

國立臺灣大學理學院大氣科學所

碩士論文

Department of Atmospheric Sciences

College of Science

National Taiwan University

Master Thesis



應用三維卷積神經網路在濕對流的參數化

The representation of moist convection using 3D

Convolutional Neural Networks

鄒適文

Shih-Wen Tsou

指導教授：吳健銘 博士

Advisor: Chien-Ming Wu, Ph.D.

中華民國 108年 2月

February 2019



## 謝辭

首先，感謝父母在各種方面的支持，在我碩一寒假想要休學的時候鼓勵我繼續努力，並且相信我的選擇。謝謝兩年前的自己有好好地撐下去，沒有被壓力擊垮。感謝健銘老師以及維婷老師不管在研究上或是人生目標上各種的建議，讓我真心感覺在大氣系做研究是很幸福的，順帶一提，研究做累了可以去老師辦公室泡咖啡、吃個零食真的是很舒壓的事情。感謝游董在研究、工作方向上給的建議，沒有您的建議與幫忙，或許我的研究還沒辦法完成。感謝蘇老師在論文方向上的建議。感謝俊彥學長在撰寫論文上的各種幫忙。感謝李弘毅老師在學校開設了扎實的機器學習、深度學習課程，讓我也能夠學到最前端的技術。感謝女朋友友廷在生活中的陪伴，不管是在我低潮的時候帶我去吃大餐、研究做不下去的時候帶我去網咖紓壓、晚上睡不著覺的時候陪我聊天，都是我繼續做研究的動力。感謝柏鈞爾偶陪我打遊戲、聊個天，雖然我們在遊戲上的年齡都老了，但是我們的觀念能夠彌補操作上的不足。感謝琇榕常常找我去重訓，微量的運動維持著我健康的身心。感謝介安偶爾找我去打球、喝飲料，雖然真的很偶爾，但是跟老同學一起出門心情總是很好。

雖然我不善言辭也不善表達，但在當研究生的兩年半中，我真心感謝我遇到、陪著我度過這段日子的人事物，我一個人是沒有辦法這麼順利的完成研究的。



## 摘要

濕對流過程參數化是造成氣候模式和數值天氣模式不確定性的主要原因。隨著硬體和技術的發展，前人嘗試使用機器學習藉由高解析雲解析模式的結果求得特定網格解析度下的濕對流效應(機器學習濕對流參數化)，並期待未來能夠取代現存模式中由積雲參數化所診斷的濕對流效應。

而另一方面，氣候模式及數值天氣模式的發展潮流為增加水平解析度，以獲得更局部的氣象資訊並期望更準確的預測濕對流的發展過程。然而現存的積雲參數化在模式解析度逐步上升至~ $O(10\text{km})$ 尺度時，濕對流過程該由積雲參數化處理或是由網格尺度的物理過程處理變得模糊不清(grey zone resolution)。在這樣的水平解析度下，網格內可能不再同時包含對流以及其環境，意即決定網格對流效應的環境將延伸至相鄰網格。

然而，至今仍未有研究評估在 grey zone resolution 下以相鄰網格作為決定對流效應之環境對於模擬結果的影響。因此本研究嘗試利用高解析雲解析模式(VVM)的資料，並使用一種可掃描區域內參數梯度的 3D 卷積神經網路，在 Grey zone resolution 利用各個網格及其相鄰網格的資訊，預測各個網格的對流效應。為了證實在 grey zone resolution 鄰近網格資訊對於預測對流效應的助益，本研究呈現多元線性回歸(Multiple Linear Regression)及深度神經網路(Deep Neural Networks, DNN)的結果作為對照組，它們分別是機器學習模型的最低標準，及最基礎的深度學習模型。研究結果指出，由 3D 卷積神經網路所預測的對流效應在對流位置、對流強度高度分布及對流效應對可降水量的反應皆較另兩種模型有更好的表現，其中在對流效應對可降水量的反應的預測中，3D 卷積神經網路甚至在環境濕度較高的環境下的預測，可以捕捉到 5 公里處 freezing level 非常細節的對流特徵，這是其他模型較難掌握的現象。

**關鍵字：** 機器學習、深度學習、三維卷積神經網路、濕對流參數化、雲解析模式、VVM



## Abstract

The parameterization of moist convection contributes to the uncertainty in climate modeling and numerical weather prediction. As the improvement of computational resources and the theory of machine learning, building parameterization of moist convection based on machine learning method becomes feasible nowadays. In the meanwhile, the grid size used by GCMs become finer in order to represent processes at finer scale, in which more accurate representation of moist convection is expected. However, the ambiguity between subgrid-scale and resolved scale processes is introduced which arouses the grey zone problem. In such resolution  $\sim O(10\text{km})$ , the relationship between moist convection and its environment may not be distinguishable upon the same grid box. This study tries to apply the machine learning method to evaluate the model that considers the adjacent grid boxes in grey zone resolution. We use 3-Dimension Convolutional Neural Networks (3D-CNN) as our model to predict the effect of moist convection according to the meteorology parameters from both the target grid box and its surroundings. The high-resolution data is obtained from the results of a vector vorticity equation cloud-resolving model (VVM) under various environmental conditions. The unique aspect of 3D-CNN is that it can detect the gradient of each fields between adjacent grids within a certain domain. The result based on two other models, which are Multiple Linear Regression and Deep Neural Network is also shown in this study to manifest our findings. The result shows that 3D-CNN predicts more reasonable spatial distribution and vertical distribution of the effect of moist convection, and the response of the effect of moist convection to total precipitable water compared to the other models.

**Key words :** Machine Learning, Deep Learning, 3D-Convolutional Neural Networks, representation of moist convection, cloud-resolving model, VVM

# 目錄



謝辭 .....	i
摘要 .....	ii
Abstract .....	iii
目錄 .....	iv
圖目錄 .....	vi
1. 前言 .....	1
2. 資料與方法 .....	4
2.1 資料 .....	4
2.2 機器學習模型之輸入及輸出 .....	5
2.3 三維深度卷積神經網路(3D Deep Convolutional Neural Networks, 3D-CNN) .....	6
3. 機器學習模型評估 .....	8
3.1 對流位置 .....	8
3.2 對流垂直強度分布 .....	10
3.3 網格可降水量與對流強度的關係 .....	12



4. 結論與討論 .....	15
圖片 .....	18
參考文獻 .....	43
附錄 .....	46
機器學習及監督式學習 .....	46
多變量線性回歸(Multiple Linear Regression).....	48
深度學習(Deep Learning) .....	49
深度神經網路(Deep Neural Networks).....	50
激勵函數(Activation function) .....	51
對流可用位能(CAPE) .....	51
垂直積分水氣輻散(MDC) .....	51
濕靜能(MSE).....	52
主成分分析(Principal components analysis, PCA) .....	52
附錄 reference .....	53



## 圖 目 錄

圖 1 平均示意圖。將原始 256*256 格(512*512 公里)的網格每 8*8 格(16*16 公里)做水平平均，將平均後的資料作為機器學習模型的輸入及輸出。.....	18
圖 2 (a)大尺度之冷卻、濕化效應，(b)水平風，(c)位溫，(d)比濕的初始場。位溫、比濕和大尺度強迫是取自 GARP Atlantic Tropical Experiment(GATE) phase III 的平均場。水平風為 TOGA COARE 之觀測資料，綠色為有風切之實驗、藍色為無風切之實驗。模擬中有 $2Kd - 1$ 的輻射冷卻。本圖取自 [Tsai and Wu, 2017]。 .....	19
圖 3 實線表示原始資料集之總、中、低層網格可降水量分布，分別以黑色、紫色、橘色代表；虛線代表測試集之抽樣結果，分別以黑色、紫色、橘色代表不同層總、中、低層網格可降水量之頻率分布。.....	20
圖 4 (a)傳統全連接神經層及(b)卷積層。傳統全連接層的會將輸入連接到下一層的每個神經元，因此每個神經元都會接收到全部輸入的使用特徵，而卷積層只會使用到部分的特徵，如上圖(b)的n3只會與 $x_2, x_3, x_4$ 的輸入有連結。..	21
圖 5 (a)卷積核(kernel) (b)全連接層。卷積核之掃描範圍通常比輸入之圖像小很多，並且在掃過整張圖的過程中，都使用相同的權重。同時具有局部連接之特性，意即不一次輸入整張圖片。傳統全連接神經網路則一次輸入整張圖片。...	21
圖 6 本研究使用的 3D-CNN 架構示意圖。其利用 3D-CNN 將原始模式輸入之網格間的資訊，利用卷積層抽取為特徵圖後，以池化層壓縮，最後將特徵拉直，藉由全連接層的轉換，將最後的特徵轉為對流的結構輸出。.....	22
圖 7 可降水量之分布，紅色線為某時間下可降水量分布，綠色線為整組測試集可降水量的分布。由左至右為由乾到濕環境的分布。圖中橫軸為可降水量單位為毫米(mm)，縱軸為分布之機率大小(frequency) .....	23



圖 8 高度 3 公里垂直濕靜能紊流通量水平分布，顏色代表其量值大小，單位為(K m/s)。	24
圖 9 3D 卷積神經網路(3D Convolutional neural networks)模型預測之高度 3 公里垂直濕靜能紊流通量，顏色代表其量值大小，單位為(K m/s)，線條為真實模式輸出之對照。	25
圖 10 多元線性回歸(Multiple Linear Regression)預測之高度 3 公里垂直濕靜能紊流通量，顏色代表其量值大小，單位為(K m/s)，線條為真實模式輸出之對照。	26
圖 11 深度神經網路(Deep Neural Networks)模型預測之高度 3 公里垂直濕靜能紊流通量，顏色代表其量值大小，單位為(K m/s)，線條為真實模式輸出之對照。	27
圖 12 真實垂直濕靜能紊流通量之隨高度分布頻率圖(CFADs)，橫軸為(K m/s)之量值，縱軸為高度座標(m)，顏色為機率，某高度總和之機率為 1。	28
圖 13 3D 卷積神經網路預測之垂直濕靜能紊流通量之隨高度分布頻率圖(CFADs)，橫軸為(K m/s)之量值，縱軸為高度座標(m)，顏色為機率，黑色線條為真實模式輸出之對照，某高度總和之機率為 1。	29
圖 14 深度神經網路預測之垂直濕靜能紊流通量之隨高度分布頻率圖(CFADs)，橫軸為(K m/s)之量值，縱軸為高度座標(m)，顏色為機率，黑色線條為真實模式輸出之對照，某高度總和之機率為 1。	30
圖 15 多元線性回歸預測之垂直濕靜能紊流通量之隨高度分布頻率圖(CFADs)，橫軸為(K m/s)之量值，縱軸為高度座標(m)，顏色為機率，黑色線條為真實模式輸出之對照，某高度總和之機率為 1。	31
圖 16 真實模式對流強度隨著網格總共可降水量的關係，顏色由冷色到暖色為網	



格可降水量少到多。其 x 軸為對流效應的強度  $w'h'$  (K m/s)，y 軸為高度座標 (m)。其中紅線為與機器學習模型比較的橫切面，黑色箭頭為與機器學習模型比較的水氣量選取。 ..... 32

圖 17 真實模式對流強度隨著網格中層水氣的關係，顏色由冷色到暖色為網格可降水量少到多。其 x 軸為對流效應的強度  $w'h'$  (K m/s)，y 軸為高度座標(m)。

..... 33

圖 18 真實模式對流強度隨著網格低層水氣的關係，顏色由冷色到暖色為網格可降水量少到多。其 x 軸為對流效應的強度  $w'h'$  (K m/s)，y 軸為高度座標(m)。

..... 34

圖 19 垂直濕靜能紊流傳送與網格水氣量在 6 公里高度下的關係。黑色線為模式真實對流、紅色線為 3D 卷積神經網路、綠色線為深度神經網路、藍色線為多元線性回歸模型所預測的結果。 ..... 35

圖 20 在環境較濕時的極端對流效應對總水氣量反應之垂直結構，其中黑色線為模式真實對流、綠色線為 3D 卷積神經網路、黃色線為深度神經網路、紅色線為多元線性回歸模型所預測的結果。 ..... 36

圖 21 環境較濕時的極端對流效應的對低層水氣反應之垂直結構，其中黑色線為模式真實對流、綠色線為 3D 卷積神經網路、黃色線為深度神經網路、紅色線為多元線性回歸模型所預測的結果。 ..... 37

圖 22 環境較濕時的極端對流效應的對中層水氣反應之垂直結構，其中黑色線為模式真實對流、綠色線為 3D 卷積神經網路、黃色線為深度神經網路、紅色線為多元線性回歸模型所預測的結果。 ..... 38

圖 23 環境較乾時的對流高度結構，其中黑色線為模式真實對流、綠色線為 3D 卷



積神經網路、黃色線為深度神經網路、紅色線為多元線性回歸模型所預測的 結果。 .....	39
--	----

圖 24 本研究的資料前處理流程。首先挑選特徵、計算對流效應(標籤)，接著將原  
始 2 公里解析度的資料做平均後得到 16 公里解析度的特徵、標籤。同時將  
標籤利用 PCA 降為 5 維。將資料依時間的維度打亂(方便後面的分析)並將其  
分為訓練集、測試集，並且做資料的標準化。最後將資料塞入 3D 卷積神經  
網路之中，本資料將包含目標(標籤位置的網格)及其鄰近網格的特徵作為訓  
練預測對流的輸入。 ..... 40

圖 25 本研究的 3D 卷積神經網路模型架構。 ..... 41

圖 26 本研究的驗證流程。本研究分成三個部分驗證機器學習模型的預測與真實  
模式對流效應的差別，分別為對流的位置、對流垂直強度分布以及網格可降  
水量與對流強度的關係。而其中對流效應在不同濕度下的反應與其強度的垂  
直分布能夠被 3D 卷積神經網路很好的預測，這是傳統對流參數化很難做到  
的。 ..... 42



## 1. 前言

在全球環流數值模式中，有許多物理過程的作用尺度小於常用的網格解析度大小(約 100 公里)，因此需要設計各種參數化來表達這些物理過程對網格的效應。然而由於缺少觀測資料的佐證，我們對於這些小尺度的物理過程並無法詳盡的掌握，使得參數化的設計過程中會有其盲點，進而造成各種面向的模式誤差。然而在近年中研究中，我們在濕對流過程的掌握還是存在著瓶頸[Randall et al., 2003]，因此 O'Gorman 等人嘗試使用機器學習的方法[O'Gorman and Dwyer, 2018]，從機器學習的角度來解決傳統濕對流過程的問題。

另一方面，隨著硬體資源進展，加高全球環流數值模式的水平解析度在近年變的可行，像是 NCEP 和 ECMWF 作業用的天氣模式之解析度在近年都發展至約 10 公里，然而在模式加高解析度的同時，新的問題又產生了：設計在水平解析度為百公里的積雲參數化並不能適用於解析度為 10 公里的網格[Arakawa and Schubert, 1974]。定義適用積雲參數化的網格大小應足夠大到包含一整組由淺至深的積雲，並應足夠小以代表大尺度擾動中的一部分。此定義隨著解析度從原本的百公里尺度上升至~ $O(10\text{km})$ 尺度時很明顯將不再適用。在這樣的高解析下~ $O(10\text{km})$ ，濕對流過程該由積雲參數化處理或是由網格尺度的物理過程處理變得模糊不清，因此前人將~ $O(10\text{km})$ 解析度定義為 grey zone resolution [Arakawa and Wu, 2013; Wu and Arakawa, 2014; Arakawa et al., 2011]。

在積雲參數化中，決定對流效應強度大小的方式為假設網格之中對流消耗不穩定度的速率等同於大尺度過程產生不穩定度的速率(即網格達到對流準平衡, quasi-equilibrium)，然而當在 grey zone resolution 時，同個網格可能不再同時包含對流以及其環境，決定網格對流效應強度大小的環境將不再侷限於相同的網格。本研究的目的為了解在 grey zone resolution 下，對流環境若包含鄰近網格是否會改進對流效應的預測。由於傳統積雲參數化在技術上較難考慮鄰近網格對對流效應



強度大小的影響，因此本研究嘗試利用機器學習的方式來建立模型。

機器學習是一種從大量資料中自動分析、學習，從資料之中獲得規律，並利用其規律對於未知資料進行預測的演算法，而為了利用機器學習考慮鄰近網格對對流效應強度大小的影響，我們必須在訓練模型的時候加入鄰近網格的資訊。在數值模式中，網格間溫度、風向、濕度等等參數的梯度是影響當地天氣系統發展的重要指標，雖然各種機器學習的模型在訓練時都能加入鄰近網格的資訊，但不是每一種模型都能考量到相鄰網格間參數的梯度。我們認為網格資料空間中的關係若使用 3D 卷積神經網路[Shuiwang et al., 2013]是最適合的，因為它有卷積層(convolutional layer)的特性以及池化層(pooling layer)，可以很好的將鄰近網格的資訊抽取、歸納，意即能考量到相鄰網格間參數的梯度。

由於對流效應在垂直方向上是連續的，因此在訓練機器學習的模型時我們需要使用結構化學習(structured learning) [Lee, 2018]。結構化學習的定義是，模型的輸入和輸出皆是一種有特殊結構的物件，如機器翻譯的輸入輸出為不同語言的一句話、物件偵測的輸入為一張圖像而輸出為圈住偵測物的方框、搜尋引擎(retrieval)的輸入為關鍵字而輸出為一串網址清單，而在大氣中也有相似的例子，例如給機器看一張包含天氣系統的衛星雲圖，機器自動幫輸出框住各種不同的天氣系統。結構化學習另一常見的應用方式為 Video to Text [Venugopalan et al., 2015]：給機器看一段影片，讓機器自動生成該影片的敘述文字。Video to Text 利用卷積神經網路將影片中的時間方面的資訊轉換為另種格式，再藉由遞迴神經網路(Recurrent neural network, RNN)[Jim et al., 1996; Hochreiter and Schmidhuber, 1997; Gers et al., 2000]將其輸出為一段能解釋影片的敘述，這種不同資料格式的轉換，比較類似於我們的研究。

本研究的研究目的為評估利用機器學習的方式，在 grey zone resolution 下考量相鄰網格的資訊，所預測的各個網格的對流效應。本研究使用 3D 卷積神經網路來建構模型，並結合結構化學習，藉以利用一特定空間的氣象參數來預測目標網



格之對流效應的垂直結構。同時本研究也使用了多元線性回歸(Multiple Linear Regression)及深度神經網路(Deep Neural Networks, DNN)[Bengio and LeCun, 2007; Delalleau and Bengio, 2011; Pascanu et al., 2014a; Montufar et al., 2014]來建構模型，作為 3D 卷積神經網路的對照組。它們分別是機器學習模型的基準及基礎的深度學習模型，這兩種模型使用回歸(Regression)的方法來預測對流效應，意即在模型的建構上，是使用個別高度層的資訊來分別預測各層的對流效應。

