

結合人工智慧 (AI) 技術發展「體對點」 (Volume-to-Point) 的降水預報

Shih-Hao Su¹, Ting-Shuo Yo², Jung-Lien Chu³, Chiao-Wei Chang¹, Hung-Chi Kuo²
Chinses Culture University¹
National Taiwan University²
National Science and Technology Center for Disaster Reduction, Taipei, Taiwan³

摘要

現行主要的降水預報方式，一是以數值模式的輸出為主，將網格上的預報降水內插到所欲預測的位置；一是以雷達回波強度為基礎，進行定量降水推估 (Quantitative Precipitation Estimation, QPE)，透過經驗公式將欲知降雨位置的雷達回波強度，轉換成該位置的降雨。無論是數值模式或是雷達資料為基礎，現行的降水預報都是基於「點對點」 (point-to-point) 的方式：對於一個定點的降水，以該定點上能取得最佳的資料及經驗公式推斷之。

近年來人工智慧 (AI) 與機器學習 (machine learning) 技術的發展，讓我們在處理大規模資料的能力上有長足的進步。以往我們需要針對個別地點與特定天氣系統，以人為的方式分析其特徵，再加入到經驗公式中對降水預測加以修正；現在透過機器學習演算法，我們可以將大範圍、多時間的資料疊加在一起，成為高維度的資料體 (data volume)，然後利用人工智慧進行特徵抽取 (feature extraction) 與統計建模 (learning statistical model)，直接預報出特定位置的降水。以雷達資料為例，我們將每十分鐘一筆的雷達回波資料堆疊在一起，透過機器學習的演算法，可以直接進行目標位置的定量降水推估。具體來說，透過這樣架構所建立的『體對點』雷達降水推估模組，我們使用 2013 年至 2016 年台北 45 個氣象站的每小時降水量數據進行評估。與作業化的 QPE 相比，『體對點』雷達降水推估模組在一般的情境下情況下，對於定量降雨推估的表現良好，並且在偵測暴雨事件方面表現出色。透過體對點降水預報的架構，我們可以建立以資料驅動 (data-driven) 的資料特徵，除了特徵定義上的客觀性之外，並可透過與過往專家所定義氣象要件做比對，更進一步的理解影響區域降水的物理因素。

關鍵字：體對點降水預報，定量降水推估

Artificial intelligence (AI) based Volume-to-Point QPE Approach

Shih-Hao Su¹, Ting-Shuo Yo², Jung-Lien Chu³, Chiao-Wei Chang¹, Hung-Chi Kuo²
Chinses Culture University¹
National Taiwan University²
National Science and Technology Center for Disaster Reduction, Taipei, Taiwan³

Abstract

In this study, we proposed a volume-to-point (VTP) framework for quantitative precipitation estimation (QPE) based on the QPESUMS Mosaic Radar dataset. With a data volume consisting of the time series of gridded radar reflectivities, we used machine learning algorithms to establish QPE modules in each weather stations. The statistical model extracts spatial and temporal features from the input data-volume and then associates the features with the location-specific rainfall. In contrast to the Z-R relation QPE module, we used the machine learning algorithms to automatically detect the existence and movement of weather systems and convections, and also associate these patterns to specific topographic attributes. We evaluated VTP framework with the hourly rainfall data of 45 surface rain gauges data in Taipei during 2013 ~ 2016. The VTP framework performed comparably well in general cases and excelled in detecting heavy-rainfall events.

