

應用季長期天氣預報推估高雄地區登革熱流行趨勢

徐葦茵¹ 王玉純² 呂金璋³ 李明旭¹

國立中央大學水文與海洋科學研究所¹ 中原大學生物環境工程學系²

摘要

近年來登革熱一直為全球熱帶地區、亞熱帶地區相當嚴重的病媒蚊傳染性疾病，至1980年為止，全球亞熱帶地區，有活動性登革熱病毒傳播的國家多達61國，所涵蓋總人口數約有15億之多。台灣地處亞熱帶，高溫多雨，是病媒蚊傳染疾病的繁殖溫床，近年來由於氣候異常，若高溫多雨將促使登革熱病毒及病媒活化，進而導致登革熱疫情之擴大。台灣地區近十年來的登革熱病例流行，以2002年高雄地區為最主要的病例大爆發年，主因為2001年登革熱病例持續發生且跨年而導致該年的大流行，其中出血性登革熱病例高達240例，死亡率更高達百分之八點七。因此，預先研判登革熱在未來的可能發展趨勢為一重要的防疫課題。本研究利用中央氣象局發布之季長期預報資料結合週期性流行指數，利用ARIMA(Autoregressive Integrated Moving Average, 自迴歸移動平均)時間序列模式預測登革熱病例在高雄地區的流行趨勢。結果顯示利用此模式所預測之登革熱病例數與歷史事件比較，所建立模式能掌握登革熱病例發展趨勢，因此可利用此模式結合未來之預報資料預測可能之病例發展趨勢。

關鍵字：登革熱、高雄地區、季長期預報、時間序列

一、前言

近來登革熱一直為全球熱帶地區、亞熱帶地區相當嚴重的病媒蚊傳染性疾病，至1980年為止，全球亞熱帶地區，有活動性登革熱病毒傳播的國家多達61個，涵蓋總人口約有15億之多；1990年Halstead學者於其研究報告中指出：從1970至1980年代，每年約有25萬人感染出血性登革熱(dengue hemorrhagic fever, DHF)；由於都市化以及日益增加的國際貿易造成登革熱病毒及病媒蚊的傳播速度增加，1998年登革熱出現了前所未有的大流行，共有56個國家向世界衛生組織報告了120萬的病例數，自此之後登革熱及登革出血熱流行有明顯上升趨勢，拉丁美洲及東南亞國家有千萬人受到感染，每年向世界衛生組織(World Health Organization, WHO)報告的平均病例在過去40年中，每10年幾乎增長1倍。世界衛生組織在2009年提出的文章指出：全球約有2.5億人口是生活在受登革熱病毒威脅的區域中。

台灣地處亞熱帶，高溫多雨，是病媒蚊傳染疾病的繁殖溫床。近年來由於氣候暖化，促使登革熱病毒及病媒活化，導致台灣地區登革熱病例數有明顯增加趨。回顧台灣過去登革熱流行史，於日據時期曾多次發生過登革熱流行，自1987年後幾乎年年都會出現規模不等的本土疫情。

近十年來的登革熱病例流行，以2002年為最主要的病例大爆發年，主因為2001年登革熱病例持續發生

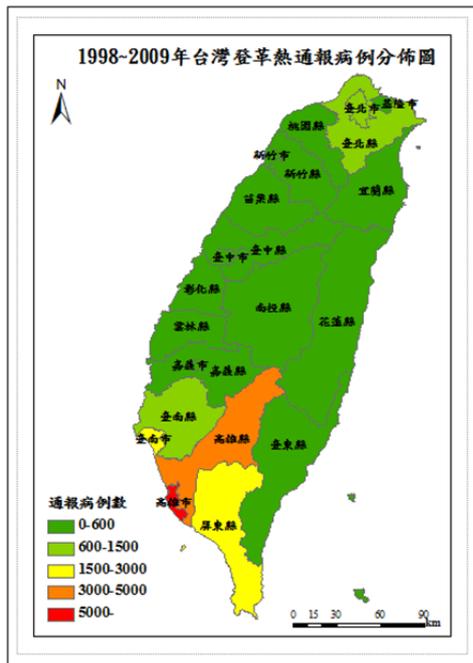
且跨年，而導致該年的大流行，2002年登革熱的確定病例有5,345例，其中登革出血熱(dengue hemorrhagic fever, DHF)或登革休克症候群(dengue shock syndrome, DSS)的病例數高達240例，導致死亡數有21例。此次主要流行於高雄縣市、屏東縣及臺南縣市，為自1987年來最嚴重的一次流行。

