# 侵台颱風個數機率分布季預報

林昀靜　盧孟明

中央氣象局科技中心

## 摘　要

颱風是台灣地區夏季天氣的主要現象，每年都會有數個颱風侵襲，而造成大量的降水及生命財產的災害損失，對台灣人民的生活造成極大的影響，因此若能利用統計方法提前預報侵台颱風的多寡，就能提早做好防災的準備，減少颱風所帶來的災害損失。季節預報有很高的不確定性，因此以或然率表達的預報結果比較合理。據此，本文提出以貝氏定律(Bayesian approach)為框架的機率回歸模式，做為西北太平洋颱風以及侵台颱風的年頻率機率預報方法。

## 一、　前言

台灣是個位於亞熱帶地區的海島國家，在台灣地區天氣與氣候的主要現象，在夏季是颱風，在冬季則是東北季風及寒潮，通常東北季風及寒潮帶來的損害較少，但由颱風所造成的豪雨與損害則相當巨大。因此若能利用統計方法提前預報侵台颱風的數量，就能提早做好防災的準備，也就可以減少颱風所帶來的災害損失。

統計預報是利用歷史資料建立預報對象與預報因子間的關係，而歸納出統計模式，並用來預報未來的氣候。Chu 等人(2007)以 5 月的海平面溫度、海平面氣壓場、可降水量、850Pa 渦度場及垂直風切等五種環流場當作預報因子，利用 LAD (Least Absolute Deviation Regression Method)模式預報侵台颱風個數的研究。由於預報模式考慮的範圍相當有限，還有許多影響預報對象的因素沒有考慮到，因此預報的結果一定存在有難以確定的誤差，也就是預報的不確定性。如 LAD 這種決定性的預報無法提供預報結果不確定性的資訊，唯有機率性預報才能彌補這個缺點。

貝式方法(Bayesian approach)是以機率為概念的方法，因為比較符合預測不確定的精神，已有相當多的研究採用貝式方法進行區域尺度的氣候預測。Tebaldi 等人(2004a,b)利用貝式方法定義一個統計模式，得到在不同情境下模式系集平均的預測溫度或降水之機率分布型態。Greene 等人(2006)亦採用貝式方法定義統計模式，在各情境下進行模式系集平均的季節性預測。Furrer 等人(2007)結合了 GCM 的格點輸出於機率的推估，可得到氣候變遷下機率分布型態在區域尺度所造成的改變。Elser 和 Jagger (2004)以層次式的貝式方法(Hierarchical Bayesian Approach)及 Poisson 回歸方式建立推估美國東海岸每年颶風發生次數的模式，以 NAO 及 ENSO 作為預報因子，利用馬可夫鍊蒙地卡羅法(Markov chain Monte Carlo, MCMC)整合模式。Chu 與 Zhao (2007)亦利用貝式回歸模式，預報中太平洋地區熱帶氣旋的個數。

## 二、　資料

本研究所選擇的颱風資料來自 Data: Japan Meteorological Agency / RSMC (Regional Specialized Meteorological Center) Tokyo-Typhoon Center 的最佳路徑資料。由於本研究將針對侵台與西北太平洋兩個地區的颱風個數做預報，因此在侵台方面選擇的是 1979-2007 年的颱風最佳路徑資料，將經過台灣附近（東經 119-125 度、北緯 21-26 度）且最大風速大於 35 kt 的颱風視為侵台颱風，統計每年 6-11 月侵台的颱風個數。西北太平洋則是統計 1980-2008 年每年最大風速大於 64 kt 的颱風。

預報因子的大尺度環流資料為 NCEP/DOE 再分析資料(NCEP R2)，在侵台部份選擇的是 1979-2007 年當年五月份月平均的海平面氣壓、850hPa 相對渦度、垂直風切(200、850hPa)以及海平面溫度(ERSST)。而西北太平洋則是選擇前一年四月到當年三月共 12 個月的海平面氣壓場及海平面溫度。

