

中央氣象局110年天氣分析與預報研討會

臺灣北部區域應用自組織映射圖分析歷史短延時雨量之分群樣態

李文生¹ 洪國展¹ 郭雅燼² 于芃¹
多采科技有限公司¹ 多采工程顧問有限公司²

摘 要

臺灣的洪水或淹水災害主要驅動因子是降雨，即使地表因人為開發或相關治理工程的建設，可能改變淹水狀況，但發生和以往導致水災的類似降雨分布和量值，仍需關注並注意可能對地表造成的影響。

本研究蒐集民國87~107年之颱風豪雨事件雨量資料，以及期間內大台北地區19場淹水紀錄資料，探討淹水事件6小時短延時累積降雨之空間分佈狀況。針對歷史雨量事件採用SOM(Self Organizing Map, SOM, 自組織映射圖)影像分群智慧化技術，依照歷史雨量事件所提取的特徵進行分群，做為判斷淹水之參考。在即時應用上，以目前雨量事件配合空間分佈相關係數的比較，挑選出空間分佈相似之歷史雨量事件分群，進而找出歷史上空間分佈相近的樣本事件。

109年研究歷史資料分群以全臺灣範圍進行計算，分群結果同時受到南部或東部的降雨分佈影響，因此實際發生於北部之降雨事件未能搜尋得最佳分群。本研究將目標空間由全台縮小至臺灣北部區域，使其能夠凸顯北部區域降雨特性，提出更適用且符合北部之降雨分群結果。由108年0722豪雨及106年0602豪雨之歷史淹水事件模擬比較可知，降雨分群範圍縮小為目標區域，可更佳凸顯北部區域降雨特性，並由類似事件得出可能致災區域，提供防災參考。

關鍵字：SOM，自組織映射圖，淹水，短延時

一、前言

臺灣的洪水或淹水災害主要驅動因子是降雨。降雨落於地表上來不及排水(如：下水道容量不足、都市開發過度等)，則直接在地表上漫淹。

暴雨淹水發生區域與深度推估，常見方式可分為以下兩種：(1)建置物理模型，透過質量和力學平衡合理推演水流傳輸狀況。模型參數採用有限歷史事件檢定驗證，參數描述能力受限於檢定之歷史事件，但可適度反映物理機制；(2)透過統計、類神經等技術推測，計算時間短、效率高，但由於歷史淹水事件有限，可供統計或訓練之資料不足，若模擬結果不符期待，難以採用物理機制協助判斷和調整。

考量降雨為淹水主要影響因子，李文生等(2020)^[5]將智慧化技術應用於降雨趨勢研判分析上，針對歷史颱風事件降雨資料進行智慧化模式訓練，透過智慧化辨識歷史相似降雨趨勢，找出降雨特徵及相關性。降雨空間分布以SOM(Self Organizing Map, SOM, 自組織映射圖)影像智慧化技術，進行降

雨特徵分群，並以目標雨量挑選出空間分佈相似之歷史雨量事件分群及樣本事件。歷史資料分群以全臺灣範圍進行計算，在目標區域為大台北地區時，分群結果因為受到南部或東部的降雨分佈影響，因此實際發生於北部之降雨事件未能搜尋得最佳分群。

為探討目標區域採用之合適歷史分群資料，本研究將歷史資料分群由全臺灣範圍縮小至臺灣北部區域，使其能夠凸顯北部區域降雨特性，提出更適用且符合北部之降雨分群結果，並配合樣本事件所預先進行如淹水數值模式的計算結果，快速組合提供淹水潛勢機率，降低物理模式計算耗時問題，提供更佳之淹水潛勢結果。