Key word: Volume-to-point, QPF

一、前言

世界銀行的世界自然災害熱點及風險分析 (Natural Disaster Hotspots: A Global Risk Analysis, 2005) 報告中顯示，臺灣是少數同時面臨多種自然災害的區域，有高達 73% 以上之人口暴露在三種以上自然災害的極高風險區域。若是關心暴露兩種以上天然災害影響的高風險區域，那臺灣有超過 90% 以上的人口都暴露在風險之下。進一步針對這些影響臺灣的自然災害事件進行分析，可以發現對於臺灣而言頻率最高，且會導致大規模災損之自然災害事件，就屬於伴隨劇烈降水現象導致的洪水事件。在過去針對臺灣的極端降水事件進行系統性的分析研究工作中顯示，這些劇烈降水事件在時間與空間上呈現非均勻的分佈，而這些時空分佈特徵又與局部地區複雜地形和所伴隨之天氣型態有所關係 (Su et al., 2012)。我們發現臺灣地區的降水分佈與天氣系統有密切的關聯，像是夏季颱風與梅雨季鋒面系統可合計貢獻近 80% 以上的極端降水事件。這類型極端降水事件在臺灣發生的頻率不低，主要是因為臺灣的地理位置位於副熱帶的東亞季風區域，氣候上本就偏向高溫且潮濕，多雨就是臺灣氣候主要的特徵之一。

先前的研究中顯示，在過去 60 年的統計資料中臺灣各地都有發生極端降水事件的紀錄。我們同時注意到臺灣地區的極端降水事件有所謂發生的熱區，而發生頻率較高的時間與位置和季節有關。我們注意到，在影響臺灣地區的極端降水事件中，短延時極端降水 (short-duration extreme rainfall) 事件，對於社會民生經濟會有重大的衝擊；尤其伴隨短延時極端降水事件之暴洪 (flash flooding) 現象，相較於一般伴隨強綜觀天氣系統之長期劇烈降水導致的淹水事件更加危險。

而過去對於導致極端降水事件的相關研究中，有足夠的科學證據顯示，大氣的環境條件變化，如短時間內氣溫變化趨勢，可在短時間 (數十分鐘至數小時) 內導致降水強度的增加，

也影響降水的時間分佈特性 (Haerter and Berg, 2009; Lenderink and van Meilgaard, 2009; Haerter et al., 2010; Berg et al., 2013; Westra et al., 2013)。Su et al. (2012) 的研究中，也說明颱風季的極端降水事件發生頻率最高，但非颱風的強降水事件發生的頻率依然佔了極端降水事件的很大一部分。而這些事件的降水強度與颱風降水相比也相當，且其中一部分的極端降水事件是屬於短延時強降水類型之事件。在這些極端降水事件方面，我們也注意到一部份的強降水事件是在大氣環境無明顯綜觀天氣系統的情況下，由局地的熱對流系統導致短延時強降水事件的發生。這些極端降水事件由於缺少明確的天氣系統作為預報的參考依據，所以無論在天氣分析與預報作業上都十分的困難。

在過去十數年間，學界投入了大量的精力開發定量降水預報 (quantitative precipitation forecast, QPF) 系統。從精準定量降雨預報的觀點出發，現行作業化的定量降水預報系統中，可以依據預報的預警時間長短區分為兩種不同的主要型態。在 1-6 小時內的降水預報，常利用雷達回波資料計算對流系統移動方向加以推估，而較長期的定量降水預報則主要使用數值模式預報之產品。這兩種預報方法都有長久的發展歷史與理論基礎，也是目前作業單位最為依賴的降水預測模型，但這兩種方法也各自存在著其無法克服的限制條件。

若改以增加對於極端降水事件的預警時間為出發點，學界一直嘗試使用數值模式來協助進行極端降水事件的預警。但是受限於有限的計算資源，模式中的物理過程與網格解析度都受到一些限制，而這都會導致使用數值模式的模擬降水進行預報上會產生誤差。在數值模式中受限於模式解析度、物理過程之時空尺度或因缺少詳細量測資料作為依據的因素下，模式中部分的物理過程必須進行簡化或假設。這樣過於簡化的物理參數化方法或物理模組加上對於初始觀測資料的誤差等不確定性，在經過長時間積分後，此一誤差會被累積且逐漸放大，最終造成降水模擬的偏差。在先前研究中，我

