# 一、研究問題與資料來源

觀察臺灣歷史颱風資料,學者們將過去的颱風依照 行進路徑進行統計與分類,如圖1所示,將98%的颱風 路徑歸類為九種。而在風災的應變過程中,決策者也會 參考歷史颱風資料,如:各地累積雨量、災情等,自經 驗中擬定災防對策,是以有效地搜尋歷史中相似的颱風 是風災應變中的關鍵。為此,本研究欲開發一互動式操 作介面,利用使用者輸入點位與搜尋半徑等參數,回傳 其搜尋條件下,颱風歷史事件當中最為相似的路徑排 名,藉此輔助決策者相關政策之規畫。而本研究使用之 路徑比對方法,主要基於中國土木水利工程學刊第30期 [2]的討論,並針對其未多加著墨的時間權重延伸探討。

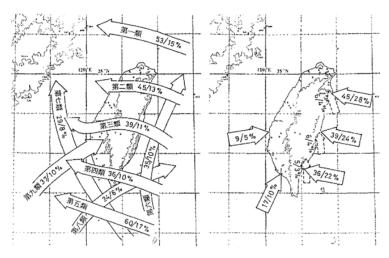


圖 1 1897-1996 年間侵台颱風分類、發生次數及登陸地點統計[1]

資料來源方面,自日本気象庁公開 API 接收 1951 年至今,所有気象庁管轄範圍內熱帶氣旋之資料,如圖 2。經統整後,近70年的西北太平洋熱帶氣旋資料總計 約有 1,700 筆,並包含約 65,000 筆點資料。

91092706 002 5 325 1293 935 30180 0140 30400 0260 095 圖 2日本気象庁提供的最佳路徑資料,包含: header 與 data

# 二、分析方法與演算法

本研究颱風路徑比對之演算法主要基於「最短距 離」、「時間加權」與「基數排序(radix sort)」。首 先,由使用者輸入欲搜尋的颱風路徑點與其相對應的搜 尋半徑,若歷史颱風路徑點與使用者輸入點之間的最短 距離小於搜尋半徑,則賦予時間加權。其次,針對每一 筆歷史颱風路徑,總合計算加權總分。最後,依總分、 時間加權、通過點個數與時間進行基數排序,統整出路 徑相似程度較高的歷史颱風資料。

**墾泓恺** 廖昱嘉 何承諭



# 西北太平洋歷史颱風 路徑之比對搜尋系統

## 2.1 最短距離

將資料庫中H筆歷史颱風路徑以陣列P儲存,第i筆歷史颱風之  $N_i$  個路徑點  $(X_{ij}, Y_{ij})$  以陣列  $P_i$  儲存,以  $P_i = \{P_{i1}, P_{i2}, P_{i3}, ..., P_{iN_i} | P_{ij} = (X_{ij}, Y_{ij})\}$ 表示。另外,使用 者輸入M個欲搜尋之颱風路徑點, $(X_k,Y_k)$ 儲存於陣列 U,以 $\{U_1, U_2, U_3, ..., U_M | U_k = (X_k, Y_k)\}$ 表示,而第 i 筆 歷史颱風資料的第 / 個點與使用者輸入之第 / 個點的距 離以  $D_{ijk} = |P_{ij} - U_k|$  表示。將所有歷史颱風的所有點 與所有使用者輸入點的距離,儲存於矩陣D,如方程式 1所示。而本演算法尋找的即是使用者輸入之第 k 個點 與每一個歷史颱風路徑點之間的最短距離,如方程式2

#### 2.2 時間加權

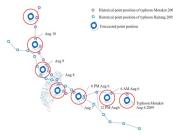
若歷史颱風路徑點與使用者輸入點之間的最短距離  $d_{ik}$  不大於搜尋半徑  $r_k$  即該歷史颱風路徑通過使用者輸 入點的搜尋環域,則賦予時間加權 kw,其中 k 代表使用 者輸入的第 k 個點,如最短距離小於使用者輸入之第 2 個點的搜尋半徑,賦予其時間加權2w。而w為外生時 間權重,後續有更詳細的討論。理論上,使用者越晚輸 入的點(越接近現在),較使用者早期輸入的點重要, 亦即歷史路徑與現在的相似度比早期的相似度更為重 要,而時間加權的設計即是為了實現此想法。因此每筆 歷史颱風路徑的總分能以  $\sum_{k=1}^{M} (1+kw)$  計算,  $\sum_{k=1}^{M} 1$ 為考量歷史颱風路徑通過使用者輸入搜尋環域的個數,  $\sum_{k=1}^{M} kw$  則是前述時間加權的累積。例如:使用者輸入 5個點 (k=5) ,某一歷史颱風路徑通過前三個點的搜 尋環域,其總分能以3+6w表示。其後,重複前述操 作,計算每筆歷史颱風路徑之總分,以陣列 D 儲存,如 方程式3所示。

$$v = \left\{ v_1, v_2, v_3, \dots, v_H \,\middle|\, v_1 = \sum_{k=1}^{M} (1 + kw) \,\middle|\, d_{ik} \le r_k \right\}$$
 方程式 3[2]