二、研究方法

(一)選用模式

本研究選用之影像群集分類智慧化技術SOM為前饋式、非監督式學習的特殊形式類神經網路，由Kohonen於1982年提出，優點為可將N維度的資料映

射(mapping)至二維空間上，並維持資料中的拓樸特性^[3]。其次，可透過調整權重係數，促使神經網路收斂。

SOM神經網路是兩層前饋神經網路，僅含輸入層與輸出層，兩層間各神經元是連接的，網路中沒有隱含層，輸入層含有 m 個神經元，神經元個數與輸入樣本的特徵數相同，而輸出層有 n 個神經元，神經元間與鄰近的神經元兩兩相接，架構如圖1。SOM演算^[4]之主要精神如圖2所示，包含三個重要程序：競爭(competitive)、合作(cooperative)及調整(adaptive)。圖2中，藍色區塊為資料實際分布情況，方形網格為自組織映射圖的神經元權重，各網格都代表一個神經元權重。首先，在競爭過程中，可針對某一輸入資料(圖2中白點)，找出與該輸入資料最近似的神經元，視之為優勝神經元(圖2中黃點)。接著，在合作過程中，根據前述過程找出之優勝神經元，定義鄰近拓撲區域中神經元間相互的鄰近關係。最後，在調適過程中，依據合作過程中決定之拓撲鄰域，調整拓撲鄰域中所有神經元的權重值，使優勝神經元權重與拓撲鄰域神經元權重均將朝該輸入資料的方向接近。

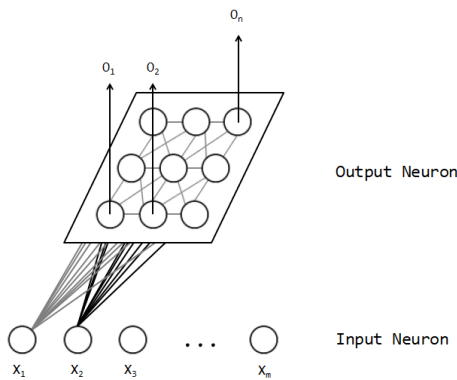


圖 1 SOM 架構

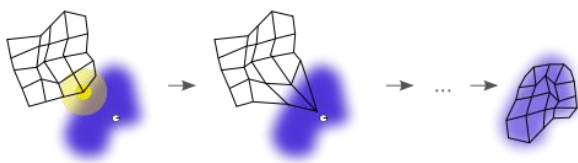


圖 2 SOM 原理示意

(二)國內相關研究及應用

王潔如等(2018)[6]以2012年蘇拉颱風之實際案例，完成以SOM為基礎之系集定量降水預報群集分析技術的測試，結果顯示，使用不同設定下SOM獲取之群集分析結果，均有助於預報參考依據。使用2012年至2017年20場劇烈降兩個案資料進行敏感度測試與系統性分析，結果顯示，群集分析技術可正確分辨不同特徵之系集定量降水預報，明確且合理地將不同降雨事件的系集定量降水預報資料分至不同類別，且使用群集分析加值後某一群集之系集平

均，相較於未分類前之所有系集成員平均，有八成機率得到較佳之定量降雨預報結果。

陳儒賢等(2011)^[7]應用SOM網路與線性動差法於台灣地區乾旱頻率分析之研究，選用全台灣81個雨量測站，將其乾旱特性資料與地文因子經由SOM網路分成六個群集，再經由以線性動差為基礎的不一致估量及異質性估量，評估結果同一群集之乾旱資料皆具有一致性及均一性，最後，分別推估各群集重現期距分別為5、10、20、50及100年之頻率乾旱量及乾旱延時。

陳宜欣(2011)^[8]以SOM為基礎建置日雨量繁衍模式，使用大氣環流模式(GCM)大尺度大氣因子模擬結果作為預測因子，推估未來氣候情境下小尺度地面降雨變化趨勢。首先使用SOM群聚分析法對大氣因子作分類，所分出之每個類別即為一種天氣型態，再計算各天氣型態的統計特性，以建構雨量繁衍模式。

