

「適當氣候平均」統計模式應用於台灣季節性預報

李清騰、盧孟明

中央氣象局

摘要

本研究利用氣候平均(Optimal Climate Normals; OCN)的統計模式來進行台灣地區季節性之溫度預報的可行性，使用測站資料為氣象局 12 個局屬測站之長期季平均資料(1941-2006 年)。利用過去的歷史資料來進行統計分析，並且找尋在歷史資料中最佳及最合適的年數(K 年)是 OCN 的核心步驟。分析結果得到台灣地區之溫度季預報，在 4 個季節(MAM, JJA, SON, DJF)當中應有不同的預報年份(K year)，其中春季(MAM)的預報技術得分明顯偏低，所以無法得到合適的 K 值，夏季(JJA)的 K 值為 13，秋季(SON)之 K 值為 23，冬季(DJF) K 值則為 16。利用此 OCN 選取出 K 值可當作台灣地區溫度之季預報的年份，並且其預報技術得分比持續法(前一年資料)或是氣候值(前 30 年平均)有更高的預報得分技術，顯示出此 OCN 對於季預報有相當程度的掌握程度。

一、前言

大氣是一個混沌系統，天氣可預報度的理論上限值在兩週內，兩週以外的長期天氣預報其可預報度需靠緩慢變化的外作用力(External Forcing)影響和大氣低頻擾動，前者可以用聖嬰現象(ENSO)來代表，而後者則是用 Madden Julian Oscillation(MJO)或 Arctic Oscillation(AO)為代表。近年來許多研究顯示出大氣與海洋都存在著週期十年以上的年代際變化，在大氣方面可用 Pacific Decadal Oscillation(PDO)代表，海洋方面則可用 Atlantic Multi-decadal Oscillation(AMO)代表。這些大尺度的變化特徵對不同的區域氣候有什麼影響？是否能藉以改進區域天氣的長期變化之預報？都是氣象學家想要努力解決的問題。

目前常用的統計預報模式有迴歸趨勢模式、正交經驗函數 EOF、奇異值分解法(SVD)、正準相關分析(CCA)等(Barnston 1994, 陳 1998, Lee 2002)，然而這些的統計模式之預報過程較為複雜，且得到的預報結果

並沒有顯著的技術得分，相對於所花費的計算資源並不成正比。由 Huang, J. et al, (1996)提出一個簡單的統計預報方法，適當氣候平均(Optimal Climate Normals; OCN)，此統計方法指出在某些特定季節利用靠近預報時間 K 年的氣候平均值可以得到技術得分不差的預報結果(Wilks, D. S. 1996)。OCN 的主要核心步驟是在找尋在過去資料中最佳及最合適的年數(K 年)，一旦 K 年確定後，則可利用此 K 年的月或季之平均值來當做下一年的預報值。

然而 OCN 因為方法上的限制，所以不能用任意抽出一年當作目標之交錯檢驗(Cross Validation)方

法來驗算其預報技術(Michaelson, 1987)，則是必須要使用一段時期作為發展模式的訓練期而用另一段時期作為預報期，若當資料長度不足時則無法建立 OCN 統計預報模式。在這種的資料長度的限制下，若有 10 年以上時間尺度的年代際氣候型(Climate regime)的轉變，對預報結果的適用性是一大考驗。儘管有這些限制，OCN 由於方法簡單，預報時間可領先長達 12 個月，仍是一個值得開發、利用的統計方法。本文介紹 OCN 統計方法應用於台灣地區溫度之季節性預報，以及其技術得分與檢驗。