由於現今觀測技術的限制，我們無法藉由觀測儀器來獲得夠多樣本的高解析對流資料，因此本研究利用高解析雲解析模式(Vector Vorticity equation Model, VVM)[Wu and Arakawa, 2011]之模擬結果來代表真實對流發展的情況，並將其重新水平平均之後得到 grey zone resolution 的環境場以及對流效應，分別作為本研究中機器學習模型的特徵(feature)和目標(target)。本文的第二章將介紹本研究中所使用的資料、神經網路，以及機器學習模型的訓練流程。第三章將評估由機器學習模型所預測的對流效應，第四章為結論與討論。



## 2. 資料與方法

本研究利用雲解析模式之高解析模擬(2km)結果代表真實對流發展的情況，將其做水平 16 公里的平均之後得到 grey zone resolution(16km)的環境場以及對流特性如圖 1，用以作為機器學習模型輸入的訓練參數，以及預測的對象。

### 2.1 資料

為了訓練出能夠各種不同環境下的對流效應，我們必須選擇一個涵蓋了各種不同環境特徵的資料集，因此，本研究的資料取自於 [Tsai and Wu, 2017]，它們為了瞭解環境如何影響組織性對流的發展情形，利用高解析雲解析模式(Vector Vorticity equation Model, VVM)設計了一系列的實驗，其實驗包含了各種不同環境的對流特徵，很適合用來做理想的濕對流參數化實驗，因此我們將他們的資料集用於訓練我們的機器學習模型。他們利用水平均勻的大尺度強迫之環境提供了適合發展對流的環境，其大尺度強迫為典型的 Global Atmospheric Research Program(GARP) Atlantic Tropical Experiment(GATE) Phase III 之平均垂直剖面(如圖 1a)。

環境濕度的量值從最濕到最乾是將大尺度的濕化效應乘以+6 至-6(區間為 1.5，共 9 個)，接著以有垂直風切、無垂直風切依照上述環境濕度分為兩組，共 18 個不同濕度、有垂直風切、無垂直風切的實驗。模擬之模式解析度為 2 km，使用了 VVM 在海溫固定為 299.8K 的熱帶洋面上，大氣輻射冷卻率為  $2 Kd^{-1}$  的自由對流模擬，圖 2 為此實驗初始場的設置，其中(a)大尺度冷卻及濕化效應、(b)水平風、(c)位溫、(d)比濕，其中位溫、比濕和大尺度強迫是垂直剖面的平均場，水平風的垂直剖面取自 TOGA COARE，其中有風切的實驗以綠色標記，無風切的實驗則以藍色標記。



## 2.2 機器學習模型之輸入及輸出

我們將代表大尺度環境場的 9 個參數，分別為水平風( $u, v$ )、垂直風( $w$ )、水氣( $qv$ )、溫度( $T$ )以及高度( $z$ )，對流可用位能(CAPE)、垂直積分水氣輻散(MFD)、濕靜能(MSE)，作為輸入模型的特徵，其中後三者具有網格的對流訊息，例如對流可用位能是來自氣塊在該環境下浮力的積分，可以用來評估大氣在垂直方向上是否穩定、可作為對流發展與否即強度的指標；垂直積分水氣輻散同樣是用來評估是否容易發展對流之水氣參數；濕靜能為一個描述氣塊狀態的熱力參數。

在輸出的方面，根據[Pan and Randall, 1998]，一個網格的平均能量、水氣的收支方程式為：

$$\rho \frac{\partial}{\partial t} \bar{s} = -\overline{\nabla \cdot (\rho s \mathbf{V})} - \frac{\partial}{\partial z} (\rho \bar{s} w) + Q_1$$

$$\rho \frac{\partial}{\partial t} \bar{q} = -\overline{\nabla \cdot (\rho s \mathbf{V})} - \frac{\partial}{\partial z} (\rho \bar{q} w) + \frac{Q_2}{L}$$

上式中  $\rho$  為空氣密度、 $\bar{s} \equiv c_p T + gz$  為乾淨能( $T$  為溫度、 $g$  為地球引力常數、 $c_p$  為比熱)、 $q$  為水蒸氣混和比、 $w$  為垂直方向速度、 $\mathbf{V}$  為水平方向速度、 $L$  為蒸發潛熱項。其分別包含了水平平流項、垂直平流項以及對流相關效應，其中對流的效應  $Q_1, Q_2$  使用[Yanai et al., 1973]的定義：

$$Q_1 \equiv Q_R + LC - \frac{\partial}{\partial z} (\rho \overline{w' s'}) \quad (1)$$

$$Q_2 \equiv LC + L \frac{\partial}{\partial z} (\rho \overline{w' q'}) \quad (2)$$

$$Q_1 - Q_R - Q_2 = - \left( \frac{\partial}{\partial z} (\rho \overline{w' s'}) + L \frac{\partial}{\partial z} (\rho \overline{w' q'}) \right) \quad (3)$$

其中  $Q_R$  為輻射加熱率， $C$  為凝結速率，以及本研究中最關心的平均垂直紊流通量，本研究將  $Q_1$  能量的垂直紊流傳送(上式(1))， $Q_2$  水氣垂直紊流傳送(上式(2))兩效應



合併成  $\overline{w'h'}$  作為對流效應(上式(3))，其中  $h \equiv c_p T + gz + Lq_v = s + Lq_v$ 。而此處的'代表某參數的紊流項，— 則代表網格的平均。本研究將  $\overline{w'h'}$  作為機器學習模型的訓練目標。

最後，為了能夠驗證我們訓練好的模型以及方便分析，我們將原始 18 個不同的實驗依時間維度打亂並將其 80% 作為訓練模型用的訓練集，20% 作為測試模型的測試集。前人研究指出對流的效應主要由中層水氣的對流關係主導[Holloway et al., 2009]，圖 3 顯示了我們測試集的挑選不論在總共、中層、低層的可降水量(Column Water Vapor, CWV) 分布都非常接近原始的資料集的分布。其中，實線代表原始資料集的分布，虛線代表測試集的資料分布。黑色線代表網格垂直積分總共的可降水量(CWV)的分布，橘色線代表網格中層水氣垂直積分的可降水量(850-200hPa) 的分布，紫色代表網格低層水氣垂直積分的可降水量(sfc.-950hPa) 的分布。

### 2.3 三維深度卷積神經網路(3D Deep Convolutional Neural Networks, 3D-CNN)

卷積神經網路(Convolutional Neural Networks)[LeCun, 1989]，是深度學習的經典結構之一，在圖像處理上取得了很大的成功。

以下兩段整理自[Goodfellow, 2016]。卷積神經網路因為有著特殊的結構，卷積層(Convolutional layer)、池化層(Pooling layer)，它的特點是可以直接輸入原始的圖像，避免了前處理之大量工作，可以自動提取特徵，是其相對於傳統的圖像處理方法之優點之一。

其中卷積層有局部連接(Locally Connected, 圖 4)、權重共享(Share Weight, 圖 5)的性質。所謂局部連接，是其相較於傳統的神經網路而言，卷積層包含許多小的核(kernel)，而每個核的運算只與部分的輸入連接(通常比輸入小很多)，而不與全部的輸入相連；權重共享，意即每一個核的權重是共享的，在核掃過整張圖的過程中，核的權重是不變的。卷積神經網路藉由卷積層的這兩種特性，相較於傳統



神經網路全連接的性質，可以節省大量的運算資源，並保留了應有的特徵，也因為其局部連接的特性，特別適合處理圖像類的資料(圖像中鄰近的像素相關性較高)。

池化層就是利用平均(meanpooling)或取最大值(maxpooling)等方法把權重縮減的方法，它的好處是可以節省運算空間、讓地區空間部分偵測出的特徵更有說服力。

藉由上面兩種特殊結構，使卷積神經網路相較於傳統神經網路在處理圖像或類似型式之特殊資料時，可以避免因全連接造成參數巨大導致神經網路難以訓練的情況，同時節省了很多的運算資源。

而本研究中使用的 3D 卷積神經網路則是以此為基礎，擴展到三維的一個模型，其中網路架構參考[Shuiwang, 2013]的研究設計，圖 6 為本研究神經網路架構的示意圖。3D 卷積神經網路中卷積層的運算也特別適合大氣模式網格資料的性質，藉由上面的介紹我們可以知道，它非常適合用來偵測鄰近網格間相互影響的效應。本研究中的 3D 卷積神經網路模型如圖 25，其包含了兩個大卷積區塊，每個卷積區塊包含三個 3D-Convlutional 層分別為 32、64、128 個 kernel，其使用的激活函數為 ReLU，並且隨後有一個 Maxpooling 層，最後再將輸出拉平，將 output 轉為長度為 5 之一維向量之 Dense 層，如表一。其中因為了簡化我們的輸出，並且減少訓練的時間，我們利用主成分分析(Principle Component Analysis, PCA)將我們的輸出從原始的 33 維(長度 33 的向量)降為 5 綴的向量(長度 5 的向量)，因此我們的輸出層為 5 個神經元。PCA 在數學定義上是一個正交化線性變換，也就是把數據變換到一個新的坐標系中，使這一個數據的任何投影第一大變異數在第一個座標(稱為第一主成分)上，第二大變異數在第二個座標(第二主成分)上，以此類推。本研究中使用前五個變異數，其可解釋的變化量為 95%。



### 3. 機器學習模型評估

本研究的研究目的為利用機器學習來了解在 grey zone resolution 下，對流環境若包含鄰近網格是否會改進對流效應的預測。在本章節中，我們分別從三個不同的面向來評估模式效能的差異，分別為對流發生的位置、對流效應的垂直分布，以及對流效應對網格可降水量的反應。除了評估由 3D 卷積神經網路所建構的模型以外，我們呈現了另兩個沒有使用鄰近網格資訊的機器學習模型作為比較，且這兩者是使用回歸而不是結構化預測的方式來預測對流效應。兩個模型的建構方式分別為多元線性回歸(Multiple Linear Regression)以及深度神經網路(Deep Neural Networks)。

#### 3.1 對流位置

為了檢驗機器學習模型是否能在正確的位置上預測對流的發生，我們分別呈現真實情況以及三種模型在不同濕度環境下高度 3 公里處的對流效應水平分布。

圖 7 為不同濕度環境下之瞬間與整個測試集之水氣分布圖，紅色線為某瞬間的網格水氣分布、綠色線為整個測試集的網格水氣分布，由左到右為由濕到乾的環境。

圖 8 為真實情況下高度 3 公里處之對流效應水平分布，依照圖 7，由左至右依序呈現由乾至濕的環境。由圖中可以看到在越濕的環境下對流效應介於 0.1-1 (K m/s)的面積較大，且大於 4(K m/s)的發生頻率也較多。

圖 9 為 3D 卷積神經網路所預測之高度 3 公里處之對流效應水平分布，依照圖 8 之順序分別呈現由乾至濕的環境，色塊為模式預測的高度 3 公里處之對流效應，紫色、黃色線條分別為量值 0.01(K m/s)、2(K m/s)之真實情況下高度 3 公里處之對流效應。我們可以看見 3D 卷積神經網路不管在任何濕度的環境下，對於對流位置的預測都有非常好的表現。在真實情況下對流效應大於 2K m/s 的地方 3D 卷積



神經網路幾乎都能很好的掌握，但是在圖中藍色色塊的區域，3D 卷積神經網路有高估的情況，在真實情況為白色色塊的部分很多都被預測為藍色的色塊，也就是沒有能量傳遞的位置被預測為向下傳送。

圖 10 為多元線性回歸所預測之高度 3 公里處之對流效應水平分布，從圖 10 中可以看見多元線性回歸不管在任何濕度的環境下，對於對流位置的預測表現與真實結果的差異較大。只有在真實對流效應大於  $2 \text{ K m/s}$  的地方多元線性回歸預測較好，在真實對流情況很多為白色色塊的部分也都被預測為淺紅色的色塊，也就是它將沒有能量傳遞的位置預測為了微弱的向上傳送。此結果呈現了多元線性回歸模型的模式參數過少，且模型為線性模型，沒有辦法掌握複雜並且為高度非線性的對流效應，也就是我們輸入的參數會互相影響的狀況，例如網格的溫度不只影響了網格的對流效應，同時也影響了鄰近網格的溫度，同時其他網格的溫度也會影響目標網格的對流效應。

圖 11 為深度神經網路所預測之高度 3 公里處之對流效應水平分布，從圖 11 中可以看見深度神經網路不管在任何濕度的環境下，對於對流位置的預測表現也不佳。同樣只有在真實對流效應大於  $2(\text{K m/s})$  的地方才掌握得比較好，在真實對流情況很多為白色色塊的部分也都被預測為淺紅色的色塊( $0.1\text{K m/s}$ )，也就是它將沒有能量傳遞的位置預測為了微弱的向上傳送。另外在真實情況應該為藍色色塊的部分都沒有被它預測出來，我們推測可能是因為在深度神經網路中的激活函數為 ReLU，因此向下傳送的對流效應都被此函數抹平了。

藉由以上的比較，我們發現 3D 卷積神經網路相較其他模式能夠更好的掌握真實情況的對流效應水平分布，它在對流強度為  $0.1(\text{K m/s})$  及更高強度的對流預測的掌握非常好，這是另兩種模型沒辦法做到的，同時也不會像深度神經網路將能量向下傳送的網格抹平的問題。根據分析的結果，證實加了鄰近並使用結構化垂直預測對流效應，在預測對流的位置上是有幫助的。



### 3.2 對流垂直強度分布

在 3.1 節的討論中，各種模式在對流強度上的預測有很大的差異，並且強度預測對於氣候模式的影響也非常重要，因此本節將以統計的角度討論模式對於垂直分層的對流強度的預測。我們將垂直濕靜能紊流通量依高度分層，將它發生的頻率以 Contoured frequency by altitude diagrams (CFADs) 呈現，以下將垂直濕靜能紊流通量稱為對流效應。

圖 12 為真實對流效應強度隨高度之分布，由於大氣對流的稀有事件的特性，大部分網格的是沒有對流效應的，也就是其對流效應強度在接近於零的發生機率非常的高(圖中冷色部分)。然而在發生頻率較低的對流效應部分卻影響了模式中許多物理效應(圖中暖色部分)。因此本節主要關注於有對流效應、強對流效應的網格的對流隨高度的強度分布，也就是圖中的深紅色、紅色、橘色、黃色的形狀隨高度的分布，它們分別為發生頻率區間 0.001-0.01(%)、0.01-0.1(%)、0.1-1(%)、1-5(%)，在模式真實的對流效應中它們對流效應強度分布隨高度的極大值分別發生在高度為 8 公里、4 公里、4 公里以及兩公里的高度，強度分別為 8、6、4、1(K m/s)。

圖 13、14、15 為模式預測的對流效應強度隨高度的分布，圖 13 為 3D 卷積神經網路、圖 14 為多元線性回歸、圖 15 為深度神經網路，圖中的黑色線條為真實對流效應我們想關注的強度區間。

圖 13 為 3D 卷積神經網路圖 13 在我們關注的對流效應區間之對流效應強度的預測，其非常接近真實的分布，在色塊為深紅色、紅色、橘色、黃色的部分幾乎是與真實對流效應重合的，其差異用肉眼幾乎看不出來。在沒有對流效應的強度區間的發生頻率是有低估的(向對流效應強度小於零的方向偏移)。

圖 14 為多元線性回歸在我們關注的對強度區間之對流強度的預測，在這些區間的極值的預測都被預測得太高，並不像真實的對流效應強度垂直特徵的橘色部



分強度極值在 4 公里高度、深紅色色塊的強度極值在 8 公里高度，它預測的強度垂直分布並沒有明顯的最大值，在垂直分布上較真實狀況平滑。在沒有對流效應的強度區間的發生頻率也被低估，它將許多完全沒有垂直能量傳送的部分預測為有垂直能量傳送，有將頻率平滑分散的預測傾向。

圖 15 為深度神經網路在我們關注的強度區間之對流強度的預測，它在強度的極值高度上每個區間的預測都略高於真實的分布，如黃色、紅色及深紅色的色塊。在它的預測中的對流效應完全沒有對流效應小於零的分布。

藉由以上的比較，可以看到 3D 卷積神經網路在對流效應垂直強度分布的預測是非常驚人的，在對流效應較強的預測上，不只極值的位置非常接近真實對流情況，也沒有如深度神經網路有抹平能量向下傳送的特性。但在弱對流效應的預測有改進的空間，推測是因為我們在訓練模型時只使用了 PCA 的前五個特徵值，因此無法將對流效應的細節很好的預測。



### 3.3 網格可降水量與對流強度的關係

前人的研究指出，降雨的反應在環境的水氣達到一個特定域值之後，降雨的量值會隨著水氣量的增加迅速變大[Holloway et al., 2009]，這種特性在本研究的資料集中也可以被看到(圖 16)。另外在 Holloway 等人於 2009 年的研究中也提到，對流強度的反應主要由中層水氣影響，與低層水氣的相關性比較小。

本節中呈現了對流效應強度與總、中層與低層水氣的反映，藉以檢測我們的模型是否能預測出這種對流效應與水氣的相關性。我們也藉由挑選不同濕度的環境，檢測我們的模型能否在不同的濕度環境下預測出對流細緻的垂直結構。

圖 16 是真實對流效應與水氣量的反應，在網格水氣量小於 60 毫米時，對流強度隨著水氣量的增加是緩慢的，但在對流強度大過這個域值時，對流強度隨著水氣量的增加是快速上升的。其中為了很好的觀察對流效應與水氣量的反應，在後面的機器學習模型比較中，我們在高度 6 公里處切了水平剖面圖，以便我們觀察不同模型輸出與真實的差異。在圖中黑色箭頭標示處，可以看到在環境較濕時，其對流垂直強度及值約在 8 公里高度，並且在 5 公里高度有個轉折，推測為凝結點。而在環境較乾時對流效應主要集中在低層，其在 1 公里處有極值。

圖 17 是真實對流效應與中層(850-200hPa)水氣量的反應，圖 18 是真實對流效應與低層(sfc.-950hPa)水氣量的反應，在兩圖中可以看到相對於低層水氣，對流強度對於中層水氣的改變是比較有反應的。

圖 19 為模式真實對流情況及各種機器學習模型預測在 6 公里高度下對流強度對水氣的反應。黑色線為真實對流情況、紅色線為 3D 卷積神經網路的預測、綠色線為深度神經網路的預測、藍色線為多元線性回歸的預測。在圖中我們可以看到其中綠色、黃色、黑色線的對流強度對於水氣量的反應以 60 毫米作為分界，在水氣量小於 60 毫米時對流對水氣量的反應是不敏感的，而在水氣量大於 60 毫米之後，對流的強度隨著水氣量的增加的改變是很大的，而紅色線在水氣量小於 60