# 三、　方法

颱風個數之機率分布以卜瓦松(Poisson)分布最為適合，因此建立一個以 $\lambda$ 為參數（$\lambda$ 為平均發生次數），在某一段時間內發生 $h$ 次侵台颱風個數的機率分布模式。

$$P(h|\lambda)=\exp(-\lambda)\frac{(\lambda)^h}{h!} \quad (1)$$

利用在大尺度環流因子發生之條件下每年發生 $h$ 次侵台颱風個數之卜瓦松分布建立一個線性回歸模式。假設有 $K$ 個大尺度環流因子(predictors)各有 $N$ 年的觀測資料，針對每年發生極端降雨事件的次數 $h_i$ 定義一個潛在的變數向量 $\mathbf{Z}$，$Z_i=\log\lambda_i$，$i=1,2,\Lambda,N$，潛在變數與大尺度環流因子間的關係可以表示為 $Z_i=\mathbf{X}_i\boldsymbol{\beta}+\varepsilon_i$，其中 $\boldsymbol{\beta}=[\beta_0,\beta_1,\Lambda,\beta_k]'$，$\varepsilon_i\sim\text{Normal}(0,\sigma^2)$，$\mathbf{X}_i=[1,X_{i1},X_{i2},\Lambda,X_{iK}]$ 為大尺度環流因子向量，因此可以建立一個線性回歸模式如下：

$$P(\mathbf{h}|\mathbf{Z})=\prod_{i=1}^{N}P(h_i|Z_i) \quad (2)$$

其中，$h_i|Z_i\sim\text{Poisson}(h_i|e^{Z_i})$，$\mathbf{Z}|\boldsymbol{\beta},\sigma^2,\mathbf{X}\sim\text{Normal}(\mathbf{Z}|\mathbf{X}\boldsymbol{\beta},\sigma^2\mathbf{I}_N)$，$\mathbf{I}_N$ 為單位矩陣。

對貝式推論來說馬可夫鏈蒙地卡羅法是個相當有效的演算法，通常貝式分析包括了對事後機率期望值的描述，如下所示：

$$E[\alpha|\mathbf{h}]=\int_{\boldsymbol{\theta}}\alpha(\boldsymbol{\theta})P(\boldsymbol{\theta}|\mathbf{h})d\boldsymbol{\theta} \quad (3)$$

$\alpha(\boldsymbol{\theta})$ 可以是任何關於模式參數 $\boldsymbol{\theta}$ 的方程式，但是此一積分式在計算上有相當的困難度，因此可以採用蒙地卡羅繁衍法，利用數值方法來計算此期望值。

$$E[\alpha|\mathbf{h}]\approx\frac{1}{L}\sum_{i=1}^{L}\alpha(\boldsymbol{\theta}^{[i]}) \quad (4)$$

馬可夫鏈蒙地卡羅法主要是將聯合後驗分布(joint posterior distribution)當成目標機率分布，利用馬可夫鏈來模擬參數的聯合後驗分布，再利用 Gibbs sampler 重複迭代出近似目標的分布。

對於 $\boldsymbol{\beta}$ 及 $\sigma^2$ 來說由於沒有任何可信的事前資訊，因此將 Gibbs Sampler 繁衍法應用於颱風個數之貝式推論時選擇無資訊之事前分布，式子表示如下：

$$P(\boldsymbol{\beta},\sigma^2)\propto\sigma^{-2} \quad (5)$$

這式子雖然不是一個適當的分布方程式，但是可以得到一個適當的事後機率分布。

當得到一新觀測的大尺度預報因子 $\tilde{\mathbf{X}}=[1,\tilde{X}_{i1},\tilde{X}_{i2},\Lambda,\tilde{X}_{iK}]$，則潛在變數 $\tilde{Z}$ 及颱風個數推估值 $\tilde{h}$ 之預測分布為

$$P(\tilde{Z}|\tilde{X},\mathbf{X},\mathbf{h})=\iint_{\boldsymbol{\beta},\sigma^2}P(\tilde{Z}|\tilde{X},\boldsymbol{\beta},\sigma^2)P(\boldsymbol{\beta},\sigma^2|\mathbf{X},\mathbf{h})d\boldsymbol{\beta}d\sigma^2 \quad (6a)$$

$$P(\tilde{h}|\tilde{X},\mathbf{X},\mathbf{h})=\int_{\tilde{Z}}\frac{\exp(-e^{\tilde{Z}}+\tilde{Z}\tilde{h})}{\tilde{h}!}P(\tilde{Z}|\tilde{X},\mathbf{X},\mathbf{h})d\tilde{Z} \quad (6b)$$

由於上式為一積分方程式無法直接計算，因此可採用下列方程式計算以替代其積分方程式。

$$P(\tilde{Z}|\tilde{X},\mathbf{X},\mathbf{h})=\frac{1}{L}\sum_{i=1}^{L}P(\tilde{Z}|\tilde{X},(\beta,\sigma^2)^{[i]}) \quad (7a)$$

$$P(\tilde{h}|\tilde{X},\mathbf{X},\mathbf{h})=\frac{1}{L}\sum_{i=1}^{L}\frac{\exp(-e^{\tilde{Z}[i]}+\tilde{Z}^{[i]}\tilde{h})}{\tilde{h}!} \quad (7b)$$