(二)樣本資料處理

本研究依上述理論進行實作，蒐集歷史颱風事件之雨量資料，並將歷史事件之每6小時累積雨量空間分布作為一個樣本，降雨取樣概念如圖3，若該事件長度為24小時，將每6小時累積降雨分布取出作為樣本，分別為01~06時、02~07時、03~08時、...以此類推，至最後一筆樣本為19~24時，共計有19筆樣本。

此外，為能增加訓練樣本數量，亦將歷史颱風事件每6小時累積降雨，以1.5、2.0的倍數進行樣本擴充，配合所蒐集之歷史颱風事件時間段，共組成 $4,125 \times 3 = 12,375$ 個6小時降雨事件樣本。

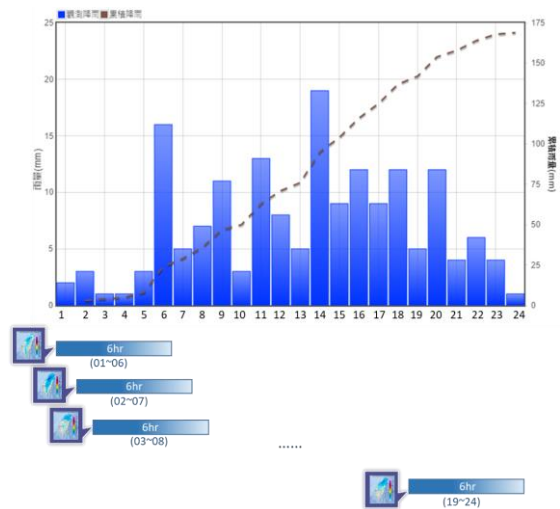


圖 3 歷史事件降雨資料之取樣概念

(三)降雨分群範圍精進

109年研究歷史資料分群以全臺灣範圍進行計算，分群結果同時受到南部或東部的降雨分佈影響，因此實際發生於北部之降雨事件未能搜尋得最佳分群。本研究將目標空間由全臺縮小至臺灣北部區域，使其能夠凸顯北部區域降雨特性，目標空間如圖4，資料範圍網格由全臺20,367個變為4,125個。

於所蒐集歷史颱風事件時間段之12,375個6小時降雨事件樣本中，定義用於分群的有效樣本為至少要有一個網格其6小時累積降雨大於等於6 mm以上，並將未符合此條件之樣本予以剔除，最後所採用的有效樣本數量為10,300筆。



圖 4 降雨分群範圍精進

三、降雨分群訓練結果

本研究以蒐集之降雨事件資料，利用智慧化技術進行分群模型的訓練。首先對降雨資料進行特徵提取，並利用所提取的特徵將具有類似空間分佈的降雨歸於同類，同類別中仍包括多筆降雨資料，降雨資料空間分佈群集分類流程如圖5。

雨量分群訓練成果如圖6，其色階表示將被分為同一群之樣本進行逐網格出現次數進行加總，並除以該分群內總樣本數，以描述在該分群所屬樣本中，各網格的降雨出現機率，用以呈現分群後的空間降雨分佈型態。若某一分群包含400個樣本，其中有360個樣本記錄某一網格具有降雨6 mm以上，則該網格便計算為0.9(360/400)，若對另一網格記錄為0個樣本，則該網格變計算為0.0(0/400)。

