

LSTM 於屏東平原地層下陷區 地下水位預測模型研究

MODELLING GROUNDWATER LEVEL FOR LAND SUBSIDENCE AREA IN PINGTUNG PLAIN BY LSTM

亞洲大學
經營管理學系
副教授

蔡存孝*
Tswn-Syau Tsay

行政院農業委員會
農田水利處
處長

謝勝信
Sheng-Hsin Hsieh

摘要

本文以資料驅動角度，建立屏東沿海地層下陷嚴重區域，地下水長短期記憶(LSTM)模型預測地下水水位，本文依據屏東沿海林邊溪南北共 11 個觀測井第 2、3、4 層含水層日地下水水位資料，分別以 3 個腳本建立 LSTM 模型。資料經過正規化後分割為訓練資料、驗證資料及測試資料，訓練過程以 Early Stopping 避免過度擬合的問題，研究發現 LSTM 模型對於測試資料 MSE 值發現，腳本 3 較腳本 1 及腳本 2 更能精確預測結果。可能因為林邊溪以南的地下水比林邊溪以北的地下水受到更多不可預知的干擾，大潭(2)、枋寮(2)、德興(2)需要再加入影響地下水位的其他特徵或再建構更優化的 LSTM 模型，才能降低測試資料 MSE 值，增加模型的精確度。大潭(2)、崎峰(2)、崎峰(3)、崎峰(4)、新埤(2)、新埤(3)、新埤(4)及德興(3)而言，無論使用單層 LSTM 模型或 3 層 LSTM 模型，30 次 LSTM 模型訓練即可訓練出 LSTM 模型讓測試資料 MSE 值在 0.001 以下，代表實際最高與最低地下水水位相差 10 公尺時，預測地下水水位與實際的地下水水位會產生 32 公分的均方差，LSTM 模型 MSE 表現最差的是枋寮(2)及崎峰(4)，表現最好的是大潭(2)及新埤(4)，整體而言，模型預測值與實際值的相關係數均在 0.98 以上。

關鍵詞：地下水，長短期記憶法，屏東平原。

* 通訊作者，亞洲大學經營管理學系副教授

41354 台中市霧峰區柳豐路 500 號，tswnsyau@gmail.com

MODELLING GROUNDWATER LEVEL FOR LAND SUBSIDENCE AREA IN PINGTUNG PLAIN BY LSTM

Tswn-Syau Tsay*

ASIA University
Department of Business Administration

Sheng-Hsin Hsieh

Council of Agriculture, Executive Yuan
Department of Irrigation and Engineering

ABSTRACT

Groundwater level prediction model by Long Short-Term Memory (LSTM) was developed for Pingtung coastal area where land subsidence problems have been for many years. There are 11 LSTM models for 3 scenarios were developed based upon daily groundwater level records for 2nd, 3rd and 4th stratified aquifers of 11 monitoring wells on the north and south of Lin-Bian Creek in Pingtung, respectively. Data set of daily groundwater level was normalized and split into training data set, verification data set and testing data set. Early Stopping algorithm was applied during training process to avoid overfitting problem. It was found that LSTM model with scenario 3 obtained more accurate predictions than the other scenarios. Furthermore, it is probably the groundwater on the south of Lin-Bian Creek received unpredicted disturb, other features were required for the 3 models of this area to increase model accuracy. For models on the north of Lin-Bian Creek, 30 times of Early Stopping model training were sufficient for obtaining LSTM model's testing MSE less than 0.001 either by 1-LSTM-Layer model or 3-LSTM-Layer model. As the groundwater level was normalized, MSE less than 0.001 indicated a root mean square error of 32 cm for the difference of maximum and minimum groundwater level of 10 meters. The worst models by LSTM are FangLiao2 and QiFeng4. The best models are DaTang2 and XinPi4 which perform very good prediction ability. Overall, correlation coefficients of predicted and actual groundwater levels are all above 0.98.

Keywords: Groundwater, LSTM, Pingtung Plain.

一、前言

由於台灣降雨時空分布不均，雖然每年高達 2,500 mm 的降雨量，仍經常發生乾旱現象[1, 2]。而就水資源的應用而言，不論民生、農業或工業用途，地下水都是補充乾旱時期用水的最佳選擇之一[3, 4]，在乾旱季節時期，地下水水位也常因為補充地表水資源而產生水位洩降，但必須注意在安全的範圍內使用地下水。台灣沿海西部地區過去即因為在短時間內集中抽取過量的地下水，因而發生地層下陷的問題[5]。其他國家的研究也發現，過度使用地下水或地下水管理不善所衍生的相關問題[6-8]，因此，除了設置觀測井以長期觀測地下水水位變化外，地下水水位預測模型也有助於地下水的管理。由於複雜的水文地質系統、邊界條件與不確定性，地下水水位與水文循環的各項因素呈現高度非線性關係，由數值方法建立的地下水模型對於水文地質參數、邊界條件有高度的要求，才能獲得比較可靠的結果，雖然達西定律及連續方程式在控制良好的條件下能獲得完美的結果，但對於實際環境上的模擬的不確定性，造成地下水模型預測上的其他問題[9-12]。

機器學習及人工神經網路 (Artificial Neural Network, ANN) 為資料科學的分析方法，可就長期蒐集的資料(如雨量、地下水水位、溫度、河川水位、水庫水位等)，作為模擬預測地下水的輸入特徵(features)，文獻指出，ANN 整合週雨量、蒸發散量、河川水位、地表水位及抽水量預估未來一週的地下水水位變化[13]，為評估乾旱及水資源管理依據；樹木年輪資料結合 ANN 可以精確的預估地下水水位的變化[14]。機器學習也整合了地理資訊系統，建立地下水潛能圖[15]。由於地下水水位受到地表水的入滲的影響，學者比較 ANN 及其他機器學習法，應用降雨、溫度、地下水水位及湖泊水位等資料，作為建立預測美國佛州地下水水位模型依據，並比較不同方法的優劣[16]。學者也在加拿大魁北克省應用 ANN 作為預測短期地下水水位的模擬工具，證實 ANN 可以獲得精確的預測成果[17]。新加坡學者成功應用湖泊水位及雨量，以 ANN 作為淡水沼澤森林區地下水水位的預測工具[18]。機器學習也應用於農業區地下水水位的模擬預測，尤其無法取得水文地質資料的區域，可以做為評估地下水的替代模擬工具[19]。

在人工智慧的領域上，處理時間序列問題方面時，經常使用長期短期記憶法(Long Short-Term Memory, LSTM)，LSTM 改善遞迴神經網路(RNN, Recurrent Neural Network)梯度消失(Vanishing Gradient)問題，可以有效的處理序列型資料[20]，並且被大量應用於語言建模[21, 22]，語音辨識[23, 24]，證券預估[25, 26]，降

雨及乾旱預估[27, 28]及地下水水位預測等[18, 29]，近年來因為 GPU 在平行計算上的進步，使得類神經網路的研究及應用快速的發展[30-32]。LSTM 被開發用於中國西北河套灌區的地下水水位模擬[33]，Bowes 等(2019)則應用遞迴神經網路(RNN)與長期短期記憶(LSTM)於模擬美國加州海岸地區暴風雨事件對地下水的反應[29]，相關研究指出，LSTM 可以獲得相當不錯的預測模型[34]。在 GPU 計算普及之前，ARIMA (AutoRegressive Integrated Moving Average)是處理時間序列問題的重要方法之一，由於撒哈拉以南非洲地區的應用地下水做為灌溉為當地經濟成長和減貧的重要資源，ARIMA 被應用於地下水水位的預測[35]，ARIMA 也被應用於巴西含水層的地下水水位預測[36]，近期的研究顯示，在模型的預測能力上，LSTM 的表現優於 ARIMA [37, 38]，但也各有優缺[39]。

從前述的文獻回顧可以歸納，即使機器學習或 ANN 在模擬上存在一些缺點如過度擬合問題、缺乏一般性、使用錯誤特徵及缺乏物理推論的理論根據等問題[40]，但機器學習及 ANN 在地下水水位的模型建立上，僅需要基本的觀測資料做為訓練的特徵，就可以獲得不錯的成果[18, 29]，且可快速的建立模擬模型，更由於電腦硬體效能的提升，更快速且平行倒傳遞計算，使得深度 ANN(如 RNN)也能在計算效能上大幅提升。

