

國立臺灣科技大學
工業管理系
111 學年度實務專題報告

專題名稱：基於集成學習方法之電力輔助市場
調頻備轉交易量預測建模

專題編號：TR-110-01-104

指導老師：林希偉教授

研究組員：B10701008 洪 瑋
B10801007 簡加雯
B10801019 黃湘宜
B10801020 李 諾

中 華 民 國 1 1 1 年 6 月 1 3 日

摘要

隨著碳排放減量政策提出、電動車興起，以及公投案中能源的議題，顯示能源政策與電力市場越來越受重視。而近年發生的（分區）停電事件等，亦顯示台灣電力公司（以下簡稱台電）供電不穩可能對民生或產業造成極大影響。分散式能源概念下的微電網（micro-grid）因能夠有效經由輸配設備匯集並共用區域性的電力資源，因而被納入台電電力交易平台的電力輔助市場機制，以整合民間與國營的電力資源。過去微電網的研究主要是針對電力技術的分配與調節，然而，市場機制的管理效率亦可能嚴重影響輔助市場中調頻備轉供需的不確定性，造成供電可靠性下降。因此，本研究計畫採用台電公開交易平台上的交易資料，結合政府公開資料中的氣候資料，透過 SARIMAX 與倒傳遞網路（Back-Propagation Network, BPN）的混合模型，進行輔助交易市場中調頻備轉交易量的預測。研究結果指出集成模型可順利將準確率提升至 69.44%，顯示結合線性與非線性模型在優化預測效果的表現。若將預測標的僅聚焦於供電狀況存在與不存在缺口的高與低需求，其準確率可提升至 72.98%，表示對於市場是否存在供需缺口，本專題的模型能夠提供良好預測。本研究亦透過調頻備轉預測數據提出重要管理意涵，包括鼓勵廠商在不同時間參與競標，以及建議台電優化在電力需求強度不同之時段的（投標）價格上限，以期能透過誘因機制來促成電力平衡配置，提昇電力穩定性。

關鍵字：微電網，集成學習，台電輔助市場，SARIMAX，倒傳遞類神經網路

目錄

摘要.....	i
圖目錄.....	iii
表目錄.....	iv
1、緒論.....	1
1.1 研究背景與動機.....	1
1.2 研究目的.....	2
1.3 研究貢獻.....	3
1.4 專題報告之架構.....	4
2、文獻探討.....	5
2.1 國內電力文獻探討.....	5
2.2 預測方式與集成學習.....	6
2.3 綠電相關文獻.....	8
2.4 輔助服務市場參與機制.....	9
3、研究方法.....	11
3.1 研究步驟與架構.....	11
3.2 研究數據.....	12
3.3 資料預處理.....	12
3.4 研究方法與程序.....	13
4、研究結果與分析.....	19
4.1 資料獲取.....	19
4.2 資料預處理.....	20
4.3 資料圖像化.....	21
4.4 研究結果與分析.....	24
(一) SARIMAX 步驟.....	24
(二) BPN 步驟.....	30
(三) SARIMAX-BPN 模型.....	31
4.5 預測結果績效評比.....	32
5、結論.....	33
5.1 研究結論.....	33
5.2 管理意涵.....	34
5.3 研究限制與未來研究建議.....	34
參考文獻.....	36

圖目錄

圖 1	各年典型春季之鴨子曲線（台電電力交易平台，2020）	8
圖 2	日前輔助服務市場運作之具體時間（台電電力交易平台，2020）	10
圖 3	日前規劃排程與實際規劃排程狀況（台電電力交易平台，2020）	10
圖 4	分析流程圖	11
圖 5	類神經網路圖	15
圖 6	民間企業供給量波動	21
圖 7	國營企業交易量	21
圖 8	每小時市場價格變動	22
圖 9	24 小時內分別的總收購容量	22
圖 10	平日資料	23
圖 11	假日資料	23
圖 12	2021 年 11 月資料趨勢	25
圖 13	2021 年 12 月資料趨勢	25
圖 14	2022 年 1 月資料趨勢	25
圖 15	2022 年 2 月資料趨勢	25
圖 16	運用 MSE 來尋找合適的 SARIMAX 參數	27
圖 17	One-step-ahead forecast 預測結果與原始數據之比對圖	29
圖 18	2021 年 11 月 1 日至 2022 年 2 月 28 日之預測殘差	29
圖 19	SARIMAX 之 Cost 圖	31
圖 20	SARIMAX 之準確率圖	31
圖 21	SARIMAX-BPN 模型的 Cost 圖	32
圖 22	SARIMAX-BPN 模型的準確率圖	32

表目錄

表 1	台電電力交易相關數據（台電電力交易平台，2021-2022）	19
表 2	自然變數相關數據（農業氣象觀測網監測系統，2021-2022）	20
表 3	初步數據整理	20
表 4	基於第一四分位數和第三四分位數分割之總交易量資料	21
表 5	使用 Dickey Fuller 檢定法之價格分析結果	26
表 6	使用 Dickey Fuller 檢定法之得標容量分析結果	26
表 7	各變數加入 ARIMA 模型後得出之 P 值	28
表 8	刪除表 7 中 P 值 > 0.05 的變數（風速、氣溫）後再次運算之 P	28
表 9	刪除表 8 中 P 值>0.05 的變數（太陽日照時數）後再次運算之 P 值	28
表 10	訓練參數篩選	30

1、緒論

1.1 研究背景與動機

2021 年 12 月 18 日臺灣舉行公投，四大議題中有兩項能源相關的議題，包括核四是否重啟，以及第三天然氣接收站是否遷址，顯示未來台灣可能面臨供電效能緊繃、電力短缺的問題。管嫫媛與黃靖萱（2021）指出，對於上述的問題，民眾直覺上會想到的解決方案往往存在執行上的落差。通常被優先提出的方案為增加再生能源或綠色發電的量能，但是目前再生能源發電占比不到 6%，若要在短期內提升產能仍存在許多待解決的問題。第二個解決方案則是盡快蓋好三接以提高供電量，然而，2020 年天然氣機組僅有 61% 的產能利用率，一年近有四成的閒置時間。而調高電價、以價制量也是常被提出的解決方案，而過去十年，台灣工業用電大幅成長，其中電子業占六成漲幅，這些工業用電大戶並不擔憂電價上漲，只擔心供電不足，加徵電費未必能抑制工業用電需求並平衡用電之供需。

2022 年 3 月 3 日上午 9 時，因興達電廠設備故障，導致台灣各地出現停電事故，管嫫媛（2022）提出，若改用區域與智慧型電網，將發電來源改為分散式中小型電廠的再生能源，可讓發電區靠近用電區，減少長程輸配電的風險及耗損。台達電子永續長周志宏於受訪時也表示，此次停電主因並非缺電，而是供電不穩定所造成，現在台灣處於能源轉型的階段，再生能源發電比重逐漸提升，相關儲能與基礎建設等配套措施應盡快完善，否則電網缺乏韌性，就很有可能再次發生停電事件（商業週刊，2022）。