毫米時也一直有微弱的反應，且在水氣量大過 60 毫米之後，其對流強度上升的幅度不及其他模型所預測。

圖 20 為在環境較濕時的對流效應垂直分布，其中黑色線為模式真實對流、綠色線為 3D 卷積神經網路、黃色線為深度神經網路、紅色線為多元線性回歸模型所預測的結果。從圖中可以看到 3D 卷積神經網路雖然在強度量值上的預測略強於真實結果，且甚至可以捕捉到在高度 5 公里處的凝結點之轉折。深度神經網路在極值的垂直分布上的預測略弱於真實結果，且無法捕捉到在 5 公里高度強度的轉折。多元線性回歸模型在強度的預測很在 4 公里以下的高度預測蠻接近真實的對流結構，但在其他高度的預測與真實結果相距較多。

圖 21、22 為在低層、中層網格水氣較濕時的對流效應，為了顯現出對流垂直結構的細部差異，橫軸的尺度是不一樣的。藉由低層與中層水氣量與對流關係的分別繪製，可以看到每個模型都可以預測出中層水氣與對流效應的強度較有反應，但在對流的垂直結構上唯有 3D 卷積神經網路可以預測出與真實對流最相似的結構，無論在中層或是低層水氣影響的對流效應，它都可以很細節的捕捉到在模式 5 公里處的轉折，並且在對流效應強度的預測都與真實對流效應相當接近。其中多元線性回歸無論在高層水氣與低層水氣相關的對流反應的預測與真實結果差距較大。深度神經網路在低層水氣的對流效應的低層有強度的高估、高層有強度的低估，中層水氣對流效應的部分則在任何高度層都有些微的強度低估。

圖 23 呈現了環境較乾的時候之真實對流高度結構與各種模型預測的結果，在這種環境時對流對於中層、低層與整層的水氣的反應是差不多的，因此我們只呈現對流對總水氣的反應。3D 卷積神經網路在這個濕度的環境下預測的不好，在 1 公里處的量值有低估，並且隨著高度上升，整個對流效應的預測有不正常抖動，不像真實對流效應那樣平滑，我們推測是因為在訓練它的時候利用 PCA 簡化了輸出，因此在對流強度較弱的時候其沒辦法學到完整的對流結構。深度神經網路在這個環境下的預測非常接近真實的對流效應，但在 2 至 6 公里高度處有些微強度

上的高估。多元線性回歸在 1 公里的強度有低估，且在 1 公里之上的強度都是高估。





## 4. 結論與討論

現存的積雲參數化在模式解析度逐步上升至~ $O(10)$ 公里尺度時，濕對流過程該由積雲參數化處理或是由網格尺度的物理過程處理變得模糊不清(Grey zone resolution)。在這樣的水平解析度下，網格內可能不再同時包含對流以及其環境，意即決定網格對流效應的環境將延伸至相鄰網格，然而，至今仍未有研究評估在 Grey zone resolution 下以相鄰網格作為決定對流效應之環境對於模擬結果的影響。因此本研究嘗試使用一種可考慮掃描區域內參數梯度的 3D 卷積神經網路(3D Convolutional Neural Networks, 3D-CNN)，在 grey zone resolution 利用各個網格及其相鄰網格的資訊，預測各個網格的對流效應。為了證實在 grey zone resolution 鄰近網格資訊對於預測對流效應的助益，本研究呈現多元線性回歸(Multiple Linear Regression)及深度神經網路(Deep Neural Networks, DNN)的結果作為對照組，它們分別是機器學習模型的最低標準，及最基礎的深度學習模型。研究結果指出，由 3D 卷積神經網路所預測的對流效應在對流位置、對流強度高度分布及對流效應對可降水量的反應皆較另兩種模型有更好的表現，其中在對流效應對可降水量的反應的預測中，3D 卷積神經網路甚至在環境濕度較高的環境下的預測，可以捕捉到 5 公里高度處 freezing level 非常細節的對流特徵，這是其他模型所辦不到的。

在第三章機器學習模型的討論中，我們看到 3D 卷積神經網路不管在對流位置的掌握、對流強度的預測以及對流強度對水氣量的反應，相比於我們對照的模型，都有比較好的表現，這不只說明了 3D 卷積神經網路的特性較另外兩種模型更適合本研究的資料，也說明了增加鄰近網格的資訊以及使用結構化預測的方式，對於預測網格對流特性是有正向幫助的，以下我們分別討論各種模型的差別。

3D 卷積神經網路：在我們各種預測的檢驗上都非常接近真實模式輸出的情況，但在對流強度較弱時的預測以及低層預測上與真實結果差異較大，我們推測為我們訓練模型時，將其輸出利用 PCA 簡化，因而造成了對流強度較弱與低層的對流



預測效果與真實結果差距較大。這個問題可以藉由減少我們資料降維的幅度來解決，但由於資料維度的上升，在計算資源的需求也會隨上升。

多元線性回歸模型：從 3.2 以及 3.3 節的結果中可以看到，在預測複雜的對流垂直結構以及強度的頻率分布上，它趨向將其預測的量值以及強度做比較平滑的預測，也就是它趨向把強度較大或發生頻率較大的部分分給強度較小或是發生頻率較低的地方，使得對流的預測分布是較為平緩的。筆者認為是因為它的參數量比較小以及為線性的模型，無法藉由它的參數量預測對流複雜的非線性特徵。網格間互相影響的非線性效應，例如水平風的平流可能會造成垂直風場及網格溫度的改變，接著以各種複雜的物理過程影響對流的效應。

深度神經網路：從 3.2 及 3.3 節中可以看到它能夠在對流垂直結構及強度上的預測出大致的樣貌，但是沒有辦法捕捉到細部的結構，例如圖 16 中對流強度在 5 公里高度的轉折。同時在第三章的結果中可以看到它無法預測出對流效應的強度小於零的部分，而這問題需要更多的實驗來了解為什麼會有這樣的狀況發生。

綜合以上的討論，我們認為在 grey zone resolution 下考慮鄰近網格的資訊對於預測對流效應是很重要的，我們藉由 3D 卷積神經網路能夠考慮這種網格間梯度的差異，並在與其它模型的結果比較中證實了這點。

本研究中建立了一個可行的流程，包含了前處理(圖 24)、訓練模型(圖 25)以及驗證(圖 26)：利用高解析雲解析模式的資料當作真實對流情況，首先選擇輸入特徵以及訓練目標(grey zone resolution 下的對流效應  $\overline{w'h'}$ )，將其平均之後得到我們原始的資料。同時利用 PCA 將對流效應  $\overline{w'h'}$  降維之後作為我們的標籤。利用這樣的資料訓練我們的 3D 卷積神經網路。並且用模型沒有看過的資料，驗證其表現結果。然而，我們在眾多實驗中發現，即使模型的損失函數較小，也不一定表示模型在對流效應的預測結果比較好，因此沒有辦法簡單的用正確率(accuracy)或是損失函數(loss)來評價我們的結果。是以我們在第三章利用原始資料與機器學習



模型的預測比較，根據對流效應的性質，我們分別比較了對流位置、對流垂直強度分布以及網格可降水量與對流強度的關係。

在未來我們將利用高解析雲解析模式(VVM)在更多不同的環境、解析度的資料上做更多的測試，將 3D 卷積神經網路調整的更為完善，並且期待未來能將其應用在氣候模式的積雲參數化上。

## 圖片

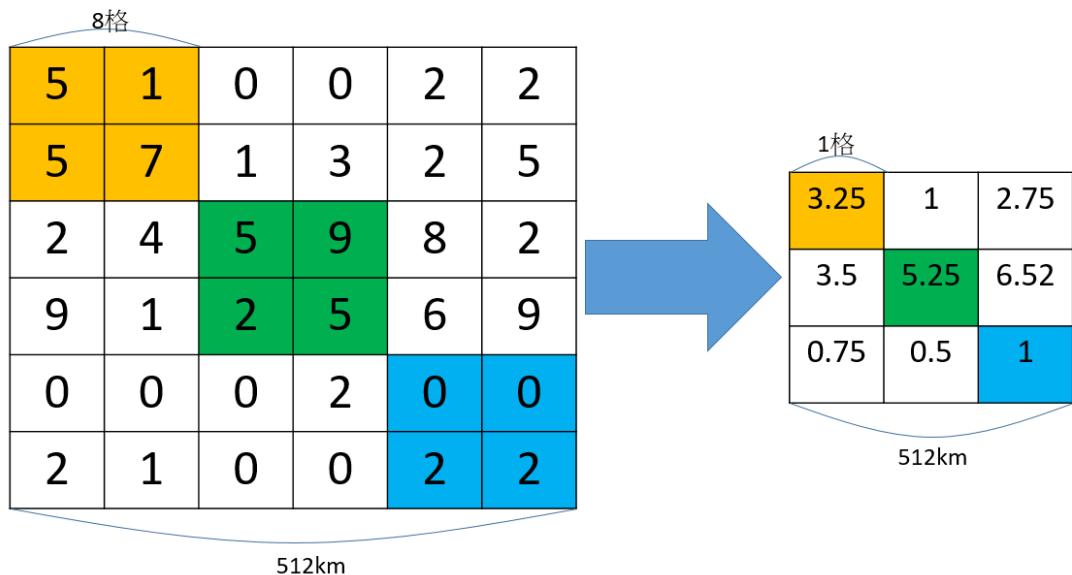


圖 1 平均示意圖。將原始 256\*256 格(512\*512 公里)的網格每 8\*8 格(16\*16 公里)做水平平均，將平均後的資料作為機器學習模型的輸入及輸出。

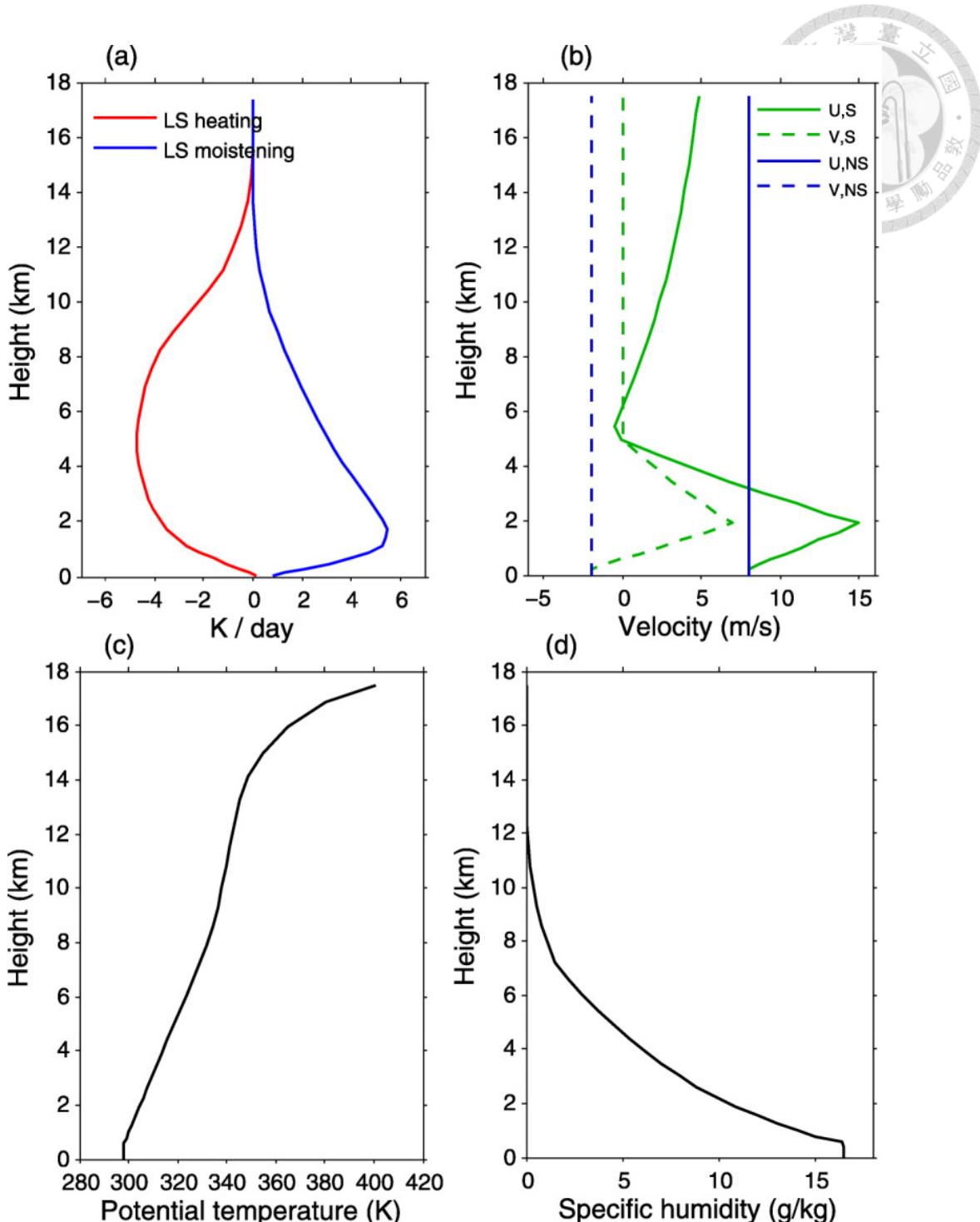


圖 2 (a)大尺度之冷卻、濕化效應，(b)水平風，(c)位溫，(d)比濕的初始場。位溫、比濕和大尺度強迫是取自 GARP Atlantic Tropical Experiment(GATE) phase III 的平均場。水平風為 TOGA COARE 之觀測資料，綠色為有風切之實驗、藍色為無風切之實驗。模擬中有  $2 \text{ K} \text{d}^{-1}$  的輻射冷卻。本圖取自 [Tsai and Wu, 2017]。

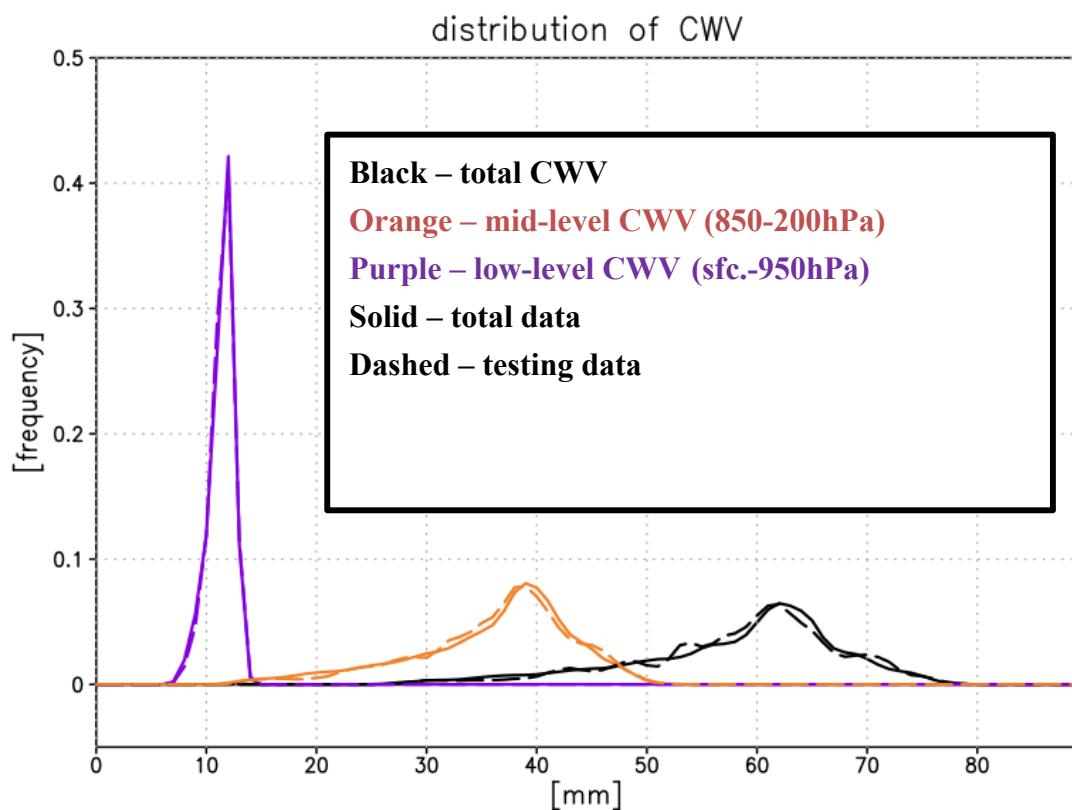
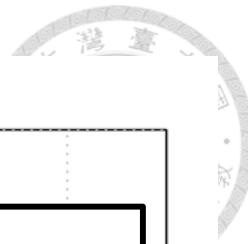


圖 3 實線表示原始資料集之總、中、低層網格可降水量分布，分別以黑色、紫色、橘色代表；虛線代表測試集之抽樣結果，分別以黑色、紫色、橘色代表不同層總、中、低層網格可降水量之頻率分布。

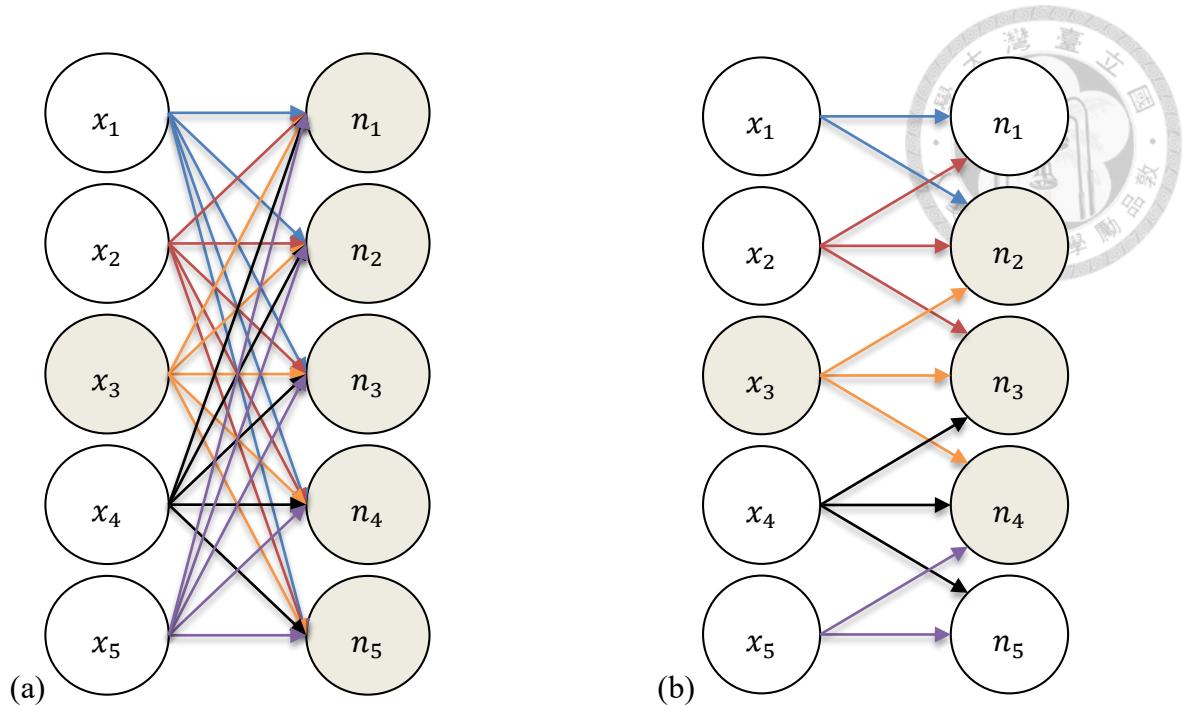


圖 4 (a)傳統全連接神經層及(b)卷積層。傳統全連接層的會將輸入連接到下一層的每個神經元，因此每個神經元都會接收到全部輸入的使用特徵，而卷積層只會使用到部分的特徵，如上圖(b)的 $n_3$ 只會與 $x_2, x_3, x_4$ 的輸入有連結。

