

中央氣象局110年天氣分析與預報研討會

應用機器學習估計客觀颱風強度

陳冠儒¹ 章鶴群² 張育承² 周鑑本²

林賢宏³ 羅時宜³ 張凱瑞³ 藍秀仁⁴ 蔡宜真⁴

¹中央氣象局第四組 ²中央氣象局衛星中心

³樺鼎商業資訊股份有限公司 ⁴DataRobot台灣分公司

摘要

熱帶氣旋的強度與其結構特徵息息相關，美國氣象學家：Vernon F. Dvorak基於這個概念，配合自身多年的經驗創造了Dvorak颱風強度估計法（Dvorak Technique），透過同步氣象衛星雲圖，觀察並分析氣旋系統的雲系相關特徵，不需直接進行風速量測，便能得到可靠的颱風強度指標。由於它的分析過程依賴「主觀辨識」，氣象人員需要具備大量的實務經驗，且不同的分析人員可能訂出不同的估計值。藉由DataRobot機器學習平台引進自動化機器學習（Automated Machine Learning, AutoML）之圖形辨識技術，可在Dvorak法的概念之下，由機器學習模組對衛星雲圖進行辨識，提供預報員客觀的颱風強度指標，同時大幅降低AI模型開發的門檻與時程。

針對2020年颱風的實測數據來看，表現最佳的模型為「Light Gradient Boosted Trees Regressor with Early Stopping」，RMSE為6.24m/s，平均絕對誤差為4.38m/s，顯示機器學習能提供可靠的颱風強度資訊。而分析個別颱風後發現，機器學習針對中颱以下的強度估算與中央氣象局官方發佈之颱風強度非常接近，但由於強烈颱風的訓練樣本不足，使得極端強烈的颱風強度估計不盡理想。但整體而言機器學習技術確實能夠有效應用在天氣分析的議題，並為預報人員提供可靠且客觀的颱風強度資訊。

關鍵字：颱風強度、機器學習、Dvorak法

一、前言

2020年是自1964年以來首次沒有颱風直接侵襲台灣的年份，雖然民眾少了很多生命財產的損失，但另一方面卻讓台灣在2021年面臨嚴重的缺水危機，顯見颱風不僅是臺灣主要的天然災害之一，在水資源方面亦扮演著重要地位。

如何掌握確實的颱風強度、位置與路徑等資訊，是一門相當重要的課題，儘管我們可以利用飛機或船舶等方式進行直接觀測，但風險過高且難以取得連續性的資訊；氣象雷達能夠提供精確的風場及降雨資訊，但由於量測距離的限制，遠洋資訊難以取得。因此，熱帶氣旋接近陸地之前，氣象衛星是非常有用且難以取代的分析工具。

熱帶氣旋在發展初期，僅由許多對流雲簇組成，結構鬆散且氣旋式環流不明顯。隨著熱帶氣旋持續發展，對流雲系增強伴隨雲頂亮溫下降，逐漸形成

渦流以及螺旋雨帶。當氣旋發展到足夠強度時，中心將會因下沉氣流而出現颱風眼，並在外側形成風雨最強的眼牆結構。

由此可知，熱帶氣旋的強度與其結構特徵息息相關。美國氣象學家：Vernon F. Dvorak基於這個概念，配合自身多年的經驗創造了Dvorak颱風強度估計法（Dvorak Technique），此方法透過同步氣象衛星之紅外線與可見光衛星雲圖，觀察並分析氣旋系統的雲系特徵，以估計出颱風強度（Dvorak 1975, 1984）。此分析方法發表於1984年，並在3年後正式由世界氣象組織通過使用，並沿用至今（圖一）。

Dvorak颱風強度估計法的優勢在於不需要直接測量風速，便可以得到可靠的颱風強度資訊，但Dvorak颱風強度估計法並非沒有弱點。由於它的分析過程非常倚賴「主觀辨識」，這不僅代表氣象人員本身需要具備大量的實務經驗，不同人員在分析相同的熱帶氣旋時，也可能因為主觀差異而推算出