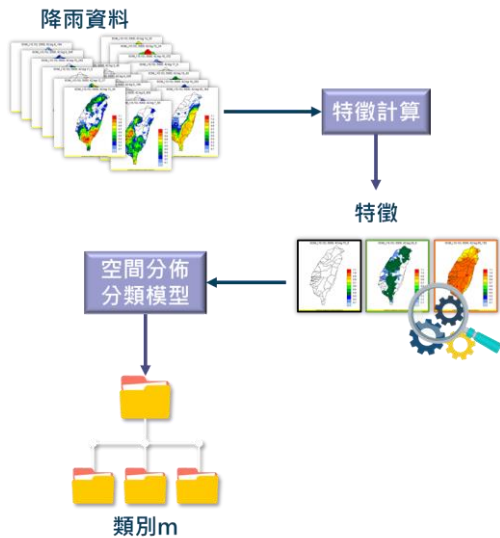


圖 5 降雨資料空間分佈群集分類流程

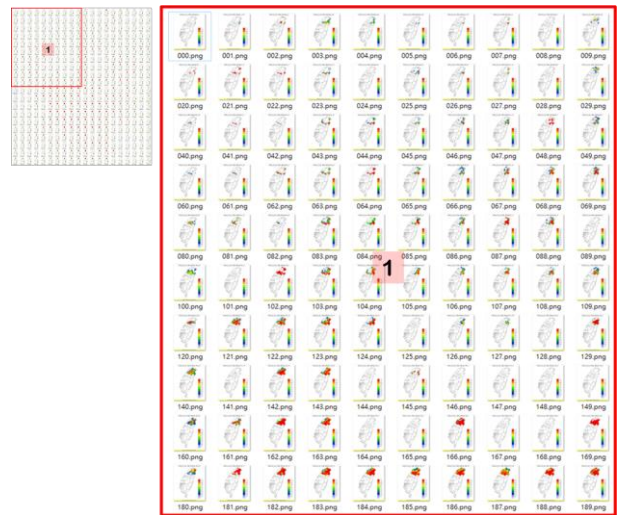


圖 6 降雨分群訓練結果(1/4)



圖 6 降雨分群訓練結果(2/4)



圖 6 降雨分群訓練結果(3/4)

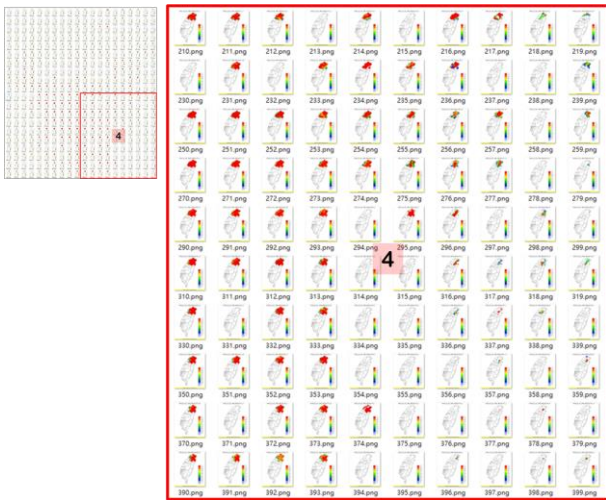


圖 6 降雨分群訓練結果(4/4)

四、歷史事件模擬比較

以目標資料與分群結果之空間分佈相關性做為比對標準。以北部轄區共有4,125個網格點，在目標資料上，將其利用6 mm為門檻值的方式轉換為二元值，亦即大於等於6 mm的網格記錄為1，否則記錄為0，並將此4,125個二元值，依照由北向南、由西向東的方式組成一個長度為4,125的一維陣列；同樣地，對於分群結果亦由北向南、由西向東取出其降雨出現機率，將相關係數自大至小排序，以決定此目標資料與預先分群結果中於空間分佈上最為相近的分群，並取出相似事件對應之多個淹水模擬結果，組合出淹水機率。

本研究根據近年有較大降雨且有災情之108年0722豪雨及106年0602豪雨，進行事件淹水模擬比較，比較結果說明如後。

(一)108年0722豪雨

民國108年7月22日因大氣環境不穩定，午後對流雲系發展旺盛，中央氣象局於當日16時~17時針對大臺北地區發佈豪雨特報，當日較大降雨發生於15時~17時，於大臺北地區造成積淹水災情，大安區多站時雨量破百，最大時雨量(136.5mm/hr)發生於大安森林站，本研究以大安區之淹水點位進行比較。

降雨分群目標範圍精進後之淹水模擬比較結果如圖7，上圖為109年採用全台灣作為降雨分群目標，並依降雨分群結果所對應之淹水模擬結果，下圖為本研究精進採用北部轄區進行降雨分群目標所對應之淹水模擬結果。