基於這些時空因素與民生產業經濟上的需求，台灣的電力產業也因此出現一個新興的輔助電力市場——「微電網」輔助服務。其中，由台灣電力公司（以下簡稱台電）架設的電力交易平台以輔助服務及備用容量作為電力交易機制之推動對象，形成一個小型微電網。由於發電及儲電科技的進步，現在的產業發展趨勢是人人都能夠成為電力市場的供應商，民間和企業可以一起合力供電。而透過微電網所建立的電力輔助服務，就可以使系統發電及負載得以維持平衡、確保系統穩定。

電力輔助服務調頻備轉（regulation reserve）是本專題關注的焦點，這個服務的主要用途為修正電力系統頻率偏差，或用以減緩系統頻率的變動幅度，因

應再生能源高變動性所導致之瞬時發電變化及系統偶發事故。舉例來說，台積電等晶圓製造廠作為台灣科技業的主力，需要穩定的供電以保證產品的品質，此時電力的調頻備轉就顯得相當重要。調頻備轉作為最即時也最重要的備用電力，其標價（售價）往往也是最高的，因此是供電業者主要競爭的領域。

根據經濟部能源局發佈的智慧電網總體規劃方案核定本（經濟部能源局，2020），台灣近年的電力願景為供電系統的強化與維新，亦即從當下的強化運轉彈性、中期（2025 年）強化電網韌性，與長期（2030 年）落實電業改革的最終目標。根據台電電力交易平台資訊（台電電力交易平台，2022），我國電力現況在 2025 年將發生系統性衝擊，首先，春夏季風力發電的變異較大，顯示台灣需要更多的備轉容量，此外，根據太陽能發電中呈現實際負載及淨負載之間落差的鴨子曲線（duck curve）指出，當近期再生能源總裝置容量大幅增加時，亦需要更多能多次啟停且反應快速的發電能源。隨著儲電設備的進步，價格越來越低，容量也逐漸增大，電力共享的概念也開始被採納，台電利用共享經濟的概念，建立輔助服務交易平台，讓民間電力資源投入市場，也將分散於民間的電力資源納入台電電力系統。許志義與林振玄（2021）認為家庭可參與台電的需量反應方案，自行生產的太陽光電，自用之餘賣給台灣電力公司亦可獲得相當大的收益。透過量化研究，可以讓未來參與智慧家庭能源管理的家庭以此衡量收益。

雖然在能源轉型的路上，微電網以及調頻備轉電力輔助服務的建立可以提升電力供應的效率及穩定。然而，過去微電網的研究主要是針對電力技術的分配與調節，著重在純電力技術或電力分配模型的面向，例如劉志文（2018）再探討我國再生能源管理與運轉策略時，提出的輸電級再生能源併網情境建立及供電穩定之影響評估方案。然而，除了電力技術的突破之外，市場機制的管理效率亦可能嚴重影響輔助市場中調頻備轉供需的不確定性，造成供電可靠性下降，因此本計畫希望能夠透過數據分析探討此現象，並加入合適的管理機制，共促電力市場的發展進步。

1.2 研究目的

綜合上述，本專題將結合時間序列與類神經網路方法建立調頻備轉容量之供給量能的預測模型，並依據電力數據的分析與預測，建立能夠讓廠商可以優

化收益之投標策略，並且使用最佳化模型決定台電公司可依據供給量能訂出之最佳投標價格上限。

本專題將收集台電電力交易平台的公開資料，據以建立電力交易量的預測模型。然電力交易量與供給量存在諸多影響因子，鑒於調頻備轉是供應再生能源缺口為用途，本計畫將透過其他開放資料平台，納入日照、風力等氣候資訊，並輔以季節變化、工作日等影響因子，以提昇預測模型的準確度。本研究在方法上將結合時間序列與類神經網路預測。

再者，由於目前民營市場供給量遠低於台電的需求量，所以當前的民營交易量數據更偏向供給量。因此，本研究希望能透過預測出市場的電力供給量，探討供給不足的時段，是否可以透過提高價格上限來創造進入市場誘因，吸引供電能力相對有限的民營廠商針對電力較緊缺的時段進行投標。進而提出能提高電力供給以實現平穩供電的解決方法。

我們期望建構出的預測模型能精準估計交易及需求量，提供更準確的調頻備轉狀況，同時也最佳化價格上限，避免因供需不平衡而造成需求方的無效支出，達到市場供需平衡的目標。本專題所提出的建議可供電力的供需雙方參考，讓更多民間業者投入輔助市場，增加民間在微電網上的投資，因應國內再生能源的發展以達到最終供電穩定的目的。

1.3 研究貢獻

近年來，能源轉型快速發展，具有穩定電壓與頻率發電的微電網技術也因為能夠提供智慧輸電、穩定儲能等重要功能而興起並受到重視；而一方面，經濟的活絡與科技發展使得大眾對於電能的需求持續增加，欲達到穩定的能源供應與適當運用，準確的供需預測、合宜的交易機制與價格機制也變得相當重要。本專題考量陽光與風力等的氣候因子、季節變化、工作日等時間因子，透過整合時間序列與類神經網路之模型預測（台電電力交易平台）輔助市場的交易與供給量。本專題研究透過建立 SARIMAX-BPN 混合模型預測價格，將預測價格分類為高、中、低三個區塊，提供民間業者價格參考指標可以靈活調整進入輔助市場的時機與電力容量，創造進入市場誘因，進而提出能提高供給或者平衡供需，以實現平穩供電的解決方案。

除了透過 SARIMAX-BPN 混合模型精準估計交易及需求量，更準確評估調

頻備轉狀況外，本研究亦提出最佳化之分時差異化價格上限的概念，以達到市場供需平衡的目標。

過去許多和微電網有關的研究往往著重於純電力技術本身，但本專題改從供需預測與市場機制的管理面向著手，希望透過管理的機制與手段，於再生能源供應不穩定、微電網的電力負載需求不確定等情形下，得以確保供電的平穩可靠。亦可在價格機制方面，建議管理單位在需求緊繃的時段透過提高投標價格上限來吸引更多供給者提供電力，而在需求寬鬆的時刻則適度降低該價格上限，以促成更高的社會效益。