二、資料與方法

本文使用中央氣象局之 12 局屬測站(包括台北、新竹、台中、台南、高雄、恆春、花蓮、宜蘭、大武、成功、台東、澎湖測站)之季平均(3 個月平均)溫度資料來進行 OCN 統計預報。資料時間長度為 1941-2006 年，其中以 1941-1970 來當做訓練期(training period)，而 1971-2006 則為預報期(target period)，並且利用此預報期時間來檢驗其預報技術。由於 K 是 OCN 統計預報模式最重要的一個參數，在此介紹 K 值將如何選取。假設有一時間序列 $T_i = 1, 2, 3, \dots, N$ ，其中 N 代表月資料的總年數，則 K 年的氣候平均值計算方程如下：

$$\bar{T}_{i,k} = \frac{1}{k} \sum_{j=1}^k T_{i-j}$$

其中 $k = 1, 2, \dots, 30$; $i = 31, 32, \dots, N$

由方程式可以知道，若全部資料長度為 1941-2006，則 1941-1970 年資料為訓練期，而 i 是代表 1971~2006 的預報資料(預報期)。另外 k 只介於 1 至 30 年之間一個整數，這是因為 k 超過 30 以後則

沒有明顯的預報技術可言(Huang, J. et al; Wilks, D. S. 1996)。而後將計算得到的所有 (共 30 個 k 值之時間序列)後,再經過統計方法的篩選和預報檢驗之後,若能得到 $k = K$ 的預報是有技術的,則表示未來可利用最靠近預報時間的 K 年之月或季的平均值作為預報值,例如最終選取的 k 為 10,則未來的預報值可用前 10 年的氣候平均值來當做預報。

K 的決定步驟為計算每一個 k 的預報與觀測之間的相關係數與均方誤差,然後以最高相關係數及最小均方根誤差的 k 為 K ,即為最適當的氣候平均長度。其相關係數方程式如下:

$$COR_{dep}(K) = \frac{\sum_{i=31}^n \hat{T}_i^f \hat{T}_i^{ob}}{[\sum_{i=31}^n (\hat{T}_i^f)^2 \sum_{i=31}^n (\hat{T}_i^{ob})^2]^{1/2}}$$

其中 $\hat{T}_i^f = \bar{T}_{i,k} - C_{WMO}$
 $\hat{T}_i^{ob} = T_i^{ob} - C_{WMO}$

本研究嘗試過以兩種方式計算 C_{WMO} ,第一種是採取以 10 年為等單隔的滑動 30 年氣候平均值,第二種則是採用固定的 30 年氣候平均值,結果發現兩者結果並無明顯差別,故採用前者之定義。

挑選出最佳及最合適的 K 值之後,需要評估此 K 值的預報結果是否有達到預報技術標準。所以在預報技術得分採用 Heidke Skill Score 之三分類(3:4:3)評估方法(Wilks, D. S, 1995)。若 HSS 值得到為 0 代表沒有預報技術(No Skill),而 HSS 值為 1 則是為最佳預報(Perfect Score)。此外建立一組隨機資料來執行 OCN 統計預報,藉此來評估此預報技術是否比隨機預報更好。隨機預報的結果顯示 K 年介於 1-10 年間的 HSS 值約為 0.2,而 11-30 年則 HSS 值為 0.1,因此若得到 OCN 的預報技術得分大於 0.15 以上,則明顯優於隨機預報結果。

三·溫度季預報結果與分析

圖 1 顯示 4 個季節(MAM, JJA, SON, DJF)台北測站在 1971-2006 年 OCN 預報值與觀測值之相關係數與技術得分(HSS)之 K 年序列圖,圖中實線為相關係數值,虛線則為 HSS,另外若相關係數值大於 0.42 代表此係數有通過 99% 以上的統計檢定,另外 HSS 大於 0.1 則代表此預報比隨機預報有較高的技術得分。從 K 年序列圖中可看到最高的相關係數在 4 個季節中所代表的 K 年均不太相同,不過大致可看到 K 在 10-23 間有較高的相關係數值及 HSS 值,這代表了此 OCN 的預報方法比一般季節性預報所使用之持續法(前一年的資料來當做今年的預報,也就是 $K=1$)及氣候法(前 30 年資料之平均當做今年的預報, $K=30$)均有較高的相關係數及預報技術得分。在 MAM 季節中,由此此季節相關係數明顯較其它季節偏低,但大致可看出在 MAM 季節中 K 值為 13-16 有較高的相關係數及 HSS