[1] 謝信良, 王時鼎, 鄭明典, 葉天降, & 丘台光. (1998). 百年侵台颱風 路徑圖集及其應用. 中央氣象局, 台北. [2] 中國土木水利工程學刊第 30 期

# 2.3 基數排序

欲比較歷史颱風路徑與使用者搜尋路徑之相似度,首先以陣列 v 中各個歷史颱風總分  $\sum_{k=1}^{M}(1+kw)$  排序,將路徑相似度與時間加權同時納入考量;若同分,則再依序比較時間加權的分數  $\sum_{k=1}^{M}kw$  通過點個數  $\sum_{k=1}^{M}1$  現在月份的差距與現在年份的差距。以 2009 年的莫拉克颱風與 2005 年的海棠颱風例,如圖 3 所示,莫拉克颱風的總分為 7+28w 海棠颱風的總分為 3+12w 莫拉克颱風的總分較海棠颱風為高,是以其排序會位於海棠之前。然而,未必所有排序情況皆如圖 3 的舉例顯而易見,如圖 4 所示,藍色的颱風通過使用者輸入之後兩個點,其總分 3+6w 此時,兩者排序的先後將受時間權重 w 的影響,是以接下來欲討論不同 w 對歷史颱風路徑相似度排序的影響。



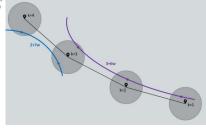


圖 3 2009 年的莫拉克颱風與 2005 年的海棠颱風[2]

圖 4 排序受 w 影響的兩條颱風路徑

# 2.4 權重影響

## 1. $w \ge \widehat{w}$

隨著使用者輸入颱風路徑點個數 M的不同,相對應的  $\hat{\mathbf{w}}$  也不同。當給定的權重不小於臨界權重,  $\mathbf{v}_{uncertain}$  中所有總分組合受到的時間加權影響最大化,亦即通過環域個數相同時,通過時間次序較晚之環域的歷史颱風排序都將較通過時間次序較早的為前,而等號的成立與基數排序優先順序有關。

#### 2. $0 < w < \hat{w}$

若給定的權重介於0與臨界權重之間,vuncertain 中的總分組合排序,將視個別路徑通過使用者輸入點搜尋環域的個數 $\Sigma 1$ 與時間加權 $\Sigma kw$ 計算結果的大小而定。

#### 3. w = 0

若w=0最終颱風路徑相似度排序將只與通過使用者輸入點搜尋環域個數 $\Sigma1$ 有關,與時間加權 $\Sigma kw$ 無關。

## 4. w = -1

若 w = -1 各個歷史颱風總分  $\Sigma(1 + kw)$  可以改寫為  $\Sigma(1 - k) \le 0$  代表相似度排序的先後為:僅通過 k = 1 的使用者路徑點再依時間先後排序、完全不通過任何點再依時間先後排序、總分(負分)由大到小排序。

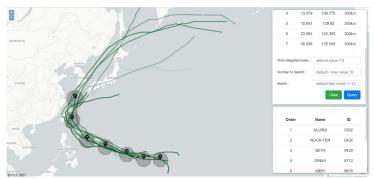
[註]受限於篇幅限制,此處無法完整呈現時間權重臨界值的推導流程,詳見: https://reurl.cc/Qx1Vb

## 三、結果與討論

將前述理論實作後,本研究最終之颱風路徑搜尋 使用者介面如圖 5所示。畫面中左側為地圖圖層與依 據颱風逐時點位繪製的路徑圖;畫面右側則是資訊欄 位,包含:使用者輸入之搜尋點位與搜尋半徑、其他 搜尋參數與最終路徑比對結果。若使用者輸入點數量 為 4 , 圖 6(左)、圖 6(中) 與圖 6(右) 分別代表將時間 權重設為1、0與-1的搜尋結果。若時間權重為1,恰 好為使用者輸入點數量為 4 時的臨界權重,表示時間 次序是這次搜尋的優先比序,即在兩颱風路徑通過環 域個數相同的情形下,通過時間次序較晚之環域的颱 風與使用者查詢路線的相似度排序必定高於另一者; 若時間權重為 0, 路徑相似度與環域時間次序無關, 因此將以通過環域個數作為該歷史颱風路徑相似度的 排序依據,而在通過環域數量相同時,再行依照歷 史颱風發生時間排序;若時間權重為-1,最終結果 為僅通過第一個環域,且避開其他三個環域,再根據 颱風發生時間排序。最後,期望藉由本颱風路徑比對 搜尋系統,決策者能更有效率、直覺地搜尋歷史颱風

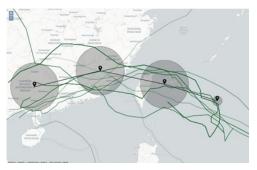
#### 颱風搜尋系統網址: https://reurl.cc/ZvD6p

的應對方針。



資料,結合過往經驗,擬定對於當下颱風災害最合適

圖 5 使用者操作介面



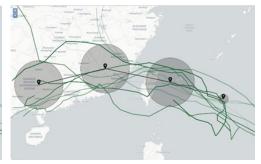


圖 6 不同時間權重下搜尋結果的差異