當 Gibbs sampler 達所需繁衍的長度後，捨去前面一段較不穩定的繁衍過程取樣值(burn-in period)，其中，$(\beta,\sigma^2)^{[i]}$ 為 Gibbs sampler 捨去不穩定過程後的第 $i$ 次繁衍取樣值，$\tilde{Z}^{[i]}$ 則是來自 $\tilde{Z}^{[i]}|\tilde{X},(\boldsymbol{\beta},\sigma^2)^{[i]}\sim\text{Normal}(\tilde{Z}^{[i]}|\tilde{X},\boldsymbol{\beta}^{[i]},\sigma^{2[i]})$ 的取樣值。

# 四、　應用與結果

(一) 貝式回歸模式建立－台灣

為了探討各個侵台颱風預報因子與颱風間的相關性，因此在各個預報因子格點資料的距平值，計算其與每年的侵台颱風個數在時間序列上的相關係數，根據統計顯著性 t 檢定，若相關性大於 0.366 可視為顯著相關。如圖(1)所示為各個預報因子在 5 月與颱風的相關係數等值圖，其中符號「•」所表示的即為顯著相關格點。將顯著相關格點計算成一個區域平均值，再將其標準化即可得到一個代表此預報因子於此年的代表值。

根據式(7)及交叉驗證法可以推估出每年發生侵台颱風個數的機率分布，若以每年發生最大機率(Mode)的推估個數作為預報個數，其與實際發生個數在時間序列上的分布如圖(2-a)所示，其相關係數為

0.649；若以每年預報機率分布的平均發生個數(Average)作為預報個數，其與實際發生個數在時間序列上的分布如圖(2-b)所示，其相關係數為 0.655。計算預報機率分布的累積機率在 25%、50%及 75%時的推估個數，若以發生機率累積 50%(Median)作為預報個數，25%及 75%作為上下界限，其在時間序列上的分布如圖(2-c)所示，其與累積機率為 50%的相關係數為 0.623，實線部份為機率累積 50%的發生次數，而虛線為 25%及 75%的發生次數。由圖中可看出，雖然無法正確的預報出實際發生的個數，但是發生個數多寡的分布型態還是有掌握到。

（二）貝式回歸模式建立－西北太平洋

交叉驗證法就是移除某一年的資料後，利用其餘年份對該年進行預測，因此在建立相關係數圖時，也應該利用交叉的方式建立，避免該年的資料對於相關係數圖有影響，西北太平洋的相關係數圖就採用 1979/04-2008/03 的預報因子與 1980-2008 年西北太平洋颱風個數建立 12 個月份的交叉相關係數圖(leave-one-out correlation map)，因此每個月份會有 29 張相關係數圖。為了方便挑選預報因子，將全球劃分成 A-T 等 20 個區域如圖(3)所示，挑出每張相關係數圖上是統計顯著相關的格點，根據每個格點在 29 年中是屬於統計顯著相關的發生次數，若區域內統計顯著相關格點的發生頻率等於 29 年的格點個數大於 25 個，則將此區域當成是一個候選預報因子。計算候選預報因子的區域平均值，並採用逐步回歸(stepwise regression)挑出對每年颱風個數影響性最大的預報因子，再利用 29 年中重複出現超過 20 年的預報因子建立貝式回歸模式。最後重複出現超過 20 年的預報因子有四個，如圖(4)所示，圖(4-a)為前一年 4 月的海溫，圖(4-b)為前一年 12 月的海溫，圖(4-c)為前一年 8 月的海平面氣壓，圖(4-d)為當年 1 月的海平面氣壓，其中陰影的部份為相關係數圖，符號「·」所表示的即為顯著相關格點。

若以每年發生最大機率(Mode)、預報機率分布的平均發生個數(Average)及發生機率累積 50%(Median)作為預報個數，其與實際發生個數在時間序列上的相關係數為 0.873、0.868 及 0.866，分布圖如圖(5)所示，結果比侵台颱風稍好。

（三）2009 年預報結果

根據 1979-2007 年歷史侵台颱風在 Poisson 累積機率的個數三分類，0~3 個颱風是屬於偏少(Below Normal, BN)，4~5 個是屬於正常(Normal, N)，而大於 6 個則是屬於偏多(Above Normal, AN)。根據 1980-2008年歷史的西北太平洋颱風則是0~13個屬於偏少，14~16 個屬於正常，大於 17 個則是屬於偏多。