不同的強度指數。

美國威斯康辛大學麥迪遜分校的氣象衛星研究合作研究所 (Cooperative Institute for Meteorological Satellite Studies, CIMSS) 以Dvorak法為基礎，建構一套完全透過電腦運算的先進Dvorak技術

(Advanced Dvorak Technique, ADT)，該演算技術對颱風衛星影像進行系統性的處理與分析，並納入微波資料、數值模式等特徵輔助資訊，且無需氣象人員介入，成功摒除因人為主觀所造成的差異，取得了客觀的颱風強度資訊，該項技術是目前世界上大多數熱帶氣旋作業單位的重要參考 (Olander 和 Velden 2007)。

ADT並非唯一的客觀方法，機器學習 (Machine Learning, ML) 也是另一種可以嘗試的工具。早期電腦受限於運算能力，難以消化機器學習所需的龐大運算量，直到近年來電腦技術的發展，加上GPU的應用大幅提高運算處理的能力，讓機器學習重新獲得重視。

機器學習屬於人工智慧 (Artificial Intelligence, AI) 中的一個分支，讓電腦從大量資料中自行尋找存在於資料彼此之間的規則，並建立一套專屬的演算流程，當下次電腦接收到相似且新產生的資訊時，便可以透過它的「學習經驗」，告訴我們正確的答案。

根據麥肯錫公司2017年的研究報告指出，高科技、汽車、電信、金融等行業較早涉略AI領域，在AI技術上與其他產業相比差距愈來愈大，而AI技術的使用可以為產業帶來更多的價值，創造更高的利潤 (Bughin等，2017)。

以金融業為例，銀行放款後若借款人無法還款而形成呆帳，將直接造成銀行帳面上的損失。假設銀行希望「判斷某借款人會不會準時繳付款項」，那麼銀行會將過去每個借款人的特徵資料：如年齡、收入、性別、工作性質、借款額度等輸入機器學習模型，再提供這些借款人最後是否準時還款的目標資料(Target)，模型便會自行分析「借款人資料」與「還款狀況」之間的規則，找出借款人具備哪些特徵時將比較容易違約。

如此一來，每當有人來銀行希望借錢時，銀行便可請對方提供相關資料，輸入機器學習模型來評估該借款人「違約的機會高不高」，加以評估是否核放貸款，從而減少呆帳增加的風險。

近年來機器學習在氣象學界與業界已逐漸受到重視，中央氣象局不斷努力，嘗試找出機器學習技術應用在氣象分析與預測上的可行性。氣象領域在機器學習方面起步較晚，不管是研究方法論或程式技術仍有非常多需要學習的地方，然而從頭培養AI人才曠日廢時，我們難以期望在短時間內跟上最新的AI技術。

因此，本研究將透過DataRobot平台所提供之自動化機器學習 (Automated Machine Learning) 功能，將模型訓練等繁瑣的過程自動化，讓氣象人員能專注於擅長的氣象問題，從而有更多心力專注於關鍵特徵的發掘，並且讓我們不需要具備高深的程式技術與經驗便能完成特定的AI應用目標，協助我們更快速地在機器學習的領域裡站穩腳步。

1984年的Dvorak技術的估算結果具備一定程度的主觀意識，而我們可以蒐集過去颱風的衛星影像，並搭配當時官方發佈的颱風強度資訊，另外再加上其他輔助性質之特徵資料，便可以利用機器學習建立一套新的颱風強度估計方法。基本假設依然建立在Dvorak方法之上，不同的是，颱風雲型與強度的判定完全交給機器學習代勞，藉此排除人為主觀辨識所造成的偏差，提供較為客觀的預測資訊。

二、 訓練資料

2015年7月份日本新一代同步氣象衛星：向日葵8號正式作業，因此本研究僅使用2015年7月起至2020年12月之颱風或熱帶氣旋資料，避免衛星輻射儀的特性差異而導致變因增加。