1.4 專題報告之架構

本論文架構如下：

第一章、緒論

主要包含研究背景與動機、目的以及研究貢獻之介紹。

第二章、文獻探討

探討與本專題相關之文獻，並指出研究中的缺口。

第三章、研究方法

說明本專題的研究方法，詳述專題中使用的 SARIMAX 時間序列模型與倒傳遞網路預測模型，同時簡述本研究所使用之台電交易平台資料及氣象資料。

第四章、研究結果與分析

說明本專題的預測結果與分析。

第五章、結論

針對研究結果提出本專題的結論，推衍重要的實務與管理意涵，並指出未來研究方向。

2、文獻探討

本專題的目的在建立調頻備轉容量之供給量或交易量的預測模型，並據以建立較佳的電力市場運作機制，包括針對不同時段的需求，提出更合理的投標價格上限。雖然本專題所提出的研究標的屬於新興的交易市場，但與電力市場有關的資料分析建模的研究已相對成熟，亦有部分專書，如 Chen et al. (2021)，已彙整電力市場中含需求預測、價格模型、投標策略等議題的分析研究結果。由於本計畫預計採用定量預測裡的時間序列模式，並結合類神經網路作為分析工具，並據以建立優化模型。因此文獻回顧主要將針對國內與電力負載或供需預測有關的研究進行回顧，同時亦簡單回顧預測建模方法。

2.1 國內電力文獻探討

李宜馨、陳彥銘與李明峯（2018）研究台灣地區之短期區域電力負載，採用迴歸方法搭配時間序列模型，建立個別小時為單位的負載預測模型，以進行各區域短期負載預測，並建立簡單易解釋且配適度良好的模型。根據他們的研究結論，可以發現影響負載的主要因素包括氣候、個別小時的變化以及特殊日或事件的影響。分別簡單說明如下：（1）氣候：氣溫的變化影響使用空調設備等高耗電產品的情形與時間多寡，是短期負載預測中廣泛使用且重要的氣象因素。在臺灣地區，氣溫與負載呈現非線性的關係，大致以 23°C 為分界。23°C 以下的負載與氣溫關係不明顯，原因為臺灣冬天氣溫相較於其他溫帶國家並不寒冷，使用暖氣空調而導致耗電的比例亦不高，而 23°C 以上負載則與溫度明顯呈現正相關，負載隨著溫度升高而增加，此現象與民眾習慣使用冷氣來達到冷房效果而增加用電量有關。（2）個別小時平均用電量：從數據以及文獻中針對不同時間區間的平均用電量的分析，可發現用電負載存在規律性，不論是以一天為週期，或是一週為週期都顯示出電力負載的規律性。發現負載在 06:00 之後，開始從夜間低點非常快速地加載，並在白天保持高位，但 12:00 到 13:00 之間出現明顯低谷，而 18:00 出現第二個高峰後才慢慢拖尾降載，略呈現駝峰的形狀。且根據經濟部 513 及 517 停電事故檢討報告（2021），觀察 2018 年至 2021 年之夏季尖峰日用電曲線，夜晚 19 時至 20 時出現用電尖峰，且因為已無太陽能發電抑抵，出現夜尖峰負載，顯示夜間負載有越來越大的缺口，若處理不好將影

響未來電力供應。此外，全臺灣一週內平均每小時的用電趨勢，星期一至星期五工作日負載較高，其中星期一與星期二的負載相較於星期三到星期五稍微偏低，而週末因有較高用電的單位處於休息狀態，使得週末負載明顯低於週間；此外，雖同是週末，星期六與星期日亦存在極大的差異。（3）特殊日或事件：由於預測模型之最大誤差多出現在特殊日期或事件，可知其對於負載預測模型的準確性也存在一定影響。整體來說特殊日的預測誤差相比平常日（週間日）較高，指出特殊日與平常日的用電行為影響因子可能存在有很大的差異。

張宏展與羅文毅（2009）基於氣象資料的配電饋線負載做預測研究，他們收集可能影響負載預測模型的變數，提出與李宜馨、陳彥銘與李明峯（2018）相近的結論。張宏展與羅文毅（2009）亦提出三個重要因素，包括（1）天氣因素：主要指因為溫度的改變而影響空調、風扇、冰箱等電器設備的使用，進而影響負載的短期預測。（2）日期類型因素：負載具有週期性的變化，且平日與假日具有不同的規律性。其研究更加入（3）歷史負載資料因素：指歷史上每小時的實際負載大小，後一時刻的負載量會受到之前的負載量影響。提出負載預測模型的具有週期性變化，且時間序列預測模型較符合資料特性。

姚文隆與尹居才（2002）以自迴歸模式建模倒傳遞網路為基礎進行即時用電需求量預測研究，他們在研究中指出短期負載預測需要考量的變數通常包含氣溫、日子類型（工作日、週末假日等）或小時等因子。同時，他們在自迴歸模式下，透過滾動建模方式來建立訓練資料，其中建模的概念來自灰預測（grey prediction）。此研究的結果證實在擁有充足且完善歷史資料的前提下，使用倒傳遞網路的預測結果相比灰預測來得更準確。

2.2 預測方式與集成學習

Makridakis and Hibon（1982）認為與複雜的統計方法相比，單純的方法在時間序列模型建模與預測的表現往往更佳，但是此一觀點遭到當時統計學者的批評，著名的 M 競賽（M Competition）便因此誕生。M 系列比賽近年來持續舉辦，以包含真實數據的數列來評估投稿預測模式之精準度，並探討其緣由。根據 Makridakis et al.（2018）所述，近期的 M4 競賽使用含有 100,000 個真實數據的時間序列且指定了預測間隔（prediction interval），並納入了所有的主流預測方式，包含近年受到重視的人工智能與機器學習方法及傳統的預測模式等。

Makridakis et al. (2018) 指出最精準的預測方法大部分結合不同的模式，且最精準的方法往往混合統計與機器學習模型，相較之下，純機器學習方法反而表現不佳。這證實了結合不同模式之預測方法較單一模式預測方法更好，並且結合統計模型與機器學習模型可大幅度的減少誤差，指出集成學習 (ensemble learning) 的優勢。

在用電需求預測建模部分，不少國內外學者也提出不同的集成學習方法，例如 Wu & Ramos (2017) 針對太陽能發電預測，提出結合長短期記憶 (long short-term memory, LSTM) 及卷積神經網路 (convolutional neural network, CNN)，並運用得出之混合演算法以氣象數據預測太陽能的長期發電量。Lee et al. (2020) 也發表了以集成學習方法預測風力發電量的相關研究。經由文獻得知，集成學習方法能夠結合各個運算模型不同的優點，進而提升預測的準確度。

葉怡成、楊耀華與張萬鈞 (2009) 提出了一個結合 ARIMA 與倒傳遞網路 (back propagation network, BPN) 優點的 ARIMA-BPN 網路。它結合了 ARIMA 整合數列值與殘差值於同一模型的特性與 BPN 建構非線性模型的優點。方法為將 ARIMA 模式的輸出作為倒傳遞網路模型的輸入值，組成一個非線性函數。此方法比起其他串列整合法、平行加權法、串列疊加法等試圖結合 ARIMA 與 BPN 原理的複合模式簡單，也避免了過度適配 (overfitting) 的問題。此研究透過實際資料來比較三種預測模式的效能，其驗證實例中包含電力負載預測，該研究運用臺灣地區 61 日 (共 1464 小時) 每小時用電量資料，將數據以季節差分、平穩化後區分為訓練範例與測試範例進行預測。預測結果顯示，ARIMA-BPN 神經網路演算法比起 ARIMA 與 BPN 這兩種單一預測模式穩定且優秀。