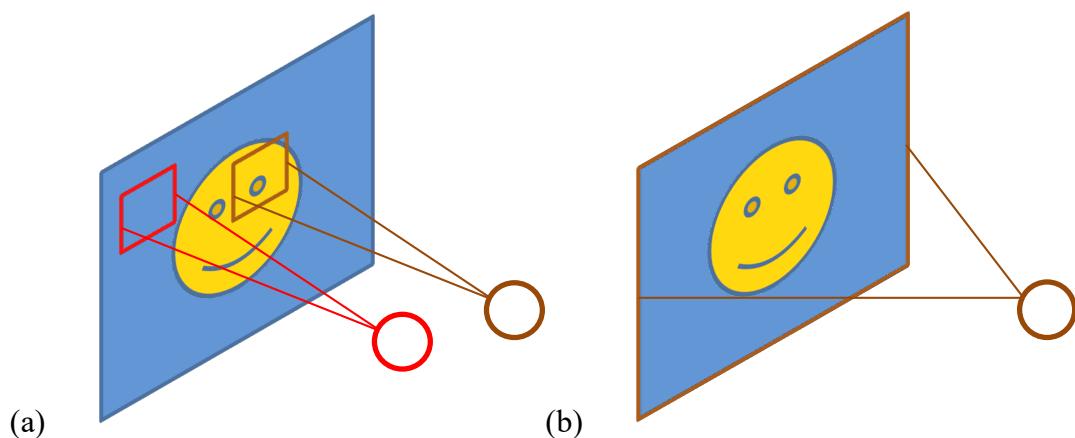


圖 5 (a)卷積核(kernel) (b)全連接層。卷積核之掃描範圍通常比輸入之圖像小很多，並且在掃過整張圖的過程中，都使用相同的權重。同時具有局部連接之特性，意即不一次輸入整張圖片。傳統全連接神經網路則一次輸入整張圖片。

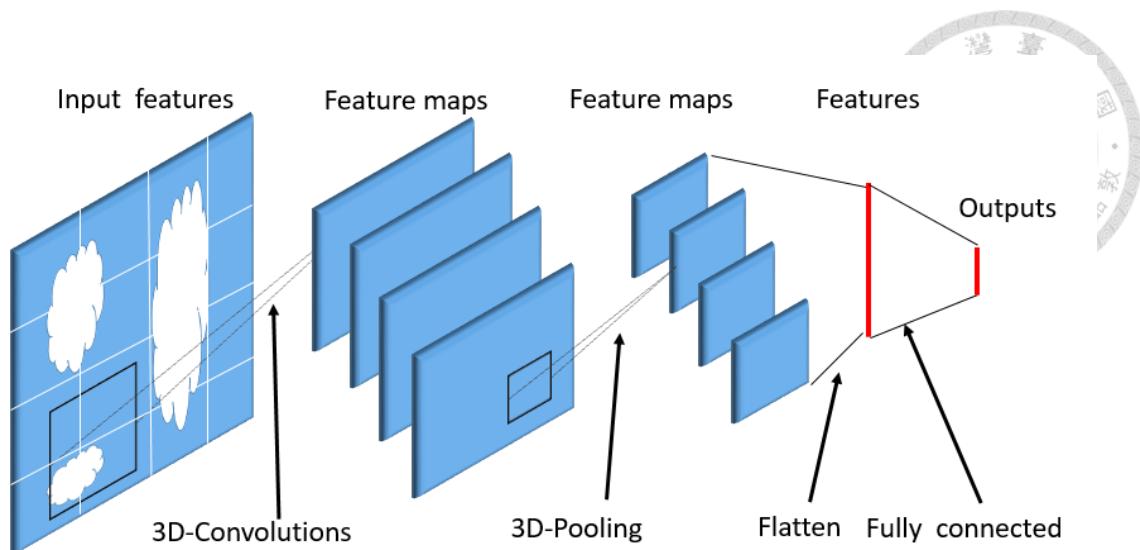


圖 6 本研究使用的 3D-CNN 架構示意圖。其利用 3D-CNN 將原始模式輸入之網格間的資訊，利用卷積層抽取為特徵圖後，以池化層壓縮，最後將特徵拉直，藉由全連接層的轉換，將最後的特徵轉為對流的結構輸出。

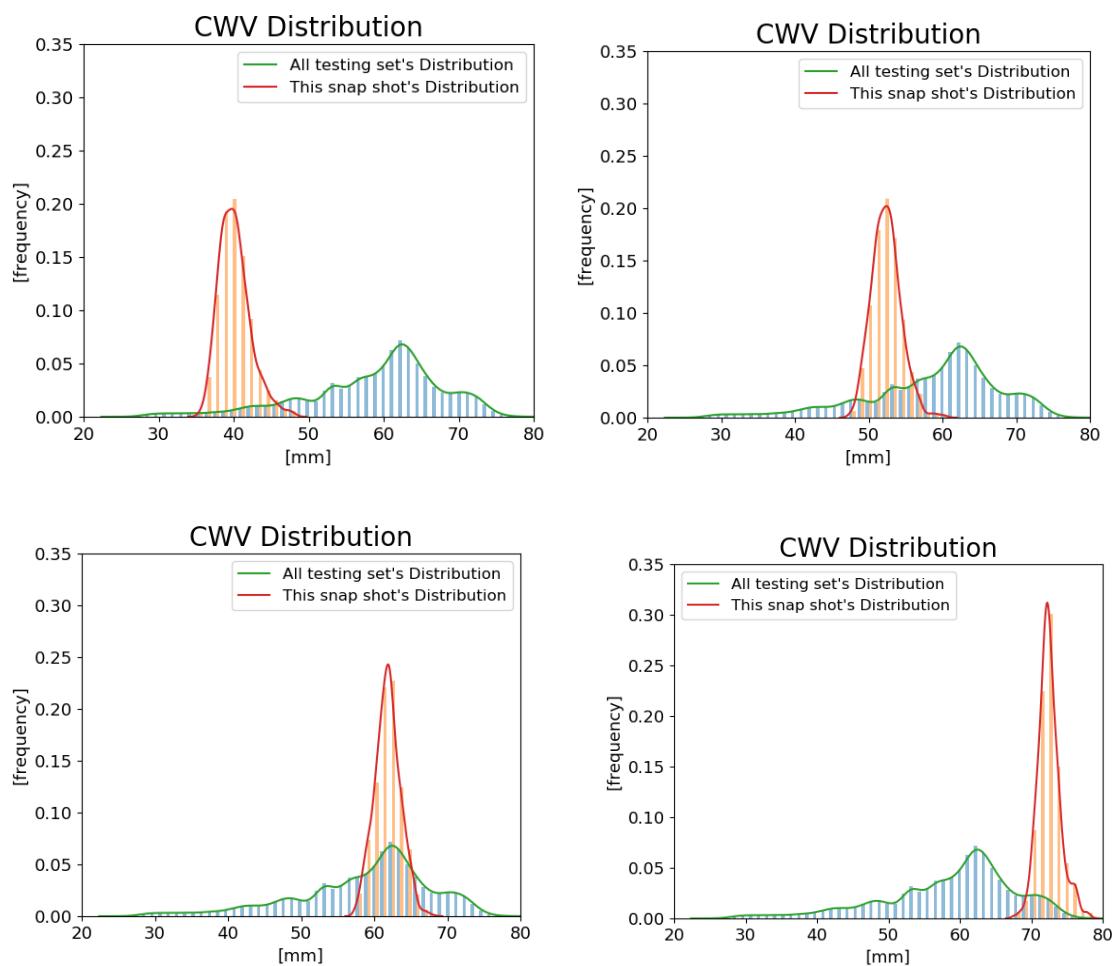
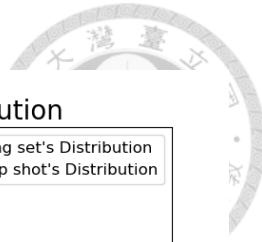


圖 7 可降水量之分布，紅色線為某時間下可降水量分布，綠色線為整組測試集可降水量的分布。由左至右為由乾到濕環境的分布。圖中橫軸為可降水量單位為毫米(mm)，縱軸為分布之機率大小(frequency)

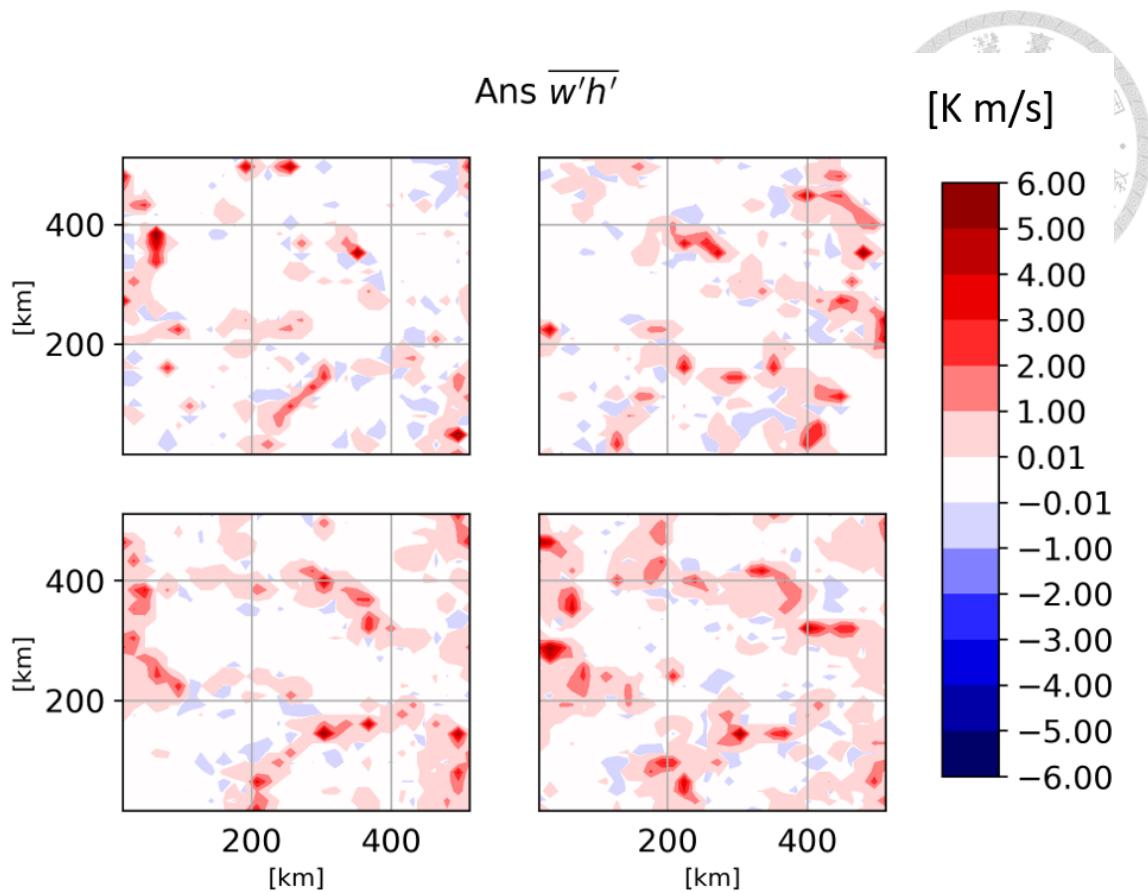


圖 8 高度 3 公里垂直濕靜能紊流通量水平分布，顏色代表其量值大小，單位為 (K m/s)。

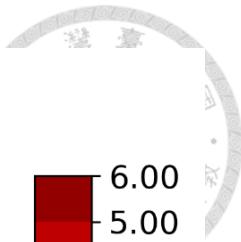
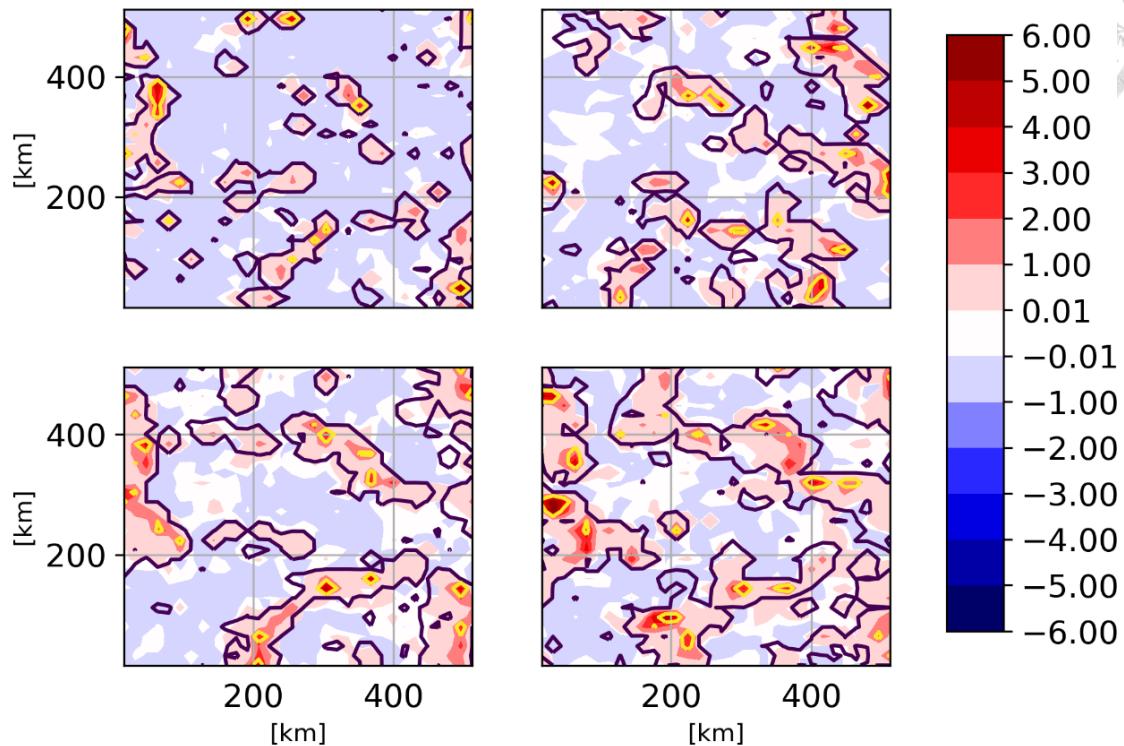
CNN9f  $\overline{w'h'}$ 

圖 9 3D 卷積神經網路(3D Convolutional neural networks)模型預測之高度 3 公里  
垂直濕靜能紊流通量，顏色代表其量值大小，單位為(K m/s)，線條為真實模式輸  
出之對照。

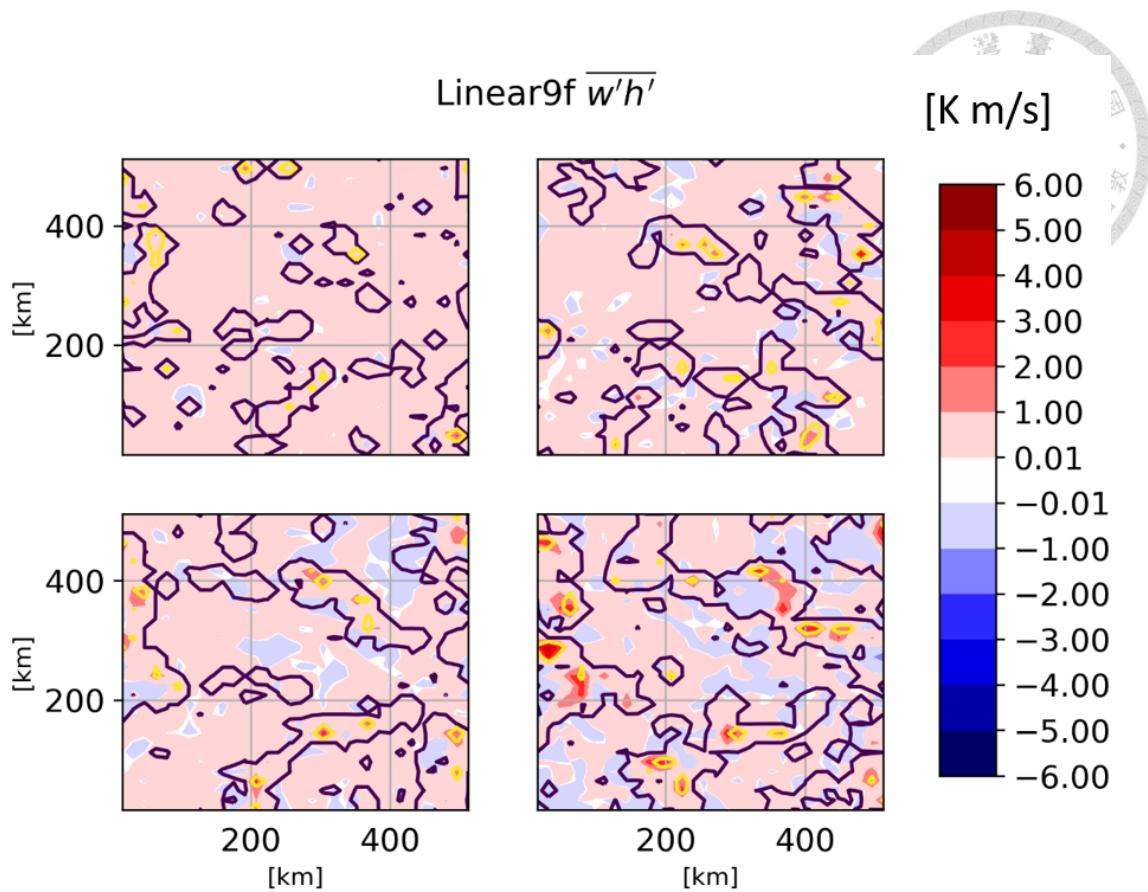


圖 10 多元線性回歸(Multiple Linear Regression)預測之高度 3 公里垂直濕靜能紊流通量，顏色代表其量值大小，單位為(K m/s)，線條為真實模式輸出之對照。

