預測		有	無	總數	觀測		預測	下雨	不下雨	總數
觀測	有	$N_1$	$N_2$	$N_1+N_2$	有	雨	8	4	12	
無	$N_3$	$N_4$	$N_3+N_4$	沒	有	雨	8	10	18	
總數	$N_1+N_3$	$N_2+N_4$	$N_1+N_2+N_3+N_4$	總	數		16	14	30	



圖五 面積預報示意圖。O 表觀測，F 表預報，OF 表預報且觀測到，NONF 表示沒有預報且沒有觀測到，T 表整個預報面積。

Fig. 5 Conceptual figure of an area forecast O is observed, F is forecasted, OF is observed and not forecasted, T is total area.



圖六 類似圖五，但觀測到之面積與預報到之面積皆遠小於預報總面積。

Fig. 6 List Fig. 5 except the area of observed and area of forecasted are all much less than total area.

表一：

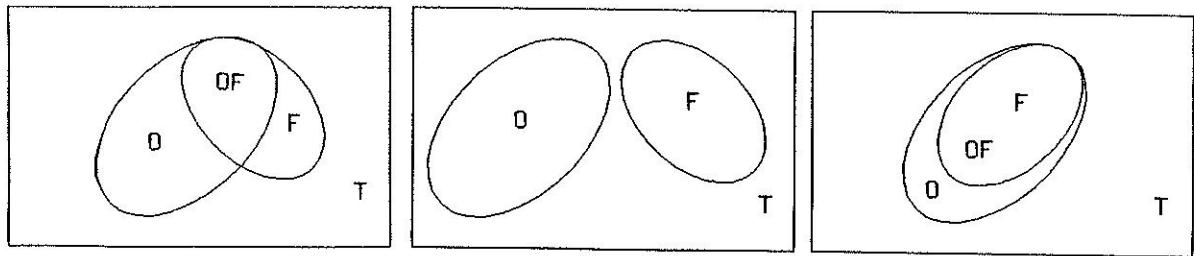
$$\begin{aligned} CP &= \frac{N_1 + N_4}{N_1 + N_2 + N_3 + N_4} \times 100\% \\ &= \frac{8 + 10}{30} \times 100\% = 60\% \end{aligned}$$

圖五：

$$\begin{aligned} CP &= \frac{OF + NONF}{T} \times 100\% \\ &= \frac{10 + 70}{100} \times 100\% = 80\% \end{aligned}$$

缺點：①範疇式預報：若某地天氣的持續率很大，則得分很高。然天氣預報重要的是希望有能力預測出天氣的轉變，此種方法無法有效的校驗出模式（或預報員）是否有預測天氣轉變的能力。

②面積預報：若預報總面積很大，而降水面積很小時，即使預報完全不正確，得分仍然很高，因為未曾預報且沒有觀測到的比例太大，如圖六。



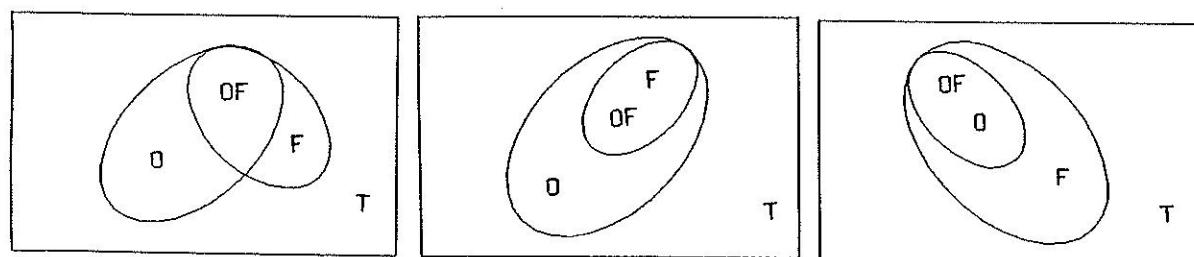
$O=25$   
 $F=15$   
 $OF=10$   
 $B=0.6$

$O=25$   
 $F=15$   
 $OF=0$   
 $B=0.6$

$O=25$   
 $F=15$   
 $OF=15$   
 $B=0.6$

圖七 降水面積預報之結果雖不同，但偏差率卻相等的例子。

Fig. 7 The example for area forecasts (precipitation) which have different forecast results but same bias.



$O=25$   
 $F=15$   
 $OF=10$   
 $TS=0.33$

$O=30$   
 $F=10$   
 $OF=10$   
 $TS=0.33$

$O=10$   
 $F=30$   
 $OF=10$   
 $TS=0.33$

圖八 降水面積預報之結果雖不同，但預兆分數卻相等的例子。

Fig. 8 The example for area forecasts (precipitation) which have different forecast results but same threat scores.