另外，根據盧展南與許元禹 (2018) 針對電力系統短期負載預測之研究，台灣電力系統在不同季節亦會有不同的用電習慣及負載特性，其研究考慮體感溫度、特殊日、季節轉換等特殊原因，將各種預測模型進行比較，其中使用混合模型，可以提供優於單獨使用類神經網路模型或半參數可加模型的預測準確度。基於上述對於不同模型的評析，本研究使用結合 SARIMAX 與 BPN 的混合模型，達到預測市場電力交易量的目的。

2.3 綠電相關文獻

許志義等人（2020）探討了太陽能及風力發電的間歇性問題，並分析臺灣出現鴨子曲線（duck curve）的可能性。許志義等人（2020）也針對鴨子曲線作出解釋，該圖形為日出時太陽能發電系統發電並快速地加入電網，使得原本淨負載下降，總負載與淨負載之間差額加大，形成像鴨子肚子凸出的區段。而日落太陽能發電驟減，形成鴨脖子的區段，淨負載需求快速上升，需要搭配其他電力組合服務升載調度。許志義等人（2020）指出台灣因地形狹長且太陽能與風力發電場域集中，所以再生能源的發電間歇性會更加明顯。再加上台灣屬於電網孤立系統，電力調度的困難可能更勝美國加州或德州。

台電交易平台中電力系統運轉與調度的參考資料也提到太陽光電受氣候、晝夜影響相當顯著（台電電力交易平台，2020）。圖 1 是各年典型春季之淨負載曲線之變化趨勢。當太陽光電占比越高，鴨肚子凸出得就越明顯，對電網的調度壓力越大，因此需要輔助市場的配合，包括可快速反應的調頻備轉、備轉容量等。在電網中加入多元且具有彈性的資源，可有效抑低其可能造成之衝擊與影響，在再生能源佔比逐漸增多時確保其穩定性。

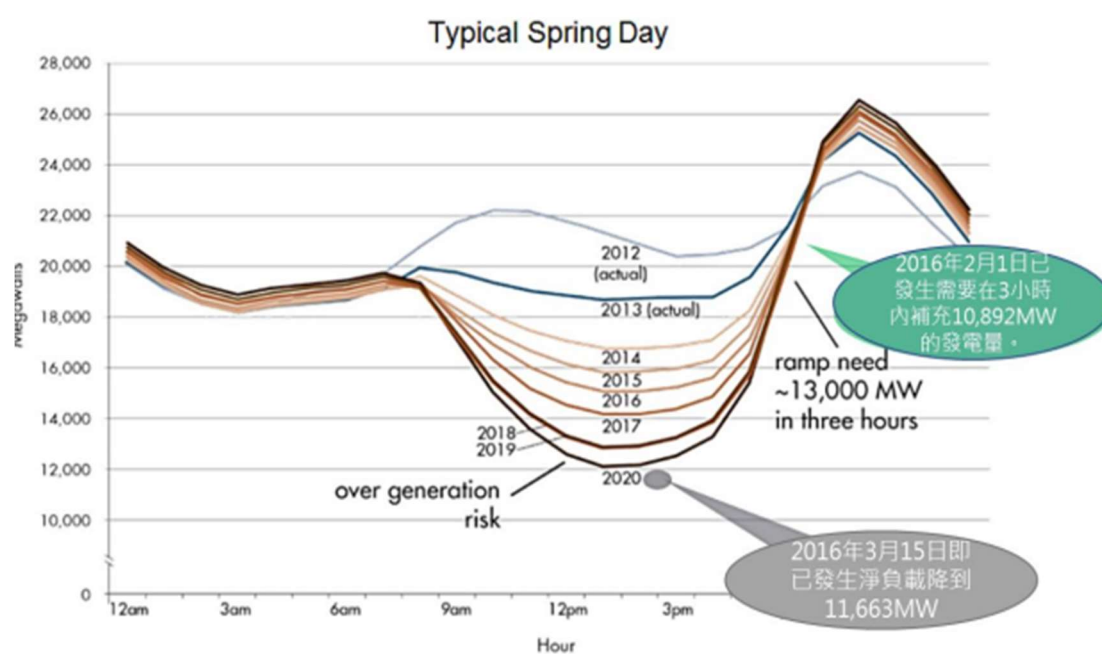


圖 1 各年典型春季之鴨子曲線（台電電力交易平台，2020）

RE100 是由國際氣候組織（The Climate Group）與碳揭露計畫（Carbon Disclosure Project, CDP）在 2014 年共同發起的能源倡議，邀請全球企業承諾 100%使用再生能源的目標，截至 2022 年 1 月，響應 RE100 的企業已有 349 家。RE100 的 2021 年報（RE100, 2021）指出，該組織旗下各企業的總電力消費量高達每年 3400 億千瓦（340TWh），數據已超越英國，足以成為全球第 12 大耗電國。目前已有許多跨國企業加入，如 Google、Apple、微軟等，承諾將於生產過程中完全使用綠電，並以同等標準要求其供應鏈廠商。

根據 RE100 Taiwan 官網（2022），台灣具營運活動的 RE100 會員有上百家，而台灣總部會員共有 18 家，包含台達電、宏碁集團、華碩等。再加上《再生能源發展條例》第 12 條第 4 項明定用電契約在一定容量以上之企業需使用一定額度以上之再生能源（經濟部能源局，2019），否則應繳納代金，這項條例讓用電大戶開始積極購買綠電。因此增加綠電供應的穩定性迫在眉睫。

2.4 輔助服務市場參與機制

為了促成電力資源地更有效利用，各國常會建立電力交易輔助服務市場，透過交易的機制來平衡電力的供需，並提高總體的社會效益。楊豐碩（2018）等人指出雖然各國的輔助服務定義及內容不盡相同，但都包含兩個項目：熱機備轉和頻率調頻，且調頻通常以過去歷史資料或某一比例的負載預測為基準。

根據輔助服務及備用容量交易試行平台第一次公開說明會簡報（台電電力交易平台，2020），目前電力輔助服務市場的交易項目分為三項：（1）調頻備轉：1 秒反應持續 15 分鐘以上（2）即時備轉：10 分鐘反應持續 1 小時（3）補充備轉：30 分鐘反應持續 2 至 4 小時以上。其中調頻備轉是本專題研究主要預測的數據，透過即時增減操作功率，修正系統頻率偏差或減緩頻率變動幅度。由於其對供電廠商的進入門檻較高，所以目前參與的民營廠商只有 16 家，其餘不足的供給量皆由國營廠商吸收。