本研究將氣象參數與病例數之間的時間序列關係式找出，然後利用歷史30年的氣象資料統計出不同區間各月份的平均溫度及雨量，再結合氣象局的季預報資料挑選出所需年份各月份之氣象資料，以SPSS 16.0統計軟體內的ARIMA(Autoregressive Integrated Moving Average, 自迴歸移動平均)時間序列模式，建立登革熱病例及氣象因子間的時間序列關係式，以預測未來登革熱病例的發展趨勢。

二、研究區域與方法

本研究區域主要以台灣南部的高雄縣市為主(縣市合併前)，由疾病管制局統計資料顯示上述兩縣市為台灣地區主要登革熱病例個案較多且每年持續有較多病例數產生的縣市。圖一為1998~2009年台灣地區登革熱通報病例分佈圖，由圖一得知台灣地區登革熱病例主要分布地區為南台灣的台南縣市、高雄縣市

及屏東縣五縣市，但主要大量的病例數還是發生於高雄縣市。



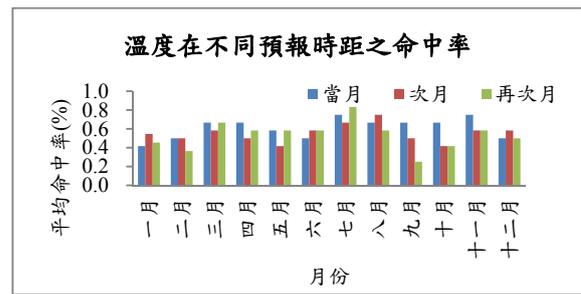
圖一 1998~2009年台灣地區登革熱通報病例分佈圖

為了確定中央氣象局所發布的季節性預報資料可以使用在本研究區域，所以必須在雨量資料上，將高雄站雨量資料與高雄縣 8 個雨量測站做區域相似性分析，以確定高雄縣市之間的氣候型態只是程度上的差異。在溫度資料上，則將高雄站之溫度資料與氣象局發布之季節性氣候預報做溫度之命中機率，來確定可以準確的應用在研究中。

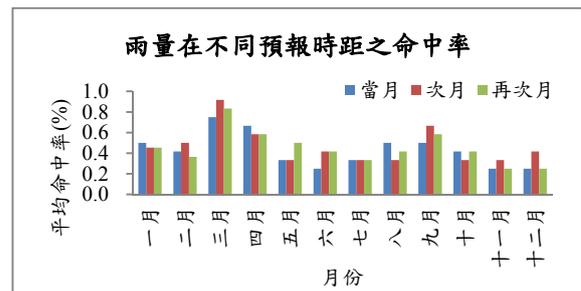
在預報命中率上，圖二、三為高雄站 1998~2009 年溫度、雨量在不同預報時距(lead time)預報命中率之結果。以溫度之命中率來看，二月份之命中率最低，不同預報時距之平均命中率為 0.455，命中率最高的月份則為七月，其值為 0.75；在雨量之命中率方面，十一月份之命中率最低，在不同預報時距之命中率僅 0.278，命中率最高的月份為三月，其值為 0.833。從不同預報時距的命中率方面來看，溫度在當月之命中率較高，其值為 0.611；雨量則在次月之命中率較高，其值為 0.468。整體而言，溫度在當月的夏季(七至十一月)之命中率較高，而雨量則僅在三月份的三個區間有較高的命中率，顯示雨量的變化較溫度劇烈，所以也較難預測準確。