$$(2) \text{ 偏差率} (B) = \frac{\text{預報的次數(面積)}}{\text{觀測到的次數(面積)}}$$

表一：

$$B = \frac{N_1 + N_3}{N_1 + N_2} = \frac{8 + 8}{8 + 4} = 1.33$$

B 大於 1 表降水預測過度。

圖五：

$$B = \frac{F}{O} = \frac{15}{25} = 0.6$$

缺點：此法僅能概念性的指示出模式有預報過度或預報不足的傾向，無法提供其他更有效的訊息。圖七為幾種不同降水面積預報，偏差率

卻相同的例子，由圖可知單憑偏差率，對預報員似乎沒有太大的幫助。

$$(3) \text{ 預兆分數} (TS) = \frac{\text{預報正確次數(面積)}}{\text{預報或觀測到的次數(面積)}}$$

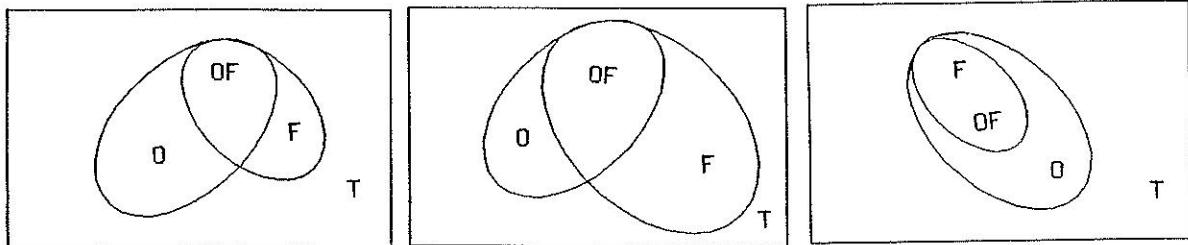
表一：

$$TS = \frac{N_1}{N_1 + N_2 + N_3} = \frac{8}{8 + 8 + 4} = 0.4$$

圖五：

$$TS = \frac{OF}{O + F - OF} = \frac{10}{25 + 15 - 10} = 0.33$$

TS 值必介於 0 到 1 之間，0 表預報與觀測完全不同，1 表預報與觀測完全相符。



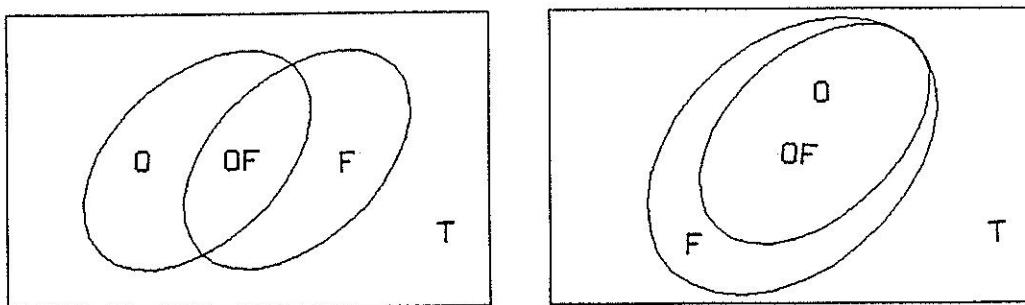
O=25  
F=15  
OF=10  
PF = 0.4  
NH = 0.6

O=25  
F=35  
OF=10  
PF = 0.4  
NH = 0.6

O=25  
F=10  
OF=10  
PF = 0.4  
NH = 0.6

圖九 降水面積預報結果不同，但前估與未命中率相等的例子。

Fig. 9 The example for area forecasts (precipitation) which have different forecast results but same prefigurations and no hit rates.



A  
O=25  
F=25  
OF=10  
PF = 0.4  
NH = 0.6

B  
O=25  
F=40  
OF=25  
PF = 1  
NH = 0

圖十 無法真正從前估和未命中率的值來判定哪個預報較好的例子。

Fig. 10 Couldn't decide which one is the best forecast by prefignration and no hit rate.

缺點：僅能用來量度在某區域內完全正確預報降水的相對頻率，所能提供有助於天氣預報的資訊仍嫌不足。圖八為幾種降水預測範圍不同，但TS值相等的例子，由圖可知TS值相等，並不代表預測好壞的程度相同。

