

# 102 年天氣分析與預報研討會

## Logistic MOS 模式與傳統 MOS 模式降水機率比較

蔡孟峰<sup>1</sup>, 陳孟詩<sup>1</sup>, 陳重功<sup>2</sup>, 羅存文<sup>2</sup>

預報中心<sup>1</sup>, 科技中心<sup>2</sup>

102.05.15

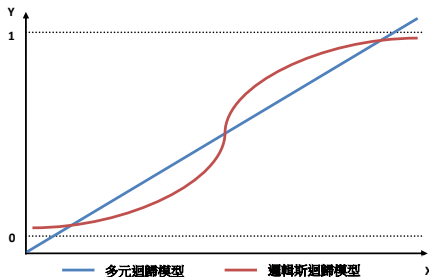




1. 動機
2. 資料與方法
  - MOS 模式
  - 資料來源
3. 校驗準則
  - Brier Score
  - ROC Curve
4. 資料分析
5. 結論



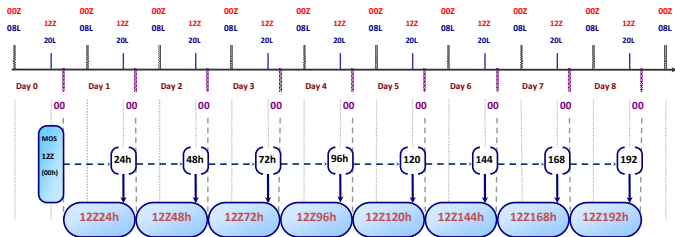
- 傳統 MOS 模式降水機率預報中，多元迴歸模型之機率預報值會超出正常範圍  $[0, 1]$  的情況。
- 在統計上，當反應變數為二元資料時，我們會將其視為類別資料，使用邏輯斯迴歸模型來建模。





## MOS模式

MOS模式主要是在於各個超前預報時段中，藉由統計模型，建立動力數值預報模式資料和對應的測站觀測資料（如溫度、降水量等）間的統計關聯性。其中，觀測資料稱為反應變數，而數值模式預報層場資料則稱為解釋變數。



## 多元迴歸模型

在降水機率預報當中，傳統 MOS 上使用的統計模型為將降水觀測資料轉換為降水事件的二元資料 ( $y$ 為1時表示降水量大於0.1mm,0則表示降水量小於0.1mm)，將其與預報因子建立線性機率之迴歸模型：

$$P(y_{ij} = 1) = \alpha + \beta_{j1}x_{ij1} + \cdots + \beta_{jk}x_{ijk} + \varepsilon_{ij},$$

其中，

- $x_{ijk}$  為觀測時間  $i$ ，超前預報時段為  $j$ ，第  $k$  個層場變數之格點內插值。
- $y_{ij}$  為在數值模式之超前預報時段  $j$  所對應的觀測資料。
- $\varepsilon_{ij}$  假設服從常態分配平均數為 0，變異數為  $\sigma^2$ 。

## 邏輯斯迴歸模型

- 定義  $\pi(X) = P(y_{ij} = 1)$

邏輯斯迴歸模型表示式如下：

$$\text{logit}[\pi(X)] = \log\left[\frac{\pi(X)}{1-\pi(X)}\right] = \alpha + \sum_{k=1}^K \beta_{jk} x_{ijk} = X\beta,$$

並可轉換成下方式子：

$$\pi(X) = \frac{\exp(X\beta)}{1+\exp(X\beta)}.$$

其中，

- Y 服從  $\text{Binomial}(n, \pi(X))$ 。



- 代表測站: 台北測站 (46692)
- 模式資料
  - NCEP GEFS(Global Ensemble Forecast System)
  - 超前預報時段至 384hr
  - 2.5×2.5 degree
- 資料期間:2009-01-01至2012-02-29
  - 配適階段:2009年1月、2月至2011年冬季
  - 校驗階段:2012年冬季



- 層場變數

- 海平面氣壓場
- 850百帕溫度場、水平風場、垂直風場
- 250百帕水平風場、垂直風場
- 500百帕、700百帕高度場
- 累積雨量場

● 預報因子：使用內插方法將鄰近台北測站之4個網格點的層場資料，加以計算至台北測站的經、緯度位置所產生。





假設超前預報時段的模型推估值為  $f_i$ 、觀測資料為  $y_i$ ，則，

$$\text{Brier Score: } \frac{\sum_{i=1}^n (f_i - y_i)^2}{n} \quad \circ$$

其中，其值介於  $[0,1]$  之間，數值愈小表示模型預報能力愈好。

## 二維列聯表

	$y = 0$	$y = 1$	
$\hat{y} = 0$	TP	FP	$N_3$
$\hat{y} = 1$	FN	TN	$N_4$
	$N_1$	$N_2$	$N$

- 靈敏度 =  $P(\hat{y} = 1|y = 1)$ , 特異度 =  $P(\hat{y} = 0|y = 0)$ 。
- 藉由每個候選臨界值, 可以建立數個二維列聯表, 得到數組靈敏度與特異度, 以其畫出 Roc Curve。