們認知到降水預報並非如過去單純考量到模式受限於解析度的因素而造成低估降水量值；在選用不同物理參數化方法時，也會因為過度反應次網格解析度的作用導致在特定條件下會發生過度預報的現象。另一方面，在近年的天氣預報作業中，系集預報的概念（黃等，2016）被廣泛利用在各種天氣預警作業之上。透過多模式預報系統與最佳化統計技術改善誤差（葉等，2016），系集預報技術在基於誤差的隨機特性假設下，確實可以透過平均降低降水模擬上的偏差。但數值模式的降水誤差並不具有完全的隨機特性，也造成使用系集預報系統並無法完全解決降水偏差的問題。在此果改以資料驅動（data driven）的角度為出發，在透過高更新頻率之觀測資料搭配合適之即時預報方式推估/預報技術的前提下，是有機會針對一定範圍內的降水事件提供較為精準之降水預測。這樣的工作在近年吸引大量資料科學家投入，如 google 的人工智慧團隊在年初便發表了利用衛星資料進行降雨預報的模式架構（Sønderby et al., 2020）。他們的工作顯示，高更新頻率的觀測資料結合深度學習模型架構，可以提供一個接近瞬間即時的降雨預報。在其結果中，我們也注意到缺乏物理背景基礎的資料驅動模型，確實可以對於區域範圍的降雨是否發生有不錯的掌握能力，但對於降雨量的預報依舊有著相當的誤差。

如果回看目前中央氣象局所使用之雷達回波推估降水方法，在其技術的背後有一基本假設是大氣環境條件相似。在此一前提下，對流系統的強度變化較低，且對流胞會依照氣流的方向向固定方向移動，也就是傳統上的運動外延預報方法。這樣的假設在大氣環境均勻且無複雜地形的情形下，其預報可信度相對較高。但大氣對流系統的發展與維持過程並非屬於穩定平衡之狀態，除環境風場會導致系統移動外，也同時存在非常多的環境因素會影響其發展或消散。這些因素都會導致使用雷達回波資料在追蹤對流系統運動時，存在著發生推估誤差的可能性。尤其在區域氣象環境差異較大與地形

因素複雜的環境條件下，上述各種機制將導致定量降水推估偏差被放大，這也是短期定量降水預報不確定性的主要來源之一。本研究將使用中央氣象局現有之高密度之雨量觀測資料與雷達整合觀測資料，配合最新之資料科學技術，發展基於機器學習方法的短延時強降雨預警工具。我們將結合資料驅動的基本概念並融入大氣科學中對於對流系統發展變化的基礎概念，提出「體對點」（Volume-to-Point, VTP）的降雨預報模型（圖一）。

二、使用資料

本研究目前使用中央氣象局所提供之 2013-2016 年間大台北地區之觀測資料，其中包含台灣地區雷達觀測網所產製的整合雷達回波資料（QPESUMS Mosaic dataset, Chiou et al., 2004）。整合雷達資料之為二維水平空間資料，資料涵蓋範圍為 $21.8875^{\circ} \sim 25.3125^{\circ}\text{N}$ / $120.0000^{\circ} \sim 122.0125^{\circ}\text{E}$ ，時間解析度為每十分鐘一筆資料，空間解析度為 $0.0125^{\circ} \times 0.0125^{\circ}$ 。合最新之資料科學技術，發展基於機器學習方法的短延時強降雨預警工具。模式中用來作為訓練模型的地表降雨資料，是使用氣象局 45 個地面氣象站之時雨量資料。

三、機器學習技術為核心的降雨預報模型

在作業化的定量降雨推估技術主要還是基於 Marshall and Palmer (1948) 所提出的粒子大小與散射強度間的物理關係所建構的冪次方程。在傳統的 Z-R 關係架構下，所用來推估反演降雨強度的資料為相同空間上的雷達回波資訊，可視為一個點對點(point to point)的模型架構。本研究則嘗試利用多重時間的二維高解析度資料組成一個高維度的資料體(Volume)，並建立起這個資料體與特定站點降雨的推估模型。換句話說，我們針對每一個單一測站的降雨推估，訓練模型將會使用包含整個觀測範圍內前六筆的所有整合雷達回波資訊。這樣的模型架構，