(4) 前估 (PF) 與未命中率 (NH)：

$$PF = \frac{\text{預報正確次數(面積)}}{\text{發生次數(面積)}}$$

$$NH = \frac{\text{發生但未預報到之次數(面積)}}{\text{發生次數(面積)}}$$

$$PF = \frac{N_1}{N_1 + N_2} = \frac{8}{12} = 0.67$$

$$NH = \frac{N_2}{N_1 + N_2} = \frac{4}{12} = 0.33$$

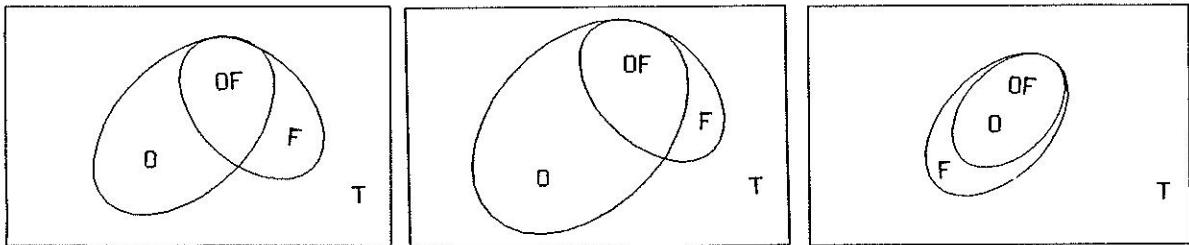
圖五：

$$PF = \frac{OF}{O} = \frac{10}{25} = 0.4$$

$$NH = \frac{O - OF}{O} = \frac{15}{25} = 0.6$$

這兩個指數之和恒為 1，其校驗之結果所代表的意義完全相同，只不過 PF 值越大越好， NH 值則

表一：



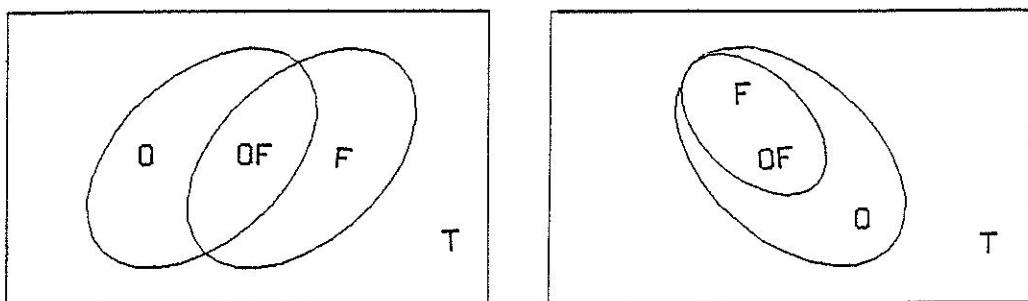
$O=25$   
 $F=15$   
 $OF=10$   
 $PA= 0.67$   
 $FA= 0.33$

$O=35$   
 $F=15$   
 $OF=10$   
 $PA= 0.67$   
 $FA= 0.33$

$O=10$   
 $F=15$   
 $OF=10$   
 $PA= 0.67$   
 $FA= 0.33$

圖十一 降水面積預報結果不同，但後符與錯誤警報率相等的例子。

Fig. 11 The example for area forecasts (precipitation) which have different forecast results but same post agreement and false alarm.



A       $O=25$   
 $F=25$   
 $OF=10$   
 $PA= 0.4$   
 $FA= 0.6$

B       $O=25$   
 $F=10$   
 $OF=10$   
 $PA= 1$   
 $FA= 0$

圖十二 無法真正從後符和錯誤警報率的值來判定哪個預報較好的例子。

Fig. 12 Couldn't decide which one is the best forecast by post agreement and false alarm.

越小越好，兩者皆介於 0 到 1 之間。其意義為某情況之發生（例如降水量小於 200 mm）與被預報到（PF）或未被預報到（NH）的比例。

缺點：對面積預報而言，只能知道預報準確的面積佔實際發生的面積的比例。圖九為幾種不同預報降水範圍與觀測結果，其前估與未命中率得分相等的例子。

若由此指數來判定預報的好壞是很危險的。如圖十（A）的 PF 值比圖十（B）低，但某些情況下，我們寧可視圖十（A）為較好的預測，因為圖十（A）僅位置稍有偏差，而圖十（B）卻是預報

過度。

(5)後符（PA）與錯誤警報率（FA）：

$$PA = \frac{\text{預報正確次數 (面積)}}{\text{預報次數 (面積)}}$$

$$FA = \frac{\text{預報但未發生次數 (面積)}}{\text{預報次數 (面積)}}$$

表一：

$$PA = \frac{N_1}{N_1 + N_3} = \frac{8}{8 + 8} = 0.5$$

$$FA = \frac{N_3}{N_1 + N_3} = \frac{8}{8 + 8} = 0.5$$

圖五：

$$PA = \frac{OF}{F} = \frac{10}{15} = 0.67$$

$$FA = \frac{O - OF}{F} = \frac{5}{15} = 0.33$$

這兩個指數之和亦恆為 1，檢驗結果代表的意義相同，但 PA 值越大越好，FA 值則越小越好。其意義為被預報而後確實發生 (PA) 或未發生 (FA) 的比例。

缺點：對面積預報而言，只能知道預報準確的佔所有預報面積的比例。圖十一為幾種不同降水範圍預報與觀測結果，其後符與錯誤警報率相等的例子。

以此二值的大小，有時亦很難判斷哪個預報較好。如圖十二 (A) 的 PA 值比圖十二 (B) 低，但圖十二 (B) 的預報不見得比圖十二 (A) 好。

(6) 技術得分 (HSS)：

$$HSS = \frac{F - E}{T - E}$$

F：預報正確次數

E：無技術預報的次數

T：預報總次數

HSS 的最大值為 1；其值越大表預報越好。所謂無技術預報，係指隨機、持續或氣候預測。此方法僅適用於範疇式預報，並不適用於面積預報。

表一：假設無技術預報為隨機預報，預報準確率為 50%。

$$HSS = \frac{\frac{N_1 + N_4 - (\frac{N_1 + N_2 + N_3 + N_4}{2})}{2}}{(N_1 + N_2 + N_3 + N_4) - (\frac{N_1 + N_2 + N_3 + N_4}{2})}$$
$$= \frac{8 + 10 - 30/2}{30 - 30/2} = \frac{3}{15} = 0.2$$

此方法所代表的意義及其缺點和正確百分比的意義與缺點差不多。不過透過此方法，可以比較預報能力是否比無技術預報更好。

以上這些方法由於較適用於範疇式預報或面積預報，對數值模式產品而言僅適用於降水一項，但目前模式所得之降水預報僅能供參考用，不能直接做為預報值，因此在數值預報之校驗系統中似較不宜採用這些方法。不過這些方法若能適當搭配，擴長縮短（例前估與後符同時使用），則倒是可以應用於預報員所做之天氣預報。

3. 平均誤差 (average error, A.E.)、平均絕對誤差 (average absolute error, A.A.E.)、標準偏差 (the standard deviation error, S.D.E.) 及均方根誤差 (root-mean-square-error, R.M.S.E.)