屏東平原地下水區面積約 1,231 平方公里，是台灣重要的地下水之一，屏東平原沖積扇頂附近地質為礫石層，從水文地質鑽探判釋結果，屏東平原下各地下水層互相連通，而扇頂為重要的地下水補注區[41]。雖然屏東平原由北至南有 4 個大沖積扇及 3 個小沖積扇作為地下水補注區，但沿海地區仍發生嚴重的地層下陷問題[42, 43]，主要原因仍與短時間內大量抽取地下水，地下水水位瞬間下降，造成含水層壓密所致[6]，因此，短期的地下水水位預估，如能提前發出地下水抽水警訊，對防止地層下陷發生極為重要。本研究從資料驅動(data-driven)的角度出發[44, 45]，以屏東平原沿海地層下陷嚴重區域為研究對象，以 LSTM 模型作為時間序列模型建立的方法，從地下水水位預測模型建立中，了解 LSTM 模型的限制及預測能力，作為應用於實務操作的參考依據。

二、研究方法

2.1 研究資料

依據經濟部提供之屏東平原民國 99 年至 100 年平均下陷速率等值線圖[43](圖 1)及地層下陷防治資訊網

之民國 83 年至 106 年下陷分析[46]指出，屏東平原的地層下陷仍以佳冬鄉、林邊鄉及枋寮鄉最為嚴重，因此，本文以其附近的觀測井作為研究對象，包含：大潭、崎峰、新埤、佳冬、大庄、枋寮及德興等各含水層之觀測井，各觀測井在各含水層的地下水位資料紀錄如

表 1，除佳冬外，其他鑽探井均有超過 20 年以上的日地下水位資料。圖 2(a-f)為各觀測井日平均地下水位歷線圖，其中，新埤、大庄、枋寮及德興第 2 含水層以下的觀測井最高及最低日平均地下水位差均超過 10 ~ 20 公尺以上，地下水位高低變化極大。

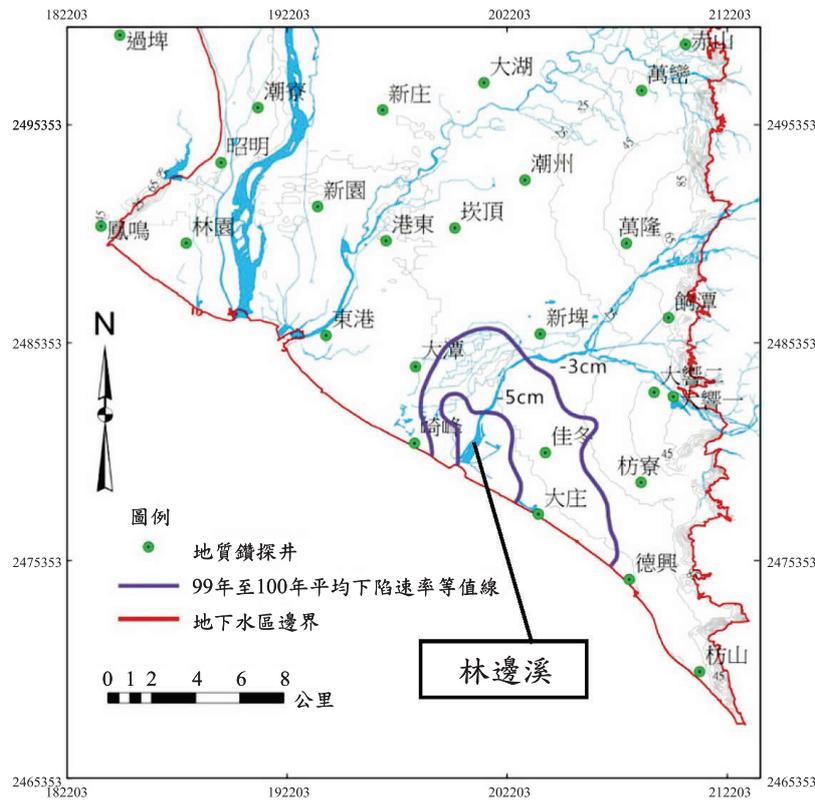


圖 1 屏東平原民國 99 年至 100 年平均下陷速率等值線圖(資料來源：經濟部)[43]

表 1 本研究屏東平原沖積扇地下水觀測井表列

觀測井	資料紀錄	備註	資料筆數
大潭(1)	1995/11/1~2018/12/31	第 1 含水層	8,462
大潭(2)	1995/11/1~2018/12/31	第 2 含水層	8,462
崎峰(1)	1997/1/1~2018/12/31	第 1 含水層	8,035
崎峰(2)	1995/1/1~2018/12/31	第 2 含水層	8,462
崎峰(3)	1997/1/1~2018/12/31	第 3 含水層	8,035
崎峰(4)	1997/1/1~2018/12/31	第 4 含水層	8,035
新埤(1)	1995/11/1~2018/12/31	第 1 含水層	8,462
新埤(2)	1995/11/1~2018/12/31	第 2 含水層	8,462
新埤(3)	1997/2/1~2018/12/31	第 3 含水層	8,004
新埤(4)	1997/2/1~2018/12/31	第 4 含水層	8,004
佳冬	無		
大庄(1)	1997/1/1~2017/12/31	第 1 含水層	7,670
大庄(2)	1997/1/1~2018/12/31	第 2 含水層	8,035
枋寮(1)	1997/2/1~2019/1/31	第 1 含水層	8,035
枋寮(2)	1997/2/1~2018/1/31	第 2 含水層	8,035
德興(1)	1998/6/1~2018/12/1	第 1 含水層	7,489
德興(2)	1998/7/1~2018/12/31	第 2 含水層	7,489
德興(3)	1998/6/1~2018/12/31	第 3 含水層	7,519

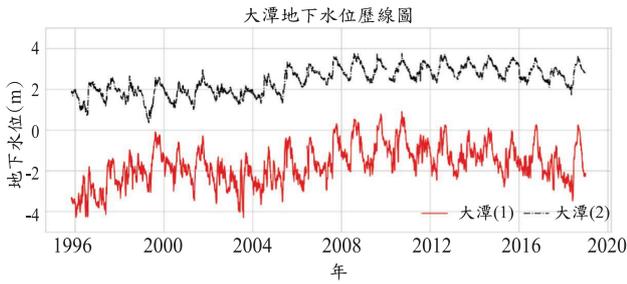


圖 2(a) 大潭站地下水水位歷線圖

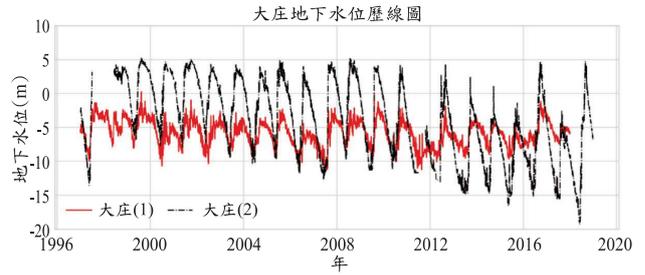


圖 2(d) 大庄站地下水水位歷線圖

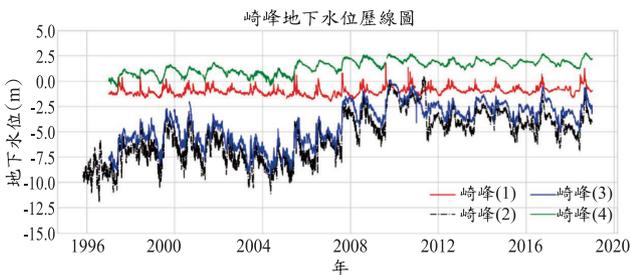


圖 2(b) 崎峰站地下水水位歷線圖

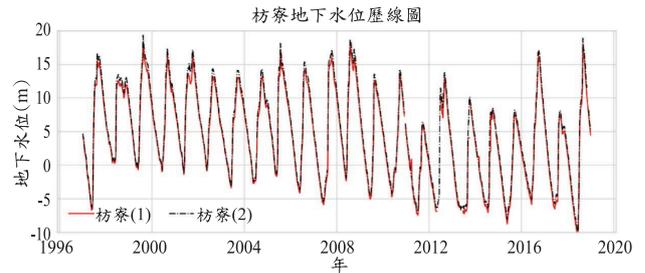


圖 2(e) 枋寮站地下水水位歷線圖

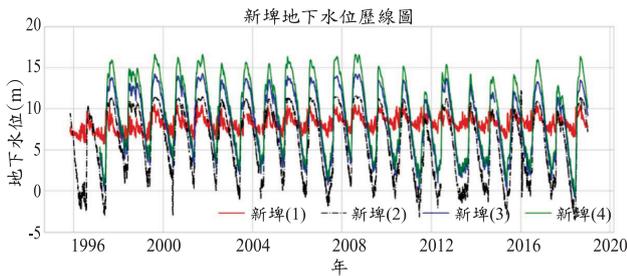


圖 2(c) 新埤站地下水水位歷線圖

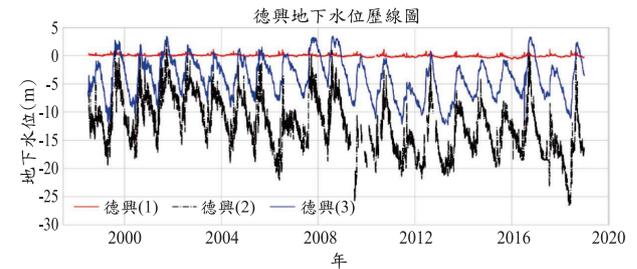


圖 2(f) 德興站地下水水位歷線圖

46

2.2 遞迴神經網路(RNN)

