

# 結合氣象資料應用在日射量預測模型建立

徐國昌<sup>1</sup> 詹麒璋<sup>1</sup> 鄭旭詠<sup>2</sup> 曾美惠<sup>1</sup> 盧明德<sup>1</sup> 陳博宥<sup>1</sup>  
工業技術研究院綠能與環境研究<sup>1</sup>  
中央大學資訊工程學系<sup>2</sup>

## 摘要

由於再生能源與智慧電網的應用漸漸興起，其中太陽能為最受矚目也是大家最深入研究的領域，其發展價值受到重視。為了評估太陽光電系統的性能，會使用和太陽光電模組裝設之傾斜角度與方位角相同位置的日射計當做參考基準。在短期日射量預測研究中，我們所使用的訓練資料只是單純參考了過去時間點的日射量量值，但在現實生活中日射量量值的多寡是受到許多外在環境因子所影響的，這些因子包含溫度、濕度、雲量、風速和季節等等因素。所以一個良好的預測模型不該只有使用日射量量值當作訓練資料，而需包含了所有影響因子的特徵資料才會有較佳的成果。因此在本論文中我們加入了一些天氣的資料來做長期的日射量預測實驗，以1天後的日射總量為例。

關鍵字：太陽光電、智慧電網、日射量預測

## 一、前言

透過日射量變化之預測與推估系統，即能幫助電廠在太陽日射量不穩定時，做好備用能源之啟動與配送，並降低轉換次數，進而降低成本及減少能源耗費[1,2,3]。為了評估太陽光電系統的性能[4,5,6]，會使用和太陽光電模組裝設之傾斜角度與方位角相同位置的日射計當做參考基準。在短中期的日射量預測，我們所使用的訓練資料單純參考了過去時間點的日射量量值，但在現實生活中日射量量值的多寡是受到許多外在環境因子所影響的，這些因子包含溫度、濕度、雲量、風速和季節等等因素。所以一個良好的預測模型不該只有使用日射量量值當作訓練資料，而需包含了所有影響因子的特徵資料才會有較佳的成果。因此在本論文中我們加入了一些天氣的資料來做長期的日射量預測實驗，以1天後的日射總量為例並加以實驗驗證。

## 二、研究方法

氣象特徵如此之多，該選用那些特徵當作特徵值的輸入？加入氣象資訊後的模型該怎麼選擇，是一個重要的課題。在本論文中的實驗主要參考了[7] The potential of different artificial neural network (ANN) techniques in daily global solar radiation modeling based on meteorological data. 所做的實驗架構和方法，並選用了不同的天氣特徵組合來做測試。

## (一) ANN預測模型

類神經網路(artificial neural network, 縮寫ANN)是模仿生物神經網路所發展出來的演算法，而「多層感知機」是類神經網路的一種，基本的多層感知機包含輸入層、(多層)隱藏層、輸出層，而隱藏層和輸出層有多個神經元，多層感知機包含以下特性:(1)每個神經元的輸出端，都有一個非線性可微分的活化函數。(2)包含一層或一層以上的隱藏層。(3)網路有高度連結性，每個神經元的輸出，都會影響下一層與之連接的神經元。

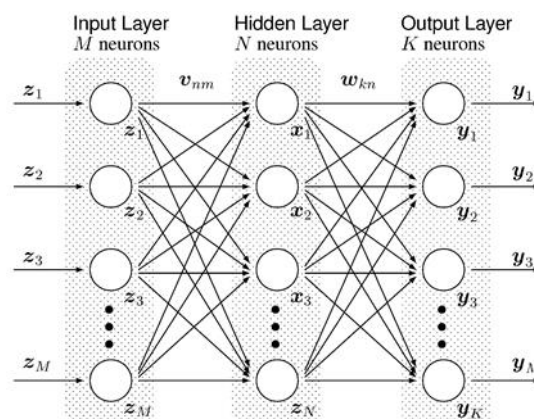


圖 1. 多層感知機示意圖

由上圖多層感知機示意圖可見，我們可以決定隱藏層的層數與每層的神經元個數 $N$ 與 $K$ ，將鍵結值與相對應的輸入相乘後做累加，之後再經過活化函數 $\sigma(x)$ ，得到一個output值，定義 $z$ 為輸入向量。