圖四為高雄站與高雄縣 8 個雨量測站(木柵、古亭坑、甲仙、多納、金山、阿蓮、美濃、旗山)之平均雨量區域相似性。8 個雨量站中以多納站與高雄站之區域相似性最低，平均相似性命中率為 0.602，其中在偏低區間及正常區間之命中率各為 0.590 及 0.513，低於偏高區間的命中率 0.704；與高雄站區域相似性較高的測站為古亭坑站，平均命中率為 0.706，偏低

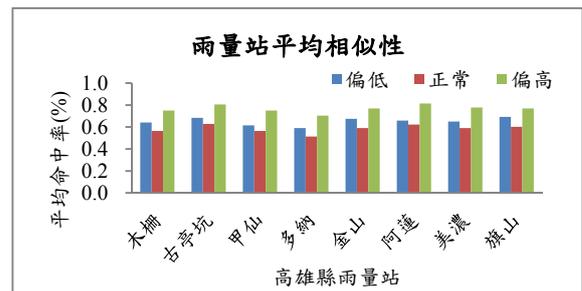
區間、正常區間、偏高區間之平均命中率分別為 0.684、0.628、0.806。



圖二 1998~2009年高雄站溫度在不同預報時距之命中率

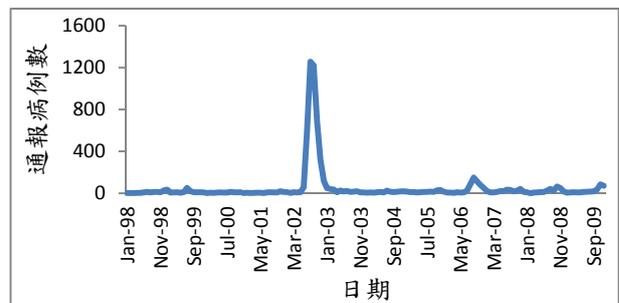


圖三 1998~2009年高雄站雨量在不同預報時距之命中率

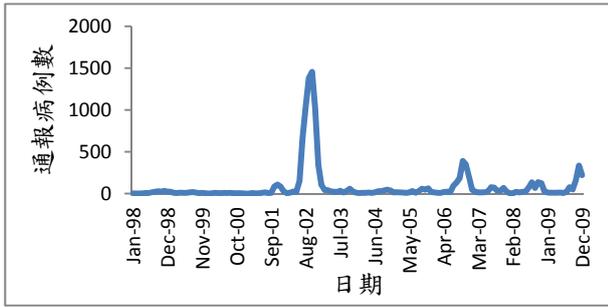


圖四 1998~2009年高雄縣雨量站平均區域相似性

圖五及圖六分別為高雄縣市在 1998~2009 年之月累積病例數，高雄縣市在 2002 年皆有疾病大爆發的發生，且在 2002 年下半年(七~十二月)高雄縣之累積病例數為 4210 例，高雄市之累積病例數更高達 5922 例，為近十年來最嚴重的一次大爆發。