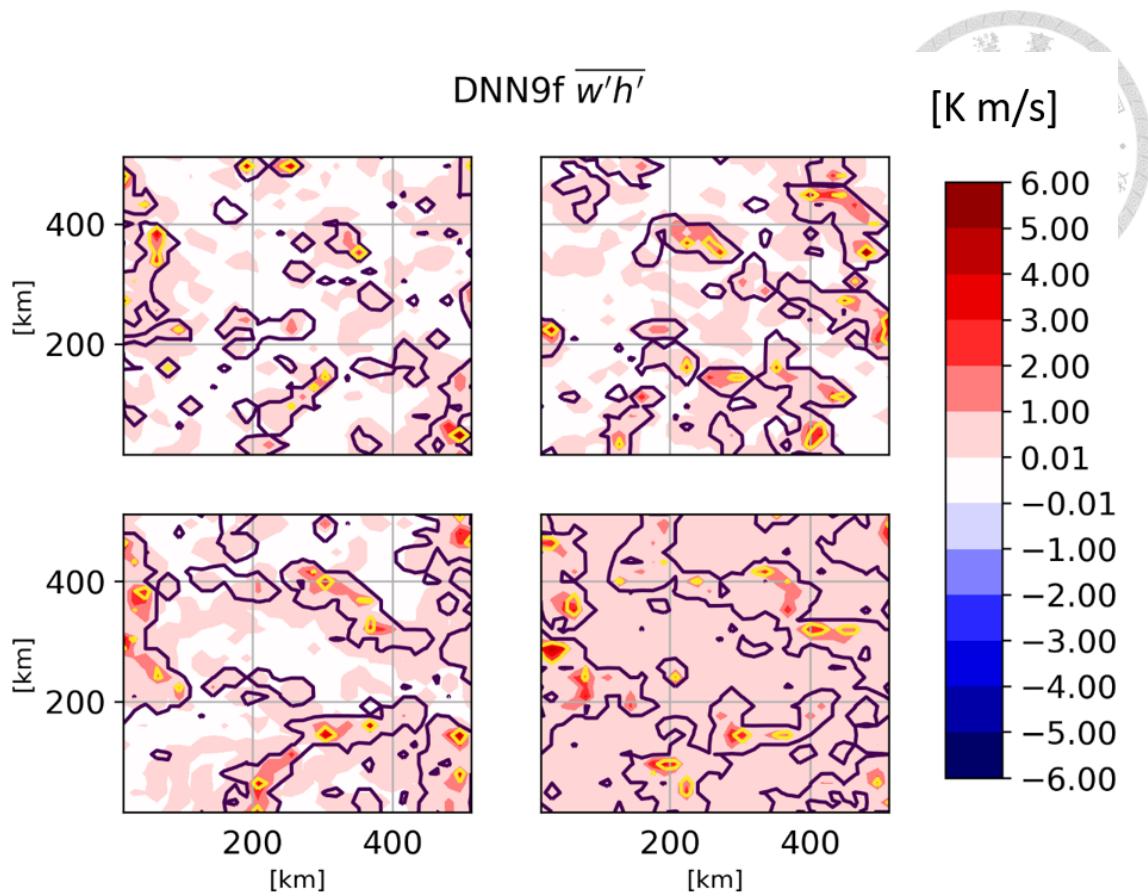


圖 11 深度神經網路(Deep Neural Networks)模型預測之高度 3 公里垂直濕靜能紊流通量，顏色代表其量值大小，單位為(K m/s)，線條為真實模式輸出之對照。

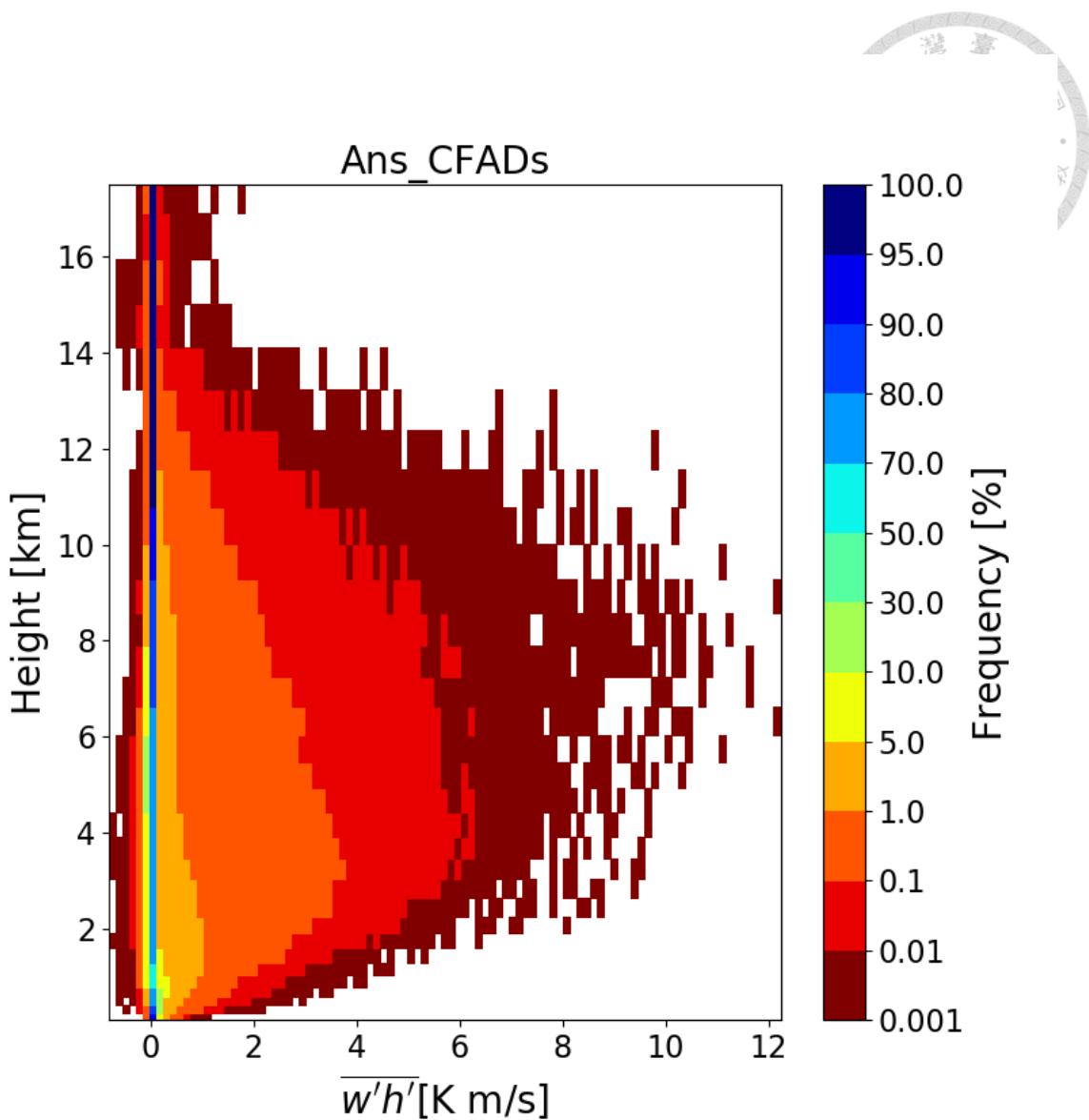


圖 12 真實垂直濕靜能紊流通量之隨高度分布頻率圖(CFADs)，橫軸為(K m/s)之量值，縱軸為高度座標(m)，顏色為機率，某高度總和之機率為 1。

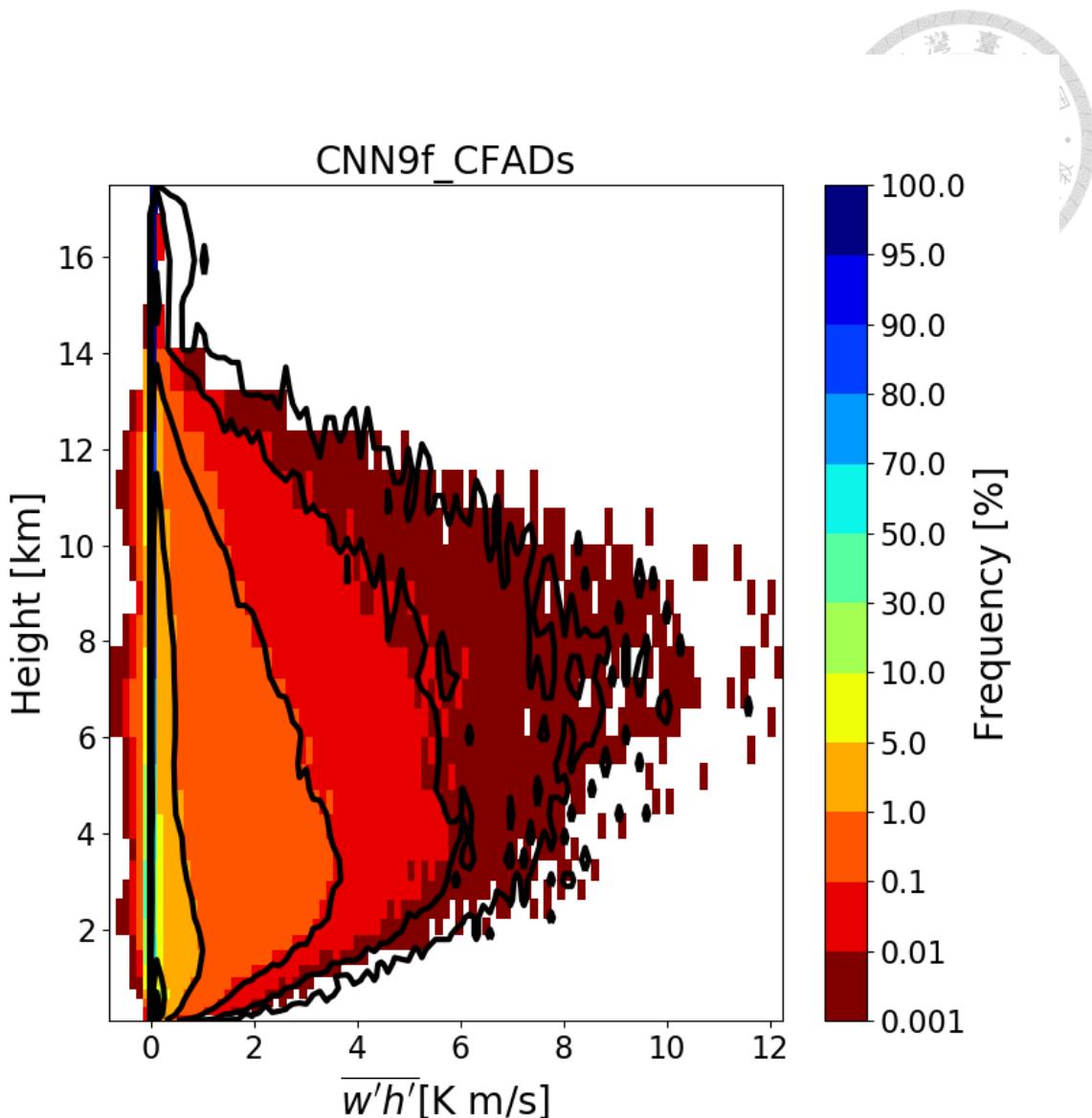


圖 13 3D 卷積神經網路預測之垂直濕靜能紊流通量之隨高度分布頻率圖

(CFADs)，橫軸為(K m/s)之量值，縱軸為高度座標(m)，顏色為機率，黑色線條為真實模式輸出之對照，某高度總和之機率為 1。

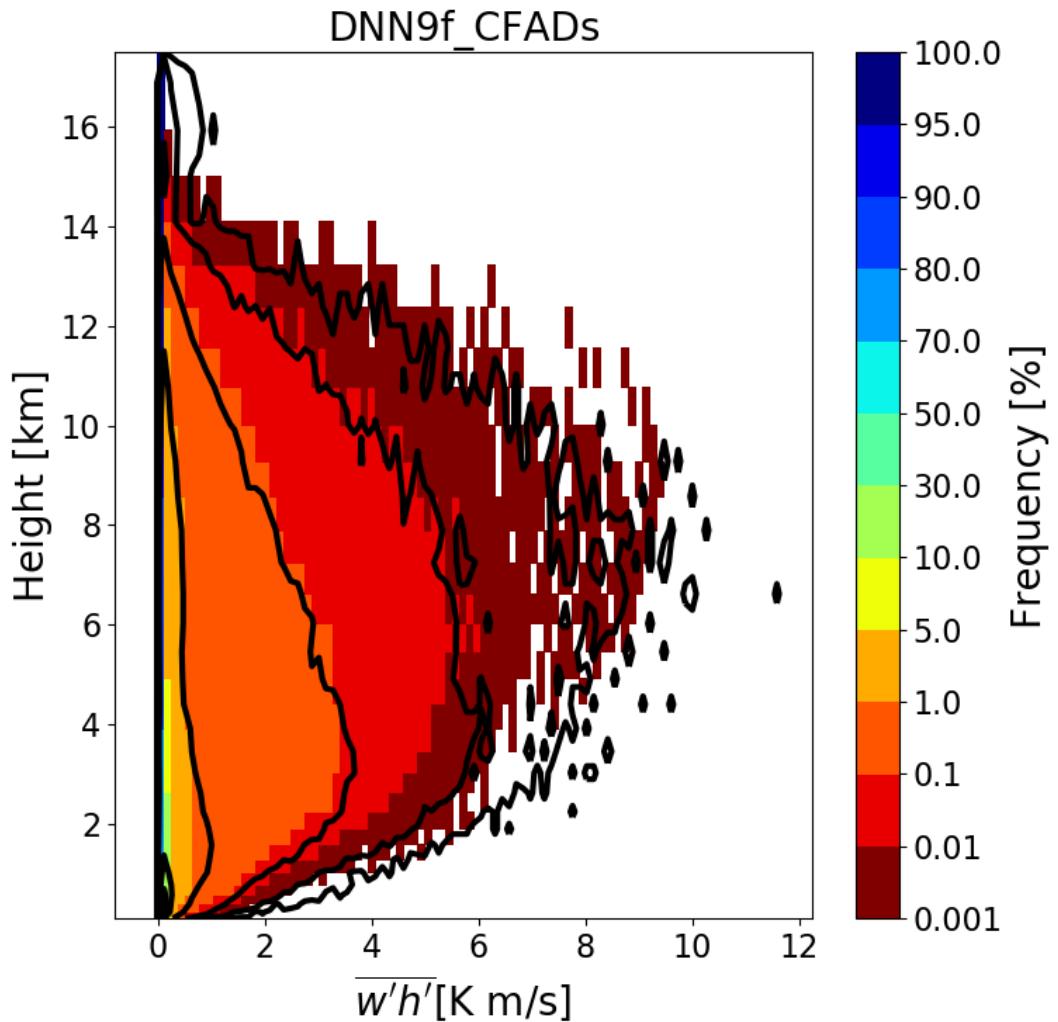


圖 14 深度神經網路預測之垂直濕靜能紊流通量之隨高度分布頻率圖(CFADs)，  
橫軸為(K m/s)之量值，縱軸為高度座標(m)，顏色為機率，黑色線條為真實模式  
輸出之對照，某高度總和之機率為 1。

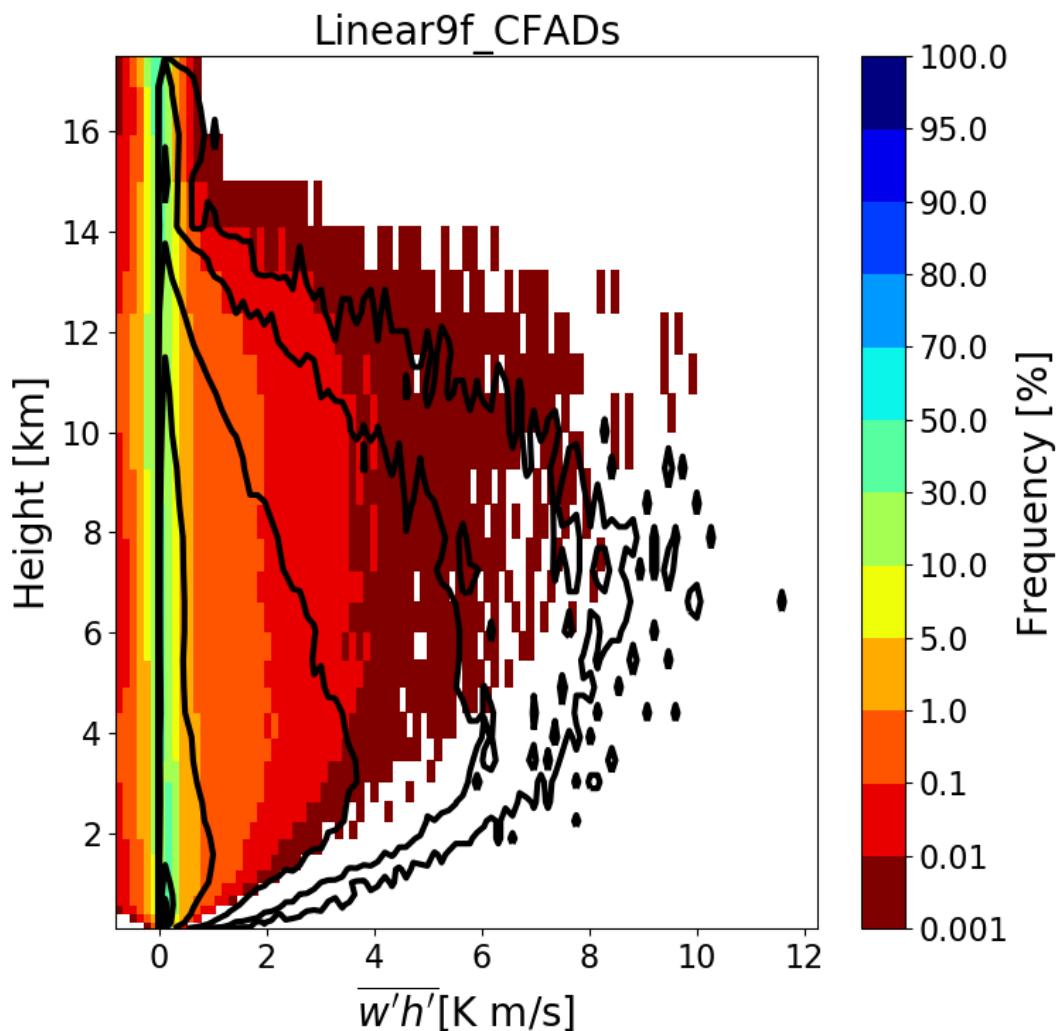
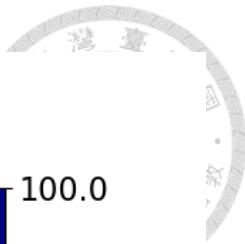