由圖7可看出，採用全台灣作為降雨分群範圍之淹水模擬於大安區範圍內，均無出現可能淹水之淹水村里範圍；而降雨分群範圍精進為北部轄區後，於大安區下方有較大淹水範圍(近4公頃)之點位(四維路至和平東路)，模擬出可能淹水，而大安區上方其他區域之淹水點位可能因淹水範圍較小(小於3公頃)，且淹水位置相對零星，故均無模擬出可能淹水。

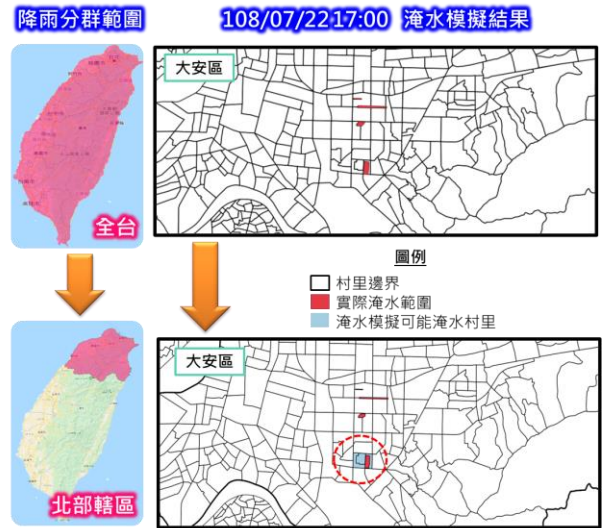


圖 7 降雨分群目標範圍精進後之淹水模擬比較結果 (以 108 年 0722 豪雨之大安區為例)

(二)106年0602豪雨

民國106年6月1日起，受西南氣流及滯留鋒面影響，造成台灣各地區發生豪大雨，且氣象局於6月1日16時發布豪雨特報，其中，大臺北地區降雨主要集中於6月2日02時~14時。

模擬比對結果如圖8及圖9，上圖均為109年^[2]採用全台灣作為降雨分群目標，並依降雨分群結果所對應之淹水模擬結果，下圖為本研究精進採用北部轄區進行降雨分群目標所對應之淹水模擬結果。

由圖8及圖9可看出，採用全台灣作為降雨分群範圍之淹水模擬村里範圍較大，多數村里模擬可能淹水，但實際卻無發生淹水情形，而降雨分群精進為北部轄區後，淹水村里位置與實際淹水位置較相符。

整理精進前後之淹水村里命中率及準確率等數據進行比較，全轄區之臺北市、新北市及基隆市村里數為1,645處，分別為456、1,032、157處村里。

依照實際淹水調查之範圍圖層進行村里套疊，有淹水的村里數為34處，降雨分群精進前之模擬淹水村里數為181處，模擬命中村里數(實際及模擬皆有淹水)為20處，故漏報村里數為14處；降雨分群精進後之模擬淹水村里數為23處，模擬命中村里數(實際及模擬皆有淹水)為16處，故漏報村里數為18處。

計算其命中率與準確率如表1，結果顯示降雨分群精進後之模擬淹水村里數量減少許多，模擬命中村里數量雖略微減少，但整體命中率及準確率皆有明顯提升情形，有效降低109年(全台灣)降雨分群誤報較多之情況。

106年0602豪雨 淹水模擬結果

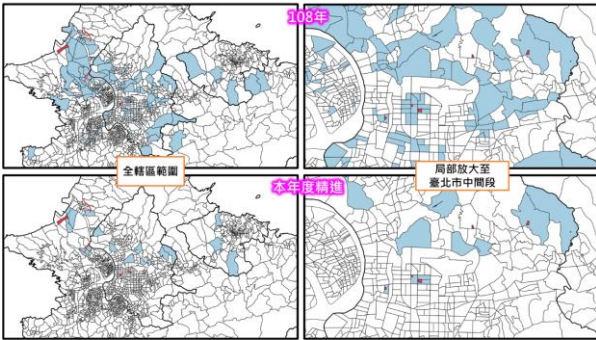


圖 8 降雨分群目標範圍精進後之 106 年 0602 豪雨淹水模擬比較結果(全轄區及臺北市)