多層感知機的隱藏層和輸出層都含有多個神經元，其學習方式與單一神經元類似，將計算輸出與修正鍵結值兩步驟，分別稱為前饋階段(Feedforward)與倒傳遞階段(Backforward)，二個階段皆以層為單位，前饋階段從輸入層開始計算每一層的輸出作為下一層的輸入，直到獲得輸出層的輸出，倒傳遞階段則相反，由輸出層開始計算差值，並將往回傳遞至輸入層為止，每一次的傳遞，都會修正當下層的鍵結值。

## (二) 倒傳遞演算法

本論文使用倒傳遞演算法步驟如下：

### 步驟一：決定網路架構

決定幾層的架構以及每一層的類神經元數目是多少，此研究使用三層隱藏層，每層神經元數量可六類神經元。

### 步驟二：初始化

隨機的方法來將初始鍵結值設定為很小的實數，並將學習次數設定為 0。

### 步驟三：前饋階段

在第 n 次學習循環時，將輸入向量送入網路，輸入向量由輸入層引入，計算位於第一個隱藏層的類神經元在第 n 次學習循環的輸出，如方程式(2)、(3)所示：

$$v_j(n) = \sum_{i=0}^P w_{ji}(n) y_i(n) \quad (2)$$

$$y_j(n) = \phi_j(v_j(n)) \quad (3)$$

以前饋方式經由隱藏層，一層一層傳至輸出層，計算出每一層的類神經元輸出，便算出網路最後輸出值。

### 步驟四：倒傳遞階段

計算位於輸出層的類神經元第 n 次學習時的區域梯度函數，如方程式(4)所示：

$$\delta_j(n) = e_j(n) \phi'(v_j(n)) = (d_j(n) - O_j(n)) O_j(n) (1 - O_j(n)) \quad (4)$$

以倒傳遞方式經由倒數第一個隱藏層，一層一層地傳導至第一個隱藏層，計算出每一層隱藏層的類神經元區域梯度函數，如方程式(5)所示：

$$\delta_j(n) = \phi'(v_j(n)) \sum_k \delta_k(n) w_{kj}(n) = y_j(n) (1 - y_j(n)) \sum_k \delta_k(n) w_{kj}(n) \quad (5)$$

### 步驟五：調整鍵結值向量

調整每一個類神經元的鍵結值向量，修正方程式，如方程式(6)所示：

$$w_{ji}(n+1) = w_{ji}(n) + \Delta w_{ji}(n) = w_{ji}(n) + \eta \delta_j(n) y_i(n) \quad (6)$$

### 步驟六：收斂條件測試

將學習循環 n 加 1，回到步驟三，直到符合收斂條件或收斂條件或疊代次數超過某一設定值則停止。

其中，

n: 第 n 次疊帶

$v_j(n)$ : 第 j 個類神經元的內部激發狀況

$y_j(n)$ : 上一層第 j 個神經元輸出

$w_{ji}(n)$ : 第 i 個類神經元連結至第 j 個類神經元的鍵結值

d: 第 j 個類神經元的期望輸出

$O_j$ : 輸出層第 j 個神經元的輸出

$\delta$ : 學習率

$\phi$ : 活化函數

依照上述步驟，我們可完成 ANN 日射模型的訓練。

## (三) 資料蒐集

實驗部分我們原本是希望能利用天氣預報資料來做訓練，但是過去的天氣預報資料已經無法取的，所以替代方案我們使用了歷史的天氣觀測資料來做訓練，資料來源我們使用中央氣象局所提供的開放天氣觀測資料查詢系統(CODIS, CWB Observation Data Inquire System)來做蒐集，下圖為 CODIS 首頁。