利用貝式回歸模式對 2009 年侵台颱風的機率預報結果如圖(6)所示，其中背景為 1979-2007 年歷史侵台颱風發生個數的 Poisson 分布，柱狀圖為預報 2009 年侵台颱風在不同個數下的可能發生機率。根據三分類的結果，在個數偏少(BN)的情況下今年的發生機率比歷史機率少了 5% ，而在個數偏多的發生機率，今年則是多了 9%，比例雖不高，但表示今年侵台颱風傾向於正常偏多。

對 2009 年西北太平洋颱風的機率預報結果如圖(7)所示，其中背景為 1980-2008 年歷史西北太平洋颱風發生個數的 Poisson 分布，柱狀圖為預報 2009 年西北太平洋颱風在不同個數下的可能發生機率。根據三分類的結果，在個數偏少(BN)的情況下今年的發生機率比歷史機率少了 19%，而在個數偏多的發生機率，今年則是多了 24%，由此可知，今年西北太平洋颱風傾向偏多。

## 五、結論與建議

本研究採用貝式回歸模式對台灣及西北太平洋的颱風進行預報，從交叉驗證結果來看，其在相關係數大約為 0.65 及 0.87 ，此模式除了可以掌握到颱風個數在時間序列上分布的型態，也可以提供颱風預報個數的機率分布。

根據貝式回歸模式預報今年侵台颱風的機率分布圖與歷史颱風個數的 Poisson 機率分布來看，今年颱風個數偏多(AN)的機率增加了 9%，而偏少(BN)的機率則減少了 5%。從個數來看大約是 4-5 個左右，落在正常的範圍，因此預測今年侵台的颱風傾向正常偏多。西北太平洋颱風(最大風速>64Kt)個數偏多(AN)的機率增加了 24%，而偏少(BN)的機率則減少了 19%。從個數來看大約是 17 個左右，落在偏多的範圍，因此預測今年西北太平洋的颱風大約是偏多。

## 參考文獻

Chu, P.S. and Zhao, X., 2007: A Bayesian regression approach for predicting seasonal tropical cyclone activity over the central north Pacific. Journal of Climate, vol. 20, p. 4002-4013.

Chu, P.-S., X. Zhao, C.-T. Lee, and 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.

Elsner, J.B. and Jagger, T.H., 2004: A hierarchical Bayesian approach to seasonal hurricane modeling.

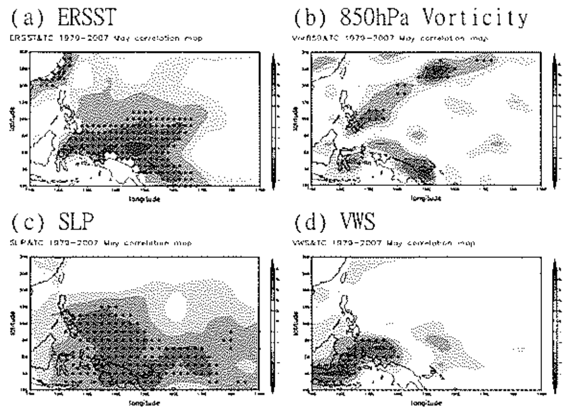
Journal of Climate, vol. 17, p. 2813-2827.

Furrer, R., S.R. Sain, D.W. Nychka, and G.A. Meehl, 2007: Multivariate Bayesian analysis of atmosphere-ocean general circulation models. Environ. Ecol. Stat., in press.

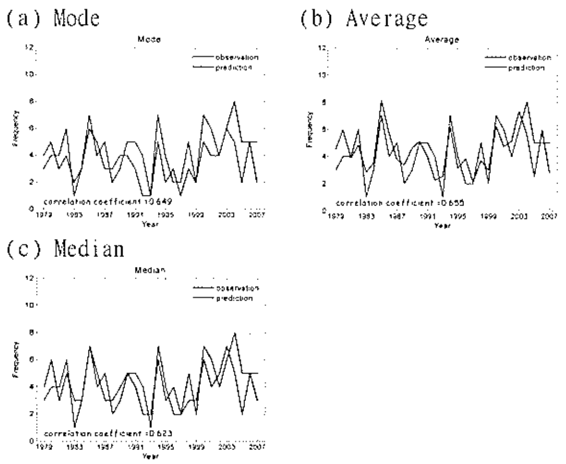
Greene, A.M., L. Goddard, and U. Lall, 2006: Probabilistic multimodel regional temperature change projections. J. Clim., 19, 4326–4343.

Tebaldi, C., L.O. Mearns, D. Nychka, and R. Smith, 2004a: Regional probabilities of precipitation change: A Bayesian analysis of multi-model simulations. Geophys. Res. Lett., 31, L24213, doi:10.1029/2004GL021276.