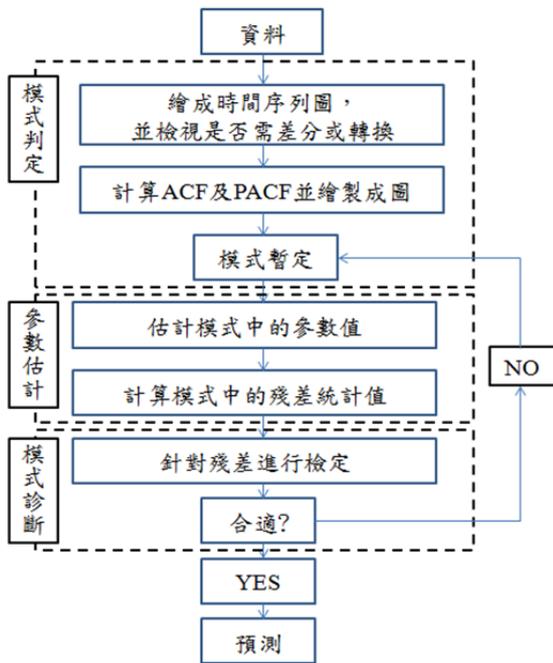


圖五 1998~2009年高雄縣各月份之累積通報病例數



圖六 1998~2009年高雄市各月份之累積通報病例數

本研究利用之時間序列模式(ARIMA)，其流程如圖七所示



圖七 ARIMA建立模型之流程圖

將所需之氣象參數及非氣象參數利用圖七之步驟建立高雄縣市登革熱病例之時間序列之後，結合氣象局季預報資訊選取歷史 30 年之氣象資料，投入建立之模式，預測預測登革熱在未來可能之流行趨勢。

三、研究結果與討論

(一) 高雄縣

表一為高雄縣登革熱病例時間序列之候選模式，Case1~3 所加入之參數及模式分別為：Case1:登革熱病例數+前一個月之流行指數(PI)+月流行指數(MPI)+月平均溫度，模式為 ARIMA(0,0,3)；Case2：登革熱病例數+流行指數，模式為 ARIMA(1,0,0)；Case3：登革熱病例數+月平均溫度，模式為 ARIMA(1,0,0)。由表一中，因 Case1 中之 Normalized BIC 最小值，因此

選取 Case1 為模擬高雄縣登革熱病例數最適模型，其迴歸式如式(1)所示。

表一 高雄縣登革熱病例候選模式擬合統計資料

	Case1	Case2	Case3
R2	0.542	0.630	0.685
RMSE	140.492	123.580	113.303
MAPE	74.420	72.503	78.018
MAE	31.462	29.587	27.006
Normalized BIC	9.229	9.826	9.604

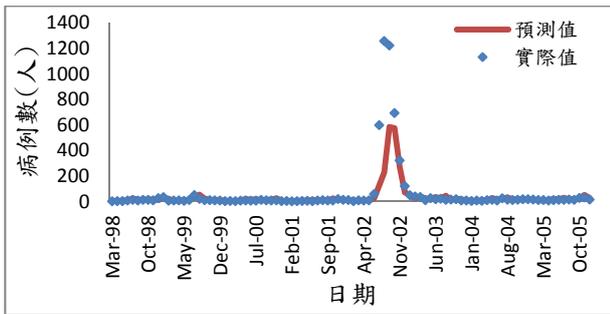
$$\ln(Y_t) = -0.518\ln(\varepsilon_{t-1}) - 0.480\ln(\varepsilon_{t-2}) - 0.589\ln(\varepsilon_{t-3}) + 0.493\ln(\text{PI}_{t-1}) + 0.735\ln(\text{MPI}_t) + 3.312\ln(T_t) - \ln(12.588) \quad (1)$$

上式中 Y_t 為所欲預測當月份之病例數， $\varepsilon_{t-1} \sim \varepsilon_{t-3}$ 為所欲預測之病例數前一個月份之實際病例數與預測病例數之誤差值， PI_{t-1} 為前一個月份之流行指數 (Population Index)， MPI_t 為當月份之月流行指數， T_t 則為當月份之月平均溫度，各月份之 MPI_t 值如表二所示。

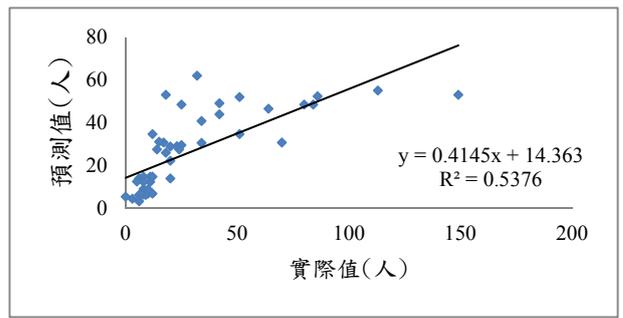
表二 高雄縣各月份之月流行指數

月份	一月	二月	三月	四月	五月	六月
指數	15.71	10.00	9.29	7.00	9.86	9.14
月份	七月	八月	九月	十月	十一月	十二月
指數	19.14	13.14	11.14	14.00	14.57	12.71