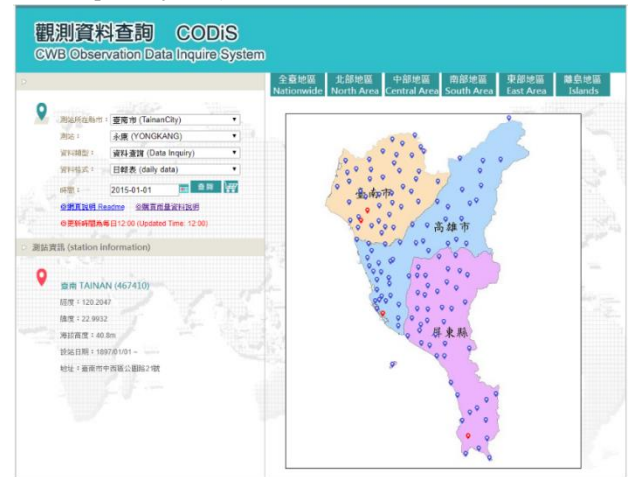


圖 2. CODIS 首頁

使用 python 爬蟲程式來爬取 2015 年 3 個離原本地區最近的觀測站的逐時觀測資料，分別是大內觀測站(大內)、永康觀測站(五王國小)和玉井觀測站(玉井)，中括弧內為工研院所提供的日射量資料的地區名稱，下圖為查詢永康區 2015/01/01 結果範例示意圖。

站名	時間	Temp	Humidity	Wind	Cloud	WDW3	WDW2	WDW1	WDW0	WDW3	WDW2	WDW1	WDW0	WDW3	WDW2	WDW1	WDW0
01	1501-0	18.2	88.0	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
02	1501-1	18.5	88.5	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
03	1501-2	18.8	89.0	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
04	1501-3	19.1	89.5	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
05	1501-4	19.4	90.0	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
06	1501-5	19.7	90.5	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
07	1501-6	20.0	91.0	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
08	1501-7	20.3	91.5	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
09	1501-8	20.6	92.0	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
10	1501-9	20.9	92.5	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
11	1501-10	21.2	93.0	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

圖 3. CODIS 查詢結果範例

可以看到永康觀測站所觀測到的欄位較為豐富，是因為該觀測站是有人工觀測的，而大部分觀測站是自動觀測站，所觀測的資訊欄位並沒有這麼豐富，

所以我們採用欄位的最大集合當作特徵欄位，產生了四個欄位分別是：最高溫、最低溫、平均溫度和相對溼度，分別以這些現有的氣象觀測資料來做實驗找取效果最好的特徵組合。

#### (四) 評估方法

欲知整體預測模型評估好壞，本計畫採取「平均絕對百分比差」(Mean Absolute Percentage Error, 簡稱 MAPE) 為評估方法。方程式(1)、(2)為其運算公式：

$$MAPE_{fc} = \frac{\sum_{k=1}^N |y_k - \hat{y}_k|}{\sum_{k=1}^N |y_k|} \quad (1)$$

$$y_k = \frac{1}{m} \sum_{i=k}^{k+m} Z_i \quad (2)$$

其中，「 $y_k$ 」為平均實際量測日射量；「 $\hat{y}_k$ 」為平均預測日射量；「 $N$ 」為資料集合(DataSet)的大小。而方程式(8)是平均實際量測日射量的計算方式，「 $z_i$ 」為預測時間點的實際量測日射量，「 $m$ 」是要將幾分鐘的資料合併平均作為平均實際量測日射量，此參數是使用者可以自行調整的。最後我們算出來的 $MAPE_{fc}$ ，也就是誤差值總和占日射量總和的多少個百分點， $MAPE_{fc}$ 值越低，代表我們的預測結果越準確，本文中提及的 $MAPE$ 皆為 $MAPE_{fc}$ 的簡稱。

### 三、研究結果

實驗流程如圖 3 多層感知機實驗流程所示，使用多層感知機訓練分別採用一層隱藏層與兩層隱藏層做實驗，其中的神經元都為 3 個，輸入的 Input 值是明天天氣觀測與預測當天天氣觀測的差值，而 Output 是預測當天相對明天日射總量的差值。實驗流程是先將原始資料進行前處理後，進行訓練、預測、評估等處理程序。