這幾種方法適用於連續場的預報，例如海平面氣壓、500 mb 高度場、溫度場等。

(1) 平均誤差 (A.E.)

$$A.E. = \frac{\sum_{i=1}^N (F_i - O_i)}{N}$$

F<sub>i</sub>：預報值

O<sub>i</sub>：觀測值（或網格點分析值）

N：測站數目（或網格點數目）

此指數亦可稱為偏差 (bias)，但與前面所說的偏差含意不同。前者為對某一特定測站（或某一個網格點）長期平均的偏差；此指數為一特定時間模式對整個空間平均的偏差。這個指數對氣象場而言，可約略看出模式的預報傾向，例如溫度預測偏高，或者氣壓值偏低。此方法用於氣壓場所涵蓋的意義不及用於溫度場。圖十三為氣壓場之預報與分析值簡例，雖然 (B) 例的平均誤差比 (A) 例大，但我們可能認為 (B) 例為較好的預測，因為我們重視系統強度與類型甚於氣壓值本身。

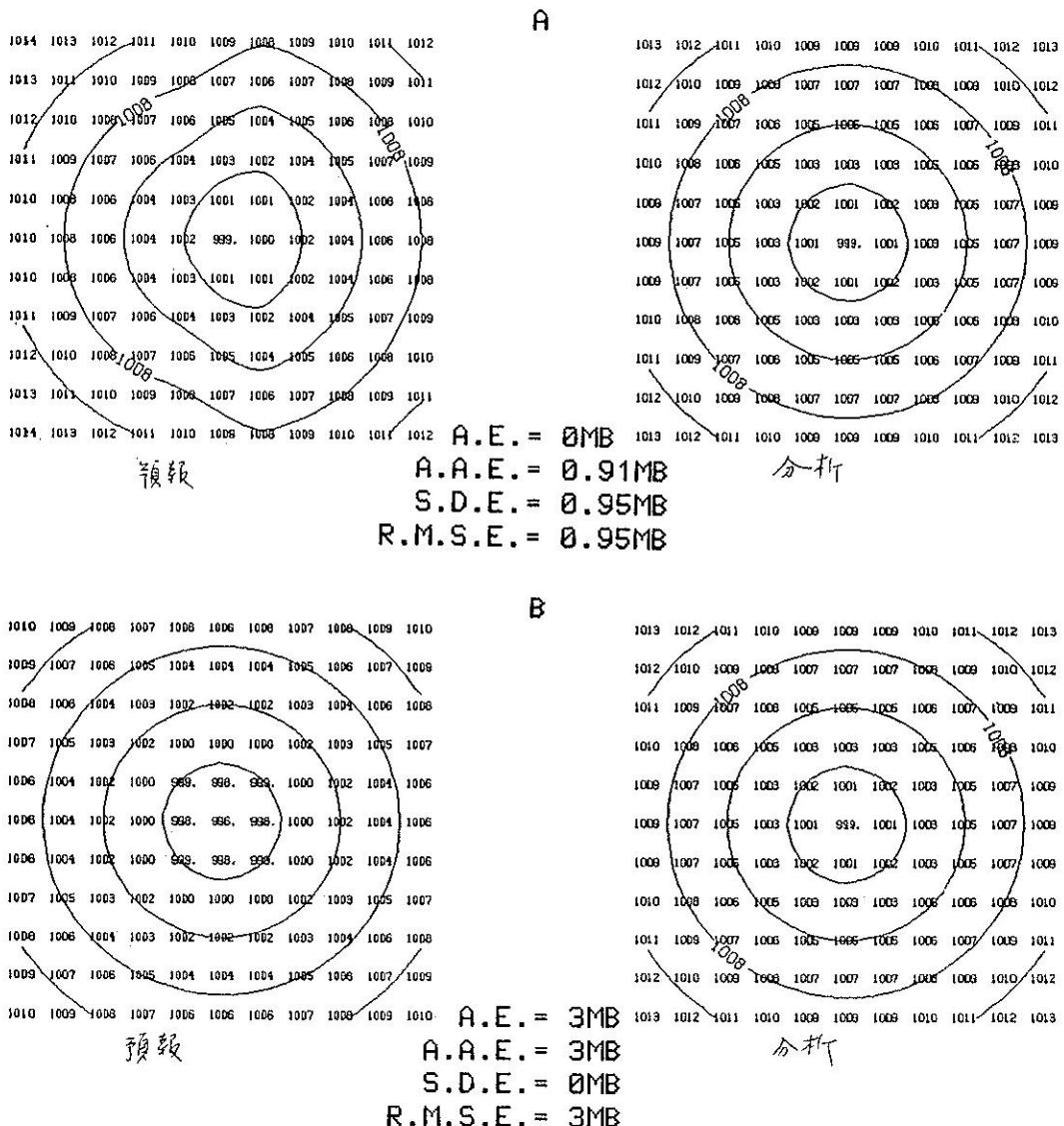
(2) 平均絕對誤差 (A.A.E.)

$$A.A.E. = \frac{\sum_{i=1}^N |F_i - O_i|}{N}$$

此指數可以知道誤差的大小（不論正負），但對於系統預報之好壞，仍沒有足夠的能力判斷。圖十三之二例亦可看出 (A) 例的平均絕對誤差比 (B) 例小，但對系統預報而言，我們認定 (B) 例預報得較好。此指數比平均誤差進步的是 (A) 例中的平均誤差為 0 mb，可能會使人誤以為其預報完全正確，平均絕對誤差指數就不會造成這種困擾。

(3) 標準偏差 (S.D.E.) 與均方根誤差 (R.M.S.E.)

$$S.D.E. = \sqrt{\frac{\sum_{i=1}^N [(F_i - O_i) - (\bar{F}_i - \bar{O}_i)]^2}{N}}$$



圖十三 (B) 例的平均誤差雖比 (A) 例大，但或可視 (B) 例為較好的預報（氣壓場）。  
Fig. 13 The average error of example (B) is larger than that of example (A), but example (B) may be the better forecast

表二 最低溫可能發生的機率

Table 2 Possibility of the minimum temperature maybe occure.