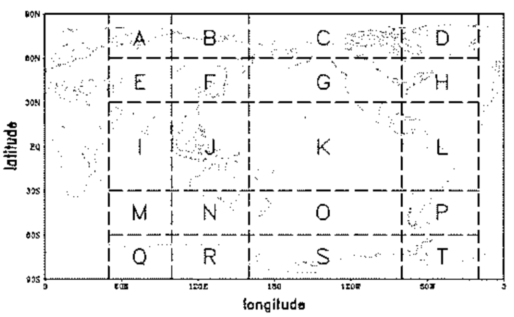
Tebaldi, C., Smith, R.L., Nychka, D., and Mearns, L.O., 2004b: Quantifying uncertainty in projections of regional climate change: A Bayesian approach to the analysis of multimodel ensembles. Journal of Climate, vol. 18, p 1524-1540.




圖(1) 侵台颱風與 5 月各個預報因子的相關係數等值圖及顯著相關格點

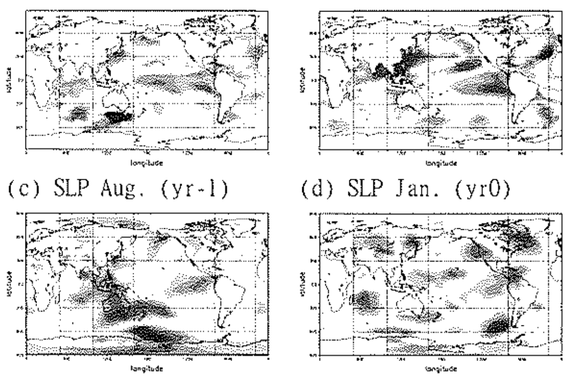



圖(2) 以不同發生機率的推估個數作為預報個數與實際侵台颱風發生個數在時間序列上的分布

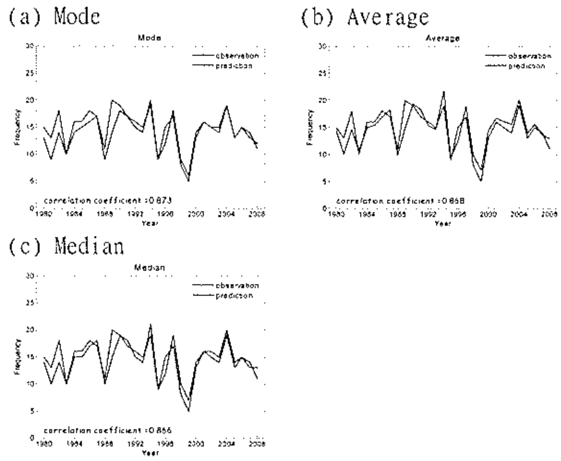



圖(3) 全球劃分成 A-T 區域以方便挑選預報因子

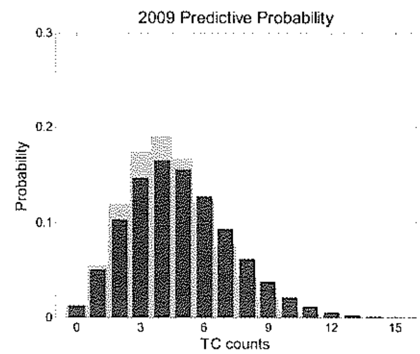
(a) ERSST Apr. (yr-1)　　(b) ERSST Dec. (yr-1)




圖(4) 對西北太平洋每年颱風個數影響性最大的預報因子並建立貝式回歸模式

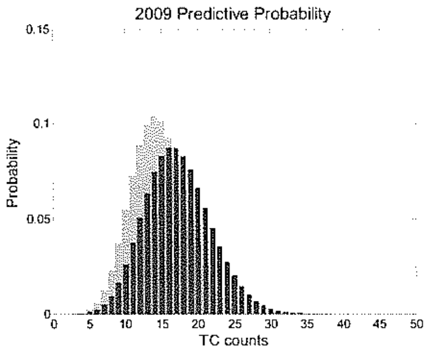



圖(5) 以不同發生機率的推估個數作為預報個數與實際西北太平洋颱風發生個數在時間序列上的分布




圖(6) 2009 年侵台颱風的機率預報結果，背景為 1979-2007 年颱風發生個數的 Poisson 分布，柱狀圖為預報 2009 年侵台颱風在不同個數下的可能發生機

率。




圖(7) 2009 年西北太平洋颱風的機率預報結果，背景為 1980-2008 年颱風發生個數的 Poisson 分布，柱狀圖為預報 2009 年颱風在不同個數下的可能發生機率。