與 Z-R 關係式相似之處在保留著雷達回波與降雨間的非線性關係，且更進一步地將空間上回波強度的差距與時間上變化等資訊也一並作為輸入的參數。當然基於系統的穩定的需求，在架構上一定會包含資料之前處理與檢驗機制、整齊資料之前處理與檢核機制（含特徵變數之提取）等步驟。舉例來說，在進入機器學習模型訓練前的整齊資料前處理與檢核機制階段中，所輸入資料為經過資料檢核後之連續六筆十分鐘解析度的 QPESUMS 雷達回波資料，同時合併個別測站的雨量或雨量類別標籤。

基於這樣的架構，我們在空間上每一個降雨值將由 $6 \times 12 \times 275$ 的資料體提供相關訊息，也就是共有 267,300 個獨立變數作為輸入值，透過輸入資料的過程訓練模型得到最佳化的迴歸關係。以大台北地區降雨推估為例，使用這種 VTP 架構降雨推估模型的優勢在於，可以將更多對流發展資訊作為模型預測的推估依據。就如先前研究中顯示，地形與對流交互作用的動力過程或是對流系同本身的合併、移動與強化等過程亦會影響降雨強度 (e.g. Kuo and Wu, 2019; and Miao and Yang, 2020)。而這些對流變化特徵的資訊，在 VTP 架構的模型下可透過連續時間的回波強度與空間位置變化，讓這些資訊在訓練資料驅動的非線型模型中得到充分的反映。但為了達成這樣的目的，我們必須由資料體中抽取出具有代表性的時空特徵值，在此我們借助於表徵學習 (representation learning) 與非線性擬合模型去達到這個目標。在測試各種不同的深度學習 (deep neural networks) 模型架構後，我們使用能有效處理空間變異資訊的卷積神經網路 (convolution neural network, CNN) 為基礎架構。在此可以想像雷達回波資料為一組高維度的圖像資料，資料輸入之後以特定的矩陣掃描資料並重新擷取圖像特徵，並且計算局部矩陣與原始資料的相似程度，此過程即為卷積。每一次卷積的過程會製造新的特徵矩陣，這些矩陣在隱藏層當中進行最大值池化 (max pooling) 計算以保留圖像當中最重要特徵，最後通過解碼器

(decoder) 輸出代表輸入資料特徵的低維度矩陣。圖二為本研究中所使用的 VTP 架構的示意圖。

四、實驗結果

我們使用 2013-2015 年的雷達回波與降雨資料進行 VTP 降雨模組進行訓練，並利用 2016 年的資料作為評估資料集。在大台北地區的 45 個降雨站的 QPE 測試中，平均的方均根誤差 (RMSE) 為 1.86mm/hr，在同一期間的 Z-R 推估模型 QPE 的平均 RMSE 1.90mm/hr。進一步分析兩種模型的 RMSE 的空間分佈特徵 (圖三)，可以發現無論使用哪一種模型的 RMSE 分佈都十分相似。但同時也注意到在資料品質較佳的區域，VTP 模型確實有較好的表現，這代表了只要有高品質的觀測資料作為輸入值，VTP 模型確可有效的改善降雨推估誤差。另一方面，如果我們針對降雨強度較高的事件 (在本研究中使用 30mm/hr 作為標準) 進行降雨推估能力的評估，我們可以發現 VTP 模型的平均命中率(hit rate)為 0.8，比單純使用 Z-R 推估模型的 0.34 表現更好，且在 45 個目標測站中的 39 站的表現都有所改善(圖片未展示)。圖四為 VTP 和 Z-R 關係兩個模型的針對強降雨事件的 ROC(the receiver operating characteristic)分析，我們可以看到 VTP 模型在命中率的改善的效益，比在誤報率(False alarm rate)的耗損來的更多，也顯示這個模型確實有改善定量降雨推估的能力。