RNN 的基本原理可以源自於人工神經網路 (Artificial Neural Network, ANN)，ANN 包含輸入層、隱藏層及輸出層，圖 3 顯示輸入層(Input Layer)有 3 個神經元(x_1, x_2, x_3)，隱藏層(Hidden Layer)有 2 個神經元(①與②)，輸出的結果為(h_1, h_2)，最後再連接輸出層(圖 3 未顯示)，則

$$h = \varphi(W^T x + b) \dots\dots\dots (1)$$

式(1)中， W 為權重參數， W^T 為 W 的轉置矩陣 (Transpose)， b 為偏移值，其中，

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, W = \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \\ w_{31} & w_{32} \end{bmatrix}, b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, h = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \dots\dots (2)$$

φ 為激勵函數(activation function)。

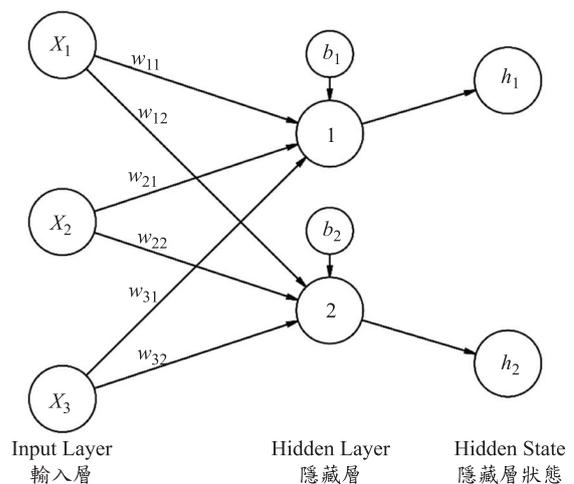


圖 3 ANN 結構範例圖

h 為這個神經網路的輸出結果，在 RNN 中， h 可視為隱藏層狀態(hidden state)[47]，由於 RNN 在計算時將上一個時間($t-1$)的隱藏層狀態作為這一個時間(t)的輸入，加上時間變數後如圖 4。

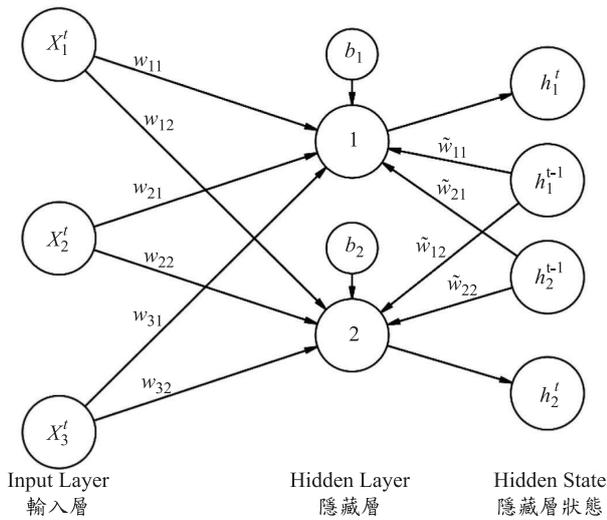


圖 4 RNN 結構範例圖

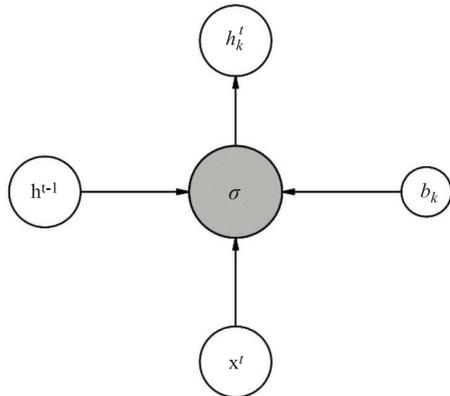


圖 5 簡化後 RNN cell 示意圖

其中 x^t 為時間 t 的輸入特徵(input feature)，以圖 4 而言，共有 3 個輸入特徵(x_1^t , x_2^t 及 x_3^t)，①與②稱為 RNN cell(包含輸入 h^{t-1} 、 x^t 、 b_k 及輸出 h_k^t)，輸出 2 個隱藏層狀態(h_1^t 及 h_2^t)，因此， h_1^{t-1} 及 h_2^{t-1} 為前一個時間 $t-1$ 的隱藏層狀態， W_x 為 x^t 的輸入權重參數， W_h 為 h^{t-1} 的輸入權重參數，表示為：

$$x^t = \begin{bmatrix} x_1^t \\ x_2^t \\ x_3^t \end{bmatrix}, W_x = \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \\ w_{31} & w_{32} \end{bmatrix}, W_h = \begin{bmatrix} \tilde{w}_{11} & \tilde{w}_{12} \\ \tilde{w}_{21} & \tilde{w}_{22} \end{bmatrix} \dots (3)$$

從 RNN cell 的角度來看，圖 4 可以簡化為圖 5， σ 為激勵函數，第 k 個 RNN cell 的輸出，即隱藏層狀態為

$$h_k^t = \sigma(W_h^T h^{t-1} + W_x^T x^t + b_k) \dots (4)$$

其中， σ 為 sigmoid 函數，定義為：

$$\sigma(x) = \frac{1}{1 + e^{-x}} \dots (5)$$

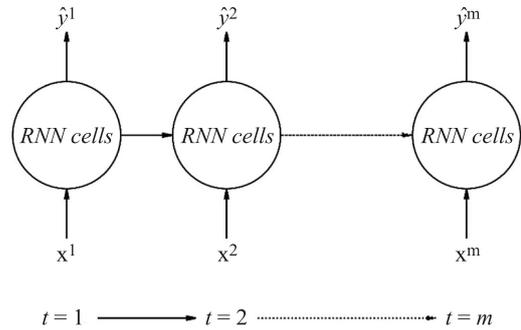


圖 6 RNN 訓練結構圖

每一個 RNN cell 僅輸出一個隱藏層狀態，而且為純量，式(2)中 W_x 、 W_h 及 b_k 均與時間無關，表示 RNN 的計算假設，相同的 RNN cell 就有相同的權重參數 (W_x^T 與 W_h^T) 及偏移值 (b_k)。當時間 $t=0$ 時，初始的隱藏層狀態假設為 0。 h^t 並不是 RNN 的最後輸出結果，而是在訓練過程的一個結果， h^t 會再接上其他的神經網路(可能是輸出層)，計算預測的輸出結果 (\hat{y}^t)。RNN 訓練的過程，利用梯度下降法(Gradient Decent)不斷的調整每一個 RNN cell 的 W_x 、 W_h 及 b_k ，求取 \hat{y}^t 與實際的結果之間的最小差異或損失，求得最佳的 RNN cell 的 W_x 、 W_h 及 b_k ，就是 RNN 訓練的結果。因此，RNN 的訓練結構可以表示為圖 6，類似一個深度學習神經網路。

2.3 長期短期記憶(LSTM)

圖 6 的 RNN 訓練中，如果使用梯度下降法訓練，求取最佳的權重參數與偏移值，會耗費許多訓練的計算時間，而倒傳遞演算法(Backpropagation)[48, 49]可以大幅提高計算效率，因為 RNN 的結構也類似深度學習的神經網路，在 RNN 中使用倒傳遞演算法(Backpropagation Through Time, BPTT)時，RNN 在時間上會產生很深的網路，進行演算中，當參數小於 1 時，BPTT 會產生梯度消失的問題，LSTM 則是解決遞迴神經網路 BPTT 的梯度消失問題的方法之一[20]。LSTM 的方法是增加 3 個控制閥，分別為忘記閥(f_k^t)、輸入閥(i_k^t)及輸出閥(o_k^t)及一個 Cell State (C_k^t)，提高遠端(long term)記憶的影響。計算過程如下：

$$f_k^t = \sigma(W_{fh}^T h^{t-1} + W_{fx}^T x^t + b_{fk}) \dots (6)$$

$$i_k^t = \sigma(W_{ih}^T h^{t-1} + W_{ix}^T x^t + b_{ik}) \dots (7)$$

$$o_k^t = \sigma(W_{oh}^T h^{t-1} + W_{ox}^T x^t + b_{ok}) \dots (8)$$

W_{fh}^T 、 W_{fx}^T 、 W_{ih}^T 、 W_{ix}^T 、 W_{oh}^T 及 W_{ox}^T 為權重參數矩陣，

b_{fk} 、 b_{ik} 及 b_{ok} 為偏移值，對於忘記閥、輸入閥及輸出閥而言，輸出結果需要介於 0 與 1 之間，因此激勵函數使用 sigmoid 函數(式(5))。 C_k^t 的計算應用忘記閥、輸入閥、上一個時間的 cell state (C_k^{t-1})及 \tilde{C}_k^t ：

$$C_k^t = f_k^t C_k^{t-1} + i_k^t \tilde{C}_k^t \dots\dots\dots (9)$$

其中，

$$\tilde{C}_k^t = \tanh(W_{Ch}^T h^{t-1} + W_{Cx}^T x^t + b_{Ck}) \dots\dots\dots (10)$$