最低溫度 ( °C )	10	11	12	13	14	15	16
可能發生的機率	0.05	0.12	0.15	0.25	0.30	0.12	0.05

$$R.M.S.E. = \left\{ \frac{\sum_{i=1}^N (F_i - O_i)^2}{N} \right\}^{1/2}$$

此二指數所代表的意義和平均絕對誤差差不多，其缺點也很類似。不過以圖十三（B）之預報來說，其標準偏差為 0，而均方根誤差為 20 mb。若是僅考慮系統的強度與類型，而不考慮絕對的氣壓值大小，標準偏差為 0 是可以被接受的。換言之，在這種情況下，標準偏差或可視為比均方根誤差更具意義。

以上四種指數，除可用於校驗連續場外，亦可用於校驗時間序列之預報，只需將公式中的網格點數換成預報次數即可。不過同樣的這些校驗方法亦有其特具之缺點。我們以表二之溫度預報為例說明之。

若以平均絕對誤差校驗之，以預報 13°C 為最好，因為其平均絕對誤差總和為最小。若以均方根誤差校驗之，則預報 13、15 °C 最好，由此可見應用不同校驗方法，會產生不同的“最好”預報，這也就是早期校驗工作之重要性受人懷疑的原因。

#### (4) 相關係數 (correlation coefficient, C.C.)

$$C.C. = \frac{\sum_{i=1}^N F_i O_i}{\sqrt{\sum_{i=1}^N (F_i)^2} \sqrt{\sum_{i=1}^N (O_i)^2}}$$

F：表預報

O：表觀測

此指數可以求得預報場與觀測場（分析場）的相關。若預報完全正確，則相關係數為 1。此指數不僅可用於預報場與觀測場，也可用於預報值之變化與觀測值之變化的相關，即將預報值和觀測值減去初始場再求相關即可。此指數同樣的也是對整個預報區域求平均，對於天氣系統之預報是否準確助益不大。

#### (5) S1 分數 (Teweles and Wobus, 1954)

$$S1 = \frac{\sum_{i=1}^N |e_g|}{\sum_{i=1}^N |G_L|}$$

$e_g$ ：氣壓差（梯度）的預測誤差

$G_L$ ：觀測或預報氣壓差（梯度）較大者

這個指數專用於氣壓梯度場（高度梯度場），由於大氣狀態近似地轉平衡，故由此指數亦可獲得

某些風場的資訊。S1 分數越小越好（表誤差較小）。用這種校驗方法有許多好處，其一為由於是對氣壓梯度做校驗，對於系統強度之指示較具代表性；此外，分母  $G_L$  有減少季節變化所影響的效果。不過若系統較強且移動較慢。則 S1 分數會較小；反之，系統弱且移速快時 S1 較大；換言之，利用 S1 分數做校驗，對系統強且移速慢的較為有利。

#### (6) 主觀評價法 (Subjective evaluation method) (劉等, 1984)

此法是以三個主觀性評分點 (1, 0, -1) 就每天的預報圖與分析圖加以評斷，其計分如下：

1：極為可信——預報圖與分析圖的類型及天氣系統極相似。

0：略為可信——預報圖與分析圖的類型及天氣系統略同。

-1：極不可信——預報圖與分析圖的類型及天氣系統有極大錯誤，足以導致預報員預報錯誤的天氣現象者。

此法是由人工校驗，主觀因素過重，相同的預報結果由不同的人進行校驗可能有不同的校驗得分，結果極不客觀，故不適宜列入日常作業的校驗系統中。

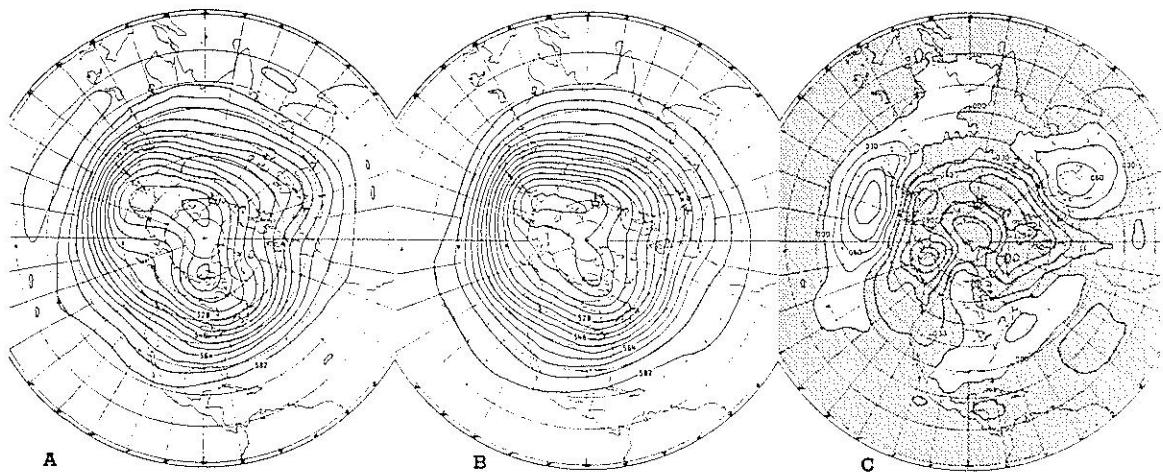
以上是各種以指數或分數來表示預報之好壞的校驗方法。綜觀上述之討論，我們可以發現，這些方法能提供我們有利於預測之改進的資訊仍然有限。因此，現今的檢驗工作，不能僅止於上述的幾種方法，而必須更進一步做誤差特性的分析以積極提供有利於模式預報產品之應用或模式改進所需的訊息。