五、結論

在本研究中，我們利用整合雷達回波資料建構了一個『體對點 (VTP)』的定量降水推估模型架構。與過去方法不同之處在於，VTP 架構的 QPE 模型不再是基於空間上單點的信號進行推估，而是使用覆蓋更大區域和更長時間的資料體組為模型推估的依據。這樣的資料體可以提供有關天氣系統的特徵和對流系統發展、強化與移動等資訊，並更能反映出各項雲動力與熱力特徵的變化。我們透過 CNN 的方法來抽

取資料體中的特徵，因為 CNN 具有從大量數據集中提取空間相關基本特徵的能力。實驗中使用 2013~2015 年的數據對模型進行訓練，並利用獨立的 2016 年資料評估 VTP 模型的 QPE 績效。

與傳統單純使用 Z-R 關係的 QPE 模型相比，VTP 在一般的情境下有良好的表現，並且在偵測強降雨事件方面表現出色。就降雨推估誤差修正的空間分佈而言，VTP 在人口稠密地區和主要分水嶺上的表現尤其出色，這在實際防災應用中具有重要意義。而上述兩個表現優異的區域受惠於五分山雷達所提供的高品質資料，也暗示了更加的資料體（觀測資料）有助於改善 VTP 系統的 QPE 能力。目前的工作可視為未來發展的 VTP 架構 QPE 模型的基準，可用來評估未來進一步擴展 VTP 架構的資料體後的 QPE 成效，例如納入更多不同的觀測資料或數值資訊。目前所使用的資料體僅包含一小時的整合雷達回波資料，而 VTP 架構的 QPE 方案在執行效率與成效上都已經有很好的表現。在未來只要下為調整此一架構的資料體，或進一步整合不同資料類型，亦可將這樣 VTP 的架構應用在劇烈降水事件的定量降雨預報（QPF）方案上。

六、參考文獻

黃椿喜、葉世瑄、呂國臣、洪景山，2016：系集定量降水預報方法之探討與分析－系集平均、機率擬合平均與超越之定量降水預報。大氣科學，44，173-196。

葉世瑄、林沛練、洪景山、黃椿喜，2016：機率擬合之系集定量降水預報後處理方法。大氣科學，44，83-111。

Berg, P., Moseley, C., and Haerter, J. O., 2013, Strong increase in convective precipitation in response to higher temperatures. *Nature Geosci.*, **6**, 181-185.

Chiou, Y.-M., Chang, F.-J., Jou, B.J.-D., Lin, P.-F., 2004. Quantitative Precipitation Estimation using multiple sensors.

Presented at the 8th Conf. on Atmospheric Sciences, Tao-Yuan, Taiwan.

Haerter, J. O. and Berg, P., 2009, Unexpected rise in extreme precipitation caused by a shift in rain type? *Nature Geosci.*, **2**, 372-373.

Haerter, J. O., Berg, P., and S. Hagemann, 2010, Heavy rain intensity distributions on varying time scales and at different temperatures. *J. Geophys. Res.*, **115**, D17102.

Kuo, K.-T., and C.-M. Wu, 2019: The precipitation hotspots of afternoon thunderstorms over the Taipei Basin: Idealized numerical simulations. *J. Meteor. Soc. Japan*, **97**, 000-000, doi:10.2151/jmsj.2019-031.

Lenderink and van Meilgaard, 2009, Linking increases in hourly precipitation extremes to atmospheric temperature and moisture changes. *Environ. Res. Lett.*, **5(2)**, 025, 208.