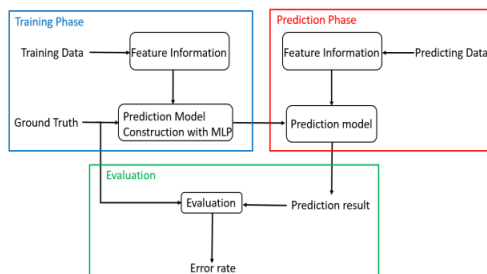


圖 3. 多層感知機實驗流程

實驗主要針對不同的天氣特徵組合來做測試，組合包含：

1. HH (High Temperature, Humidity)

2. AH (Average Temperature, Humidity)
3. HLH (High Temperature, Lower Temperature, Humidity)
4. HAH (High Temperature, Average Temperature, Humidity)
5. 對於上述組合做雜訊過濾 HH\_S, AH\_S, HLH\_S, HAH\_S

並且對於多層感知機(MLP)使用一層隱藏層和兩層隱藏層做實驗。實驗結果第一部份會使用一層隱藏層的多層感知機。第二部份會使用兩層隱藏層的多層感知機，在以上部份皆使用 Fixed model 即使用一、二、三月的資料來做訓練四到十二月的資料當作測試資料第三部份是藉由第一和二部份找到的最好天氣特徵組合來做 Sliding model 即使用要預測的月份前三個月的資料來做訓練。對於圖表中 BaseLine 的方法我們是使用前一天的日射總量當作明天的日射總量預測值，可以想成都不做任何預測就把前一天的值當作預測值，要達到至少能好於 BaseLine 方法。實驗說明如下：

第一部份：

下圖為使用一層隱藏層去訓練不同天氣特徵組合的直線圖，縱軸為 MAPE 橫軸為使用的特徵組合，從圖中可以看到大部分的訓練組合對於預測都是有幫助的，在圖中可以看出使用 High Temperature 和 Humidity 的特徵效果最佳，與之前資料視覺化時的分析滿吻合的，High Temperature 與日射量呈現著很明顯的正關係，而使用 Low Temperature 或不採用 High Temperature 的組合效果並不是這麼好。

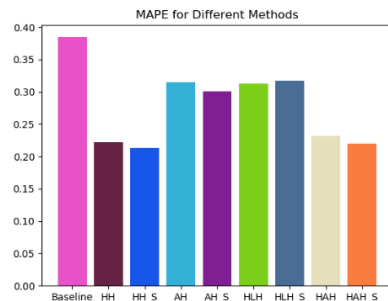


圖 4. 一層隱藏層配上不同特徵結果圖

第二部份：

下圖為使用一層隱藏層去訓練不同天氣特徵組合的直線圖，縱軸為 MAPE 橫軸為使用的特徵組合，在 HAH 部份可以看到有明顯降低 MAPE，HH 部份可能因為兩層導致一些過擬合而上升些許，不過基本上其實都是差不多的，所以我們在最後的 Sliding Model 還是暫時先採用一層隱藏層和 HH 的天氣特徵做實驗，採用一層隱藏層是因為效果與兩層隱藏層差不多，所以選一層隱藏層來防止過擬合。

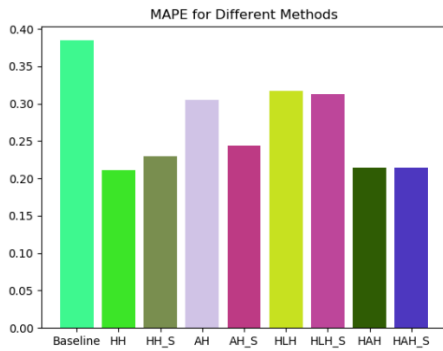


圖 5. 兩層隱藏層配上不同特徵結果圖

第三部份：

下圖分別採用前面所描述的 Sliding Model 和 Fixed Model 做測試，Sliding Model 對於要預測的月份使用最近的歷史資料來做訓練可以增進一些精準度。圖 6 為使用 Sliding Model 和 Fixed Model 的 4~12 月的平均 MAPE，而圖 7 為使用 Sliding Model 和 Fixed Model 的 4~12 月的每月 MAPE 數據。