#### (二) 誤差分析

Wallace 和 Woessner (1981) 指出模式誤差之研究可分為兩類，第一類為利用模式積分至某一特定預報時間，例如 24 小時或 36 小時預報。然後將模式預報的氣象場，例如：500 mb 之高度場和實際觀測值（或分析場）來做比較。另一類則為針對某些綜觀天氣系統如地面高、低壓系統等之預報誤差做分析，下面我們將就此二類誤差之研究進行討論。

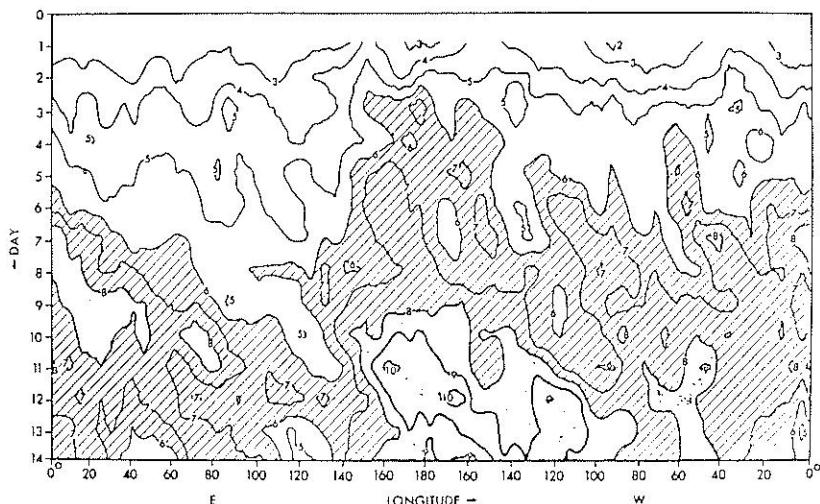
##### 1. 對整個預報場之誤差分析

(1) 這類誤差研究分析方法，最常用的即是前面曾提過的平均誤差。所不同的是，這裡是做時間的平均（例如一年、一季或一個月等）而非空間平均，並且所得到的，不是一個值，而是誤差的空間分



圖十四 一個大氣環流模式 500 mb 高度場之誤差分析，12 個冬季個案的平均 (1964 ~ 1969)。 (A) 觀測 (B) 預報 (C) 誤差 (預報 - 觀測) 之空間分佈圖 (Miyakoda 等, 1972)。

Fig. 14 The spatial distribution of error for 500 mb height prediction by a general circulation model. It is an average of 12 cases in winter (1964-1969). (A) observed (B) prediction (C) error (Prediction - observed) (Miyakoda et al., 1972).



圖十五 一個大氣環流模式 500 mb 溫度場之均方根誤差隨預報時間與經度的變化，12 個冬季個案  $30^{\circ}\text{N} \sim 45^{\circ}\text{N}$  的平均 (1964 ~ 1969) (Miyakoda 等, 1972)。

Fig. 15 The time-longitude chart of the root mean square 500mb temperature error by a general circulation model. It is an average of 12 cases in winter (between 30 N and 45 N, 1964-1969) (Miyakoda et al., 1972).

佈，如圖十四，其他亦可以均方根誤差來表示。

平均誤差之空間分佈，可以了解誤差隨空間分佈的情形，譬如隨緯度的變化、隨經度的變化等等。均方根誤差則顯示誤差之大小（振幅）隨空間的分佈，不能表示誤差之正負。

(2)另一種誤差分析方法則是針對某個緯度（譬如 $45^{\circ}\text{N}$ ）或某個緯度帶平均（譬如 $35^{\circ}\text{N} \sim 45^{\circ}\text{N}$ ），分析其誤差隨經度或預報時間的變化，可以平均誤差、均方根誤差或其他之校驗指數與分數表示。圖十五即為一例，由圖中可知誤差隨預報時間之增長而增大，且在 $160^{\circ}\text{E} \sim 160^{\circ}\text{W}$ 間的誤差為最大。

(3)若是希望找出誤差的來源，有時前兩種方法不一定能滿足需求，於是可以在先將誤差做尺度分離（scale separation），再依前述方法做各種誤差分析。這種對誤差做尺度分離的手續，對於追究誤差來源極有用處。一般做尺度分離有兩種方法，一為傅氏分析（Fourier analysis）另一則為中頻帶過濾法（band-pass filter）。前者適於全球或半球模式，後者適用於區域模式。

圖十六即為先經尺度分離後再求標準差的情形

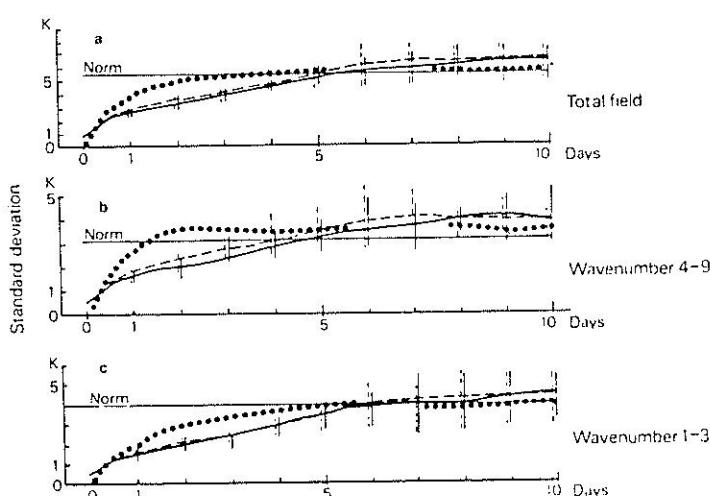
，由此圖可得長波與中長波的誤差隨預報時間的變化，圖中為兩個預報模式與兩個無技術預報（持續性預報、氣候預報）的誤差變化。