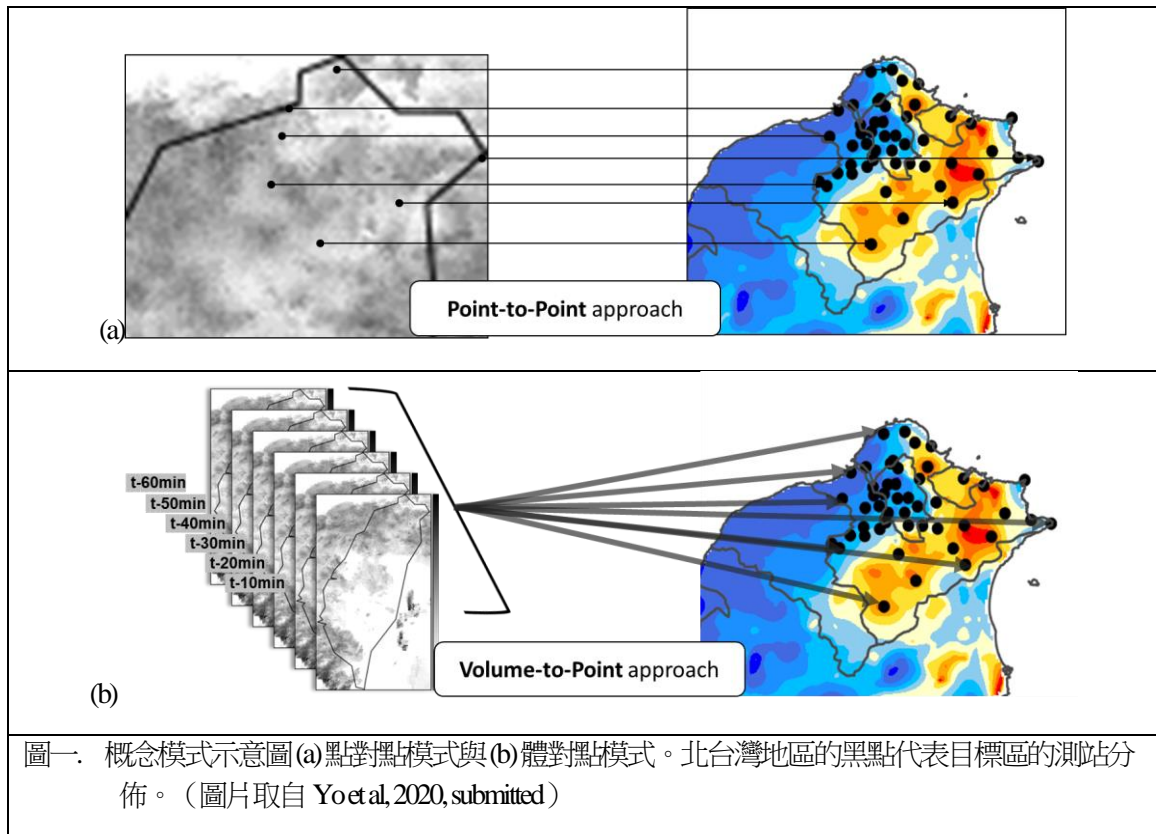
Miao, J.-E., Yang, M.-J., 2020. A Modeling Study of the Severe Afternoon Thunderstorm Event at Taipei on 14 June 2015: The Roles of Sea Breeze, Microphysics, and Terrain. *Journal of the Meteorological Society of Japan. Ser. II* **98**, 129-152. <https://doi.org/10.2151/jmsj.2020-008>.

Sønderby, Casper Kaae, et al. "MetNet: A Neural Weather Model for Precipitation Forecasting." arXiv preprint arXiv:2003.12140 (2020).

