

台灣極端降雨氣候事件辨識方法之分析

林昀靜 盧孟明

中央氣象局科技中心

摘要

極端降雨氣候事件辨識方法是建立在降雨事件的統計特性也就是水文頻率分析的基礎上。陳與盧(2007)利用皮爾森第三類分布進行雨量資料機率分布模擬時,在較為極端的雨量部份有低估的情形,顯示有必要分離極端部分的雨量利用極端值統計分析法進行頻率分析。本文將根據中央氣象局 25 個測站的時雨量資料以箱型計數法組成 1、3、6、12、24、48 及 72 小時延時的降雨事件,再根據盒鬚圖法(box-plot)挑選出七種不同延時之極端事件。利用線性動差比圖的分析結果討論廣義極端值分布(Generalized Extreme Value, GEV)和廣義柏拉圖分布(Generalized Pareto Distribution, GPD)的適用性。結果顯示以 GPD 來模擬極端降雨事件的母體機率分布是較為恰當的。以 GPD 推估重現期距為 1.5、2、5 及 10 年的雨量閾值,發現重現期距大於 2 年之降雨事件即可視為有致災可能的降雨事件。

一、前言

氣候變遷的議題在近年來受到相當的重視,氣候變遷可能造成溫度的升高和海平面高度的上升,這將可能造成水資源缺乏及沿海地區與島國面臨土地被淹沒的危機,這都直接影響到我們的生活。造成氣候變遷的因子相當多,國內外也有許多的學者針對這個議題進行長期且詳細的研究,政府單位亦成立相關的部門,希望社會大眾也可以重視此項課題。

本文的主要重點在探討台灣地區的極端降雨,以極端降雨氣候事件的統計特性建立辨識方法。由於陳與盧(2007)利用皮爾森第三類分布進行模擬時,在較為極端的雨量部份有低估的情形,顯示有必要分離極端部分的雨量來進行極端值的水文頻率分析。

一般而言,極端值理論有極值序列(extreme value series, EVS)及部分延時序列(partial duration series, PDS)選用法兩種主要的模型。其中極值序列選用法,通常選用固定一段期間內的最大值(如年最大值, annual maximum series, AMS)後再利用廣義極端值分布(Generalized Extreme Value Distribution, GEV)來表示,此一分布相當適合用來描述這些區間極大值。而部分延時序列選用法為選擇超過某一門檻值的資料序列,通常用廣義柏拉圖分布(Generalized Pareto Distribution, GPD)來表示,其分布的核心在於描述超過門檻值的餘額尾部分布。(Madsen et al., 1997)

二、資料

陳與盧(2007)詳細說明了利用皮爾森第三類分

布對台北、台中、台南、花蓮及台東等五個測站,進行極端降雨的頻率分析及其判定方法,本研究為了改進其分析結果,除了上述五個測站外,亦另外分析了中央氣象局的 20 個氣象觀測站。各站所採用的時雨量資料從 1951 年開始至 2005 年止,若測站紀錄不是從 1951 年開始,則採用的時雨量紀錄從各站有紀錄開始至 2005 年止。同樣的,本研究亦利用箱型計數法(Olsson et al., 1992)將各測站之時雨量時間序列組合成 7 種延時(1hr、3hr、6hr、12hr、24hr、48hr、72hr)之降雨事件並進行分析。為了避免小雨事件對極端事件分析的影響,因此亦濾除了時雨量小於 5mm/hr 的降雨事件。

三、方法

(一) 極端值的挑選

盒鬚圖(Box plot)是一個精確顯示分布資料的繪圖方法,其可用以檢驗資料的極端量數及分配的型態,其示意圖如圖 1 所示。盒鬚圖主要是繪製中位數(即第 50 的百分位數)、第 25 的百分位數(25%,即下四分位數)、第 75 的百分位數(75%,即上四分位數)等,盒子的下界限是第 25 的百分位數(Q1),上界限是第 75 的百分位數(Q3),盒子的長度即是代表第 75 的百分位數與第 25 的百分位數之差值($L=Q3-Q1$)。盒中的星號 * 代表中位數,通常盒中包含有 50%的觀測值,因此盒子越大,則表示觀測值散布越大。由盒子上下界所延伸出的線,即稱之為鬚(whisker),是用以連接觀測值(非界外、極端值)的最大值與最小值。此外,盒鬚圖中亦可顯示出資料的界外值(outlier)與極端值(extreme)。通