本研究以中央氣象局官方發佈之「颱風近中心最大風速」為機器學習的訓練目標，訓練樣本使用2015年7月至2019年間共計4329個時間點，每個時間點均包含向日葵8號衛星16個頻道之衛星雲圖以及其他輔助性質之特徵資料，然後將這些特徵資料與訓練目標輸入DataRobot機器學習平台進行一系列自動化的建模處理。例如，資料預處理、特徵工程、演算法選擇、參數調整、模型整合、模型交互評比，評比出最佳的演算法模型。

本研究經過眾多嘗試後，發現以下衛星頻道影像與輔助資料之搭配，能讓機器學習具備最佳表現：

- 紅外線窗區頻道 (10.5微米)
- 高層水氣頻道 (6.3微米)
- 短波紅外線頻道 (3.9微米)

- 紅外線窗區頻道減去高層水氣頻道
- 颱風經緯度
- 太陽日
- 月份

表現最佳的機器學習模型為「Light Gradient Boosted Trees Regressor with Early Stopping」，其訓練資料交叉驗證之RMSE為4.11m/s。

三、 2020年驗證資料

驗證資料採用2020年生成於西北太平洋的25個熱帶氣旋，有效資料時間點共計551筆。由整體統計數據來看，表現最佳的模型：「Light Gradient Boosted Trees Regressor with Early Stopping」其RMSE為6.24m/s，平均絕對誤差為4.38m/s，表示機器學習能提供可靠的颱風強度資訊。第二名的模型：「eXtreme Gradient Boosted Trees Regressor with Unsupervised Learning Features」之表現也相當接近最佳模型RMSE值為6.24m/s，平均絕對誤差為4.41m/s。

接著進一步比較不同等級颱風的表現（表一），「Light Gradient Boosted Trees Regressor with Early Stopping」模型對於估計熱帶擾動與輕度颱風強度的RMSE分別為4.5m/s及4.7m/s，差異並不明顯。但是當強度來到中度颱風時，其RMSE上升為8.9m/s、強烈颱風則進一步來到16.7m/s。顯示當颱風來到一定強度時，偏差將會顯著地增加。

如此顯著差異可能來自於訓練資料樣本數的差異。機器學習必須仰賴大量訓練資料以確保演算法能有效學習到資料間的規則，在訓練資料中，熱帶擾動等級之訓練資料共有1362筆，但是強烈颱風僅有261筆，明顯較少，因此學習效果便會差強人意（表二）。

為了進一步分析不同等級颱風強度估計品質之差異性，我們從驗證資料中分別選出輕度、中度與強度各2個颱風做為代表，將它們整個生命週期的近中心最大風速數據繪製成時序圖。

圖二為2020年輕度颱風：辛樂克與蓄蜜進行強度估算之驗證結果，藍線為中央氣象局發佈的颱風近中心最大風速值，橘線為最佳機器學習模型所估算出來的風速。由圖可以看出，機器學習針對輕

的強度估算與官方強度估計非常接近，顯示針對輕度颱風估計具備相當高的可信度。

圖三則為中度颱風的驗證結果，可以看到機器學習與官方估計之間的偏差較輕度颱風大，但並沒有偏離太多，「巴威颱風」的強度變化趨勢相當一致，「哈格比颱風」與官方颱風強度相比，在8月3日有較為明顯的浮動。

強烈颱風的強度估計則出現顯著的差異性。由圖四可見機器學習評估「海神颱風」的強度時雖有所偏差，但就強度趨勢而言表現還算穩定，但對極端強烈的「天鵝颱風」強度估計時，機器學習在強度顛峰時的估計值與官方估計值間出現蠻大的落差，天鵝颱風在11月1日來到強度顛峰，預報中心所發佈的近中心最大風速達到65m/s，但機器學習的結果最高僅有50m/s左右。

如前文所述，訓練資料的數量直接影響機器的學習成效，儘管整體訓練資料超過4000筆，但51m/s以上的訓練資料僅有261筆，使得機器學習模型對強烈颱風可能尚未學習完整，當出現極端強烈的颱風個案時，機器學習可能無法順利地掌握颱風強度。