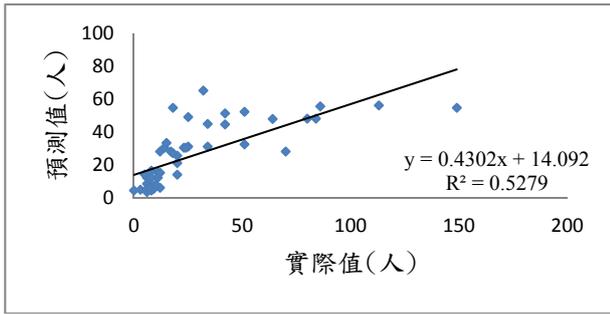
圖八為檢定年份實際值與預測值之時間序列圖，此模式在檢定年時僅在 2002 年高峯值有較大之誤差，但以整體來看其模擬之趨勢與實際值相同。圖九為利用 SPSS16.0 軟體所預測出高雄縣 2006~2009 年之登革熱病例數，實際值與預測值之 R2 值為 0.5279，顯示模式所預測出之病例數與實際病例數為高度相關。圖十為高雄縣 2006~2009 年登革熱實際值與預測值之時間序列圖，其中實線部分為在模式中利用歷史月平均溫與其他參數所預測出之登革熱病例數，而標號為疾管局實際通報病例數，利用此模式所預測出之病例數值與歷史登革熱病例較高的時間點皆有相同的高值出現，整體來看預測值之趨勢與實際值一致。



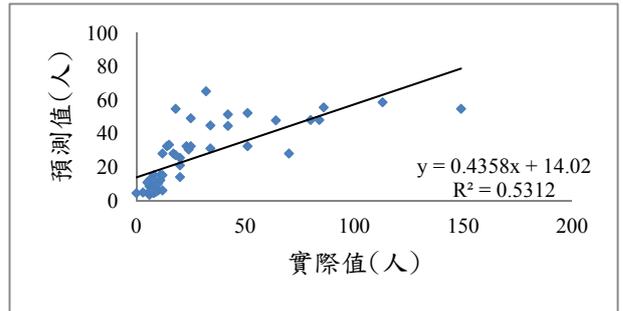
圖八 高雄縣1998~2005年模式檢定年實際病例與模式預測值時間序列圖



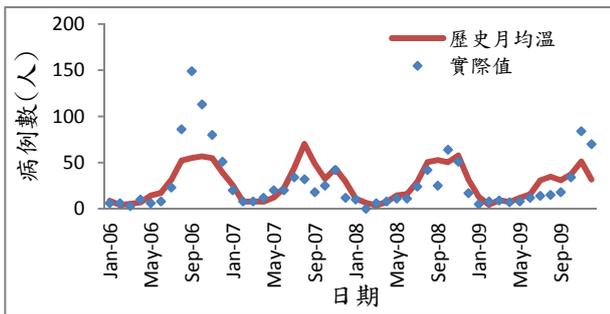
圖十一 高雄縣2006~2009年實際病例數與預測值之相關性(預報當月)



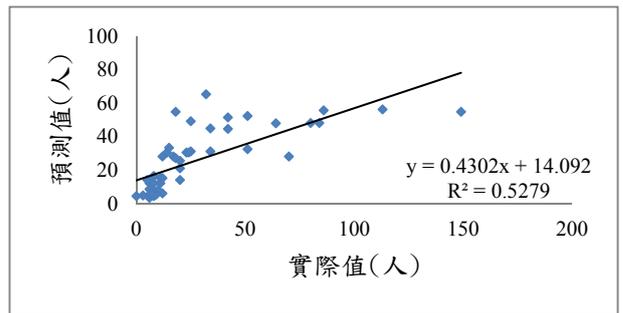
圖九 高雄縣2006~2009年實際病例數與歷史月均溫模擬登革熱預測值相關性



圖十二 高雄縣2006~2009年實際病例數與預測值之相關性(預報次月)