值,所以可以得到此 K 值為 13-16 年可代表台北測站 MAM 之合適 K 值,不過 MAM 中的相關係數值並沒有到達 99% 以上的統計檢定。在 JJA 季節中,相關係數值在 K 年中並無明顯差別,而從 HSS 值中可以得到 K 值大於 13 以上都是合適的 K 值。SON 季節中在 K 值為 22-26 有明顯的相關係數高值,且此時的 HSS 也較高,所以 SON 之合適 K 值為 22-26,而冬季中 K 值在 13-16 均有最高的相關係數值與 HSS 值,顯示此合適之 K 值為 13-16。

另外在其它的測站分析方面(圖略),由於測站的地理位置及氣候背景條件不盡相同,所以在各個測站及各個季節均有不同之範圍 K 值。為了找尋可以代表整個台灣地區各季節之合適 K 值,圖 2 顯示 12 個局屬測站之相關係數與 K 年之分佈圖,圖中只顯示測站之相關係數值有達到 99% 以上統計顯著性,由圖 2 可看到 MAM 季節相關係數有統計顯著性的只有高雄測站,且此高雄測站在 K 值在 2-4 與 21-26 有較高的相關係數值,由 HSS 分佈圖來看(圖略),顯示高雄測站最高的 HSS 值在 21-26,故 MAM 季節中 K 值 21-26 代表高雄站合適之 K 值,然而這與台北站(圖 1)之合適 K 值有差異,所以 MAM 季節中仍無法有效找到可代表台灣之合適 K 值。在 JJA 季節中,可以看到 K 值在 10-15 間有較多的測站顯示有較高的相關係數值,而從 HSS 分佈圖中(圖略)也是呈現出 K 值 10-15 間是有預報技術得分。SON 季節則明顯 K 值在 22-26 有最高的相關係數值,且此時 HSS 也有較高的值(圖略),代表 K 值 22-26 為 SON 合適之 K 值。在 DJF 季節,可以看到 K 值在 13-16 有高的相關係數值及 HSS 值(圖略),顯示 DJF 之合適 K 值為 13-16。

由以上的分析可以得到 MAM 季節中無法得到合適 K 值, JJA 季節之合適 K 值為 10-15, SON 則是 22-26,而 DJF 合適 K 值為 13-16。為了更進一步得到合適之 K 值,圖 3 顯示在 4 個季節中,針對 12 個局屬測站中來計算其相關係數有通過統計顯著性的測站次數之 K 值序列圖,例如 MAM 圖中 K 為 1 之相關係數所得到的測站次數是 1,代表此 12 個局屬測站中只有 1 個測站在 K 為 1 時有通過統計的顯著性,此外圖中也加入 HSS 達統計顯著性之測站次數。由圖中可看到 MAM 季節之相關係數與 HSS 的測站次數明顯比其它季節偏少,不過在 K 為 25 時有最多測站數,但此測站數目偏少,所以仍無法得到此季節之合適 K 值。JJA 季節中,可以看到 K 值 10-13,其相關係數有達到統計顯著性之測站數較多,且 K 值為 13 也有最多的測站次數之 HSS 達到顯著性,所以可以得到 JJA 季節之合適 K 值為 13。另外在 SON 季節中,也可得到合適之 K 值為 23。另外在 DJF 季節中, K 值約 4-6 間有最多的測站次數之相關係數通過統計顯著性,但 HSS 達顯著性的次數卻是偏少,反而在 K 值為 16 中 HSS 的測站次數達到最高,其相關係數次之,然而 OCN 最主要目的是找尋合適 K 值來代表整個台灣