由於經過尺度分離的程序後，仍需利用其適用於連續校驗場的校驗方法來校驗，所以尺度分離並非一種校驗方法，而是校驗誤差前的一道手續，這道手續對誤差來源的追蹤極有用處。

## 2. 對天氣系統之誤差分析

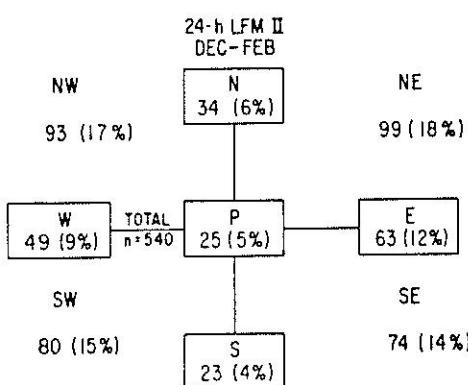
對天氣系統（例如高壓、地壓、鋒面、颱風等）的誤差分析，主要即在於研究其強度、移動速度、移動方向等的誤差，是以系統本身為參考點，歸納模式對系統預報的快慢、強弱。圖十七即為一種系統位置誤差的表示方法。圖中P表預報位置，E、SE、S、SW、W、NW、N、NE分別表示觀測所得之氣旋位置位於預報位置之東、東南、南、西南、西、西北、北和東北其下之數字表個數及百分比。

圖十八為一颱風預報方法所預報之颱風移速誤差的百分比分配。圖中橫坐標為誤差值，縱坐標為百分比，由圖可知模式對颱風之移速誤差的分佈情形，及平均位移誤差。



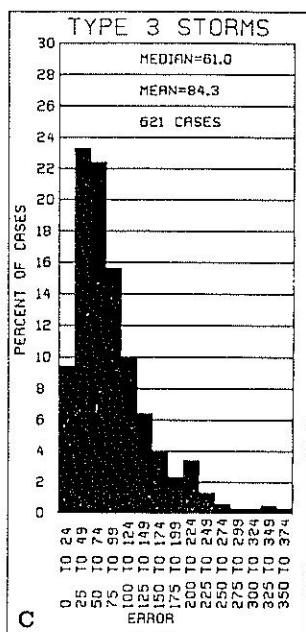
圖十六 二月之 $850\text{ mb}$ 溫度場的標準差的水平平均（7個個案平均）。粗實線與虛線為兩個不同預報模式。點線為持續性預報，細實線為氣候預報（Hollingsworth 等，1980）。

Fig. 16 Horizontal mean of standard deviations of the  $850\text{ mb}$  temperature field for seven February cases, Solid line and dashed line and forecasts from two different model, thin solid line is climate forecast (norm). (Hollingsworth et al., 1980).



圖十七 1978年12月及1979年1月、2月之NMC LFM-II 24小時誤差，觀測氣旋相對於預報氣旋位置 ( Silberberg and Bosart , 1982 )。

Fig. 17 LFM-II 24h forecast position error showing observed positions of cyclones relative to the forecast position ( Silberberg and Bosart , 1982 ).



圖十八 HURRAN ( hurricane analog ) 技術預報颱風的24小時位移誤差上方為誤差中位數、平均誤差及個案數 ( Neumann 和 Hope , 1972 ) 。

Fig. 18 Frequency distribution of HURRAN 24 h mean vector error. ( $n$ ,  $M_i$ ) ( Neumann and Hope , 1972 ).

除上述列舉的幾種誤差分析方法之外，其他有關的分析方法尚有許多，端視個人所感興趣之間題的不同，而有不同的應用，本文所要強調的是，若要追溯誤差來源，改進預報模式，“誤差分析”是絕對必要的過程。

## 五、結論與建議

中央氣象局之全球預報模式已於77年7月1日正式上線，加入天氣預報的行列。其他之區域模式、中尺度模式及颱風模式亦將陸續於78年年底前正式作業，為使這些數值模式能發揮最大效益，本局亦隨之展開了一些校驗工作。

校驗工作可分為兩大類，一為日常作業校驗每天之預報結果，另一為非日常作業，校驗及分析模式在長時間裡的誤差傾向。後者可由對於模式之研究發展有興趣的人，依其需要，自行設計，大致皆必採用誤差分析的過程，以達追溯誤差來源的目的。至於前者，則是本局真正必需重視且積極建置的。

日常作業之校驗系統，基於電腦資源與實際之需要，既不可能也不必要將前節所述之各種方法完全納入，只需應用幾種有效的方法即可發揮其應有的效能。本文在此對於本局目前已採行的校驗方法做一概略介紹，同時亦對本局尚未採用，但或有助益之方法提供一些建議，給校驗系統之擴充與發展做個參考。