常其界定是以任何觀察值與盒端的距離介於 1.5 至 3 倍盒長 (1.5~3L) 之間者，即稱之為界外值，而超過 3 倍的盒長 (>3L) 者，即稱之為極端值。

(二) 頻率分析

1. 基本概念

頻率分析是按統計機率分布推估某一特定水文氣象量值在一段時期所發生機率之方法，根據王與易 (1999) 所述，頻率分析之程序包含以下四點：(1) 水文資料之處理、(2) 資料序列之選擇、(3) 重現期距、(4) 機率點繪。

頻率分析之主要目的為決定某指定量值之事件發生之機率或其重現期距 (return period)，而其核心在於建構各種理論機率分布，統計機率分布的種類甚多，在過去常用之統計分布方法有廣義極端值分布、極端值第一型分布、廣義柏拉圖分布、皮爾森第三類分布及對數皮爾森第三類分布等。其參數推估法常用的有最大概似法 (maxima likelihood estimation)、動差法 (method of moments) 及線性動差 (L-moment) 等，因此統計機率分布的選擇及參數的推估是相當重要的。(盧與陳，2005；Koutsoyiannis and Baloutsos, 2000；Katz, R. W. et al., 2002；Madsen et al., 1997)

重現期距習慣以排序定點的方法估計，本研究採用常用之威伯法 (Weibull) 估計，如公式 (1) 所示，其中 T 為重現期距， N 為紀錄年數， m 為事件大小順序 ($m=1$ 為最大)。

$$T = \frac{N+1}{m} \quad (1)$$

2. 廣義柏拉圖分布

廣義柏拉圖分布其累積機率分布與等分位函數為下式所示：

$$F_{\xi, \sigma, \mu}^{\xi}(x) = 1 - \left(1 + \frac{\xi(x - \mu)}{\sigma}\right)^{-\frac{1}{\xi}} \quad (2)$$

$$x(F) = \mu - \frac{\sigma}{\xi} \left[1 - (1 - F)^{-\xi}\right] \quad (3)$$

其有三個分布參數， ξ 為形狀參數 (shape parameter)、 σ 為尺度參數 (scale parameter)、 μ 為位置參數 (location parameter)，其中位置參數 μ 相當於門檻值。

對於極端值分布而言，當門檻值 μ 夠大時，超過 μ 的餘額分布可由 $G_{\xi, \sigma}(y)$ 近似，其中

$y = x - \mu$ ，即

$$G_{\xi, \sigma}(y) \approx G_{\xi, \sigma}(x - \mu) = F_{\xi, \sigma, \mu}^{\xi}(x)。$$

$$G_{\xi, \sigma}(y) = 1 - \left(1 + \frac{\xi y}{\sigma}\right)^{-\frac{1}{\xi}} \quad (4)$$

3. 機率點繪

點繪法為水文頻率分析常用方法之一，主要是利用一個適當的點繪公式將觀測資料與理論機率分布點繪於適當機率紙上，其用意在於檢視理論機率分布與採用之參數是否適合描述樣本母體的統計特徵。首先須由實際樣本資料排序定點得到的重現期距關係和反推之理論值。本研究中實際樣本資料之重現期距以常用的威伯法估計，為了反推求在此重現期距中之廣義柏拉圖分布理論值，因此需要經由重現期距與累積機率的轉換關係，再代入式 (3)，即可得到在此重現期距中之廣義柏拉圖分布理論值。由於本研究之廣義柏拉圖分布不是採用一年一個最大值進行分析，因此須將資料長度轉換成以年為單位，其重現期距與累積機率的關係如式 (5) 所示，其中 M 為總事件數， N 為記錄年數。(Madsen et al., 1997；Wilks, 2005)