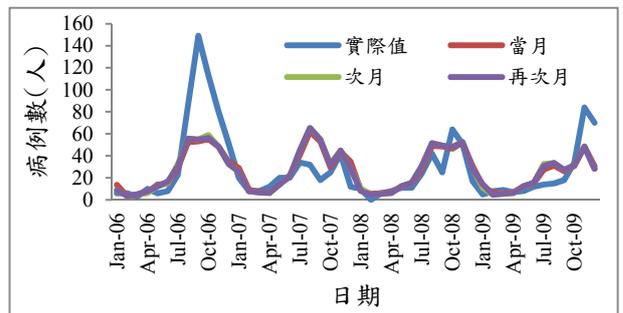


圖十 高雄縣利用歷史月均溫模擬2006~2009年之病例數時間序列圖



圖十三 高雄縣2006~2009年實際病例數與預測值之相關性(預報再次月)

圖十一~圖十三為高雄縣 2006~2009 年登革熱實際值與氣象局季預報不同時距所挑選之歷史氣象資料投入 Case1 模式所預測之病例數之間的相關性。預報當月、次月、再次月所預測登革熱病例之決定係數分別為 0.5376、0.5312 及 0.5279。在利用不同預報時距預報值所預測的病例數時間序列變化趨勢與實際病例數時間序列變化皆相似，皆能在實際病例數為高峯值時有準確的上升趨勢，如圖十四所示。



圖十四 高雄縣2006~2009年不同預報時距登革熱病例之時間序列圖

(二) 高雄市

表三為高雄市登革熱病例時間序列之候選模式，Case1~3 模式所加入之參數及模式分別與高雄縣相同。由表中，Case1 之 Normalized BIC 最小值因此選取 Case1 為模擬高雄市登革熱病例數最適模型，其迴歸式如式(2)所示。

表三 高雄市登革熱病例候選模式擬合統計資料

	Case1	Case2	Case3
R2	0.630	0.892	0.804
RMSE	162.031	86.209	115.490
MAPE	77.111	62.276	75.606
MAE	46.905	27.441	36.887
Normalized BIC	8.814	9.107	9.643

$$\ln(Y_t) = -0.656\ln(\epsilon_{t-1}) - 0.419\ln(\epsilon_{t-2}) - 0.395\ln(\epsilon_{t-3}) + 0.912\ln(PI_{t-1}) + 0.582\ln(MPI_t) + 2.283\ln(T_t) - \ln(9.216) \quad (2)$$

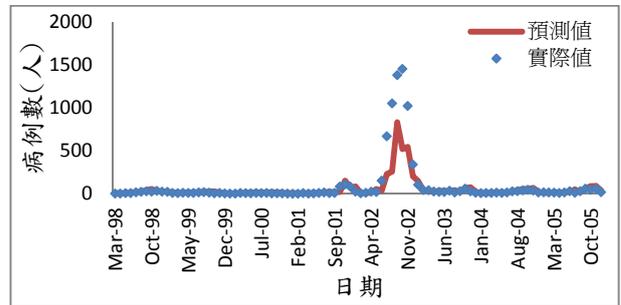
上式中 Y_t 為所欲預測當月份之病例數， $\epsilon_{t-1} \sim \epsilon_{t-3}$ 為所欲預測之病例數前一個月份之實際病例數與預測病例數之誤差值， PI_{t-1} 為前一個月份之流行指數 (Population Index)， MPI_t 為當月份之月流行指數， T_t 則為當月份之月平均溫度，各月份之 MPI_t 值如表四所示。