地區最佳之季預報的年份，所以最佳的 K 值應採取預報技術得分最好的，故針對 DJF 季節而言最合適之 K 值應為 16。

針對以上所選取之各季節合適之 K 值來分析此 K 值是否合適代表台灣整個地區之 K 值。圖 4 顯示各測站相關係數與 HSS 值之空間分佈圖，圖中正方實心代表此相關係數達統計顯著性，由圖 4a 中可看到此 MAM 之相關係數達統計顯著性有宜蘭、花蓮、台中及高雄測站，但 HSS 則是只有台中和高雄測站達到顯著性，而在其它測站中明顯 HSS 偏低，顯示出此季節無法得到合適之 K 值。圖 4b 之 JJA 季節中，可看到有 7 個測站達統計顯著性，而其它測站除了澎湖及大武測站外也有較高的相關係數值，另外 HSS 值在大部份測站也有達 0.15 以上，顯示 JJA 季節之合適 K 值為 13。另外有 SON 及 DJF 季節中(圖略)，也是呈現出看 JJA 季節相似的結果，其合適 K 值分別為 23 及 16。圖 5 顯示 DJF 之 K 值為 1、13、16、23、30 之相關係數與 HSS 分佈圖，此結果的計算是將相關係數與 HSS 有通過統計顯著性之數值平均，另外 JJA 及 SON 季節則和 DJF 的結果類似，故圖省略，由圖中可看到 K 值在 16 有最高的相關係數及最高的 HSS 值，而 K 為 13 及 23 則是有次高的相關係數及 HSS 值，不過可看到在 K 為 1 及 30 之相關係數與 HSS 最低，這也代表了利用 OCN 所得到 K 年之季預報是比持續法(K=1)及氣候法(K=30)是有較高的預報技術得分。

圖 6 顯示 12 個測站平均之相關係數(a)與 HSS(b)之季節時間序列圖，圖中選取的 K 值為 1、13、16、23、30。由圖中可以看到 OCN 在 MAM 季節之相關係數最低，且預報技術得分也最低，而後隨時間增加其預報技術得分也增加，在 SON 至 NDJ 間有最高的相關係數及預報技術得分，由此可知道 OCN 的預報並非對每個季節都有很高的預報技術得分。另外由圖中也可以看到利用氣候法(K=30)之預報其相關係數及預報技術得分是最低的。

四·結論

本研究利用台灣地區季平均資料(1941-2006 年)之 12 個局屬測站溫度資料來使用適當氣候平均(OCN)統計模式進行溫度之季預報。在歷史資料中以 1941-1970 來當做訓練期，而 1971-2006 則為預報期，並且利用預報技術得分(HSS)來檢驗此預報結果。研究結果得到在 4 個季節(MAM、JJA、SON、DJF)均有不同的 K 值可代表台灣地區溫度之季預報，不過由於 MAM 的預報技術得分過低，所以 MAM 季節仍無法有效找到可代表台灣之合適 K 值。而另外在 JJA、SON 及 DJF 的季節中，其合適之 K 值為 13、23 及 16，且此 K 值均比其它的季預報方法(如持續法、氣候法)均有較高的預報技術得分。

由於目前在動力或統計模式的預報檢驗方面，都是採取使用技術得分的方式來判定模式的預報技

術，由於技術得分是根據觀測與預報計算而得到，因此預報得分的多寡並不直接代表模式的好與壞，只能代表模式無法掌握此觀測現象，若有 OCN 的統計預報方法來當作基礎點(benchmark)，任何的預報模式都應比 OCN 的預報結果有較高的技術得分，則可說明此預報模式是有預報價值，因此更可以明確的了解預報模式的真正價值。未來將分區域性來進行 OCN 之季節性預報，並且評估此區域性的預報技術是否更好。