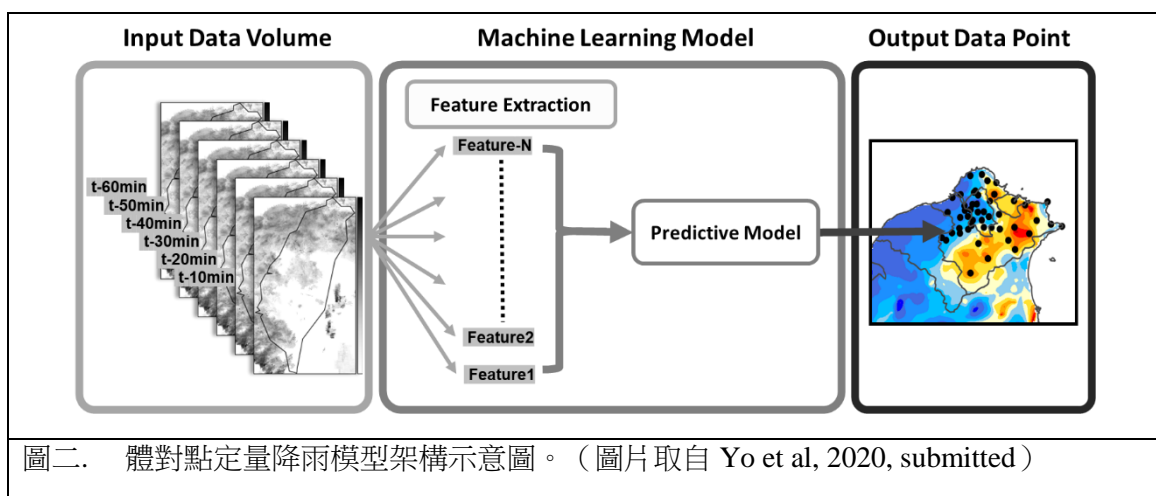
Su, -S. H., H. -C. Kuo*, L. -H., Hsu, and Y. - T., Yang, 2012, Temporal and Spatial Characteristics of Typhoon Extreme Rainfall in Taiwan. *J. Meteor. Soc. Japan*, **90**, No. 5, 721-736.

Westra, S., Alexander, L. V., and Zwiers, F. R., 2013, Global increasing trends in annual maximum daily precipitation. *J. Clim.*, 26, 3904-3918.

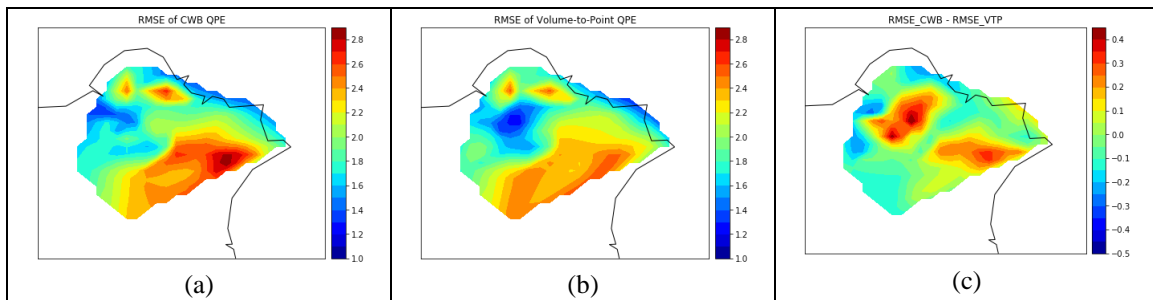
Yo, T.-S., Su, S.-H., Chu, J.-L., Chang, C.-W., & Kuo, H.-C. (2020). A Volume to Point Framework for QPE with Radar Data. *Earth and Space Science*. Submitted.