圖 15 多元線性回歸預測之垂直濕靜能紊流通量之隨高度分布頻率圖(CFADs)，  
橫軸為(K m/s)之量值，縱軸為高度座標(m)，顏色為機率，黑色線條為真實模式  
輸出之對照，某高度總和之機率為 1。

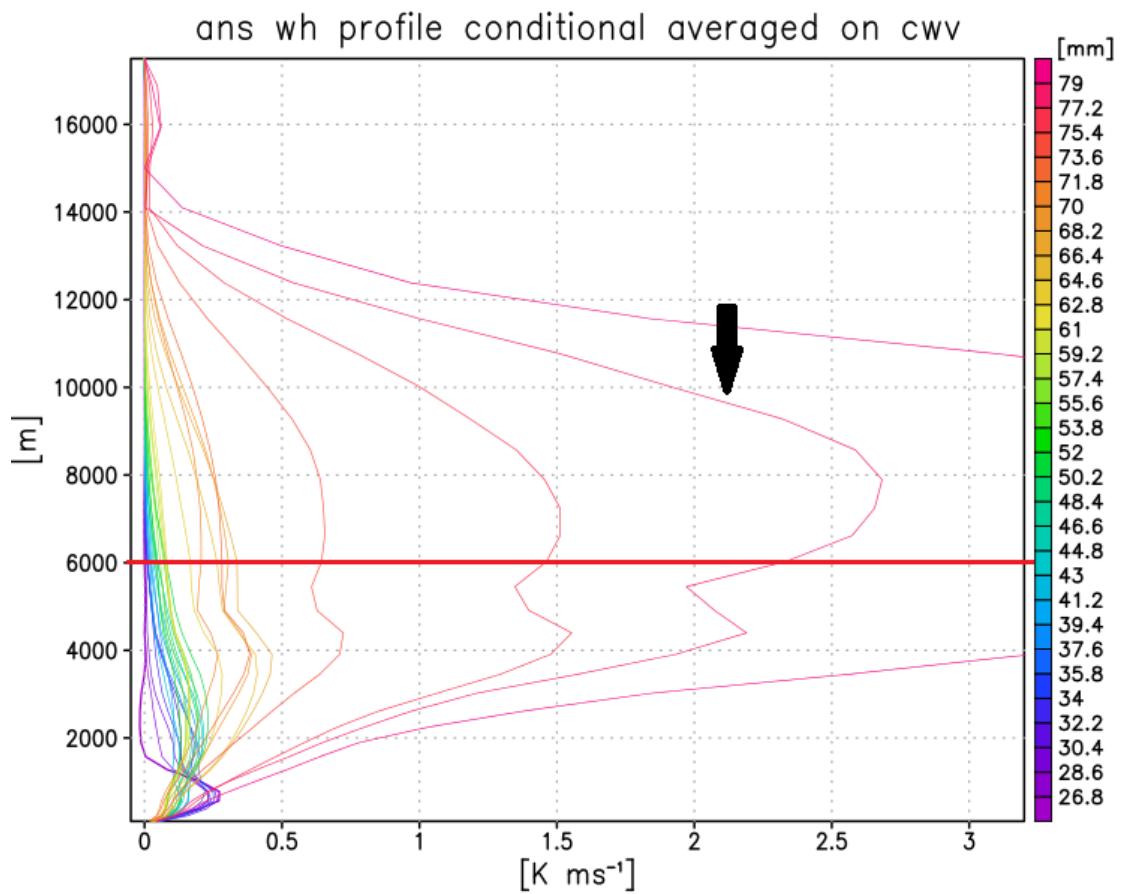
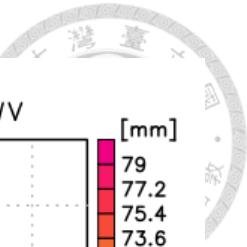


圖 16 真實模式對流強度隨著網格總共可降水量的關係，顏色由冷色到暖色為網格可降水量少到多。其 x 軸為對流效應的強度  $\overline{w'h'}$ (K m/s)，y 軸為高度座標(m)。其中紅線為與機器學習模型比較的橫切面，黑色箭頭為與機器學習模型比較的水氣量選取。

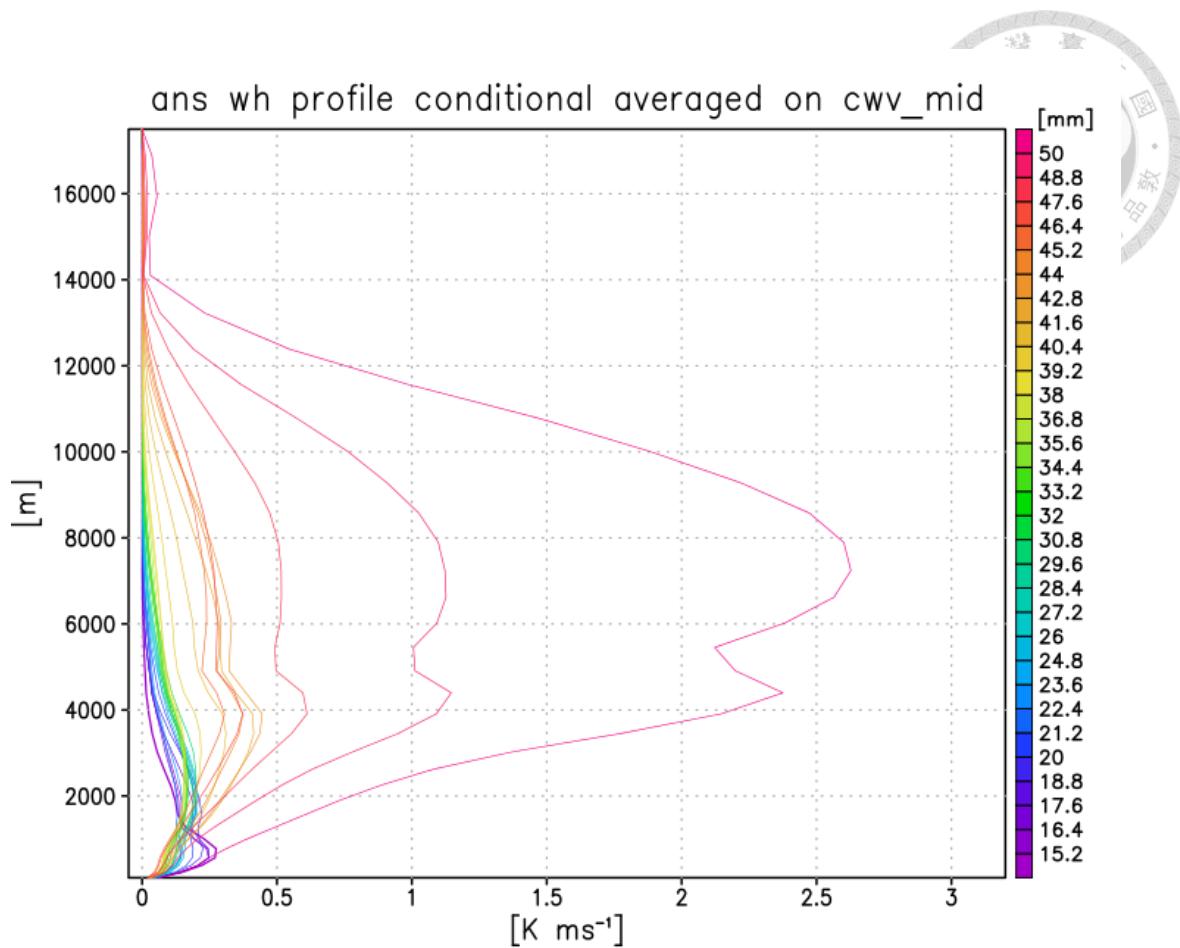


圖 17 真實模式對流強度隨著網格中層水氣的關係，顏色由冷色到暖色為網格可  
降水量少到多。其 x 軸為對流效應的強度  $\overline{w'h'}$ (K m/s)，y 軸為高度座標(m)。

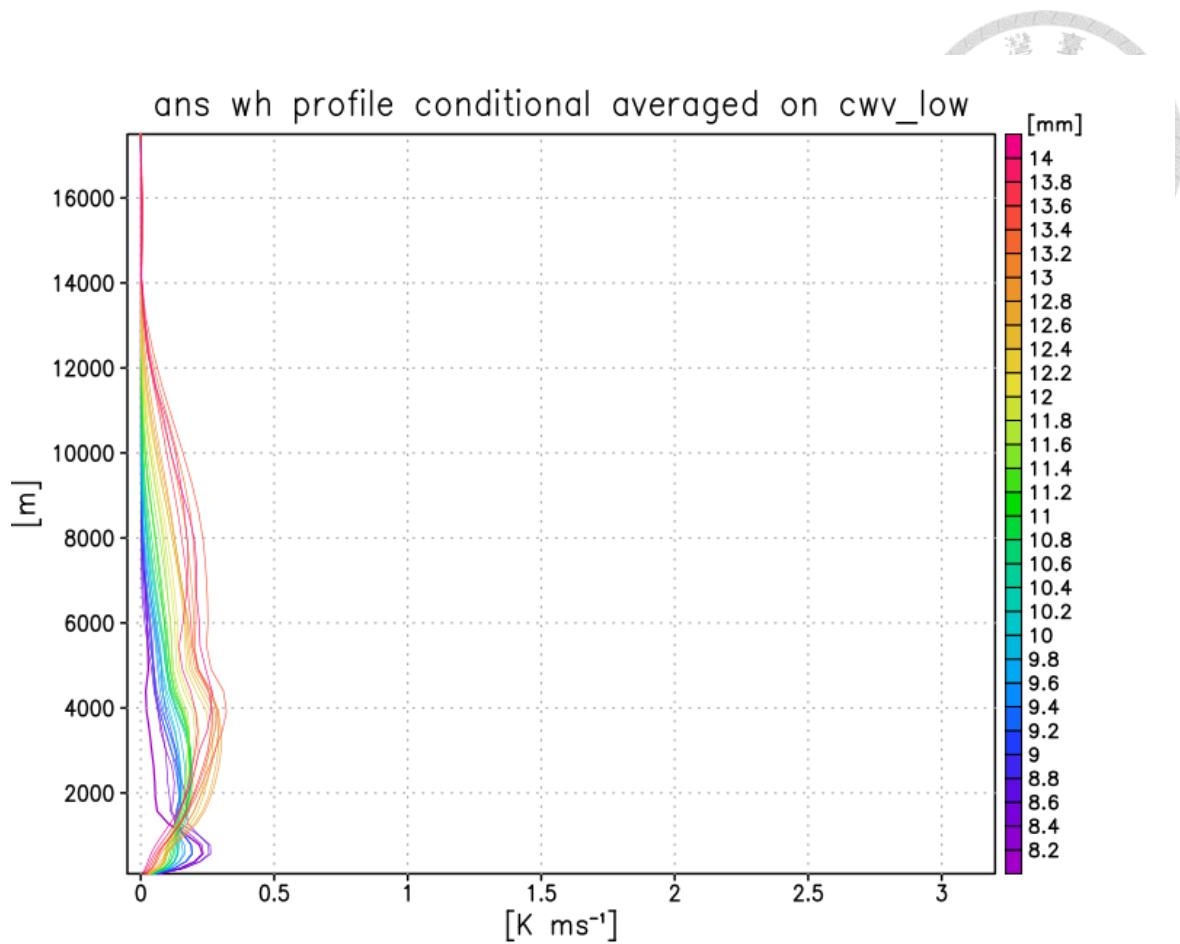
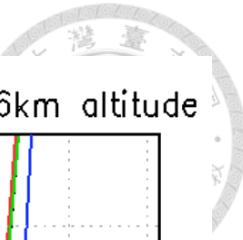


圖 18 真實模式對流強度隨著網格低層水氣的關係，顏色由冷色到暖色為網格可  
降水量少到多。其 x 軸為對流效應的強度  $\overline{w'h'}$ (K m/s)，y 軸為高度座標(m)。



Response of MSE eddy transport to CWV at 6km altitude

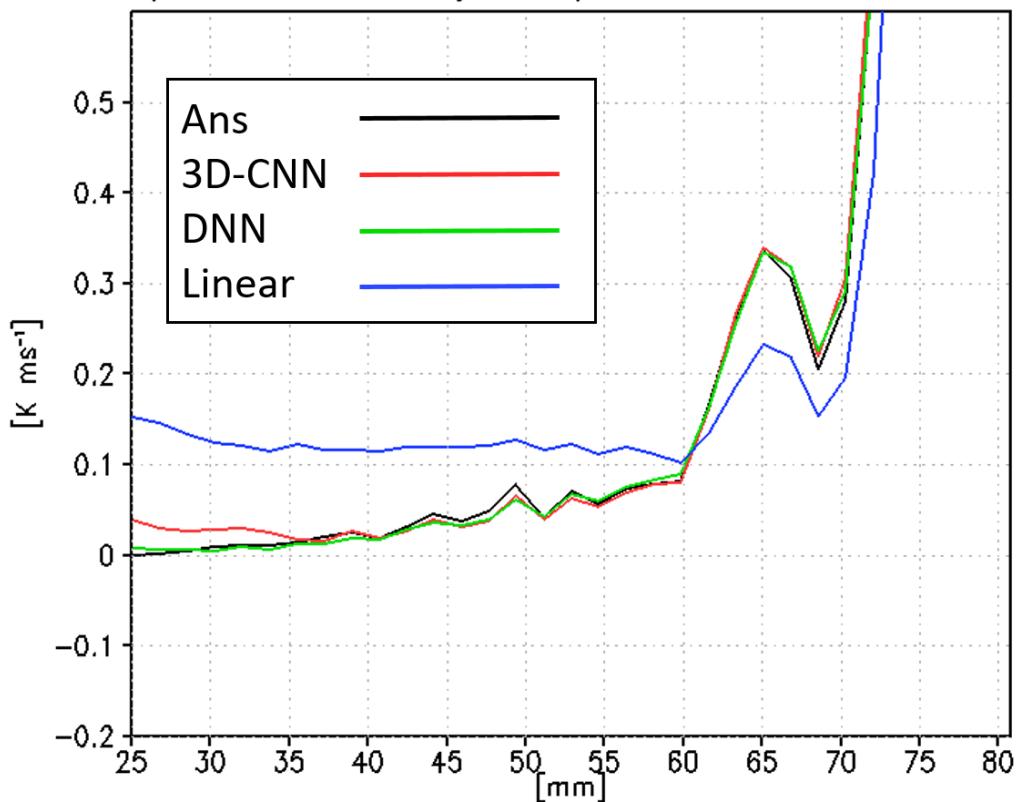


圖 19 垂直濕靜能紊流傳送與網格水氣量在 6 公里高度下的關係。黑色線為模式真實對流、紅色線為 3D 卷積神經網路、綠色線為深度神經網路、藍色線為多元線性回歸模型所預測的結果。

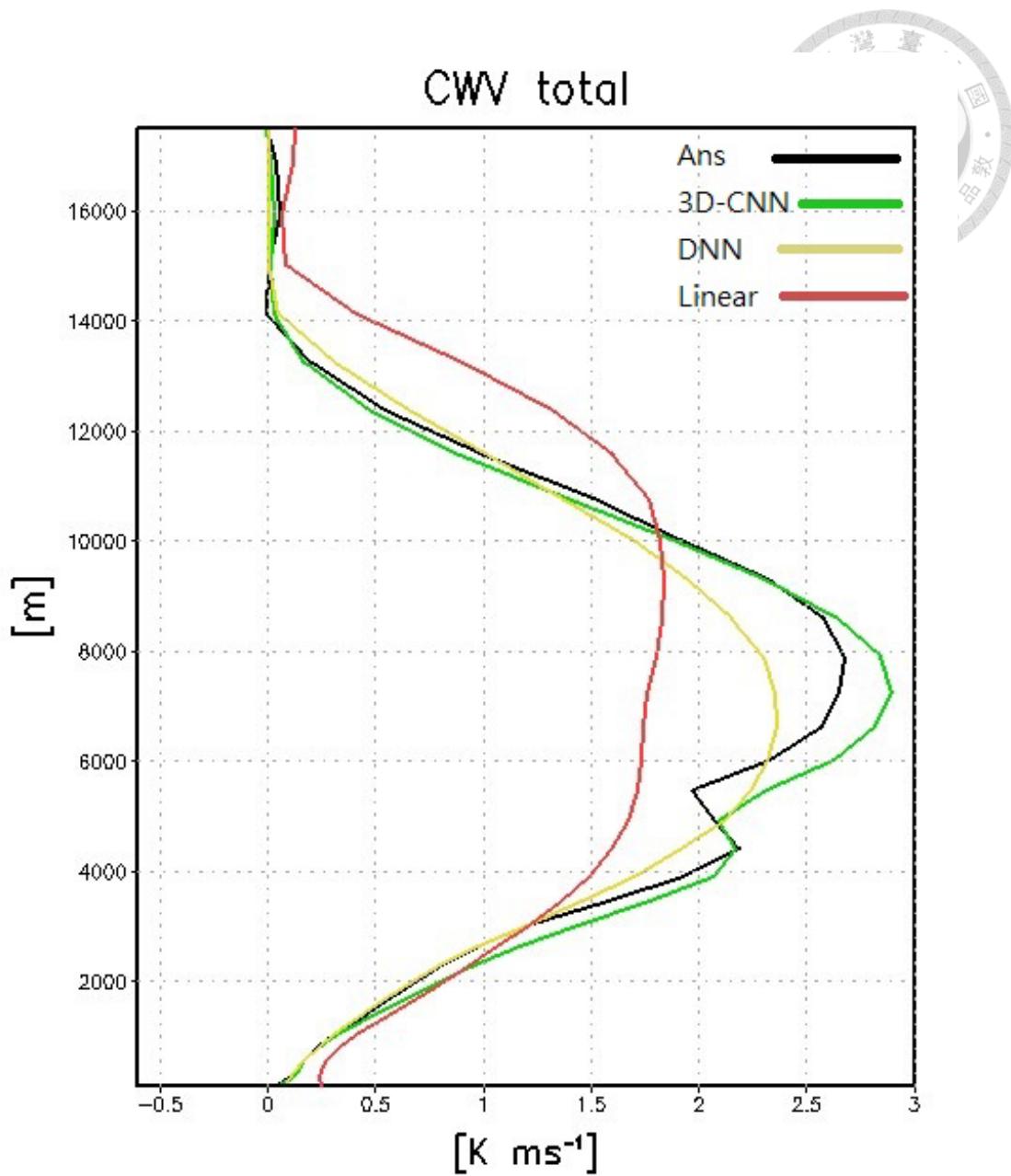


圖 20 在環境較濕時的極端對流效應對總水氣量反應之垂直結構，其中黑色線為模式真實對流、綠色線為 3D 卷積神經網路、黃色線為深度神經網路、紅色線為多元線性回歸模型所預測的結果。