W_{Ch}^T 及 W_{Cx}^T 為權重參數矩陣， b_{Ck} 為偏移值，
 $\tanh(x) = \frac{1 - e^{-x}}{1 + e^{-x}}$

最後計算時間 t 的隱藏層狀態：

$$h_k^t = o_k^t \tanh(C_k^t) \dots\dots\dots (11)$$

2.4 模型評估指標

LSTM 模型建立後，以預測地下水位(\hat{y}_t)及實際地下水位 y_t 的 MSE (Mean Squared Error) 做為評估模型精確度的指標，MSE 是指預測地下水位(\hat{y}_t)值與測試資料(y_t)值的差平方的期望值，MSE 的值越小，代表預測地下水位(\hat{y}_t)值與測試資料(y_t)越接近，即模型有更

好的精確度。

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

N 為測試資料的個數。

2.5 模型訓練流程

為求每次的 LSTM 模型訓練結果最佳化，應用 Early Stopping[50]作為模型訓練避免過度擬合的問題，詳如圖 7。神經網路計算中，將所有的訓練資料看過 1 次，稱為 1 次 Epoch，每次驗證資料 MSE 值的計算均為 1 次 Epoch，Early Stopping 的參數在 Keras 稱為 patience，patience 的次數代表訓練資料的訓練結果，應用到驗證資料所預測的損失是否持續下降的次數，patience 越大，通常需要越多的計算時間，因此，需要在不造成過度擬合的狀況下，找到合適的 patience，本文列舉 3 站地下水觀測井(德興(3)、新埤(3)及崎峰(4))，在實務模型建立經驗上，patience 通常會設定在 10 以上，本文分別以 patience=10, 15, 20, 25, 30 進行 LSTM 模型建立，分別記錄 MSE、最大預測誤差、最小預測誤差、相關係數 (correlation coefficient, γ)及計算時間(表 2~4)，最後選擇 patience=20，以獲取較佳的 LSTM 模型參數，所有的計算均以 GeForce RTX 2080 完成。

表 2 大潭(2)單層 LSTM Early Stopping patience 測試資料執行結果

patience	MSE	最大誤差	最小誤差	γ	計算時間
10	7.3709e-05	3.1837e-02	1.0530e-05	0.9986	1'27"
15	1.4503e-04	3.8793e-02	2.3203e-05	0.9985	1'23"
20	9.3530e-05	3.4252e-02	1.5043e-05	0.9986	1'25"
25	9.0540e-05	3.3383e-02	1.8853e-05	0.9986	1'32"
30	8.4320e-05	3.1445e-02	8.7245e-05	0.9986	1'28"

表 3 新埤(3)單層 LSTM Early Stopping patience 測試資料執行結果

patience	MSE	最大誤差	最小誤差	γ	計算時間
10	1.7094e-03	9.8973e-02	5.5836e-03	0.9987	0'30"
15	1.6384e-03	1.1962e-01	1.6417e-05	0.9988	0'42"
20	7.6876e-04	5.2791e-02	1.4632e-03	0.9996	1'12"
25	1.0845e-04	3.9910e-02	4.7506e-06	0.9997	1'21"
30	7.6100e-04	9.1086e-02	1.5237e-04	0.9993	2'07"

表 4 崎峰(4)單層 LSTM Early Stopping patience 測試資料執行結果

patience	MSE	最大誤差	最小誤差	γ	計算時間
10	1.1132e-03	3.5335e-01	4.9768e-05	0.9808	2'00"
15	1.1269e-03	3.5039e-01	1.1408e-05	0.9804	1'18"
20	1.1253e-03	3.5174e-01	4.2038e-06	0.9804	1'23"
25	1.0694e-03	3.3920e-01	3.2783e-05	0.9799	1'39"
30	1.0897e-03	3.4923e-01	1.0010e-04	0.9807	1'38"

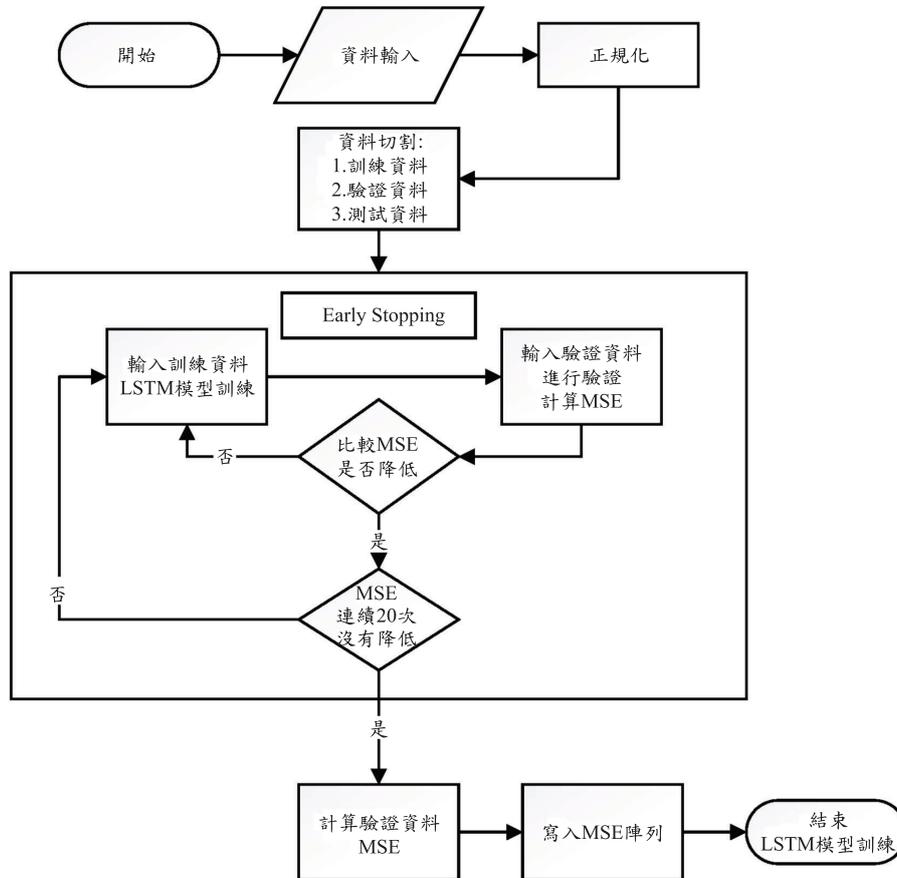


圖 7 LSTM 模型訓練 Early Stopping 計算流程圖

三、結果與討論

由於沿海地層下陷的發生原因通常因為深層地水位在幾日內的驟降引起[51-53]，本研究以第 2、3、4 層的地下水位為研究對象，包含：大潭(2)、崎峰(2)、崎峰(3)、崎峰(4)、新埤(2)、新埤(3)、新埤(4)、大庄(2)、枋寮(2)、德興(2)、德興(3)，建立 LSTM 多對多(many to many)模型，並且假設 3 個腳本，依據過去 30 天的地下水位(PD=30)，預測未來數天內的地下水位：

- (1) 腳本 1：從過去 30 天的地下水位資料，預測未來 1 天的地下水位(PD=30、FD=1)
- (2) 腳本 2：從過去 30 天的地下水位資料，預測未來 3 天的地下水位(PD=30、FD=3)
- (3) 腳本 3：從過去 30 天的地下水位資料，預測未來 7 天的地下水位(PD=30、FD=7)