四、 實際應用

受限於中央氣象局官方資料的發佈頻率，每小時最多僅有1筆官方強度估計可供校驗。但在實際應用時，向日葵8號衛星每小時高達6次的全碟掃描，能讓我們每10分鐘便可取得一次最新的客觀強度估計，大幅增加可供參考的資料密度及實用性。

本研究所建立之機器學習模組已實際應用於2021年之颱風強度估計，分析颱風完整生命週期間的強度變化，並與中央氣象局預報中心之官方發佈颱風強度進行比較。機器學習估算強烈颱風：舒力基之颱風強度隨時間的變化如圖五，橘線為機器學習估算舒力基颱風之強度指標，每10分鐘計算一次颱風強度，並以過去3小時之平均值呈現，以利分析颱風強度分析的整體趨勢；藍線為中央氣象局官方發佈之颱風強度，亦為機器學習的預測目標，兩者愈接近表示機器學習的表現越好；灰線為ADT估算之強度變化，作為另一種客觀颱風強度分析的參

考值。

比較機器學習與官方發佈之颱風強度可以發現，舒力基颱風在發展階段至強度抵達顛之時，機器學習所估計的強度偏弱且較不穩定，後期颱風逐漸衰減階段，其估計值才與官方數值較為接近。就統計數值來看，機器學習估計舒力基颱風強度的 RMSE 為 9.8m/s，平均絕對誤差為 7.0m/s。而 ADT 在初期表現不錯，但當颱風強度達到顛峰時，與官方數值相比有明顯偏高的現象。

另以中度颱風：烟花為例（圖六），由圖可見比起 ADT 所計算之颱風強度值，機器學習與官方所發佈的強度不論是數值或是趨勢均更為接近。而統計結果顯示，機器學習估計烟花颱風強度的 RMSE 僅為 3.9m/s，平均絕對誤差值為 3.2m/s，顯見機器學習對於烟花颱風的掌握度較舒力基颱風高。

就上述兩個個案的結果分析，機器學習對於中度颱風的強度分析表現相當不錯，但針對強烈颱風的強度估算存在較大的偏差，其可歸咎於強烈颱風訓練資料不足，使得極端強烈之颱風出現時，使用機器學習方法可能無法順利地掌握颱風強度。

五、 結語

人工智慧隨著近年來電腦運算能力的演進，已默默地從科幻小說中走入現實世界，並且持續蓬勃發展中。機器學習作為人工智慧的分支，以訓練演算法的方式尋找存在於資料間的規則，並為新資料做出最佳決策或預測，如今已廣泛應用於醫療、購物、商業、科技業等用途，而機器學習技術在氣象領域的應用，尚屬於剛起步的階段。

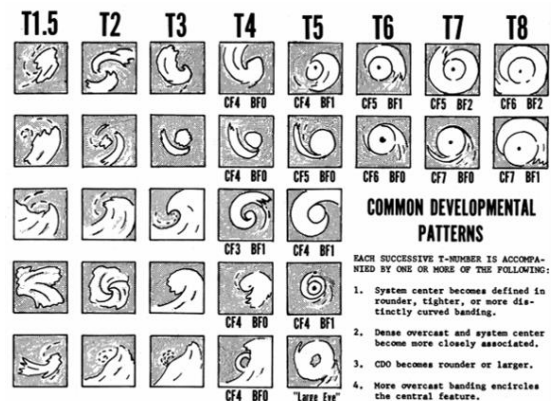
本研究嘗試以主觀辨識的 Dvorak 法為基礎，嘗試利用機器學習技術，創造客觀之颱風強度估計方法。結果顯示，機器學習技術能夠有效應用在天氣分析的議題，並為分析人員提供可靠且客觀的颱風強度資訊。

2020 年共計 25 個颱風個案之驗證結果，機器學習估計的颱風強度 RMSE 為 6.24m/s，平均絕對誤差為 4.38m/s，除了強烈颱風以外，對於其他颱風等級