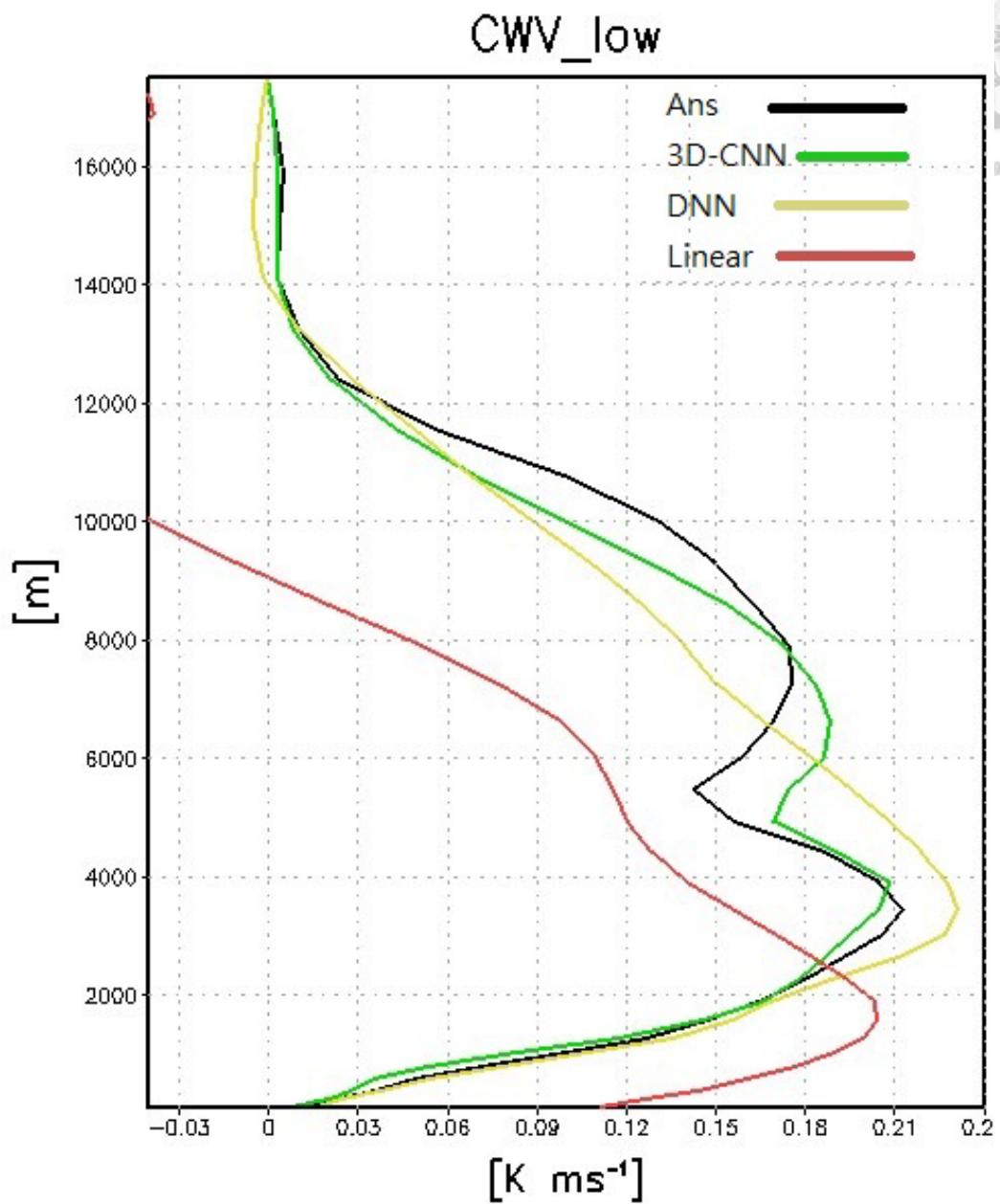


圖 21 環境較濕時的極端對流效應的對低層水氣反應之垂直結構，其中黑色線為模式真實對流、綠色線為 3D 卷積神經網路、黃色線為深度神經網路、紅色線為多元線性回歸模型所預測的結果。

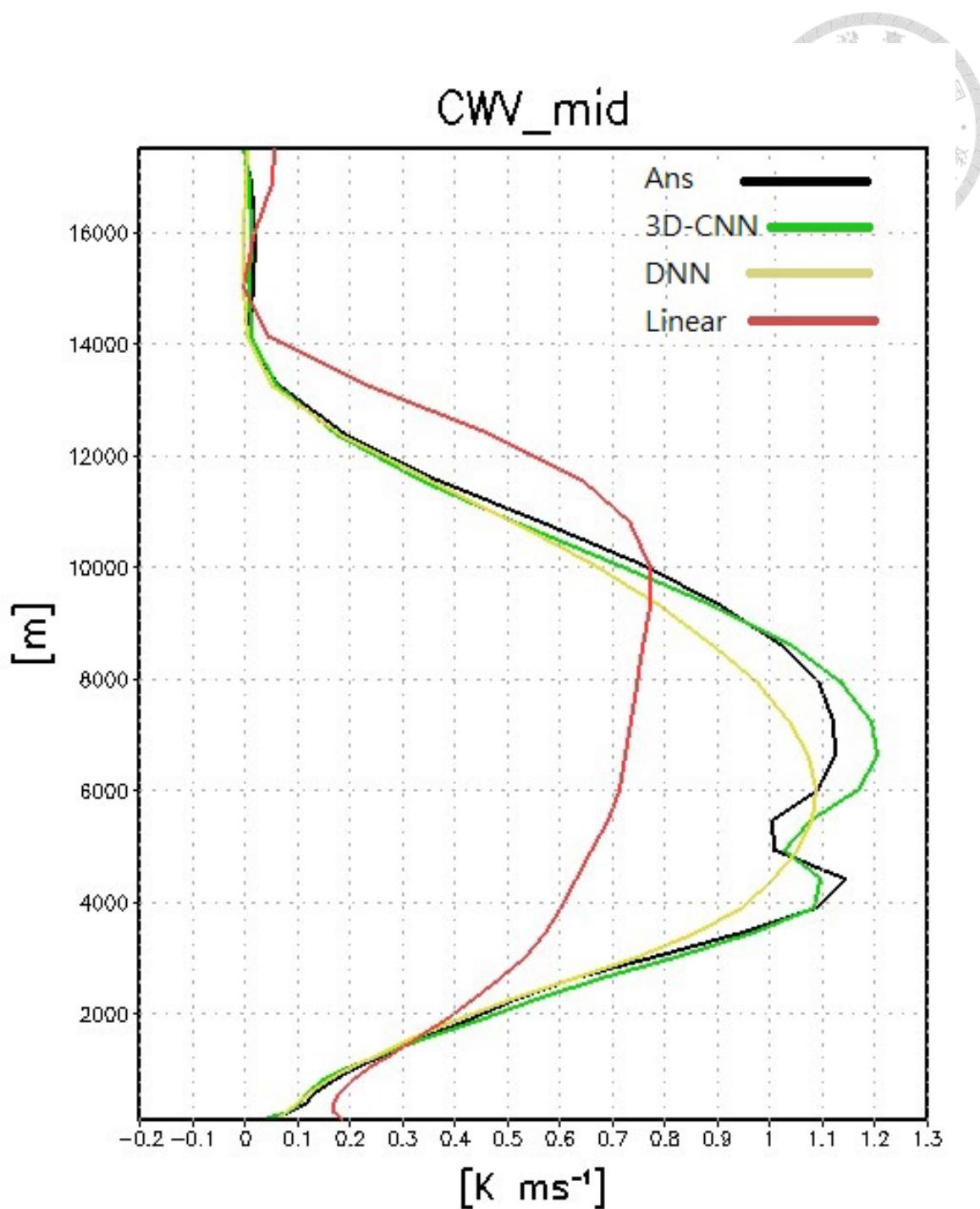


圖 22 環境較濕時的極端對流效應的對中層水氣反應之垂直結構，其中黑色線為模式真實對流、綠色線為 3D 卷積神經網路、黃色線為深度神經網路、紅色線為多元線性回歸模型所預測的結果。

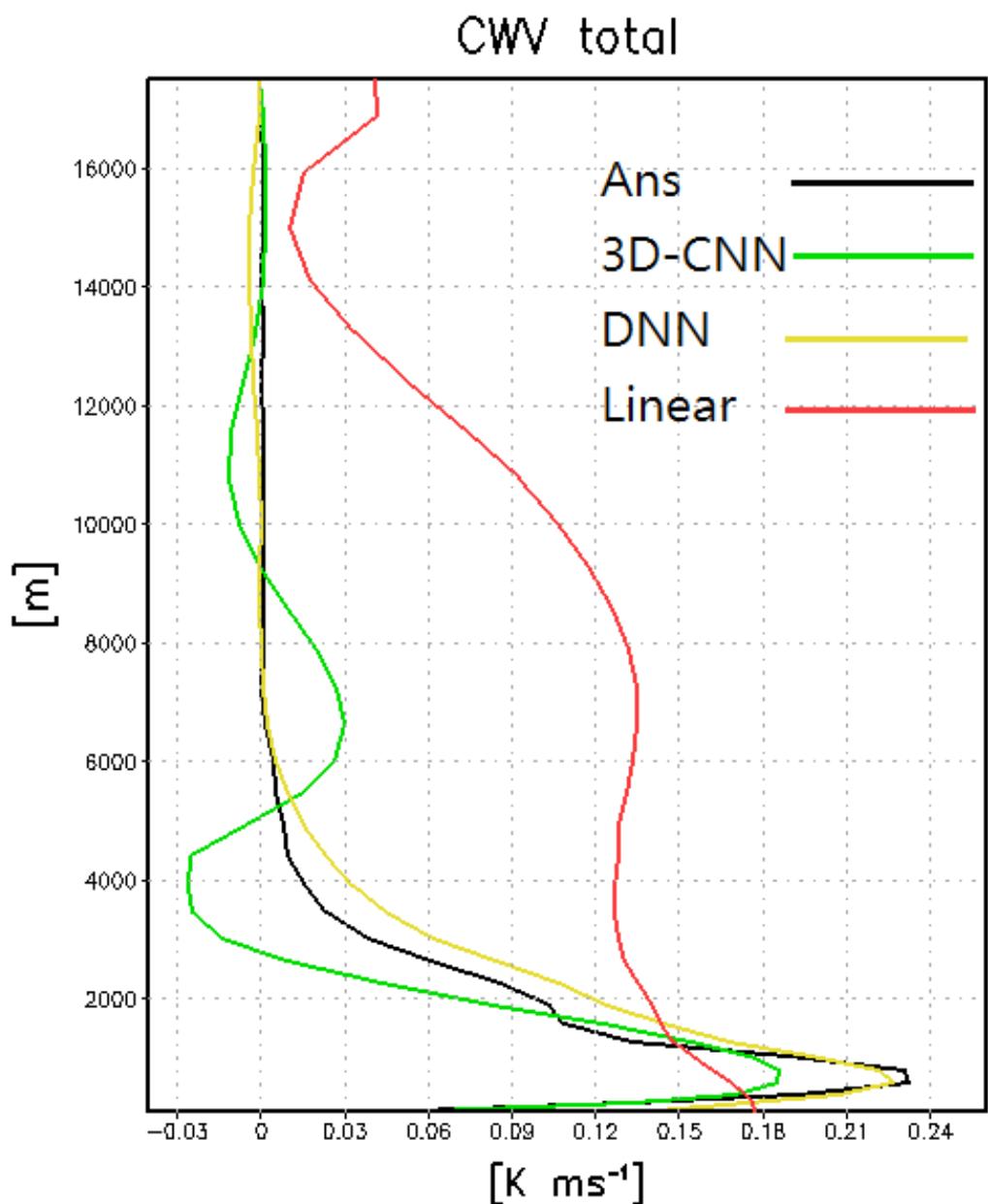
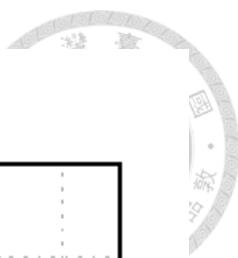


圖 23 環境較乾時的對流高度結構，其中黑色線為模式真實對流、綠色線為 3D 卷積神經網路、黃色線為深度神經網路、紅色線為多元線性回歸模型所預測的結果。

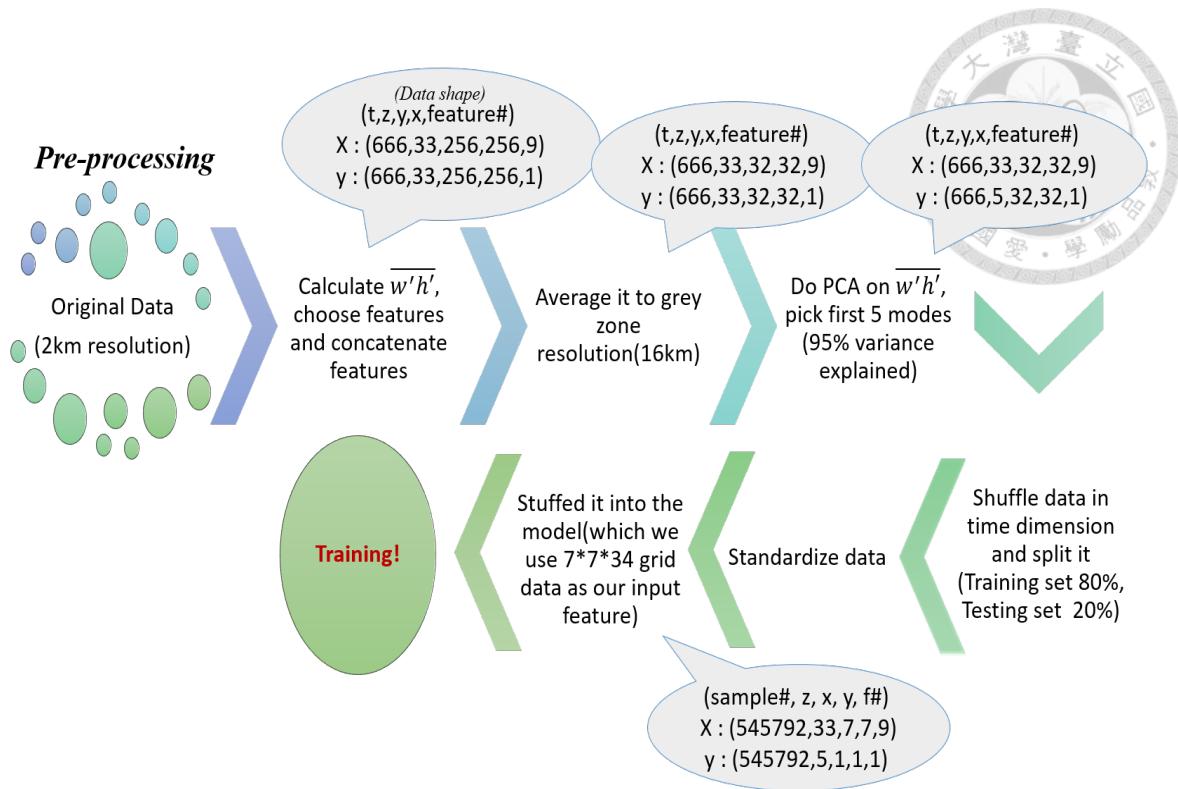


圖 24 本研究的資料前處理流程。首先挑選特徵、計算對流效應(標籤)，接著將原始 2 公里解析度的資料做平均後得到 16 公里解析度的特徵、標籤。同時將標籤利用 PCA 降為 5 維。將資料依時間的維度打亂(方便後面的分析)並將其分為訓練集、測試集，並且做資料的標準化。最後將資料塞入 3D 卷積神經網路之中，本資料將包含目標(標籤位置的網格)及其鄰近網格的特徵作為訓練預測對流的輸入。