在第二次公開說明會簡報中有對日前輔助服務市場運作之具體時間做說明（台電電力交易平台，2020）如圖 2。

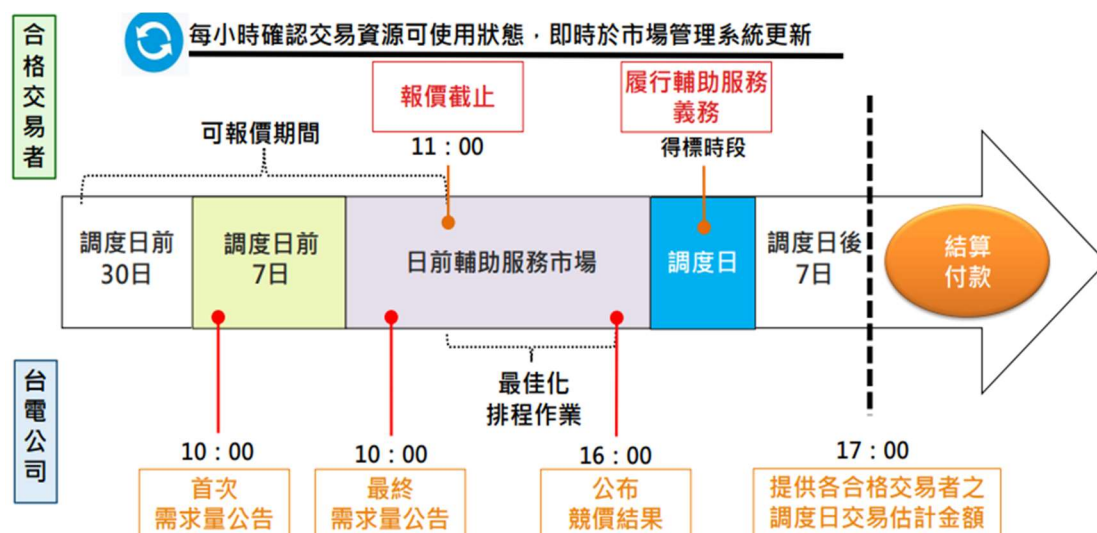


圖 2 日前輔助服務市場運作之具體時間（台電電力交易平台，2020）

由圖 2 可以看到首次需求量公告為調度日前 7 日，且最終需求量公告為調度日前一天的早上 10 點，根據台電月刊（2020），台電綜合研究所目前使用 AI 技術，建立發電量及日照量的關聯預測模型，提供整體風力及太陽光電未來 48 小時所能提供的發電量預測，供發電排程參考使用，再對供電廠商進行招標。供電廠商的報價截止時間為調度日前一天的早上 11 點，經過最佳化排程作業後，於 16 點公佈競價結果。

並且在日前市場中，實際的規劃排程會因為系統狀況調整發電量而改變，其結果和日前規劃排程不盡相同。根據電力系統運轉與調度資料（台電電力交易平台，2022），電力無法大量儲存，為了供應隨時變化的負載，發電量須隨時跟著變化，才能每分每秒維持發電與負載的平衡，所以在小時排程前，預估的發電量可能會有所變動，圖 3 即說明其運作之機制。

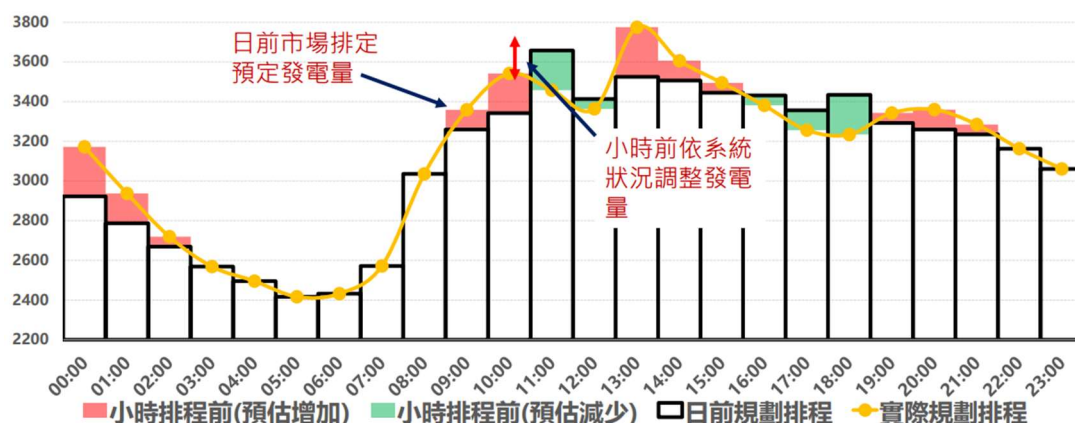


圖 3 日前規劃排程與實際規劃排程狀況（台電電力交易平台，2020）

3、研究方法

文獻回顧指出近年關於電力市場預測的研究方法往往結合傳統的時間序列統計分析模型（如 ARIMA）與機器學習模型（如類神經網路），而文獻回顧中提及的 M 競賽也證實結合不同預測模式（尤其是結合統計方法與機器學習）可以有效降低誤差，提升準確度。本計畫會在第一階段使用商業分析（business analytics）中的預測分析（predicting analytics）的時間序列及機器學習模型，在第二階段採用規範或指示型分析（prescriptive analytics）中的數學規畫方法。鑑於文獻回顧的結果，本計畫在將在第一階段的預測建模中使用組合模型的方法來預測供給量（或交易量）。

3.1 研究步驟與架構

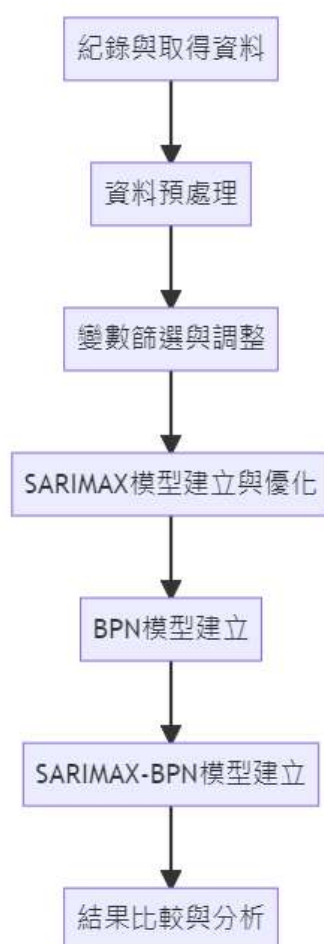


圖 4 分析流程圖

3.2 研究數據

本專題主要收集台電電力交易平台上每小時的調頻備轉得標容量與歷史投標資料，其中得標容量為國營企業容量與民營企業容量的加總，收集從 2021 年 11 月 1 日至 2022 年 2 月 28 日台電電力交易平台調頻備轉資料，總計 2880 筆資料，其中 2021 年 11 月 29 日資料缺失，共 24 筆，並結合農業氣象觀測網監測系統之太陽日照時數與風速數據進行預測。經分析後，歷史投標資料中國營企業所投標的容量大於成功交易量，而民營企業所投標的容量則是被照單全收，因此推論民營企業的市場還有再擴展的可能性。根據台電網站說明，太陽能發電大宗為台南鹽田，故太陽日照時數選用測站七股研究中心之數據；而風力發電大宗為彰工風力發電站，故風速數據選用測站臺中農業改良場之數據。