106年0602豪雨 淹水模擬結果

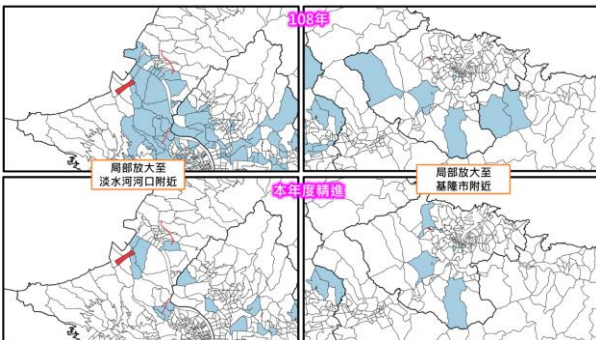


圖 9 降雨分群目標範圍精進後之 106 年 0602 豪雨淹水模擬比較結果(淡水河口及基隆市)

表 1 106 年 0602 豪雨之淹水村里命中率比較

採用模組年份	109 年 ^[2]	本研究
全轄區(臺北市、新北市、基隆市)村里數	1,645	
實際淹水村里數	34	
降雨分群	全台網格	北部轄區網格
模擬淹水村里數	181	23
模擬命中村里數(實際及模擬皆有淹水)	20	16
淹水漏報村里數(實際有淹水, 模擬無淹水)	14	18
命中率 ¹	89.4%	98.5%
準確率 ²	11.1%	69.6%

註 1：命中率計算方式為(實際及模擬皆有淹水村里數)/(實際及模擬皆無淹水村里數)/總村里數。如： $[20+(1645-181-14=1450)]/1645=89.6\%$ ； $[16+(1645-23-18=1604)]/1645=98.5\%$ 。
 註 2：準確率定義為模擬發生淹水的情況下，實際發生淹水的比例。計算方式為(實際及模擬皆有淹水村里數)/模擬淹水村里數。如： $20/181=11.1\%$ ； $16/23=69.6\%$ 。

五、結論與建議

本研究蒐集民國87~107年之51場颱風洪事件及19場淹水紀錄資料，並選用SOM技術應用於降雨趨勢研判分析上，將歷史事件之每6小時累積雨量空間分布，並以1.5、2.0的倍數進行樣本擴充，且降雨分群目標範圍由全台縮小至臺灣北部區域(4,125個網格點)，最後有效樣本數量為10,300筆，以其降雨資料進行SOM模式訓練，透過辨識歷史相似降雨趨勢，找出降雨特徵及相關性，最後訓練出400種降雨分群類別。

以近年於大臺北地區造成的淹水事件(108年0722豪雨及106年0602豪雨)做為比較案例，以實際觀測雨量做為目標資料，比對至歷史降雨事件中空間

分佈與目標資料接近之群集，並取出相似事件對應之多個淹水模擬結果，組合出淹水機率。比較結果顯示本研究建立之SOM分群法，以及配合之空間分佈相關係數計算比較，可將目標資料比對至歷史降雨事件中空間分佈與目標資料接近之群集，進而能從該群集或該等群集中，找到歷史上空間分佈相近的樣本事件。此外，降雨分群範圍由全台範圍精進為北部轄區範圍，可更佳凸顯北部區域降雨特性，有效降低預測發生淹水的誤報率，以及提高預測發生淹水準確度。