$$T = \frac{1}{\omega(1 - F(x))} \quad (5)$$

$$\omega = \frac{M}{N} (\text{yr}^{-1}) \quad (6)$$

四、應用與結果

(一) 極端值的挑選與機率分布的選取

由於本研究所採用的測站觀測雨量資料有限，若利用極端值理論當中的極值序列選用法選取年最大值的資料，其形成的集合個數就會較少，因此應用於極端事件的模擬就會有偏估的情形，但若採用部分延時序列選用法，就能更加有效的應用這些有限的雨量資料。因此本研究利用盒鬚圖法分別針對 7 種延時之降雨進行極端值之挑選，由於本研究是考慮極端大雨量的影響，因此將與盒端 Q_3 的距離大於 3 倍 L 之降雨事件挑出，即圖 1 中顯示為 extreme 的部份，為了避免所挑選出之極端降雨來自同一極端降雨事件，而重複被選擇，因此以三天為一劃分基準，若極端雨量在連續三天內皆發生，則表示極端雨量可能來自同一極端降雨事件，因此將其視為同一極端事件，並以此極端事件中之最大值作為極端事件 (extreme event) 代表值。

根據 Hosking (1997) 所提出之以線性動差比例所繪製的線性動差比圖 (L-Moment ratio)，由於不同的機率分布具有不同型態的動差比圖，因此可經由動差比圖判斷樣本資料的機率分布型態。圖 2 為根據本研究所採用的 25 個測站 7 種延時之極端降雨事件 (25 x 7 組資料) 所計算出之線性動差比例以及其平均值，將其與各種機率分布的線性動差比圖做比較，可發現其分布的型態及位置都相當接近廣義柏拉圖分布 (GPD)，因此本研究採用廣義柏拉圖分布來模擬極端降雨事件是適合的。

(二) 利用廣義柏拉圖分布定義之降雨分級閾值

若採用台北、台中、台南、花蓮及台東相同的五個測站和陳與盧(2007)利用皮爾森第三類分布模擬的結果進行比較,圖3顯示了皮爾森第三類分布對此五個測站七種延時之機率點繪圖與本研究採用的盒鬚圖法所挑選出極端值事件之門檻值,此門檻值以虛線表示,可看出在門檻值以上的降雨事件利用皮爾森第三類分布模擬的確有低估的情形。

本研究將極端降雨事件利用廣義柏拉圖分布進行模擬時,分布中之參數 μ ,採用前述盒鬚圖中Q3加上3倍L的值作為門檻值,其餘兩參數 ξ 、 σ 則可利用最大似法推估計算出,再經由機率點繪圖的方式呈現其分布的適當性與否,其機率點繪結果如圖4所示。在台北及台中測站,利用皮爾森第三類分布模擬的結果在極端降雨的部分有明顯低估的情形,而本研究採用的廣義柏拉圖分布模擬極端事件的結果,由機率點繪圖可以清楚的看出遠比皮爾森第三類分布來的更接近實際狀況。

將相同的頻率分析法應用於25個測站,可以整理出25個測站7種延時所對應之重現期距為1.5、2、5及10年之極端降雨分級閾值,如圖5所示。其中測站1~25分別代表淡水、基隆、台北、新竹、梧棲、台中、嘉義、台南、高雄、恆春、大武、台東、成功、花蓮、蘇澳、宜蘭(以上為平地測站,順序為由北至南再到東部)、竹子湖、鞍部、日月潭、阿里山、玉山(以上5個測站為高山站,順序為由高度低至高)、彭佳嶼、東吉島、澎湖、蘭嶼(以上4個測站為外島站)。圖中虛線部分是用來區隔平地、高山及外島站。

(三) 可能致災降雨事件的定義

接下來探討之問題為如何利用各測站不同的降雨分級閾值定義出可能致災之降雨事件。根據中央氣象局所定義之四種豪大雨分別為:(1)大雨事件為24小時累計降雨量達50mm以上,且其中至少有一小時雨量達15mm以上,(2)豪雨事件為24小時累計降雨量達130mm以上,(3)大豪雨事件為24小時累計降雨量達200mm以上,(4)超大豪雨事件為24小時累計降雨量達350mm以上。