表四 高雄市各月份之月流行指數

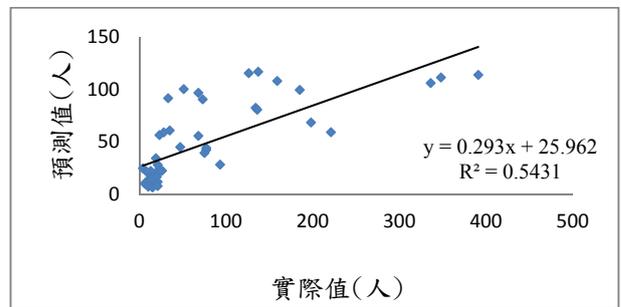
月份	一月	二月	三月	四月	五月	六月
指數	22.57	12.00	13.00	10.43	10.57	14.57
月份	七月	八月	九月	十月	十一月	十二月
指數	19.00	19.57	26.57	39.29	40.86	23.43

圖十五為高雄市在1998~2005年檢定年模式所模擬之病例數，與實際病例數僅在2002年高峯值有較大之誤差，整體看來模擬值與實際值之趨勢為一致。圖十六為利用SPSS16.0計軟體投入歷史實際月平均溫度及其他非氣象參數所預測出高雄市2006~2009年登革熱病例數與實際病例數之相關性，兩者之R2值為

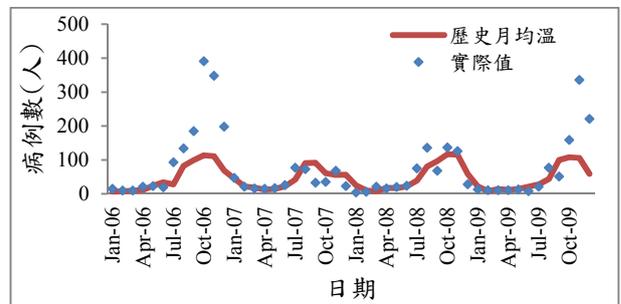
0.5431，為高度相關。圖十七為實際氣象參數所檢定之模式所預測2006~2009年登革熱病例數流行趨勢時間序列圖，其中實線部分為利用歷史氣象參與其他非氣象參數所預測出2006~2009年之病例數，標號則為疾管局實際病例數。