本文依據腳本，對所要研究的觀測井，建立單層 LSTM 模型及 3 層 LSTM 模型，分析不同腳本及不同 LSTM 模型對於測試資料的預測結果。

3.1 單層 LSTM 模型

以腳本 3 為例，資料輸入後進入單層 LSTM 模型 (1 層 LSTM 層、32 個神經元)，再連接全連結層，攤平後經過全連接層，輸出最後預測結果，如圖 8。

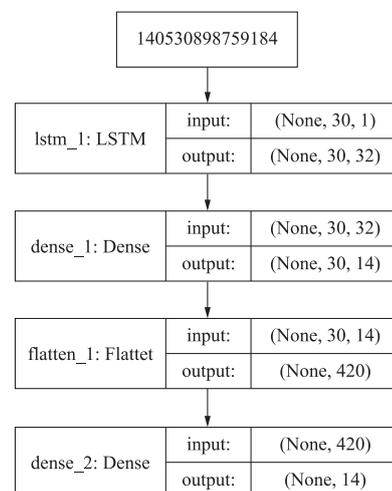


圖 8 腳本 3 單層 LSTM 模型圖

3.1.1 模型訓練

從表 1 各觀測井資料筆數，各觀測井至少有 7,400 筆以上資料可作 LSTM 訓練、驗證及測試，本研究僅取最後 5,120 筆資料，其中，前 70% 的資料作為模型訓練，後續的 20% 資料作為模型驗證，最後的 10% 作為模型的測試，訓練方法採用 Early Stopping[50]，當驗證的損失(即 MSE)經過 20 次驗證後不再下降時，終止模型訓練，以崎峰(3)腳本 3 為例，每 1 次 Epoch 的訓練資料及驗證資料的 MSE 值下降趨勢如圖 9，通常在 50 個 Epoch 後就會得到極小的驗證資料 MSE 值，但 Early Stopping 通常會再持續演算直到 20 次驗證後不再下降為止，以圖 9 為例，Epoch 約為 290 才停止模型訓練。最後進行模型測試，計算預測值(\hat{y}_i)，模型的優劣則以模型的預測值(\hat{y}_i)與實際值(y_i)的 MSE 作為評估模型標準。

在進行模型訓練時所有資料經過正規化(normalization)，以優化梯度下降法並提高精準度[54]，本研究之正規化函數為 Min-Max Normalization 函數(即內差法)，將地下水位轉換為 0~1 之間，定義為：

$$x = \frac{x - x_{min}}{x_{max} - x_{min}}$$

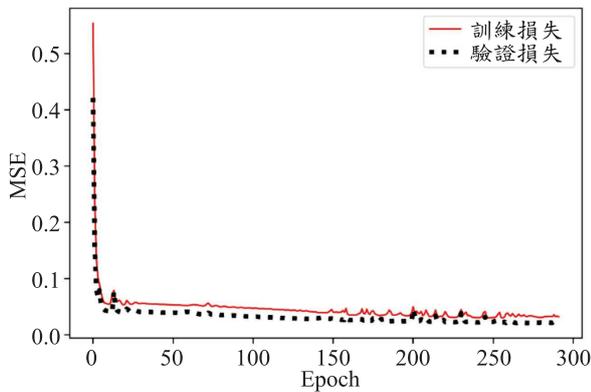


圖 9 崎峰(3)腳本 3 訓練資料及驗證資料的 MSE 值下降趨勢圖

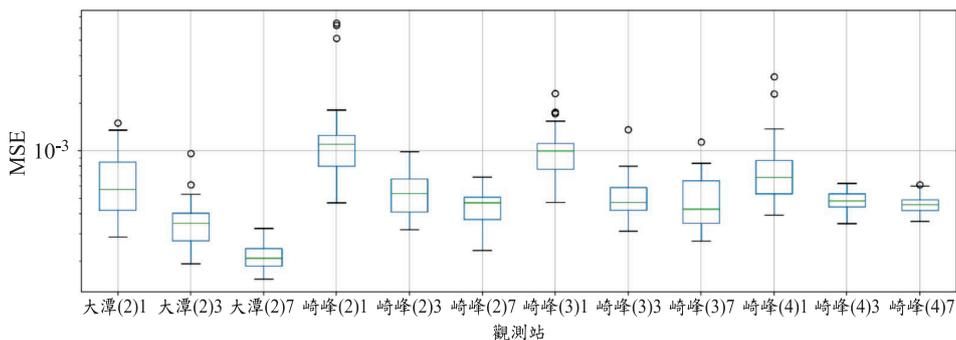


圖 10 單層 LSTM 大潭(2)、崎峰(2)、崎峰(3)及崎峰(4)MSE 箱形圖

對於測試的資料而言(y_i)包含最後 10% 資料(即第 4,609~5,120 筆資料)，以每一筆資料前 30 天資料作為模型的輸入值，計算未來 FD 天內的地下水位預測值，以 FD=7 而言，計算流程如下：

- A. 輸入第 4,579~4,608 筆資料
- B. LSTM 模型訓練，計算獲得 7 筆預測值
- C. 僅保留 B 所得的第 7 筆資料，作為第 1 次計算的預測值(\hat{y}_1)，該預測值對應的實際值為第 4,609 筆資料(y_1)
- D. 以移動視窗的方法向前移 1 天
- E. 輸入第 4,580~4,609 筆資料
- F. LSTM 模型訓練，計算獲得 7 筆預測值
- G. 僅保留 F 所得第 7 筆資料，作為第 2 次計算的預測值(\hat{y}_2)，該預測值對應的實際值為第 4,610 筆資料(y_2)
- H. 重複 D~G 步驟直到最後 1 筆資料被預測為止
- I. 計算地下水位預測值(\hat{y}_i)與地下水位實際值(y_i)的 MSE 值作為評估模型優劣標準。

3.1.2 測試資料 MSE 比較

由於每一次的 LSTM 模型訓練結果，都會得到一組與前次訓練不同的模型參數，每一組參數經過 Early Stopping 演算，再計算測試資料 MSE 值。本研究對於不同的觀測井，以同一個條件進行 30 次的模型訓練，再分析比較每一個觀測井的測試資料 MSE 值分佈，圖 10、11 及 12 分別為各觀測不同腳本測試資料 MSE 值箱形圖(大潭(2)3 代表 FD=3 大潭(2)LSTM 模型)，結果顯示 FD=7 較 FD=3 及 FD=1 更能獲得較低的測試資料 MSE 值，即代表訓練的 LSTM 模型可以更精確的預測結果，另外發現大庄(2)、枋寮(2)、德興(2)的測試資料 MSE 值較其他觀測井高(圖 12)，代表這些觀測井僅使用地下水位資料所建立的 LSTM 模型，比其他觀測井預測的準確度低，可能因為林邊溪以南的地下水比林邊溪以北的地下水受到更多不可預知的干擾，造成無法僅以其地下位作為模型建立的特徵。

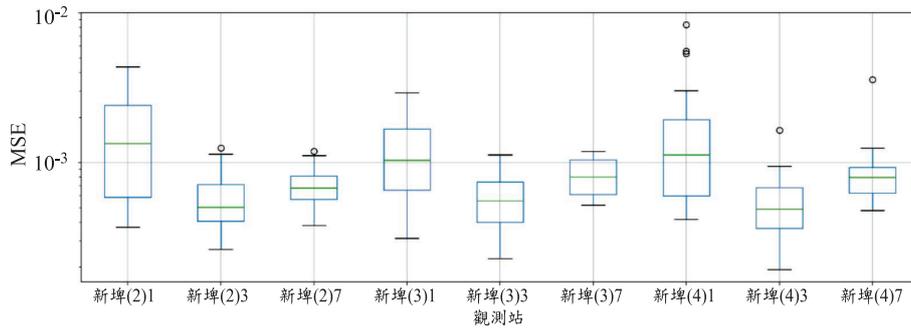


圖 11 單層 LSTM 新埤(2)、新埤(3)及新埤(4)MSE 箱形圖

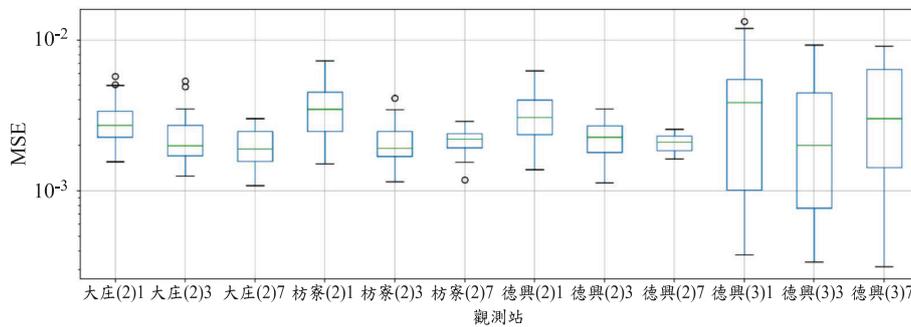


圖 12 單層 LSTM 大庄(2)、枋寮(2)、德興(2)及德興(3)MSE 箱形圖

3.2 多層 LSTM 模型

許多著名的類神經網路結構藉由加深及改變網路結構，以擷取特徵提高神經網路模型的準確性[55-58]。本文建立 3 層 LSTM，分別包含 32、64 及 32 神經元 (圖 13)，使用的訓練驗證方法與單層 LSTM 相同。因為大庄(2)、枋寮(2)、德興(2)及德興(3)在模型訓練上 MSE 值較其他觀測站高，因此以這 4 站在腳本 3 的狀況下，分別訓練 200 次單層及 3 層 LSTM 模型，並比較測試資料 MSE 值。訓練結果分別獲得單層測試資料 MSE 值及 3 層測試資料 MSE 值兩組陣列，其箱形圖如圖 14，對於大庄(2)、枋寮(2)及德興(2)而言，增加為 3 層 LSTM 沒有降低 MSE 值。以大庄(2)、枋寮(2)、德興(2)及德興(3)而言，均必須至少經過 75 次的模型訓練，才能獲得較佳的預測結果(圖 15)。