各測站24小時延時所對應之重現期距為1.5、2、5及10年之極端降雨分級閾值如圖6所示,圖中亦標出中央氣象局所定義之4種大雨事件。以平地測站而言,以中央氣象局所定義的大豪雨事件(200mm/24hr)約是在2年重現期距左右,因此定義大於兩年重現期距之降雨為可能致災的降雨事件。圖7是以各站不同延時下重現期距為1.5、2、5及10年之極端降雨分級閾值視為一個門檻值,分別檢視大於4種重現期距之極端降雨事件所發生之頻率。

影響台灣極端降雨之重要的天氣系統為颱風及

梅雨,因此首先討論在大於重現期距2年之可能致災降雨事件中,有多少個數是受到由颱風或非颱風所造成之影響,判定是否受到颱風影響之判別方法為:若是在可能致災降雨事件發生的同時有颱風侵台,則將此一可能致災降雨事件視為受颱風影響。如圖8所示,各測站極端降雨事件受颱風影響的程度反映在個數上,短延時的極端降雨事件受到颱風的影響性較小,在各測站當中又以高山站受到颱風的影響較大,外島站則是以彭佳嶼受到颱風的影響最大。

(四) 極端降雨分級閾值之應用—柯羅莎颱風個案分析

本研究提出之極端降雨事件辨識方法與分級閾值的建立,可分析各種氣候事件所造成降雨的極端程度。以2007年的柯羅莎颱風(KROSA)為例,我們擷取了在颱風發生期間(2007/10/05-2007/10/07)7種延時的極大降雨量進行分析,結果如圖9及圖10所示。

圖9是依照各個測站不同的重現期距將降雨分成5個等級,分別為(1)等級1:降雨量大於10年的重現期;(2)等級2:降雨量大於5年、小於10年重現期;(3)等級3:降雨量大於2年、小於5年重現期;(4)等級4:降雨量大於1.5年、小於2年重現期;(5)等級5:降雨量小於1.5年的重現期。由圖中可以清楚的看到柯羅莎颱風對台灣的山區及西北部地區造成較大的降雨,而對東南部的影響則較小。

圖10以台北、台中、台南、花蓮及台東5個測站作為代表,背景圖是根據各個測站在不同延時下重現期距為10、5、2及1.5年的分級閾值所繪,折線的部份即為柯羅莎颱風所帶來的雨量。

五、結論與建議

本文提出以GPD來模擬全台25個測站極端降雨事件的分布,發現利用此分布可以合理的模擬各站七種延時極端降雨事件的雨量母體機率分布,並推估各測站各延時在重現期距為1.5、2、5及10年的雨量分級閾值,作為判斷極端降雨的標準,並以2007年的柯羅莎颱風為例,探討此颱風帶來的降雨量的極端程度。未來將從梅雨、颱風所帶來的極端降雨,進行雨量和頻率分析,以建立更完善的極端降雨監測系統。

參考文獻

- 王如意、易任,1999,應用水文學(下冊),國立編譯館。
陳佳正、盧孟明,2007,台灣極端降雨氣候事件判定方法,大氣科學,35,105-118頁。
盧孟明、陳佳正,2005,豪大雨之頻率分析方法,氣

象學報, 45-60 頁。

盧孟明、陳佳正、林昀靜, 2007, 1951-2005 年台灣極端降雨事件發生頻率之變化, 大氣科學, 35, 87-104 頁。

Hosking, J.R.M. and J.R. Willis, 1997. Regional Frequency Analysis: An Approach Based on L-Moments. University of Cambridge, 224.

Katz, R. W., Parlange, M. B., and Naveau, P., 2002. Statistical of extremes in hydrology. Advances in Water Resources, 25, 1287-1304.

Koutsoyiannis, D. and Baloutsos, G., 2000. Analysis of a Long Record of Annual Maximum Rainfall in Athens, Greece, and Design Rainfall Inferences. Natural Hazards, 29, 29-48.

Madsen, H., Rasmussen, P. F., and Rosbjerg, D., 1997. Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events. 1. At-site modeling. Water Resources Research, 33(4), 747-757.