目前係以相關係數做為目標資料的分群指定，並以相關係數做為指定分群內歷史樣本資料的挑選，未來建議可評估納入降雨量值進行比對，計算資料序列的均方差或均方根差，以挑選與目標資料更為接近的歷史事件。而未來實際應用上，「六小時累積降雨事件」的組合，可以涵蓋預報資料，例如：採用過去五小時加上預報一小時的組合，並利用此組合挑選分群、相似之歷史樣本。配合所有樣本事件預先進行之淹水數值模式計算結果，結合降雨分群與淹水數值模擬，可快速組合提供淹水預測資料，降低物理模式計算耗時問題，提供更佳之淹水潛勢結果，提供防災作業單位防汛時期之預警應變作業參考。

六、誌謝

本研究承蒙『經濟部工業局智慧城鄉生活應用補助計畫-創新服務類計畫』^[1]補助，並感謝經濟部水利署第十河川局提供資料，供本研究順利進行相關研究分析探討，特此申謝。

七、參考文獻

- 「暴雨淹水潛勢創新預報技術研發計畫」計畫書，經濟部工業局智慧城鄉生活應用補助計畫-創新服務類計畫，民國 107~108 年。
- 「109 年度十河局轄區洪水預警及防汛整合作業」，經濟部水利署第十河川局，民國 109 年。
- <https://rpubs.com/jiankaiwang/som>。
- https://en.wikipedia.org/wiki/Self-organizing_map。
- 李文生, 洪國展, 郭雅燾, 于芃, 2020: “應用自組織映射圖分析歷史短延時雨量分群樣態之研究”, 中央氣象局 109 年天氣分析與預報研討會。
- 王潔如, 吳明璋, 徐理寰, 蕭玲鳳, 洪景山, 李清勝, 2018: “系集定量降水群集分析技術之測試與應用”, 中央氣象局 106 年天氣分析與預報研討會。
- 陳儒賢, 洪毓婷, 許臣王, 2011: “自組織映射圖網路與線性動差法於區域乾旱頻率分析之研究”, 農業工程學報 第 57 卷第 2 期。
- 陳宜欣, 「發展以自組織映射圖網路為基礎之雨量繁衍模式於未來降雨推估」, 碩士論文, 國立臺灣大學土木工程學研究所, 民國 100 年。

2021 Conference on Weather Analysis and Forecasting

Application of Self-Organizing Map in northern Taiwan to Analyze the Clustering Patterns of Historical Short-Duration Rainfall

Wen Sen Lee¹ Kuo-Chan Hung¹ Ya-Yun Kuo² Peng Yu¹
Manysplendid Infotech Co. Ltd¹ Manysplendid Engineering Consultants Co. Ltd²

Abstract

Rainfall is the nature main driver factor of flooding in Taiwan. Even if the ground surface may change the flooding situation due to city development or regulation engineering, the rainfall similar to historical disaster events still require attention. Similarly, if the flooding simulation numerical model simulates with the same computational conditions, the calculated flooding results will be the same.

This study collects rainfall data of typhoon and torrential rain events from 1998 to 2018, as well as 19 flooding records in the great Taipei metropolitan area, to research the spatial distribution of 6-hours short-duration accumulated rainfall that caused the flood events. For historical rainfall events, the artificial neural network technology using SOM image grouping is grouped according to the spatial distribution characteristics of rainfall. In real-time applications, the current rainfall events are compared with spatial distribution correlation coefficients to select historical rainfall event clusters with similar spatial distributions, and then find out sample events with similar spatial distributions in history.

The reference study in 2020 is based on the calculation of historical data clustering across entire Taiwan region. The clustering results are also affected by the rainfall distribution in southern or eastern Taiwan. Therefore, the best clustering results still does not show similar rainfall distribution in northern Taiwan. This study narrows the target area from the entire Taiwan to the northern part of Taiwan, so that it can highlight the rainfall characteristics of the northern region, and propose a more suitable and consistent rainfall clustering result in the northern region. From the comparison of the historical flooding event simulations of the torrential rain of 0722 in 2019 and of 0602 in 2017, the results in this study can better find the similar rainfall events of northern Taiwan, then it can find disaster areas by similar events effectively to provide reference for disaster prevention.

Key word: SOM, Self-Organizing Map, flooding, short-duration