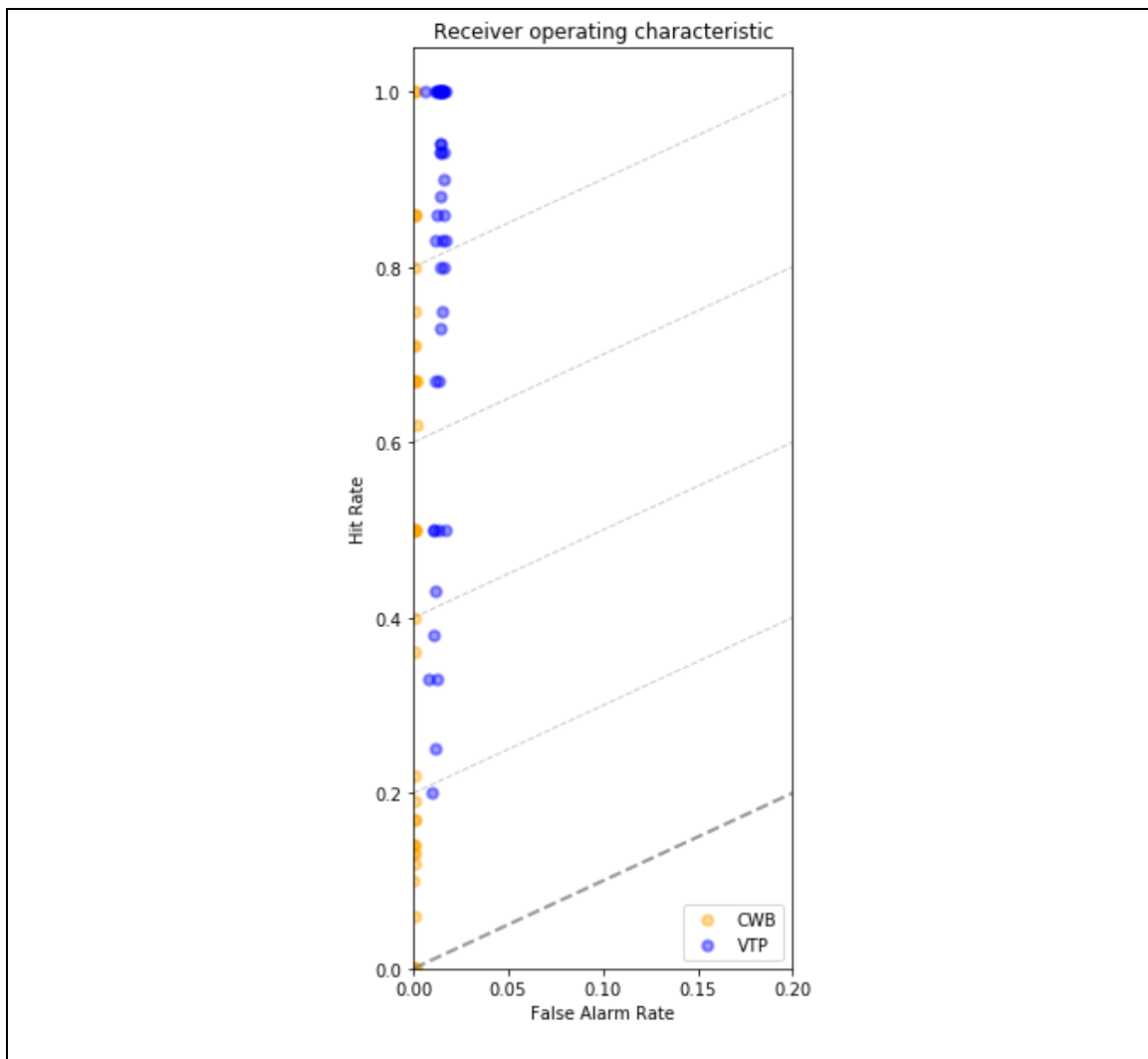
圖一. 概念模式示意圖 (a) 點對點模式與 (b) 體對點模式。北台灣地區的黑點代表目標區的測站分佈。(圖片取自 Yo et al, 2020, submitted)



圖二. 體對點定量降雨模型架構示意圖。(圖片取自 Yo et al, 2020, submitted)



圖三. 大台北地區定量降雨推估之方均根誤差的空間分佈。(a)Z-R 關係之點對點推估模型，(b)體對點定量降雨推估模型，與(c)兩者誤差相減值，單位為 mm/hr。(圖片取自 Yo et al, 2020, submitted)



圖四. 大台北地區定量降雨推估之 ROC 分析。黃點為 Z-R 關係之點對點推估模型，藍點為體對點定量降雨推估模型。圖上的斜虛線標示模型推估能力，越往左上角代表模型能力較佳。(圖片取自 Yo et al, 2020, submitted)