Wilks, D. K., 2005. Statistical Methods in the Atmospheric Sciences. Academic Press, 2nd Ed., 71-129.

為各種類分佈: E-exponential, G-Gumbel, L-logistic, N-Normal, U-uniform, GLO-generalized logistic, GEV-generalized extreme-value, GPD-generalized Pareto, LN3-lognormal, 以及 PE3-Pearson type III。OLB 為 τ_4 與 τ_3 方程式關係之整體下限。而+標記為 25 個測站七種延時資料計算結果, ×標記即為此些結果之平均值。

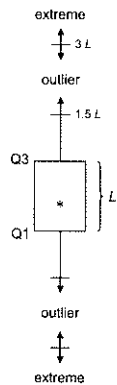


圖 1 盒鬚圖之示意圖

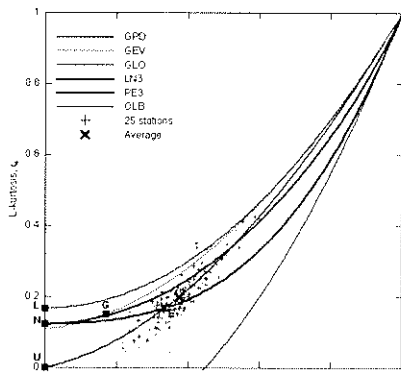


圖 2 L-moment ratio 示意圖; 圖中標示字母分別

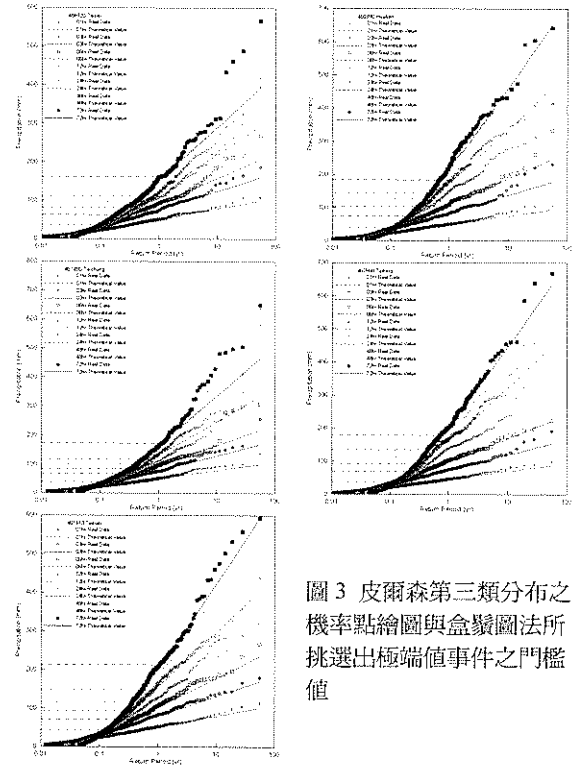
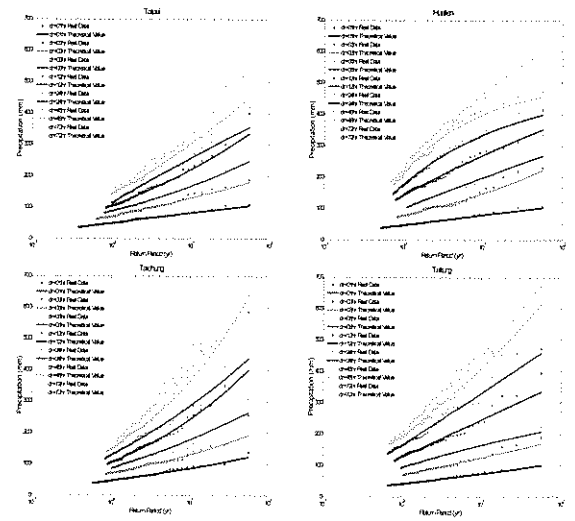