3.3 模型預測比較分析

以單層 LSTM 模型在腳本 3 的狀況下作為模型建立條件，各觀測值與預測值與實際值比較圖如圖 16。各預測值與實際值 MSE、最大預測誤差、最小預測誤差及關係係數如表 5，發現模型 MSE 表現最差的是枋寮(2)，最大預測誤差發生在崎峰(4)，相關係數(γ)表現最

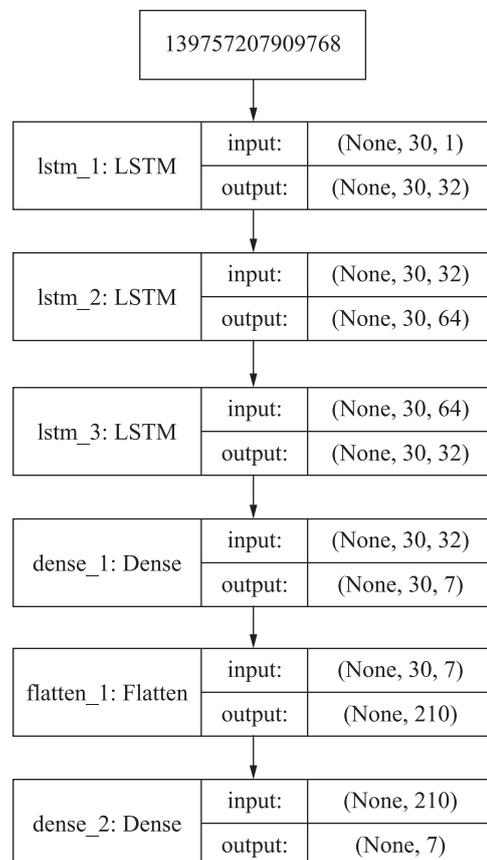


圖 13 腳本 3 · 3 層 LSTM 模型圖

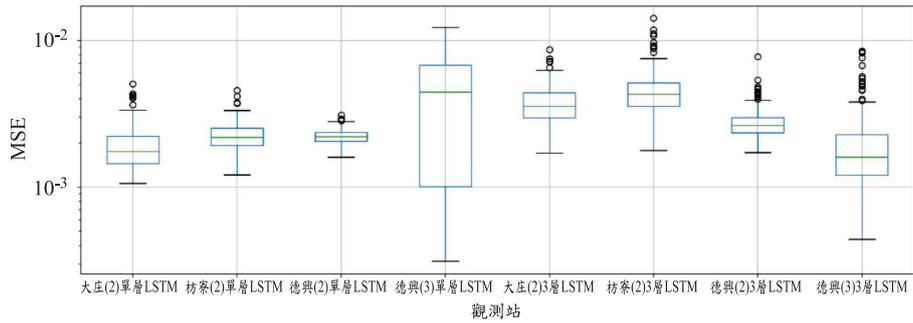


圖 14 腳本 3 不同 LSTM 結構 MSE 比較圖

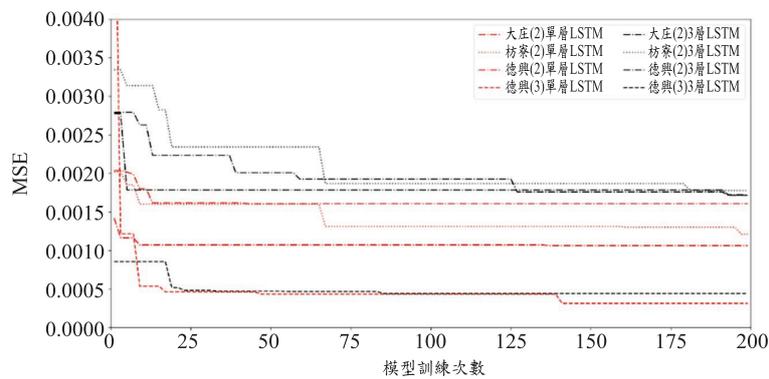


圖 15 LSTM 模型 MSE 值與訓練次數圖

表 5 預測值與實際值比較表

觀測井	MSE	最大預測誤差	最小預測誤差	γ
大潭(2)	9.3530e-05	3.4252e-02	1.5043e-05	0.9986
崎峰(2)	2.8283e-04	1.1345e-01	1.5626e-05	0.9914
崎峰(3)	3.9901e-04	1.6559e-01	4.8284e-05	0.9898
崎峰(4)	1.1253e-03	3.5174e-01	4.2038e-06	0.9804
新埤(2)	4.6347e-04	1.0226e-01	4.6983e-04	0.9985
新埤(3)	7.6876e-04	5.2791e-02	1.4632e-03	0.9996
新埤(4)	2.1770e-04	3.4980e-02	7.1694e-05	0.9999
大庄(2)	1.7413e-03	1.6727e-01	4.3114e-04	0.9962
枋寮(2)	3.0308e-03	3.4235e-01	2.2760e-04	0.9887
德興(2)	1.7788e-03	1.6789e-01	3.1833e-04	0.9806
德興(3)	8.9754e-04	6.4471e-02	5.5321e-04	0.9992

差的也是崎峰(4)，表現最好的是大潭(2)及新埤(4)，從圖 16 可以了解枋寮(2)在 2018 年 12 月到 2019 年 1 月的實際值都沒有變化，因為在資料清洗時發現枋寮(2)在這個階段均沒有資料紀錄(圖 2(e))或缺值，因此以枋寮(2)地下水位眾數取代，造成模擬的結果較差；而崎峰(4)的資料紀錄在 2018 年 10 月有驟降驟升的現象，即可能是資料誤植造成模型在預測驟降驟升的現象產生較大的誤差；對於大潭(2)及新埤(4)而言，則比較沒有資料缺值或驟降驟升的情形，因此可以獲得最佳的