3.3 資料預處理

資料預處理中包含填補遺漏值，分類數據之轉換與特徵處理。將資料調整為合適的格式提升數據的品質。

填補遺漏值：

針對 2021 年 11 月 29 日的資料缺失，鑑於平均數可能受到極端值的影響，故本專題採用填補中位數的方式來處理。利用 11 月份各時之中間值來填補，降低缺失值的影響。

分類數據之轉換：

為避免分類數據被認為有數值上的差異而影響結果，本專題使用獨熱編碼（one-hot encoding）技術，將「名目特徵」中的值轉換為新的「虛擬特徵」。

特徵處理：

鑑於各變數的區間範圍有所差距，需要對變數做處理才能放進模型中。本專題採用標準化的模式，將變數壓縮在 $[0,1]$ 之間。

3.4 研究方法與程序

本專題預計使用 SARIMAX 模型和 BPN 模型，並嘗試將兩種模型相結合，以求更高的預測準確率。本小節將分成二個子小節，以下先簡單介紹這兩種預測模式，接著說明研究的步驟。此外，本研究使 SARIMAX 模型建立時間序列模型時，考量到不同季節（即每一週期的對應時點）的資料間可能具時間相依性，並且資料可能受到其他變數的影響，因此不僅加入了外生（eXogenous）變數，同時考慮季節相依性模式（Seasonality）以構成 SARIMAX 模型，以下我們分別說明這幾個模型。

ARIMA 模型

ARIMA 模型（Autoregressive Moving-Average Models），又稱為差分自我迴歸整合移動平均模型，是由 Box and Jenkins（1970）提出的時間序列預測方式，模型一般表為 ARIMA（p, d, q），其中 p 為自迴歸項，q 為移動平均項，d 則為使序列平穩所做的差分次數，其背後的原理是將前一數據的殘差值做為參數加入預測模型中，以逐步優化預測結果。

使用 ARIMA 建立時間序列預測模型時，其函式代表時間序列受到前 p 期數值及前 q 期殘差值之影響，寫為式(3.1)：

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \quad (3.1)$$

其中：

Y_t = t 時期數列值。

ϕ_0 = 常數項。

Y_{t-n} = 前 n 期之因變數。

ϕ_n = 數列值係數。

θ_n = 殘差值係數。

ε_t = t 時期殘差值。

在函式中，數列值構成的部分為自我迴歸模型 AR（p），是當前值與歷史值之間的關係；殘差值構成的部分為移動平均模型 MA（q），為自我迴歸部分的誤差累計。差分自我迴歸整合移動平均模型 ARIMA（p, d, q）又為 ARMA 模

型的擴展，其能將非平穩的數據進行平穩化處理，並根據時間序列模型識別規則所識別出來的特徵，建立相應的時間序列模型。判斷資料是否平穩則常運用 Dickey 與 Fuller (1979) 建立的 Dickey-Fuller test。若檢驗統計量小於臨界值，則拒絕原假設，說明資料為平穩，反之亦然。

SARIMAX 模型建立

由於得標容量具季節性，且受到日照時數、風速、溫度等外生變數影響，故使用 SARIMAX 模型預測資料較符合資料性質。使用 Python 計算模型參數，利用 MSE 求得 SARIMAX (p, d, q) × (P, D, Q)。

季節性 (Seasonal)：

資料在不同月份出現以一天為單位的固定頻率，且會受到其一小時的數量影響。因為季節性違反了 ARIMA 模型的平穩性假設，因此加入 SARIMA 模型進行研究預測。每小時記錄一筆資料，具有 24 小時季節性，方程式如式(3.2)所示。

$$y_t = c + \sum_{n=1}^p a_n y_{t-n} + \sum_{n=1}^q \theta_n \varepsilon_{t-n} + \sum_{n=1}^P \phi_n y_{t-sn} + \sum_{n=1}^Q \eta_n \varepsilon_{t-sn} + \varepsilon_t \quad (3.2)$$

外生變數 (eXogenous)：

資料除了受到本身變數影響，例如：歷史交易量、價格等等，也會受到外生變數影響，透過考量外部因素與內生特徵對時間序列的影響，提高預測模型的準確性。推導出完整的 SARIMAX 方程式，如式(3.3)所示。

$$d_t = c + \sum_{n=1}^p a_n d_{t-n} + \sum_{n=1}^q \theta_n \varepsilon_{t-n} + \sum_{n=1}^r \beta_n x_{nt} + \sum_{n=1}^P \phi_n d_{t-sn} + \sum_{n=1}^Q \eta_n \varepsilon_{t-sn} + \varepsilon_t \quad (3.3)$$

類神經網路

類神經網路 (Artificial Neural Networks, ANN) 是一種模仿生物神經網路結構與功能的數學模型與計算模型，應用於辨識、決策、控制、預測等工作。它由許多高度連結的節點 (或處理單元) 所組成，模仿生物神經元彼此之間發送訊號的方式高速分類與叢集資料。

Rumelhart et al. (1985) 將類神經網路的架構分成三個層次，如圖 5：

- (1) 處理單元：類神經網路組成的基本單位。
- (2) 層：一般包括輸入層、隱藏層與輸出層，由若干個具相同作用的處理單位組合而成。
- (3) 網路：由不同作用的層集成網路，具有學習和回想兩個過程。

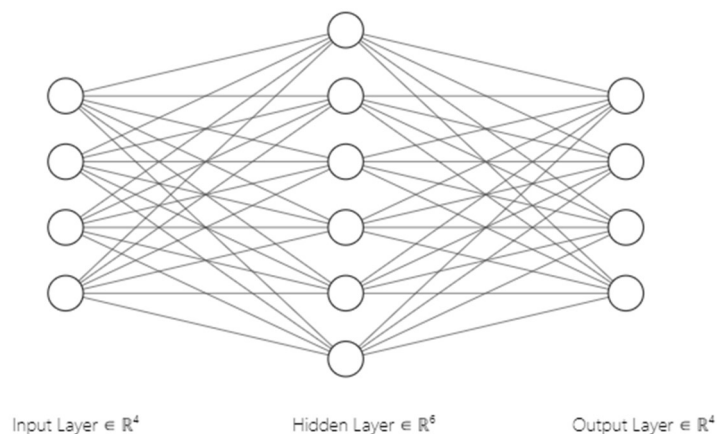


圖 5 類神經網路圖

類神經網路可以通過訓練資料來學習，由輸入的資料中歸納出隱含關係，是機器學習的一種演算法。而楊朝平與許育彰（2003）的研究中有簡略介紹，倒傳遞網路（Back-Propagation Network, BPN）為類神經網路模式中最為廣泛使用的模式之一，分為學習階段與回想階段，其優點為：