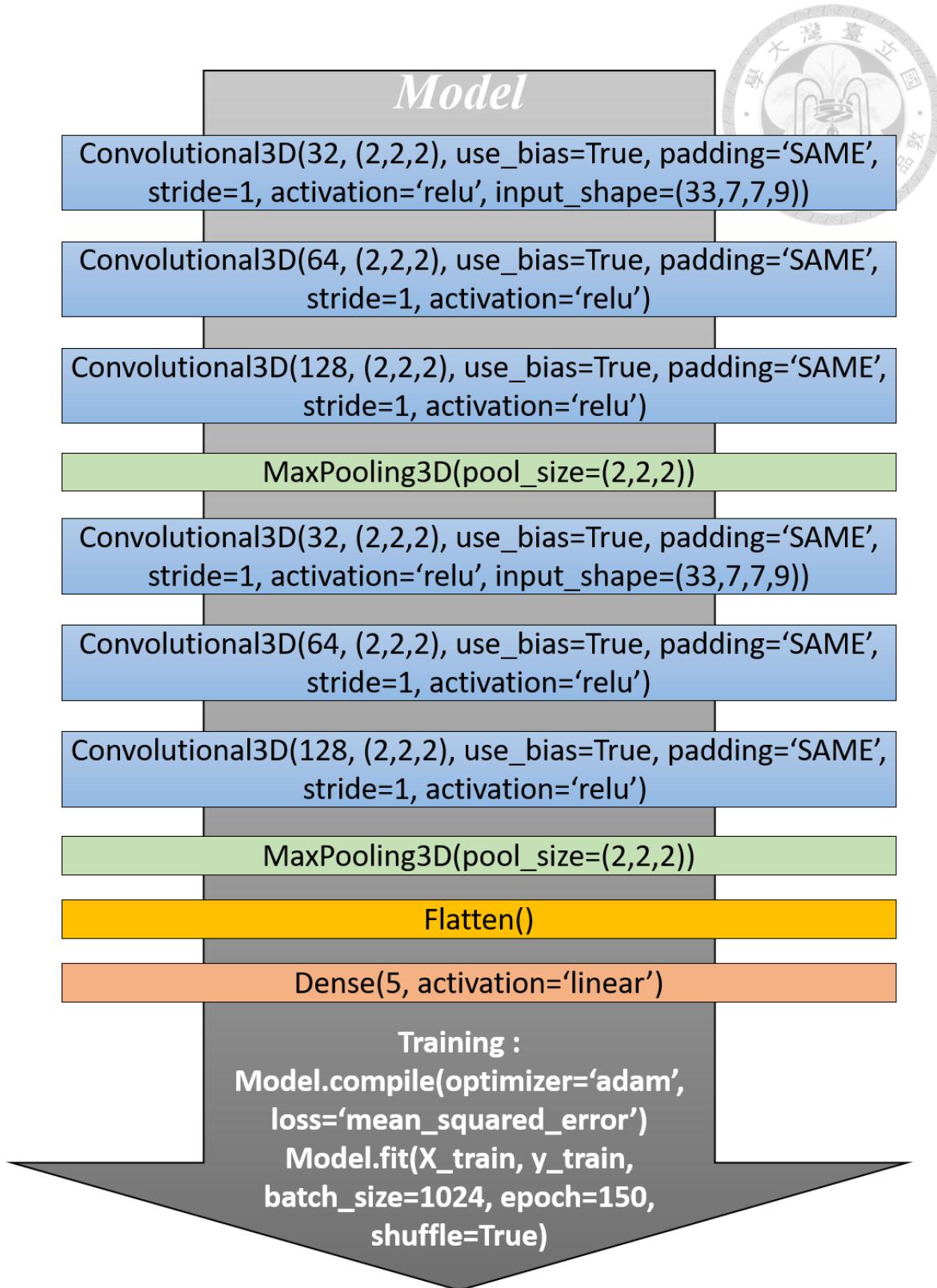


圖 25 本研究的 3D 卷積神經網路模型架構。



### *Post-processing*

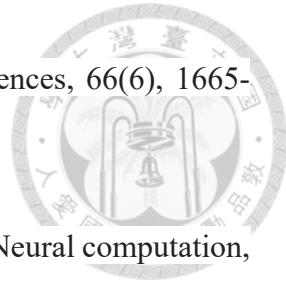
Physical evaluation metrics	3D-CNN	Linear	DNN
<b>Position of <math>w'h'</math></b>	很好的預測對流發生的位置	其預測對流效應之位置與真實結果差異較多	能夠預測對流發生的位置，但無法預測對流效應向下傳送的部分
<b>PDF of <math>w'h'</math></b>	垂直上的機率分布在極端狀況下非常接近真實對流分布	垂直上垂直上的機率分布與真實情況差距較大	垂直上的機率分布在極端狀況下與真實對流分布略有差異
<b>Moisture dependence of <math>w'h'</math></b>	能夠很好的預測對流效應對水氣量變化的反應，並且量值接近真實情況	能夠預測對流效應對水氣量變化的反應，但其量值與真實情況有些差距	能夠很好的預測對流效應對水氣量變化的反應，並且量值接近真實情況
<b>Vertical structure of <math>w'h'</math></b>	很好的預測對流效應的垂直結構，並且能預測出在高度五公里處的細微轉折	預測對流效應的垂直結構與真實情況有差距	很好的預測對流效應的垂直結構

圖 26 本研究的驗證流程。本研究分成三個部分驗證機器學習模型的預測與真實模式對流效應的差別，分別為對流的位置、對流垂直強度分布以及網格可降水量與對流強度的關係。而其中對流效應在不同濕度下的反應與其強度的垂直分布能夠被 3D 卷積神經網路很好的預測，這是傳統對流參數化很難做到的。

## 參考文獻



- Arakawa, A., & Schubert, W. H. (1974). Interaction of a cumulus cloud ensemble with the large-scale environment, Part I. *Journal of the Atmospheric Sciences*, 31(3), 674-701.
- Arakawa, A., Jung, J. H., & Wu, C. M. (2011). Toward unification of the multiscale modeling of the atmosphere. *Atmospheric Chemistry and Physics*, 11(8), 3731-3742.
- Arakawa, A., & Wu, C. M. (2013). A unified representation of deep moist convection in numerical modeling of the atmosphere. Part I. *Journal of the Atmospheric Sciences*, 70(7), 1977-1992.
- Bengio, Y., & LeCun, Y. (2007). Scaling learning algorithms towards AI. *Large-scale kernel machines*, 34(5), 1-41.
- Chien, M. H., & Wu, C. M. (2016). Representation of topography by partial steps using the immersed boundary method in a vector vorticity equation model (VVM). *Journal of Advances in Modeling Earth Systems*, 8(1), 212-223.
- Delalleau, O., & Bengio, Y. (2011). Shallow vs. deep sum-product networks. In *Advances in Neural Information Processing Systems* (pp. 666-674).
- Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). Deep learning (Vol. 1). Cambridge: MIT press.
- Gers, F. A., Schmidhuber, J., & Cummins, F. (1999). Learning to forget: Continual prediction with LSTM.
- Holloway, C. E., & Neelin, J. D. (2009). Moisture vertical structure, column water vapor,



and tropical deep convection. *Journal of the atmospheric sciences*, 66(6), 1665-1683.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780.

Jung, J. H., & Arakawa, A. (2008). A three-dimensional anelastic model based on the vorticity equation. *Monthly Weather Review*, 136(1), 276-294.

Ji, S., Xu, W., Yang, M., & Yu, K. (2013). 3D convolutional neural networks for human action recognition. *IEEE transactions on pattern analysis and machine intelligence*, 35(1), 221-231.

Jim, K. C., Giles, C. L., & Horne, B. G. (1996). An analysis of noise in recurrent neural networks: convergence and generalization. *IEEE Transactions on neural networks*, 7(6), 1424-1438.

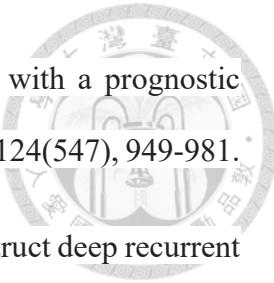
Lee, H. Y. (2018). Structured Learning. From  
[http://speech.ee.ntu.edu.tw/~tlkagk/courses/MLDS\\_2018/Lecture/rule.pdf](http://speech.ee.ntu.edu.tw/~tlkagk/courses/MLDS_2018/Lecture/rule.pdf)

LeCun, Y., & Bengio, Y. (1995). Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks*, 3361(10), 1995.

LeCun, Y. (1989). Generalization and network design strategies. *Connectionism in perspective*, 143-155.

Montúfar, G. F. (2014). Universal approximation depth and errors of narrow belief networks with discrete units. *Neural computation*, 26(7), 1386-1407.

O'Gorman, P. A., & Dwyer, J. G. (2018). Using Machine Learning to Parameterize Moist Convection: Potential for Modeling of Climate, Climate Change, and Extreme Events. *Journal of Advances in Modeling Earth Systems*, 10(10), 2548-2563.



- Pan, D. M., & Randall, D. D. (1998). A cumulus parameterization with a prognostic closure. *Quarterly Journal of the Royal Meteorological Society*, 124(547), 949-981.
- Pascanu, R., Gulcehre, C., Cho, K., & Bengio, Y. (2013). How to construct deep recurrent neural networks. arXiv preprint arXiv:1312.6026.
- Tsai, W. M., & Wu, C. M. (2017). The environment of aggregated deep convection. *Journal of Advances in Modeling Earth Systems*, 9(5), 2061-2078.
- Venugopalan, S., Rohrbach, M., Donahue, J., Mooney, R., Darrell, T., & Saenko, K. (2015). Sequence to sequence-video to text. In Proceedings of the IEEE international conference on computer vision (pp. 4534-4542).
- Wu, C. M., & Arakawa, A. (2011). Inclusion of surface topography into the vector vorticity equation model (VVM). *Journal of Advances in Modeling Earth Systems*, 3(2).
- Wu, C. M., & Arakawa, A. (2014). A unified representation of deep moist convection in numerical modeling of the atmosphere. Part II. *Journal of the Atmospheric Sciences*, 71(6), 2089-2103.
- Yanai, M., Esbensen, S., & Chu, J. H. (1973). Determination of bulk properties of tropical cloud clusters from large-scale heat and moisture budgets. *Journal of the Atmospheric Sciences*, 30(4), 611-627.

## 附錄

附錄為本文中各種名詞的解釋。



### 機器學習及監督式學習

機器學習這名詞最早在 1959 年被 Arthur Samuel[Samuel, 1959]所創造，而正式的定義是於 1997 年被 Tom M. Mitchell 在 Machine Learning[Mitchell et al., 2013.]一書中被定義，而其定義如下：

A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E.

舉個例子，若將以上定義套用在房價預測的模型之上：Task(任務)為依照房屋大小、所在地區等等「特徵」，預測房價。Experience(經驗)為觀察目前的資料，各種不同房屋之價格。Performance(表現)，在新的資料集中，真實房價與預測房價誤差的加總。

而機器學習可以依據訓練資料、產出判別的過程與結果不同，大致分為三類：監督式學習(Supervised Learning)、非監督式學習(Unsupervised Learning)、強化學習(Reinforcement Learning)。非監督式學習是訓練機器時不給機器標籤，讓機器自己在訓練資料中尋找資料特徵或者資料分群的方法。強化式學習則是讓機器本身透過觀察環境而自己決定行動，並且隨時根據新進資料修正、獲得最大報酬的方法。本研究中使用監督式學習，因此以下著重討論監督式學習的部分。

監督式學習(Supervised Learning)是從電腦標籤化(labeled)的資料中分析資料的模式，而後做出預測的一種學習方式，其中資料的標記就如同考試中的標準答案，電腦在學習的過程中透過比對誤差(Loss)，而後使用優化算法(Optimization)將誤差最小化，得到準確的預測，即是監督式學習。[Stuart and Norvig, 2016 ]



監督式學習也因為有標籤化的資料，使得其通常有較於另外兩種機器學習的方法有較高的準確率，但在現實世界的問題中往往很難取的標籤化的資料(label)，又或許需要倚賴大量的人工來將資料標籤化，不僅費時費力，在面對未知的領域時，也可能甚至完全無法使用這種方式來套入機器學習，因此本研究之資料使用三維渦度向量方程式的雲解析模式(VVM)，其被 Jung and Arakawa 在 2008 年被發明[Jung and Arakawa, 2008]，我們用其產生之原始資料，將原始資料做水平方向的平均模擬粗解析之網格資料作為機器學習模型之輸入(Input)；並利用方程式計算出平均之濕靜能傳送作為機器學習模型之輸出(Output)，將我們的積雲參數化問題轉變為一個機器學習的監督式學習之問題。

而通常監督式機器學習(Supervises Learning)可以分為兩種，回歸問題(Regression)以及分類問題(Classification)，如前述預測房屋價格的問題就是一個標準的回歸問題，若是問題之預測值並不是一個實際的量值而是類別，例如垃圾郵件分類，此類問題就被歸類為分類問題(Classification)，而本研究中使用雲解析模式之原始參數預測水氣之傳送屬於回歸問題(Regression)。



## 多變量線性回歸(Multiple Linear Regression)

本研究中使用的模型之一為多變量線性回歸模型(Multiple Linear Regression)[Neter et al., 1996]，線性回歸模型通常被認為是機器學習中回歸(Regression)問題中最基本的一種模型：即一個應變數(Y)與一個或多個自變數(X)之線性關係，而本研究中的自變數不只一個，因此我們使用的模型即稱為多變數線性回歸(Multiple Linear Regression)。所謂回歸分析(Regression analysis)即將研究的變數分為應變數(Y, 預測值)及自變數(X, 輸入值)，並建立一個由自變數映射到應變數之函數模型，也就是： $f(X) = Y$

而本研究使用之模型可寫為：

$$f(x_1, x_2, \dots, x_n) = w_1x_1 + w_2x_2 + \dots + w_nx_n \quad (w \text{ 為權重} \cdot x \text{ 為不同的特徵輸入})$$

where  $n = 6$  or  $9$  for 6features or 9features

我們藉由 2.2 節之訓練資料集訓練出其對應每個參數之權重(weights)，也就是上式之每個 w，藉由找到最小的均方根誤差來找到最適合(fit)此資料及之線性函數。



## 深度學習(Deep Learning)

深度學習為機器學習的一種方法，約從十年前開始變成機器學習(Machine Learning)研究領域中一個新的分支，同時也在許多的研討會、競賽中變得逐漸重要，如 NIPS(Neural Information Processing System)深度學習研討會、ILSVRC(ImageNet Large Scale Visual Recognition Competition)的比賽等都常見到與深度學習(Deep Learning)相關之研究。

近年來可以看到深度學習的技術被廣泛應用在各行各業之中，例如：語音辨識、影像辨識、搜尋引擎推薦系統和人類語言處理等技術，都與深度學習脫不了關係，而各大公司也競相開發有關深度學習之技術，如 Google、Facebook、Apple、Amazon 等。其中最廣為人知的深度學習應用莫過於 2016 年由 Google DeepMind 團隊所研發的下圍棋程式 Alpha Go[Silver et al., 2006]，打敗了世界頂尖之圍棋棋士李世石，從此讓世界知道人工智慧存在著超越人類的可能性，而其就是利用蒙地卡羅搜尋法結合深度學習技術所開發。



## 深度神經網路(Deep Neural Networks)

類神經網路(Artificial neural network)，是希望藉由模仿生物神經傳導的方式，其最原始的模型是由 R. Frank, 在 1957 年提出的感知器(Perceptron, Rosenblatt, Frank (1958))。感知器可以被視為最簡單的神經網路，同時也是一個二元線性分類器，其算法如下：

$$f(x) = \begin{cases} 1 & \text{if } w \cdot x + b > 0 \\ 0 & \text{else} \end{cases}$$

感知器使用特徵向量來表示的神經網路，它是一種二元分類器，把矩陣上的輸入  $x$ (實數向量)映射到輸出值  $f(x)$  上(一個二元的值)，權重  $w$  是一個實數向量、 $b$  是一個實數偏移量。 $f(x)$  的輸出是 0 或 1，用於對輸入的  $x$  值進行分類，因此感知器是一種二元分類的分類器，只要輸入乘上權重  $w$  加上偏移量  $b$  是大於 0 的數值，就會被分到 1 的類別，而若是不大於 0，就會被分到 0 的類別。而感知器就可以當作只有一個神經元(neuron)的類神經網路，而通常一個單層的類神經網路可以在一層裡有數十個到甚至幾百個神經元，然而也可以在層數上疊很多層，稱為多層神經網路(Multilayer Feedforward Network)。

所謂深度學習(Deep learning)其實就是「比較深」的類神經網路，在每個層中會有數十數百個神經元，神經元將上一層的輸入加總之後，進行一個活化函數(activation function)的轉換，當成神經元的輸出，而每個神經元的層與層之間有特殊的連結，接著再經由巨量資料的訓練，利用優化算法得到最佳的權重。而所謂深度學習的科學，就是在找尋最適合資料的架構以及每個神經元的權重。

本研究中使用的深度神經網路架構如下，其為深度總共為十層的類神經網路，每一層的神經元為 512 個，使用的激活函數(Activation function)為 ReLU。



## 激勵函數(Activation function)

在神經網路中，一個節點的激勵函數定義了該節點在給定的輸入或輸入的集合下的輸出。標準的電腦晶片電路可以看作是根據輸入得到開 (1) 或關 (0) 輸出的數位電路激勵函數。這與神經網路中的線性感知機的行為類似。

## 線性整流函數(Rectified Linear Unit, ReLU)

在神經網路中常用的一種激勵函數，其函數如下：

$$ReLU(x) = \max(0, x)$$

而其分段線性的性質能有效克服傳統較深的神經網路中梯度消失的問題，因此現今很多的神經網路都會選擇其當作激勵函數。

## 對流可用位能(CAPE)

對流可用位能為評估垂直大氣是否穩定、對流是否容易發展的指標之一。對流可用位能的單位為每公斤空氣所具有的焦耳數(J/Kg)，當其值大於零時便有發生對流的可能。一般的對流可用位能計算方式為：加總氣塊相對於環境之浮力，從自由對流高度至平衡高度為止，如下列積分式：

$$\text{CAPE} = \int g \left( \frac{T_{v,parcel} - T_{v,env}}{T_{v,env}} \right)$$

其中  $T_{v,parcel}$  為氣塊的虛溫、 $T_{v,env}$  為環境的虛溫

## 垂直積分水氣輻散(MDC)

網格垂直積分的水氣通量，其式子如下：

$$\text{MDC} = \nabla \cdot \left( \int (uq) \frac{dp}{g}, (vq) \frac{dp}{g} \right)$$

其中  $\int (uq) \frac{dp}{g}$  代表緯向的水氣流場垂直積分， $(vq) \frac{dp}{g}$  表經向的水氣流場垂直積分



， $\nabla$ 為散度運算子。

### 濕靜能(MSE)

濕靜能是一個描述氣塊狀態的熱力學參數，其定義包含了氣塊的焓(氣塊的內能和維持其體積的能量)、因高度而產生的位能和氣塊中的水氣潛熱。濕靜能在絕熱條件下上升和下降時具有保守的特性，因此適合做為研究大氣的參數。濕靜能可以被簡寫成  $S$ ，並以下列式子表示：

$$S = C_p \cdot T + g \cdot z + L_v \cdot q$$

### 主成分分析(Principal components analysis, PCA)

是一種分析、簡化數據集的技術。它通常用來減少數據集的維度，同時保留數據集中對變異數貢獻最大的特徵。由 Pearson, Karl. 於 1901 年所發明，在現今機器學習領域中常用於資料降維的技術。



## 附錄 reference

Michalski, R. S., Carbonell, J. G., & Mitchell, T. M. (Eds.). (2013). Machine learning: An artificial intelligence approach. Springer Science & Business Media.

Neter, J., Kutner, M. H., Nachtsheim, C. J., & Wasserman, W. (1996). Applied linear statistical models (Vol. 4, p. 318). Chicago: Irwin.

Pearson, K. (1901). LIII. On lines and planes of closest fit to systems of points in space. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 2(11), 559-572.

Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. Psychological review, 65(6), 386.

Russell, S. J., & Norvig, P. (2016). Artificial intelligence: a modern approach. Malaysia; Pearson Education Limited,.

Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. IBM Journal of research and development, 3(3), 210-229.

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., ... & Dieleman, S. (2016). Mastering the game of Go with deep neural networks and tree search. nature, 529(7587), 484.