- ROC 曲線下面積

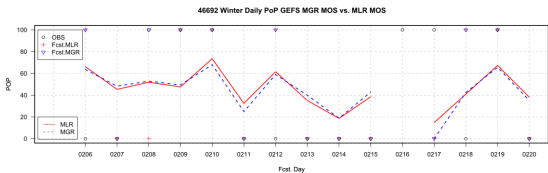
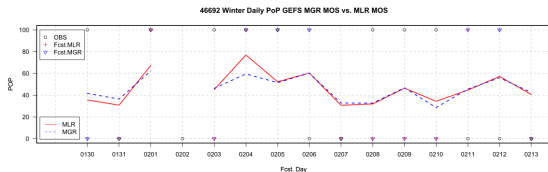
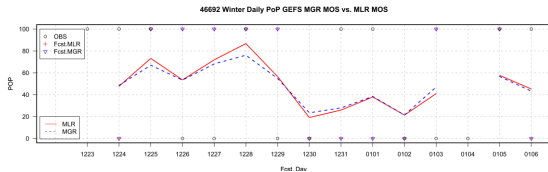
一般而論，當 AUC 大於 0.8 時，模型最具篩檢或診斷的效果。

- 準確率

$$\text{準確率} = P(\hat{y}=1 \text{ and } y=1) + P(\hat{y}=0 \text{ and } y=0)$$

- 臨界值的篩選機制

$$\text{Youdex Index} = \text{靈敏度} + \text{特異度} - 1$$





## ROC Curve

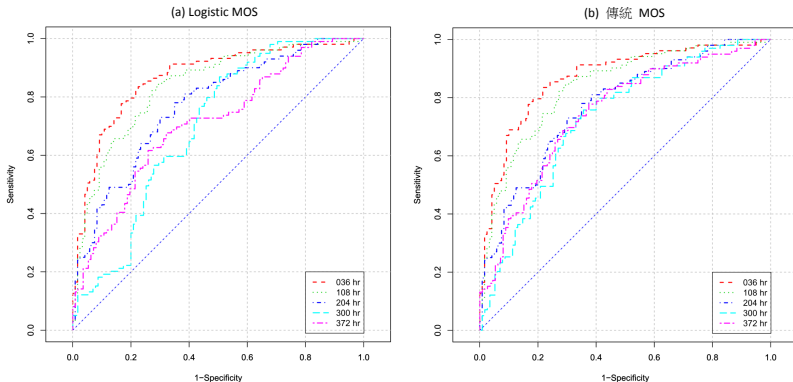


Figure 1: 兩種MOS 模式下,ROC 曲線在不同超前預報時段的變化。





## ROC Curve

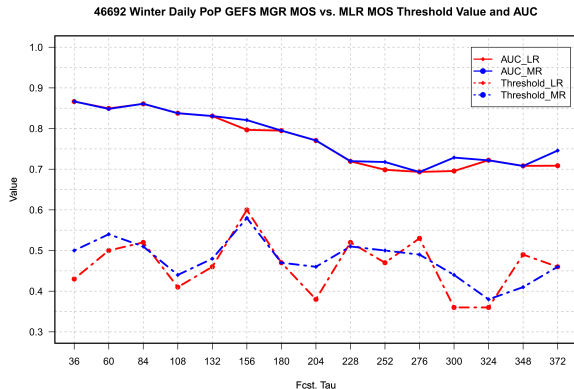


Figure 2: 在不同超前預報時段下，兩種MOS 模式之臨界值選取與 ROC 曲線下面積變化。



## Brier Score

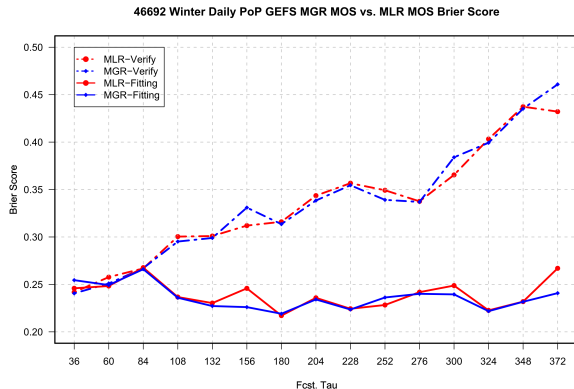


Figure 3: 在不同超前預報時段下, 兩種MOS 模式在配適階段與校驗階段之 Brier Score 變化。



## Accuracy Rate

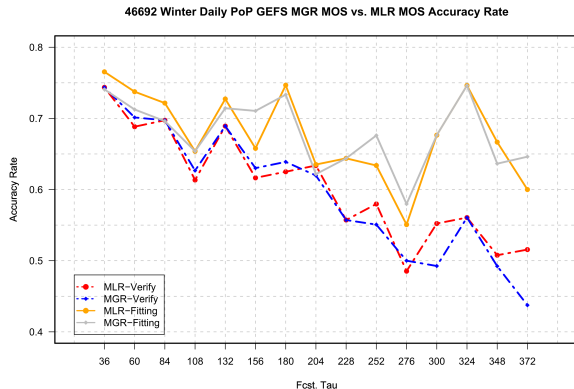


Figure 4: 在不同超前預報時段下, 兩種MOS 模式在配適階段與校驗階段之準確率變化。





- 藉由 GEFS 預報產品, 文中呈現 MOS 模式二週降水機率預報之技術發展過程。
- 在校驗成果上, 所呈現出的結果是相似的; 同時, 隨著超前預報時段的增加, 預報能力也跟著下降。
- 邏輯斯迴歸模型確實能改善傳統 MOS 在統計上的不合理性。
- 未來二週降水預報的方向, 仍可以就邏輯斯迴歸模型方向搭配其他模式資料繼續發展, 並思考如何改善在二週的預報能力。





謝謝聆聽