模擬預測成果，整體而言，模型預測值與實際值的相關係數均在 0.98 以上。

四、結論與建議

短時間內的超抽地下水，造成地下水下降為地層下陷發生的主要原因，依據近幾年地層下陷的觀測，屏東沿海林邊溪南北兩岸仍持續發生下陷的狀況，本文

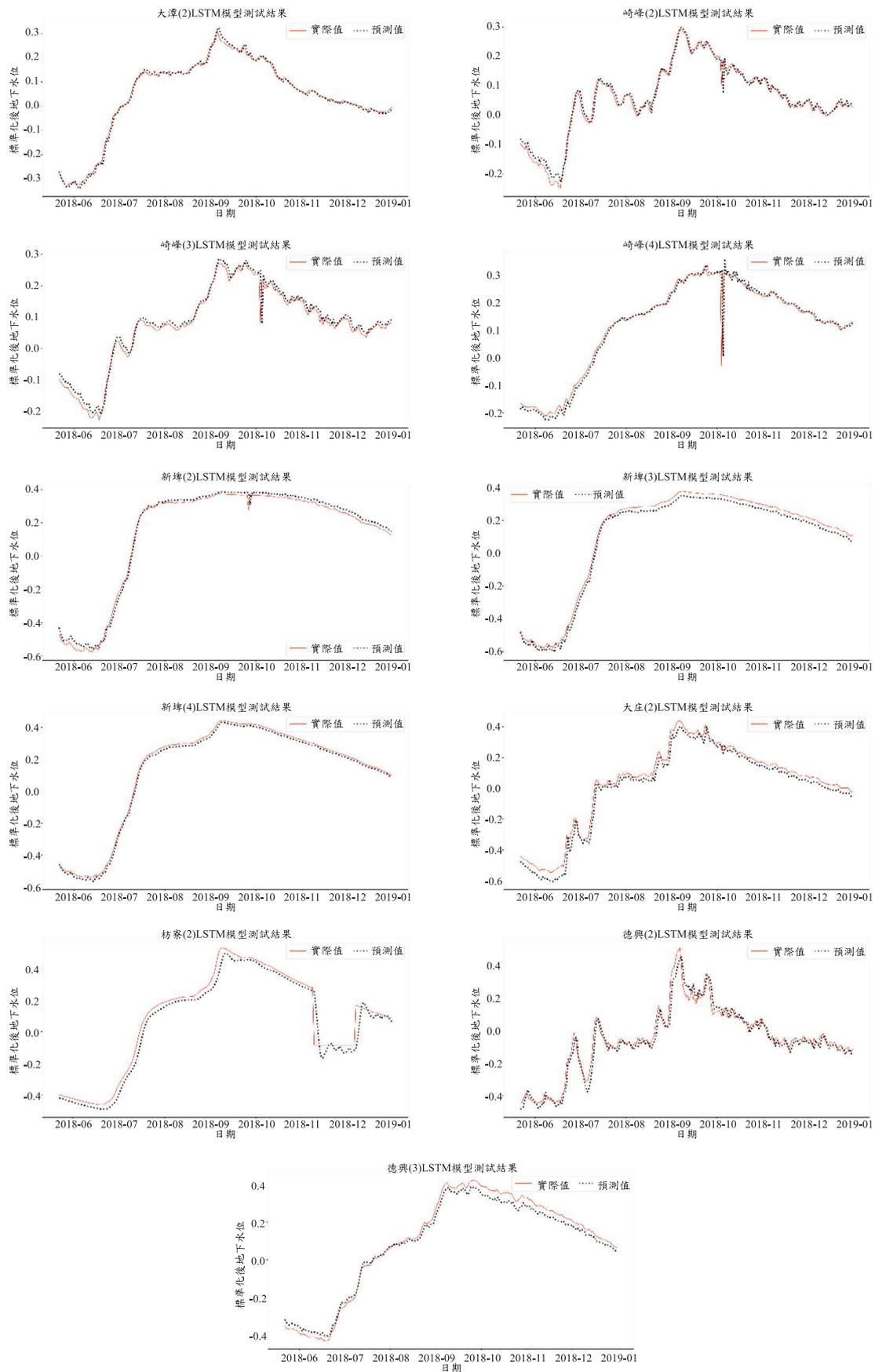


圖 16 各觀測井 LSTM 模擬預測值與實際值比較圖

純以資料驅動的角度，以屏東平原地層下陷區地下水第 2、3 及 4 層的地下水位為研究對象，從經濟部水利署蒐集大潭(2)、崎峰(2)、崎峰(3)、崎峰(4)、新埤(2)、新埤(3)、新埤(4)、大庄(2)、枋寮(2)、德興(2)、德興(3)等地下水觀測井日地下水位資料，每一口觀測井取其最近的 5,120 筆資料，建立類神經網路-LSTM 模型，在模型的訓練上，為避免過度擬合，採用 Early Stopping 演算法作為獲得模型最佳參數，本文建立單層 LSTM 模型及 3 層 LSTM 模型，以過去 30 天的地下水位資料，分別針對預測未來 1 天(腳本 1)、3 天(腳本 2)及 7 天(腳本 3)作為模型訓練腳本，從模型訓練的結果歸納以下結論：

1. 不同腳本所訓練的 LSTM 模型對於測試資料 MSE 值發現，腳本 3 較腳本 1 及腳本 2 更能精確預測結果。
2. 對於本研究模型測試而言，3 層 LSTM 模型沒有降低測試資料 MSE 值的現象，多層的 LSTM 結構尚須再經過更多的模型建構測試。
3. 研究發現大庄(2)、枋寮(2)、德興(2)的測試資料 MSE 值較其他觀測井高，代表這些觀測井僅使用地下水位資料所建立的 LSTM 模型，比其他觀測井預測的準確度低，可能因為林邊溪以南的地下水比林邊溪以北的地下水受到更多不可預知的干擾，造成無法僅以其地下水位作為模型建立的特徵。因此大庄(2)、枋寮(2)、德興(2)需要再加入影響地下水位的其他特徵，才能在 LSTM 模型建立上降低其測試資料 MSE 值，增加模型的精確度。
4. 對於大潭(2)、崎峰(2)、崎峰(3)、崎峰(4)、新埤(2)、新埤(3)、新埤(4)及德興(3)而言，無論使用單層 LSTM 模型或 3 層 LSTM 模型，30 次 LSTM 模型訓練即可訓練出 LSTM 模型讓測試資料 MSE 值在 0.001 以下，增加訓練次數也可以獲得更精確的模型，如枋寮(2)，經過約 70 次的訓練後，其測試資料 MSE 值有明顯下降，但仍未小於 0.001 以下。
5. 以測試資料 MSE 值=0.001 而言，代表均方差 (RMSE) \cong 0.032，當實際的最高地下水位與最低地下水位相差 10 公尺時，預測的地下水位均方差 (RMSE)約 32 公分，除大潭(2)及崎峰(2)、崎峰(3)及崎峰(4)外，其他深層地下水觀測井高低水位均超過 10 公尺以上，若要降低均方差到 10 公分以內(即正規化後 MSE=0.0001)，未來研究可以增加其他特徵如雨量、河川水位等，以獲取更精確的地下水位 LSTM 模型作為預測能力。
6. 以單層 LSTM 模型在腳本 3 的狀況下作為模型建

立條件，LSTM 模型 MSE 表現最差的是枋寮(2)，最大預測誤差發生在崎峰(4)，相關係數(γ)表現最差的也是崎峰(4)，表現最好的是大潭(2)及新埤(4)。整體而言，模型預測值與實際值的相關係數均在 0.98 以上。

7. 觀測井地下水位記錄沒有資料紀錄或缺值，或資料發生驟降驟升的現象，將會對 LSTM 模型的預測能力產生極大的影響。

參考文獻

1. Lo, K.Y., *et al.*, "Integrated assessment for meteorological and hydrological drought of southern taiwan," *Journal of Taiwan Agricultural Engineering*, 64(2), pp. 12-29, 2018.
2. Cheng, F.-Y. and Y. Chen, "Variations in soil moisture and their impact on land-air interactions during a 6-month drought period in Taiwan," *Geoscience Letters*, 5(1), pp. 26, 2018.
3. Gebreyohannes, T., *et al.*, "Application of a spatially distributed water balance model for assessing surface water and groundwater resources in the Geba basin, Tigray, Ethiopia," *Journal of Hydrology (Amsterdam)*, 499, pp. 110-123, 2013.
4. Shevah, Y., "Water scarcity, water reuse, and environmental safety," *Pure and Applied Chemistry*, 86(7), pp. 1205-1214, 2014.
5. Wang, H.W., *et al.*, "Assessment of Land Subsidence and Climate Change Impacts on Inundation Hazard in Southwestern Taiwan," *Irrigation and Drainage*, 67, pp. 26-37, 2018.
6. Galloway, D.L. and T.J. Burbey, "Regional land subsidence accompanying groundwater extraction," *Hydrogeology Journal*, 19(8), pp. 1459-1486, 2011.
7. Erban, L.E., *et al.*, "Release of arsenic to deep groundwater in the Mekong Delta, Vietnam, linked to pumping-induced land subsidence," *Proceedings of the National Academy of Sciences*, 110(34), pp. 13751-13756, 2013.
8. Gong, H., *et al.*, "Long-term groundwater storage changes and land subsidence development in the North China Plain (1971-2015)," *Hydrogeology Journal*, 26(5), pp. 1417-1427, 2018.
9. Xie, Y., *et al.*, "Uncertainty of groundwater recharge estimated from a water and energy balance model,"