五·參考文獻

- 陳雲蘭，1998：應用主成份迴歸分析發展長期預報。天氣分析與預報研討會論文彙編(87)，243-247。
- Barnston, A. G., 1994: Linear statistical short-term climate predictive skill in the northern hemisphere. *J. Climate*, **5**, 1513-1564.
- Lee, Yung-An, 2002: A T-EOF based prediction Method. *J. Climate*, **15**, 226-234.
- Michaelson, J., 1987: Cross-Validation in statistical climate forecast models. *J. Appl. Meteor.*, **26**, 1589-1600.
- Huang J., Huug M. van den Dool and Anthony G. Barnston. 1996: Long-Lead Seasonal Temperature Prediction Using Optimal Climate Normals. *J. Climate*, **9**, 809-817.
- Torrence, C. and C. P. Compo, 1998: A Practical Guide to Wavelet Analysis. *Bull. Amer. Meteor. Soc.*, **79**, 61-78.
- Wilks. D. S. 1995: *Statistical Methods in Atmospheric Sciences*. Academic Press, 467pp.
- Wilks. D. S. 1996: Statistical Significance of Long-Range "Optimal Climate Normal" Temperature and Precipitation Forecasts. *J. Climate*, **9**, 827-839.

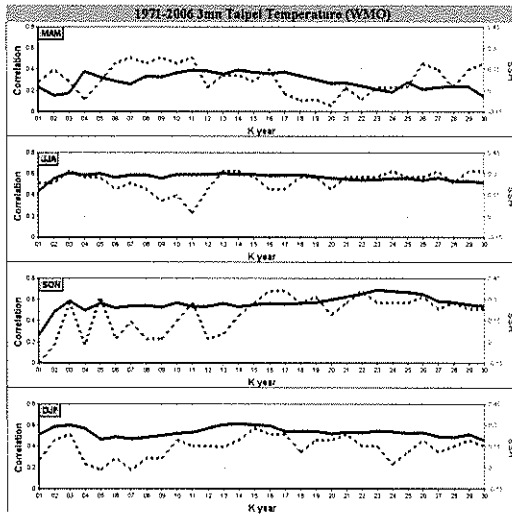


圖1. 台北測站1971-2006年4個季節之相關係數與HSS之K值序列圖。圖中實線為相關係數值，點線為HSS值，X軸代表K值，y軸代表相關係數值(左)及HSS值(右)。

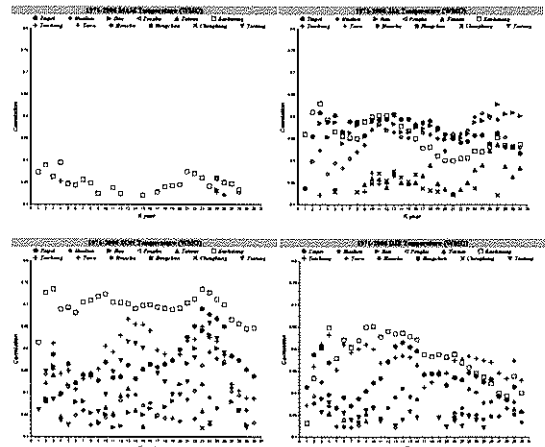


圖2. 4個季節12個局屬測站的相關係數之K值序列圖，圖中各個標誌代表各個測站。圖中X軸為K值，Y軸為相關係數值。

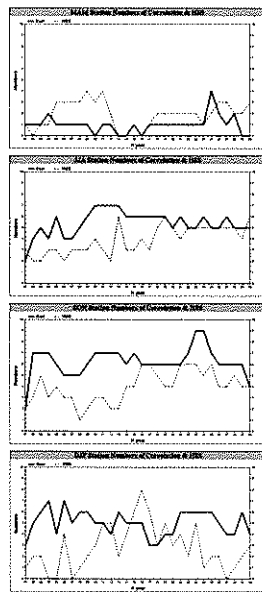


圖3. 4個季節相關係數與HSS達統計顯著性的測站次數之K值序列圖。圖中實線代表利用相關係數求得的測站數，虛線則是HSS，X軸代表K值，Y軸代表測站數。

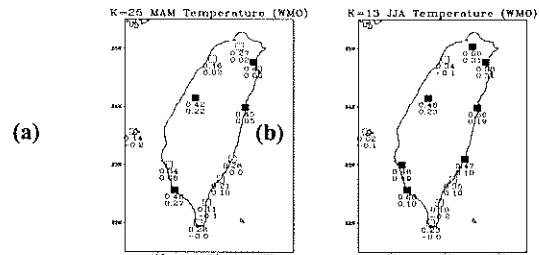


圖4. MAM(a)、JJA(b)之K值分別為25與13的相關係數與HSS分佈圖。圖中實心正方形代表此相關係數達統計顯著性，空心則代表沒有通過統計檢定，正方形下第一個數字代表相關係數值，第二個則是代表HSS值。