### (一) 本局目前之校驗方法

#### 1. S1 分數

S1分數是目前最受各作業單位及天氣學家所肯定的校驗方法，而氣壓梯度對天氣之影響遠甚於氣壓值本身，因此採用此法校驗頗具效益。不過需注意的是，系統之強弱、移速及網格點之大小對S1分數都有影響，因此若用S1分數來比較不同緯度區域或不同模式的產品，結果不一定客觀，但對於同一區域、相同網格間距的模式產品可得良好的比較結果；換句話說，我們不能拿全球預報模式的S1分數來和區域模式或中尺度模式比較，但可以由S1分數來判定全球模式之修改是否確實有效益。

#### 2. 距平相關係數

若將氣象場減去某個參考場再求相關係數，則可突顯誤差的特性。距平相關係數 ( anomaly correlation coefficient ) 是以氣候值為參考場所求得之相關係數。

### 3. 標準偏差與均方根誤差

氣壓場較重視強度與類型，溫度場則偏重其值本身。故可將標準偏差應用於氣壓場，而均方根誤差則應用於溫度場。

### 4. 平均誤差及平均絕對誤差

由平均誤差可知模式之誤差屬性（例：過暖或過冷），由平均絕對誤差可知模式誤差之大小，對模式發展人員改進模式頗有幫助。

#### (二)建議建置之校驗方法

除上述幾種本局已採用之校驗方法外，本文建議整個校驗系統亦可加入下列幾種方法以強化校驗功能。

#### 1. 趨勢相關係數 (tendency correlation coefficient)

趨勢相關係數異於距平相關係數之處為趨勢相關係數是以初始場為參考場，如此可去除因分析所造成的誤差分量，使我們真正了解模式對變化量掌握的能力，適用於三天以內的預報。

#### 2. 客觀氣旋追蹤法

這個方法基本上是屬於天氣系統之誤差分析，應用電腦以客觀、定量、快速且自動化的方式，計算出某個氣旋（也可應用於反氣旋）的移動誤差，並且立即繪圖顯示，如此預報員由氣旋初生之一、兩天（初生期）的誤差便可即時修正氣旋於第三、四天以後（成熟期、消散期）的行為，對於預報員有極大的幫助，這個方法和傳統的天氣系統誤差分析最大的差異是傳統的天氣系統誤差分析是由許多個案統計計算出一個誤差結果，客觀氣旋追蹤法則是針對個別的個案，快速地得出其誤差。這樣的結果對預報員而言遠比其他各種方法所得之結果有意義。此系統已由美國海軍環境研究中心（Naval Environmental Prediction Research Facility，NEPRF）發展完成，氣象局若能引進這套系統，相信會有很大的貢獻。

以上幾種方法，若能加入現有的校驗工作中，相信對於預報員應用模式產品及模式發展人員改進模式，都會有很大的幫助。

## 參考文獻

- Anthes, R.A., 1983: Regional models of the atmosphere in middle latitudes. *Mon. Wea. Rev.* 111, 1306-1335.
- Haltiner, G.J., and R.T. Willians, 1975: Some present

advances in numerical weather prediction. *Mon. Wea. Rev.*, 103, 571-590.

Hollingsworth, A., K. Arpe, M. Tiedtke, M. Capaldo and H. Savijarvi, 1980: The performance of a medium-range forecast model in winter - impact of physical parameterizations. *Mon. Wea. Rev.*, 108, 1736-1773.

Miyakoda, K. G.D. Hembree, R.F. Strickler and I. Shulman, 1972: Cumulative results of standard forecast experiments I. Model performance for winter cases. *Mon. Wea. Rev.*, 100, 836-855.

Neumann, F.J. and J.R. Hope, 1972: Performance analysis of the HURRAN tropical cyclone forecast system. *Mon. Wea. Rev.*, 100, 245-255.

Nitta, T., Y. Yamagishi and Y. Okamura, 1979: Operational performance of a regional numerical weather prediction, *J. Meteor. Soc. Japan*, 57, 308-331.

Rosmond, T.E., 1984: Global forecast system. Workshop on Numerical Weather Prediction (Taipei); Central Weather Bureau, 191pp.

Silberberg, S.R. and L.F. Basart, 1982: An analysis of systematic cyclone errors in the NMC LFM-II model during the 1978-79 cool season. *Mon. Wea. Rev.*, 110, 254-271.

Teweles, S. and H. Wobus, 1954: Verification of prognostic charts. *Bull. Amer. Meteor. Soc.*, 35, 455-463.

Wallace, J. and J. Woessner, 1981: An analysis of forecast error in the NMC hemispheric primitive equation model. *Mon. Wea. Rev.*, 109, 2444-2450.

Wash, C.H. 1984: Synoptic forecasting with numerical weather prediction methods. Workshop on Numerical Weather Prediction (Taipei), Central Weather Bureau, 80pp.

陳泰然, 1984 : 數值天氣預報應用。中央氣象局 數值天氣預報訓練第一階段教材, 65 頁。

劉復誠、華文達及鄭月娥, 1984 : 日本氣象廳 8L-NHM 及 10L-FLM 數值預報模式之初步校驗，氣象學報, 30, 2, 8-47。

# A Review on Verification Methods for Numerical Weather Prediction Products

*Lie-shieu Fang*

Computer Center, CWB

*Hway-jen Chen*

Computer Center, CWB

## ABSTRACT

Due to the use with finite data, incomplete physics and numerical methods, the results of numerical predictions must exist some errors in the Numerical Weather Prediction (NWP) products. How to find out the sources of errors and modify them with validation methods are the most essential and dicussing issues after establishing the numerical wether prediction system.

In this review, all the related verification methods are sorted out and have discussions with traits and defects of those methods. And some suggestions for building an operational verification system are also proposed for the NWP products of Central Weather Bureau.

Keywords: verification system.