圖十五 高雄市1998~2005年模式檢定年實際病例與模式預測值時間序列圖

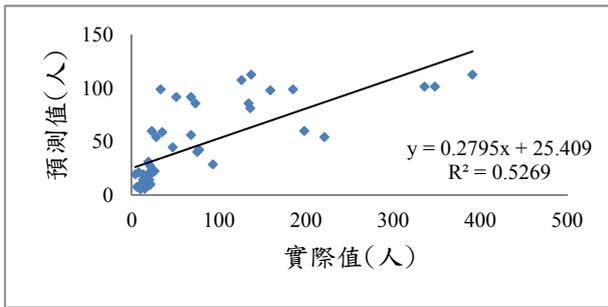


圖十六 高雄市2006~2009年實際病例數與歷史月均溫模擬登革熱預測值相關性

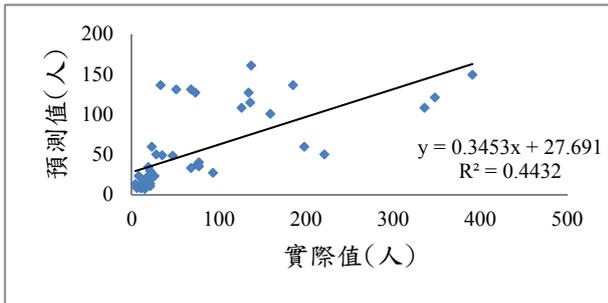


圖十七 高雄市利用歷史月均溫模擬2006~2009年之病例數時間序列圖

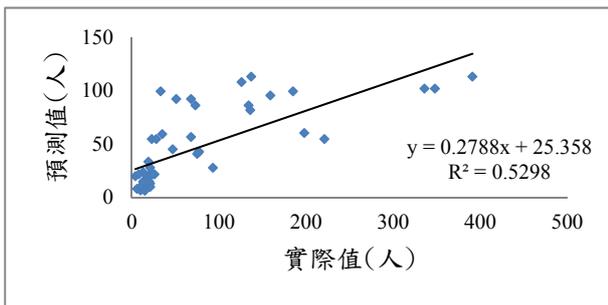
圖十八~圖二十為高雄市2006~2009年登革熱實際值與利用歷史30年氣象資料結合氣象局發佈每月季長期預報挑出每個月之月平均溫度後加入Case1之後，模式所預測出之病例值與實際值的相關性，當月、次月、再次月預報所預測登革熱病例之決定係數分別0.5263、0.4432及0.5298。在利用不同的預報時距所預測的病例數時間序列變化趨勢與實際溫度所預測之病例數時間序列變化皆相似，都能在實際病例數為高峯值時有準確的上升趨勢，如圖二十一所示。



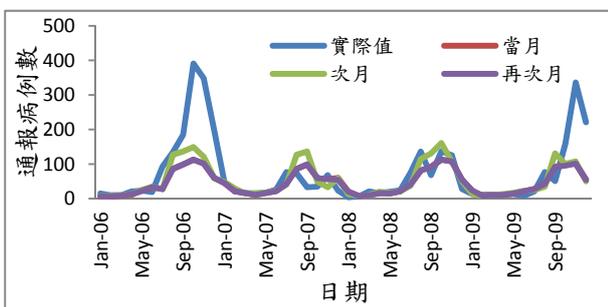
圖十八 高雄市2006~2009年實際病例數與預測值之相關性(預報當月)



圖十九 高雄市2006~2009年實際病例數與預測值之相關性(預報次月)



圖二十 高雄市2006~2009年實際病例數與預測值之相關性(預報再次月)



圖二十一 高雄市2006~2009年不同預報時距登革熱病例之時間序列圖

四、結論與建議

氣象站高雄站之溫度及雨量與氣象局所發佈預報命中率在不同時距上分別介於0.533~0.611及0.431~0.468，在月份上之平均預報命中率分別介於0.455~0.750及0.278~0.833，顯示可能雨量變化較大所以在預報上與溫度相比較會較不易。此外在區域相似性上，高雄縣八個雨量站與氣象站高雄站之雨量區域相似性平均約為0.667，偏低、正常、偏高區間相似性分別0.651、0.584及0.767，表示季長期天氣預報可套用在高雄縣八個雨量站。

高雄縣及高雄市所預測出之登革熱病例與實際值之間有相同的趨勢，在2006年及2009年的預測值中較實際值低，2007、2008年預測值與實際值較相近，誤差值較小。衛生署疾病管制局所擁有的登革熱病例通報資料始於1998年，在一開始的期間也許無法落實病例通報，所以在前幾年的數據上可能會有些許誤差，影響所建立模式的預報能力。

本研究利用月病例指數(MPI)結合季長期預報資料中之最大機率所對應之月平均溫度應用在預測登革熱病例上，未來可嘗試以不同月平均溫度選取方法或不同氣象資料，並與預測結果比較。此外對於較極端之登革熱病例的模擬，應該探討是否要將模擬之部份以區間分類，如同氣象局之預報將氣象資料區分為落入各區間之機率為何，如此便可知得往後的流行大致是落於何區間。

五、參考文獻

1. 經濟部工業局石化工業政策評估說明書，2010。
2. 楊奕農，2009，時間序列分析：經濟與財務上之應用第二版，雙葉書廊。
3. 葉小綦，1998，時間序列分析與應用，葉小綦發行。
4. 張筱玲、賴淑寬，2006，氣象資料於疾病監測之應用，行政院衛生署疾病管制局九十五年度科技研究發展計畫報告
5. Box, G. and G. Jenkins, 1970. Time series analysis: Forecasting and control, San Francisco: Holden-Day.
6. Wu, P.C., H.R. Guo, S.C. Lung, C.Y. Lin, H.J. Su, 2007. "Weather as an effective predictor for occurrence of dengue fever in Taiwan", Acta Tropica, Volume 103, Issue 1, 50-57