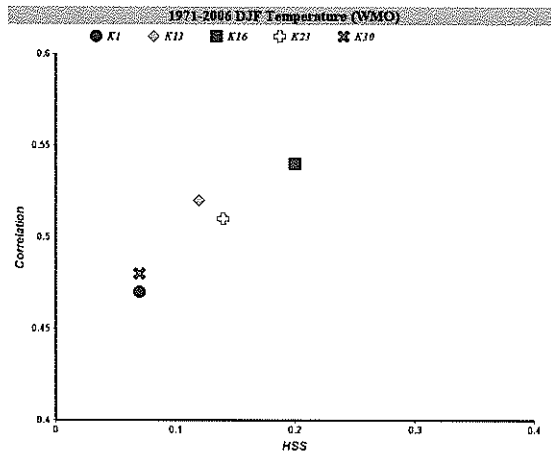


圖5. DJF季節K為1、13、16、23、30之相關係數與HSS分佈圖。圖中不同的K值有不同的標誌。

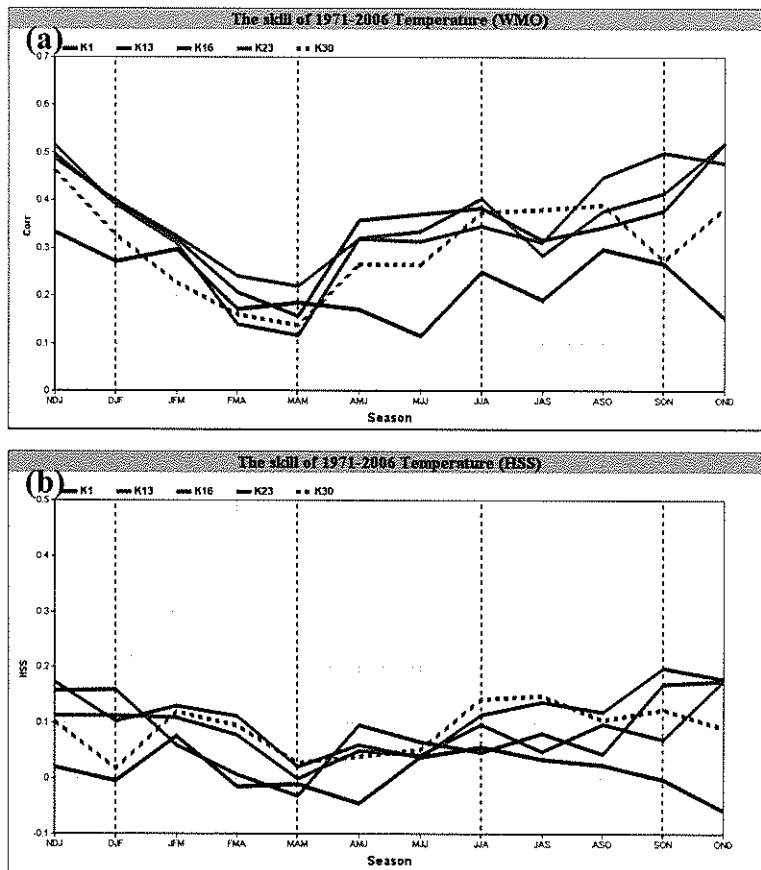


圖6. 不同K值的相關係數(a)與HSS(b)之季節時間序列圖，圖中K值為1、13、16、23、30，黑實線為K=1，紅實線為K=13、藍實線為K=16、綠實線為K=23、黑點線為K=30，X軸代表季節，Y軸代表相關係數值及HSSP值。