- (1) 添加隱藏層之使用
- (2) 改用平滑可微分的非線性轉換函數
- (3) 學習精度高
- (4) 回想速度快
- (5) 應用範圍廣

公式分為向前傳遞(3.4)-(3.10)與向後傳遞(3.11)-(3.20)兩部分介紹：

向前傳遞

1. 為「隱藏層」 $a_1^{(h)}$ 計算出「啟動單元」，方程式如式(3.4)-(3.8)所示。

$$z_1^{(h)} = a_0^{(in)}w_{0,1}^{(h)} + a_1^{(in)}w_{1,1}^{(h)} + \dots + a_m^{(in)}w_{m,1}^{(h)} \quad (3.4)$$

$$a_1^{(h)} = \phi z_1^{(h)} \quad (3.5)$$

啟動函數（sigmoid function）公式如下式(3.6)。

$$\phi(z) = \frac{1}{1 + e^{-z}} \quad (3.6)$$

其中：

$z_1^{(h)}$ = 淨輸入，為連接「輸入層」和「輸出層」的加權線性組合。

$\phi(\cdot)$ = 「啟動函數」，且必須是一可微函數才可利用「梯度下降法」學習連接到神經元的加權。

$a_i^{(l)}$ = 在第 l 層的第 i 個「啟動單元」，又以 in 來代表「輸入層」、 h 來代表「隱藏層」、 out 來代表「輸出層」。

$a_0^{(in)}$ = 偏誤單元（bias unit），設定為 1

$a_0^{(h)}$ = 偏誤單元（bias unit），設定為 1

以上公式可簡化為式(3.7)-(3.8)：

$$Z^{(h)} = A^{(in)} W^{(h)} \quad (3.7)$$

$$A^{(h)} = \phi(Z^{(h)}) \quad (3.8)$$

其中：

$A^{(in)}$ = $n \times m$ 維特徵向量加上「偏誤單元」。

$W^{(h)}$ = $m \times d$ 維的加權向量。 d 為「隱藏層」的數目。

$Z^{(h)}$ = $n \times d$ 維的淨輸入向量。

$A^{(h)}$ = $n \times d$ 維的啟動函數。

2. 同理計算「輸出層」，並應用 sigmoid 啟動函數計算類神經網路的連續輸出值，如式(3.9)-(3.10)所示。

$$Z^{(out)} = A^{(h)} W^{(out)} \quad (3.9)$$

$$A^{(out)} = \phi(Z^{(out)}), \quad A^{(out)} \in \mathbb{R}^{n \times t} \quad (3.10)$$

其中：

$A^{(h)}$ = $n \times d$ 維的啟動函數。

$W^{(out)}$ = $d \times t$ 維的加權向量。 t 為輸出單元的個數。

$Z^{(out)}$ = $n \times t$ 維的淨輸入向量。

$A^{(out)}$ = $n \times t$ 維的啟動函數。

向後傳遞

在倒傳遞類神經網路裡，需要將「錯誤」由右向左傳播。

1. 計算「輸出層」的「錯誤向量」，如式(3.11)。

$$\delta^{(out)} = a^{(out)} - y \quad (3.11)$$

其中：

$\delta^{(out)}$ = 「輸出層」的錯誤向量

$a^{(out)}$ = 啟動單元向量

y = 真實類別標籤的向量

2. 計算「隱藏層」的「錯誤向量」，如式(3.12)。

$$\delta^{(h)} = \delta^{(out)} (W^{(out)})^T \odot \frac{\partial \phi(z^{(h)})}{\partial z^{(h)}} \quad (3.12)$$

其中：

$\delta^{(h)}$ = 「隱藏層」的錯誤向量

\odot = 逐元素項相乘

$\frac{\partial \phi(z^{(h)})}{\partial z^{(h)}}$ 為 sigmoid 啟動函數的導數，如式(3.13)。

$$\frac{\partial \phi(z)}{\partial z} = (a^{(h)} \odot (1 - a^{(h)})) \quad (3.13)$$

將上述兩式整理如式(3.14)。

$$\delta^{(h)} = \delta^{(out)} (W^{(out)})^T \odot (a^{(h)} \odot (1 - a^{(h)})) \quad (3.14)$$

3. 計算成本函數的導數，如式(3.15)-(3.16)所示。

$$\frac{\partial}{\partial w_{i,j}^{(out)}} J(w) = a_j^{(h)} \delta_i^{(out)} \quad (3.15)$$

$$\frac{\partial}{\partial w_{i,j}^{(h)}} J(w) = a_j^{(in)} \delta_i^{(h)} \quad (3.16)$$

4. 將每一層中的所有節點的偏導數相加，再將下一層的節點誤差相加，如式(3.17)-(3.18)。

$$\Delta^{(h)} = (A^{(in)})^T \delta^{(h)} \quad (3.17)$$

$$\Delta^{(out)} = (A^{(h)})^T \delta^{(out)} \quad (3.18)$$

5. 加入正規項，如式(3.19)。

$$\Delta^{(l)} := \Delta^{(l)} + \lambda^{(l)} W^{(l)} \text{ (除了偏誤項以外)} \quad (3.19)$$

6. 向著梯度的相反方向對每一層 l 更新加權，如式(3.20)。

$$W^{(l)} := W^{(l)} - \eta \Delta^{(l)} \quad (3.20)$$

機器學習－集成學習

集成學習是指透過建立多個模型組合來解決單一預測問題，通常運用在統計和機器學習中，他能獲得比單獨使用一種演算法更好的預測性能。結合法可以分為自助聚合（Bagging）、提升法（Boosting）和堆疊法（Stacking）。本專題使用 Stacking（堆疊法），以不同運算方法訓練同一組數據，再使用各算法的預測結果作為附加項輸入，進行最終預測。堆疊法的作法是將前 m 個模型的輸出作為新模型的輸入。本專題將 SARIMAX 的殘差作為倒傳遞的其中一個變數加入模型中。ARIMA 多考慮了數列殘差值做為輸入值，但是只能產生線性函數，而 BPN 則可產生非線性函數，但未考慮到數列殘差值做為輸入值。因此將上述兩種預測模型做結合。結合方法為將 SARIMAX 模型得到的參差值作為變數加入 BPN 模型裡。