圖 3 皮爾森第三類分佈之機率點繪圖與盒鬚圖法所挑選出極端值事件之門檻值



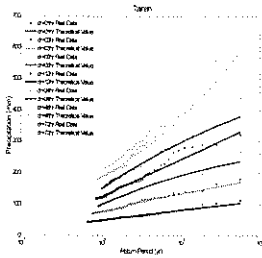


圖 4 廣義柏拉圖分布應用於台北、台中、台南、花蓮及台東之極端事件機率點繪圖

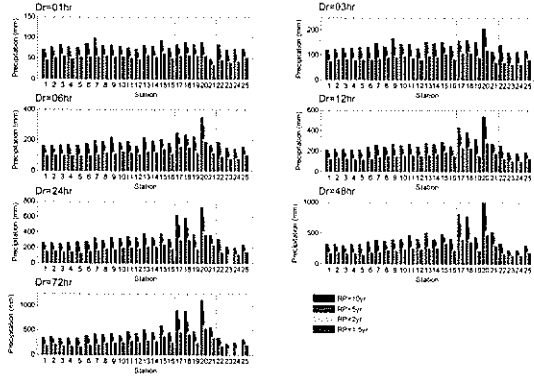


圖 5 25 個測站 7 種延時所對應之重現期距為 1.5、2、5 及 10 年之極端降雨分級閾值

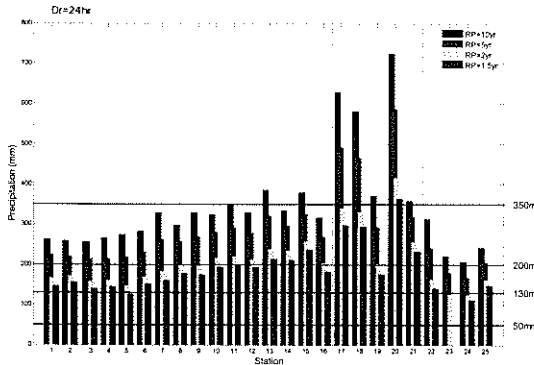


圖 6 25 個測站 24 小時延時所對應之重現期距為 1.5、2、5 及 10 年之極端降雨分級閾值與中央氣象局所定義之大雨事件

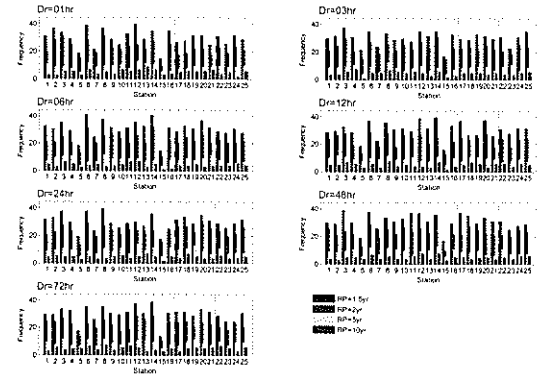


圖 7 25 個測站 7 種延時大於重現期距為 1.5、2、5 及 10 年之極端降雨發生頻率

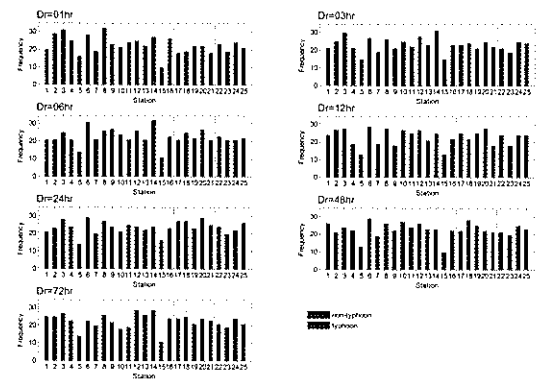
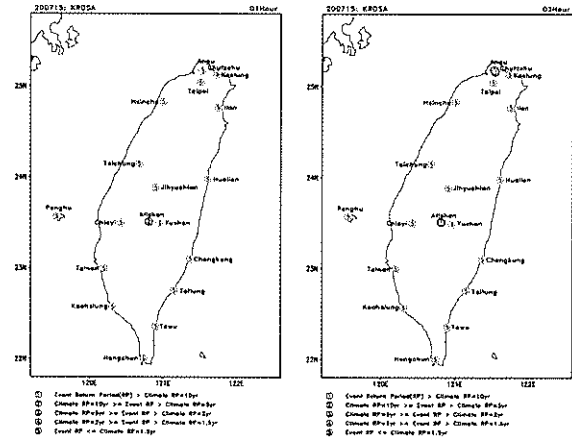