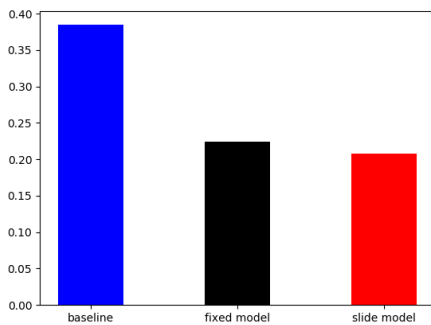


圖 6. Fixed Model 平均 VS Sliding Model 平均

最後比較一下使用 Sliding model 驗證預測 4~12 月的資料，參考圖 6 和圖 7。原本圖 6 區域總和的 MAPE 平均為 0.2368，在結合了天氣資訊後的 MAPE 來到 0.2076，假如去細看每個月的話，加入天氣資訊的 Model 是有效降低每個月的預測誤差。

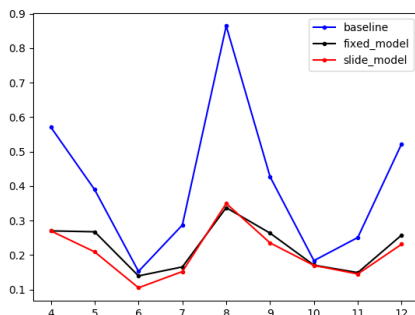


圖 7. 四月到十二月使用 Fixed Model 和 Sliding Model 預測一天後日射總量

## 四、結論與建議

本計畫實驗與其他相關研究所使用的資料庫與觀測位置皆不同，故直接比較結果並不恰當。而討論相關研究的準確率，仍可以了解現存文獻中日射量預測之成效。

目前資料使用一年當中的 3 個月當作訓練資料，其餘的月份當作驗證資料，所以模型對全部的天氣狀況可能沒有一個很完整的適應性，但對於此次研究，相比於前期沒有使用天氣資訊的模型 MAPE 為 0.2386，我們的加入天氣的 Sliding model 絕對百分比誤差(MAPE)可以降低為 0.2076，換句話說一天後的準確度接近 80%，所以加入天氣資訊的預測模型，對於一天後的預測是有幫助的，且對於天氣較穩定的月份，預測的 MAPE 更為穩定許多。

未來發展方面，我們希望能夠繼續降低 MAPE，使用更長時間的資料來做訓練，並加入太陽日照時數，蒸發量等數據，或許對模型也會有更大的幫助。

## 致謝

本文感謝經濟部能源局計畫支持，謹此致謝。  
計畫名稱：氣象資訊在綠能開發之應用服務計畫 (3/4)，計畫編號：108-D0602。

## 五、參考文獻

- [1] GK.Singh.Solar power generation by Photovoltaic technology:a review.Energy 2013;53:1-13.
- [2] H. Lund. Renewable energy strategies for sustainable development. Energy 2007; 32: 912-919.
- [3] 徐國昌、黃朝揚、曾美惠、詹麒璋、鄭旭詠，短時間性的日射量預測模型架構，「天氣分析與預報研討會」，(106)，pp.A6-7，20170912。
- [4] 徐國昌、林偉民、曾美惠、黃柏鈞、黃茂庭，台灣併網型 PV 發電系統之發電量趨勢與探討「電力電子研討會」,(第十二屆),pp.1059-1063,20131102。
- [5] IEC 61724:1998, Photovoltaic System Performance Monitoring - Guidelines for Measurement, Data Exchange, and Analysis.
- [6] 黃茂庭、徐國昌、黃秋杰、黃柏鈞，併聯型太陽光電系統性能與可靠度分析，「International Photovoltaic Science and Engineering Conference(PVSEC-23,Taiwan)」,(PVSEC-23),pp.6-P-28,20131031.
- [7] M.A. Behrang, E.Assareh, A.R.Noghrehabadi. The potential of different artificial neural network (ANN) techniques in daily global solar radiation modeling based on meteorological data., Solar Energy, Volume 84, Issue 8, August 2010, Pages 1468-1480