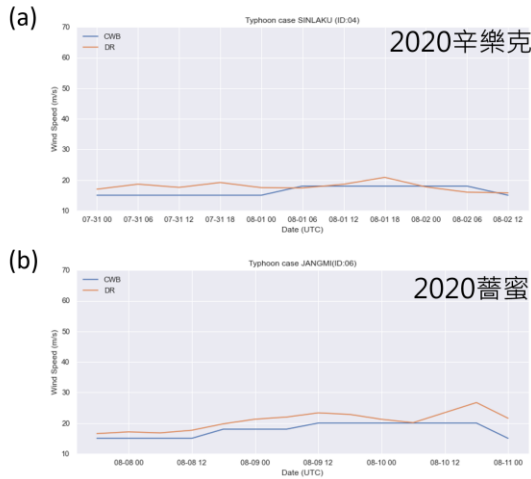
均有不錯的預測性，但隨著颱風強度增加，估計的品質與穩定性會下降。而 2021 年實際投入即時作業，同樣出現不易掌握強烈颱風強度之現象：對強颱風舒力基強度估計之 RMSE 高達 9.8m/s，但對於中颱烟花僅有 3.9m/s，表現明顯較佳。對不同強度颱風之估計值有明顯不同的準確度，可歸咎於強烈颱風訓練樣本數不足所致。

想要解決強烈颱風樣本數不足的問題，就需要增加訓練資料，本研究僅採用 2015 年 7 月至今的衛星雲圖，主要是為了避免不同世代同步衛星輻射儀的特性差異，對研究產生新的變異因子。而中央氣象局從日本 GMS-5 同步衛星開始運作以來，至今已累積了十數年的資料，只要小心處理不同衛星間資料的差異性，就能夠增加數倍的訓練資料，進而可以解決因極端強颱風的資料稀疏性所造成的預測不穩定問題。

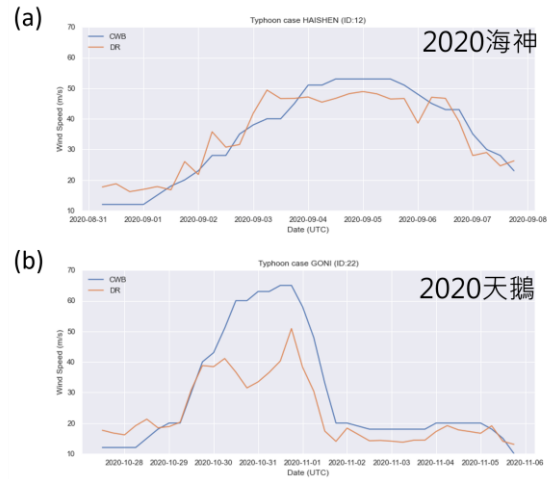
圖表



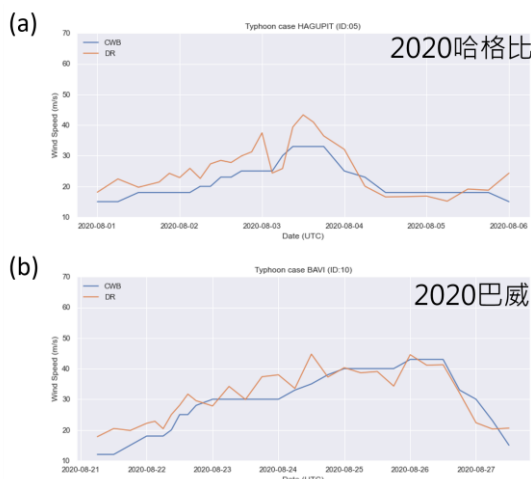
圖一、Devorak 颱風強度估計法颱風雲系分布及強度數值（T-number）之對照表。利用衛星雲圖找出颱風最符合的雲系配置，進而得到颱風強度指標。