- Journal of hydrology*, 561, pp. 1081-1093, 2018.
10. Rajabi, M.M. and H. Ketabchi, "Uncertainty-based simulation-optimization using Gaussian process emulation: application to coastal groundwater management," *Journal of hydrology*, 555, pp. 518-534, 2017.
 11. Clark, M.P., *et al.*, "A unified approach for process-based hydrologic modeling: 1. Modeling concept," *Water Resources Research*, 51(4), pp. 2498-2514, 2015.
 12. Chang, F.-J., *et al.*, "Prediction of monthly regional groundwater levels through hybrid soft-computing techniques," *Journal of Hydrology (Amsterdam)*, 541(Part B), pp. 965-976, 2016.
 13. Mohanty, S., *et al.*, "Using artificial neural network approach for simultaneous forecasting of weekly groundwater levels at multiple sites," *Water Resources Management*, 29(15), pp. 5521-5532, 2015.
 14. Gholami, V., *et al.*, "Modeling of groundwater level fluctuations using dendrochronology in alluvial aquifers," *Journal of hydrology*, 529, pp. 1060-1069, 2015.
 15. Naghibi, S.A., H.R. Pourghasemi, and B. Dixon, "GIS-based groundwater potential mapping using boosted regression tree, classification and regression tree, and random forest machine learning models in Iran," *Environmental monitoring and assessment*, 188(1), pp. 44, 2016.
 16. Gong, Y., *et al.*, "A comparative study of artificial neural networks, support vector machines and adaptive neuro fuzzy inference system for forecasting groundwater levels near Lake Okeechobee, Florida," *Water resources management*, 30(1), pp. 375-391, 2016.
 17. Khalil, B., *et al.*, "Short-term forecasting of groundwater levels under conditions of mine-tailings recharge using wavelet ensemble neural network models," *Hydrogeology journal*, 23(1), pp. 121-141, 2015.
 18. Sun, Y., *et al.*, "Application of artificial neural networks in groundwater table forecasting—a case study in a Singapore swamp forest," *Hydrology and Earth System Sciences*, 20(4), pp. 1405-1412, 2016.
 19. Sahoo, S., *et al.*, "Machine learning algorithms for modeling groundwater level changes in agricultural regions of the US," *Water Resources Research*, 53(5), pp. 3878-3895, 2017.
 20. Hochreiter, S. and J. Schmidhuber, "Long short-term memory," *Neural computation*, 9(8), pp. 1735-1780, 1997.
 21. Mikolov, T., *et al.* Recurrent neural network based language model. in Eleventh annual conference of the international speech communication association. 2010.
 22. Cho, K., *et al.*, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," *arXiv preprint arXiv:1406.1078*, 2014.
 23. Graves, A., A.-r. Mohamed, and G. Hinton. Speech recognition with deep recurrent neural networks. in 2013 IEEE international conference on acoustics, speech and signal processing. 2013. IEEE.
 24. Miao, Y., M. Gowayyed, and F. Metze. EESN: End-to-end speech recognition using deep RNN models and WFST-based decoding. in 2015 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU). 2015. IEEE.
 25. Selvin, S., *et al.* Stock price prediction using LSTM, RNN and CNN-sliding window model. in 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI). 2017. IEEE.
 26. Rather, A.M., A. Agarwal, and V. Sastry, "Recurrent neural network and a hybrid model for prediction of stock returns," *Expert Systems with Applications*, 42(6), pp. 3234-3241, 2015.
 27. Xingjian, S., *et al.* Convolutional LSTM network: A machine learning approach for precipitation nowcasting. in Advances in neural information processing systems. 2015.
 28. Le, J., *et al.*, "Application of recurrent neural networks for drought projections in California," *Atmospheric research*, 188, pp. 100-106, 2017.
 29. Bowes, B.D., *et al.*, "Forecasting Groundwater Table in a Flood Prone Coastal City with Long Short-term Memory and Recurrent Neural Networks," *Water*, 11(5), pp. 1098, 2019.
 30. Jang, H., A. Park, and K. Jung. Neural network implementation using cuda and openmp. in 2008 Digital Image Computing: Techniques and Applications. 2008. IEEE.
 31. Oh, K.-S. and K. Jung, "GPU implementation of neural networks," *Pattern Recognition*, 37(6), pp. 1311-1314, 2004.
 32. Ronao, C.A. and S.-B. Cho, "Human activity recognition with smartphone sensors using deep learning neural networks," *Expert systems with applications*, 59, pp. 235-244, 2016.
 33. Zhang, J., *et al.*, "Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas," *Journal of hydrology*, 561, pp. 918-

- 929, 2018.
34. Huang, C.-J. and P.-H. Kuo, "A deep cnn-lstm model for particulate matter (PM_{2.5}) forecasting in smart cities," *Sensors*, 18(7), pp. 2220, 2018.
35. Gibrilla, A., G. Anornu, and D. Adomako, "Trend analysis and ARIMA modelling of recent groundwater levels in the White Volta River basin of Ghana," *Groundwater for Sustainable Development*, 6, pp. 150-163, 2018.
36. de Moraes Takafuji, E.H., M.M. da Rocha, and R.L. Manzione, "Groundwater level prediction/forecasting and assessment of uncertainty using SGS and ARIMA Models: A case study in the Bauru Aquifer System (Brazil)," *Natural Resources Research*, 28(2), pp. 487-503, 2019.
37. Hrnjica, B. and O. Bonacci, "Lake Level Prediction using Feed Forward and Recurrent Neural Networks," *Water Resources Management*, 33(7), pp. 2471-2484, 2019.
38. Siami-Namini, S. and A.S. Namin, "Forecasting economics and financial time series: Arima vs. lstm," *arXiv preprint arXiv:1803.06386*, 2018.
39. Wang, B., *et al.* The forecast on the customers of the member point platform built on the blockchain technology by ARIMA and LSTM. in 2018 IEEE 3rd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA). 2018. IEEE.
40. Rajaei, T., H. Ebrahimi, and V. Nourani, "A review of the artificial intelligence methods in groundwater level modeling," *Journal of hydrology*, 2019.
41. 經濟部地質調查所，地下水補注地質敏感區劃定計畫書：G0002 屏東平原。2004，經濟部。
42. 楊萬全，"高屏溪流域和屏東平原的水資源，"國立臺灣師範大學地理研究報告，27, pp. 81-104, 1997。
43. 經濟部，地下水補注地質敏感區劃定計畫書-G0002 屏東平原 2014，經濟部。
44. Lee, G., S. Jung, and D. Lee, "Comparison of physics-based and data-driven models for streamflow simulation of the Mekong river," *Journal of Korea Water Resources Association*, 51(6), pp. 503-514, 2018.
45. Zhang, X., *et al.*, "A novel hybrid data-driven model for daily land surface temperature forecasting using long short-term memory neural network based on ensemble empirical mode decomposition," *International journal of environmental research and public health*, 15(5), pp. 1032, 2018.
46. 經濟部水利署 and 國立成功大學地層下陷防治服務團。地層下陷防治服務網：屏東 83 至 106 年下陷分析。2019。
47. Elman, J.L., "Finding structure in time," *Cognitive science*, 14(2), pp. 179-211, 1990.
48. Widrow, B. and M.A. Lehr, "30 years of adaptive neural networks: perceptron, madaline, and backpropagation," *Proceedings of the IEEE*, 78(9), pp. 1415-1442, 1990.
49. Rumelhart, D.E., G.E. Hinton, and R.J. Williams, "Learning representations by back-propagating errors," *Cognitive modeling*, 5(3), pp. 1, 1988.
50. Caruana, R., S. Lawrence, and C.L. Giles. Overfitting in neural nets: Backpropagation, conjugate gradient, and early stopping. in *Advances in neural information processing systems*. 2001.
51. 陳國彥，"屏東縣林邊佳冬枋寮地區地層下陷與海水倒灌之初步探討，"國立臺灣師範大學地理研究報告，8, pp. 179-194, 1982。
52. 康家桂 and 陳春盛，以時間序列預測麥寮地區地層下陷之研究，in 土木工程學系。2006，國立交通大學。
53. 洪偉嘉，黃大任，and 張磊，"時域相關點雷達干涉技術應用於雲林嚴重地層下陷區監測，"航測及遙測學刊，19(3), pp. 225-237, 2015。
54. Jayalakshmi, T. and A. Santhakumaran, "Statistical normalization and back propagation for classification," *International Journal of Computer Theory and Engineering*, 3(1), pp. 1793-8201, 2011.
55. Tur, G., *et al.* Towards deeper understanding: Deep convex networks for semantic utterance classification. in 2012 IEEE international conference on acoustics, speech and signal processing (ICASSP). 2012. IEEE.
56. Mollahosseini, A., D. Chan, and M.H. Mahoor. Going deeper in facial expression recognition using deep neural networks. in 2016 IEEE Winter conference on applications of computer vision (WACV). 2016. IEEE.
57. Szegedy, C., *et al.* Inception-v4, inception-resnet and the impact of residual connections on learning. in Thirty-First AAAI Conference on Artificial Intelligence. 2017.
58. Wu, Z., C. Shen, and A. Van Den Hengel, "Wider or deeper: Revisiting the resnet model for visual recognition," *Pattern Recognition*, 90, pp. 119-133, 2019.

收稿日期：民國 108 年 08 月 19 日
修改日期：民國 108 年 10 月 14 日
接受日期：民國 109 年 01 月 15 日