圖 8 重現期距大於 2 年之極端降雨事件中，受颱風及非颱風影響之發生個數



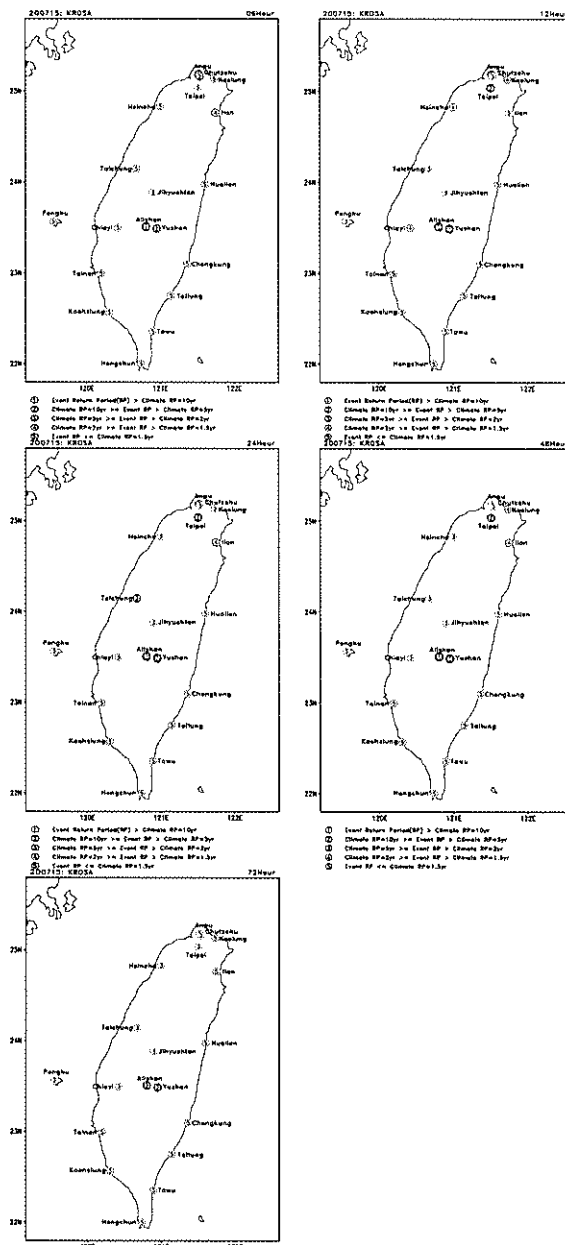


圖 9 各測站不同延時在 2007 年柯羅莎颱風侵台期間之降雨分級

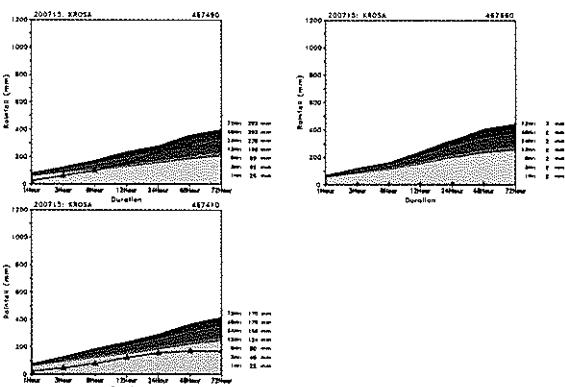
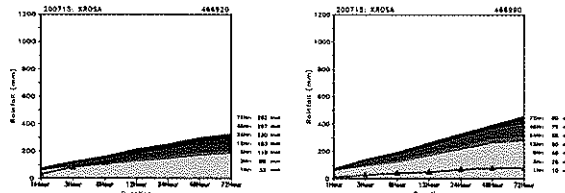


圖 10 以柯羅莎颱風的降雨量對應各測站極端降雨分級閾值