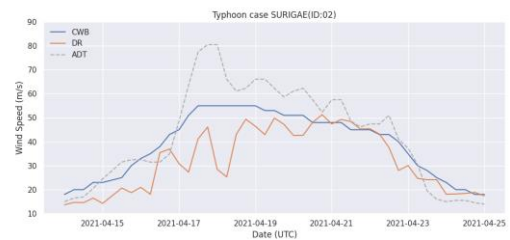
圖二、機器學習對2020年輕度颱風：辛樂克與薔蜜進行強度估算之驗證結果，橫軸為資料時間，縱軸為颱風近中心最大平均風速，單位m/s。橘線為訓練結果最佳之機器學習模型，藍線則為中央氣象局官方發布之颱風強度。由圖中顯見針對兩個輕度颱風的個案，機器學習估算的強度與官方十分接近。



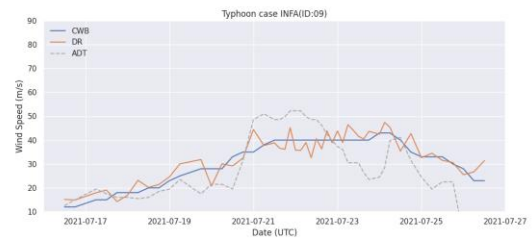
圖四、與圖二相同，但為2020年之強度颱風：海神與天鵝。研究發現機器學習針對強烈颱風下限的海神，預測雖然呈現較大之上下波動，但整體趨勢仍相當接近官方預報。但是當面對極端強烈的天鵝颱風，其強度估計顯得掌握度差強人意與明顯不足。



圖三、與圖二相同，但為2020年之中度颱風：哈格比與巴威。由圖可以發現機器學習對中度颱風的強度估計有些許的偏差，但整體趨勢仍相當接近官方預報。



圖五、2021年舒力基颱風之颱風強度估計時序圖。藍線為氣象局官方發布的強度資料，橘線為機器學習之強度估計值，灰線為ADT之強度估計值。



圖六、同圖五，但為2021年烟花颱風。由圖可見比起ADT所計算之颱風強度值，機器學習與官方所發佈的強度不論是數值或是趨勢均更為接近。

颱風強度估計之最佳機器學習模型		
Light Gradient Boosted Trees Regressor with Early Stopping		
颱風等級	RMSE	MAE
全體資料	6.24	4.38
熱帶擾動(風速 16m/s 以下)	4.46	3.25
輕度颱風(風速 17~32m/s)	4.68	3.66
中度颱風(風速 33~50m/s)	8.86	6.63
強烈颱風(風速 51m/s 以上)	16.66	13.55

表一、最佳機器學習模型在不同強度熱帶氣旋的估計表現。驗證資料為 2020 年於西北太平洋上生成的 25 個熱帶氣旋，共計 551 筆資料。

訓練資料數分布情況	
颱風等級	資料數
熱帶擾動(風速 16m/s 以下)	1362
輕度颱風(風速 17~32m/s)	1649
中度颱風(風速 33~50m/s)	1057
強烈颱風(風速 51m/s 以上)	261

表二、機器學習使用之訓練資料分布情形。不同強度的熱帶氣旋有不同數量的資料數目，越強的颱風可供訓練的資料數越少。

參考文獻

Dvorak, V., 1984: Tropical cyclone intensity analysis using satellite data.

NOAA Tech. Rep. NESDIS 11, NOAA, 47

pp., https://satepsanone.nesdis.noaa.gov/pub/Publications/Tropical/Dvorak_1984.pdf.

Olander, T. L., and C. S. Velden, 2007: The advanced Dvorak technique:

Continued development of an objective scheme to estimate tropical cyclone intensity using geostationary infrared satellite imagery. *Wea. Forecasting*, **22**, 287–298, <https://doi.org/10.1175/WAF975.1>.

Forecasting, **22**, 287–298, <https://doi.org/10.1175/WAF975.1>.

Bughin, Jacques, and Eric Hazan, Sree Ramaswamy, Michael Chui, Tera Allas,

Peter Dahlström, Nicolaus Henke and Monica Trench, 2017: Artificial Intelligence:

the Next Digital Frontier?, *McKinsey Global Institute*, June 2017