4、研究結果與分析

4.1 資料獲取

由於台電的資料不能直接獲取，數據以圖表的形式呈現在網頁上，且僅公布最近 30 天的資料，因此我們手動紀錄下從 2021 年 11 月 1 日開始的數據。本專題使用 2021 年 11 月 1 日到 2022 年 2 月 28 日的資料。而農業氣象觀測網監測系統則可以直接下載數據。

台電電力交易平台 資料獲取結果

本專題紀錄台電電力交易平台 2021 年 11 月 1 日到 2022 年 2 月 28 日共 2880 筆資料，每筆資料中包含國營交易容量、民營交易容量和每小時交易價，結果如表 1 所示。

表 1 台電電力交易相關數據（台電電力交易平台，2021-2022）

欄位名稱	內容	舉例
國營交易量	每小時成交的國營企業容量	236.4
民營交易量	每小時成交的民營企業容量	14
價格	每小時成交的競標交易價	600
時間	年-月-日-時	2021-11-01-00
特殊日	特殊節假日，類別 Bool	0

農業氣象觀測網監測系統 資料獲取結果

本專題從農業氣象觀測網監測系統 2021 年 11 月 1 日到 2022 年 2 月 28 日共 2880 筆資料，每筆資料包含臺中農改場的和七股研究中心的日照時數，結果如表 2 所示。

表 2 自然變數相關數據（農業氣象觀測網監測系統，2021-2022）

欄位名稱	內容	舉例
平均風速(m/s)	臺中農改站觀測數據	2.2
氣溫(°C)	臺中農改站觀測數據	21.4
日照時數(hr)	七股研究中心觀測數據	0

4.2 資料預處理

將時間拆成年月日欄，國營交易量與民營交易量相加成總交易量，處理完後分析所得到的數據。本專題使用 Python 的 library: numpy 與 pandas 進行整理分析，表 3 為整理結果。

表 3 初步數據整理

	count	mean	std	25%	50%	75%
國營交易量	2880	193.332569	138.063247	80	134	285
民營交易量	2880	14.433507	1.059340	14	14	15
總交易量	2880	207.766076	138.348636	94	150.6	300.8
價格	2880	594.712847	12.446056	595	596	600
風速	2880	2.711875	1.762875	1.3	2.5	3.9
太陽日照時數	2880	0.338299	0.453103	0	0	1
氣溫	2880	18.35257	3.837404	15.6	17.8	20.8
特殊日	2880	0.14965	0.356793	0	0	1

根據上表 3 得出的資訊，可以將總交易量資料依照第一四分位數和第三四分位數分為三個部分，如下表 4，方便使用類神經網路進行分類預測，也較符合本專題預測出需求量高的時段的目的。

表 4 基於第一四分位數和第三四分位數分割之總交易量資料

交易總量	編號
100 以下（第一四分位數）	0
100 到 300	1
300 以上（第三四分位數）	2

4.3 資料圖像化

將數據圖像化，觀察每個資料的走向與相關性。使用 Python Library: Matplotlib 來進行圖像化處理：

（1）圖 6 為民間企業的供給量波動，每月有逐漸上漲的趨勢，與歷史投標量也吻合，因此推測民間企業提供的容量被台電全部接收。

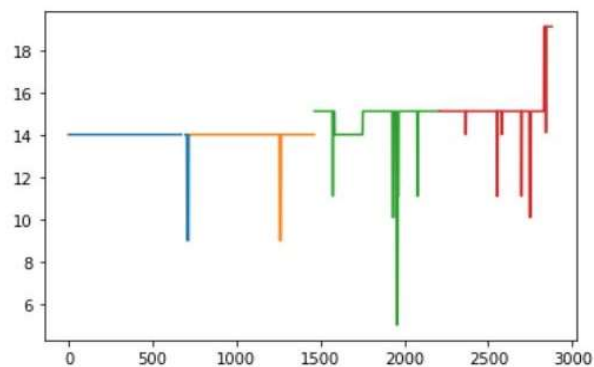


圖 6 民間企業供給量波動

（2）圖 7 為國營企業的交易量，相較民營企業有明顯的變動。

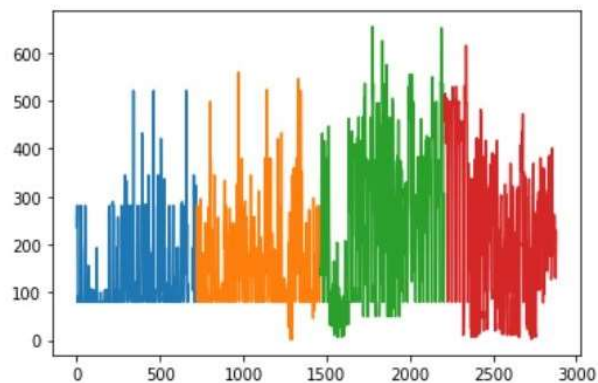


圖 7 國營企業交易量

(3) 圖 8 為市場價格每小時的變動圖，在實行輔助市場前期具有明顯波動性，此時競標功能尚且完善，但後期價格逐漸穩定，競標功能隨之效用降低，因此我們需要建立新的競標機制，也希望能透過新的競價機制為市場帶來新的參與者，例如民間企業可以自由挑選時段投標，而非一次性參與所有時段，吸引更多民間企業參與，並達到價格最佳化，故本專題將預測的對象為總交易量。

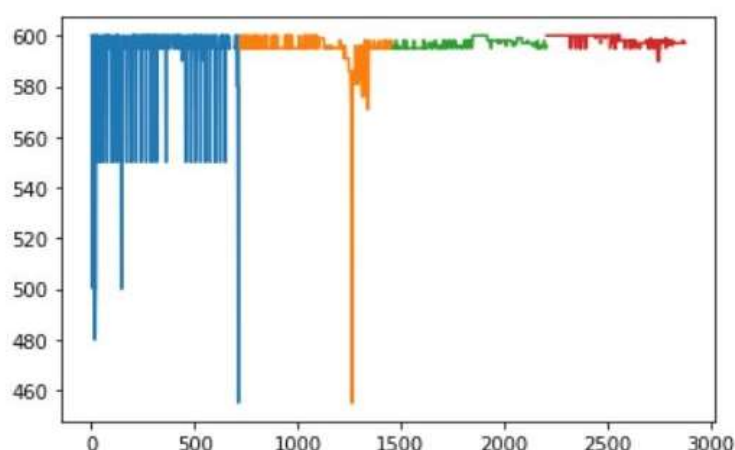


圖 8 每小時市場價格變動

(4) 圖 9 為全部資料以小時為單位的總收購容量所繪製的 Boxplot，發現 16 點到 23 點這段時間需求明顯減少，推測太陽日照可能為影響因素，故加入考量。

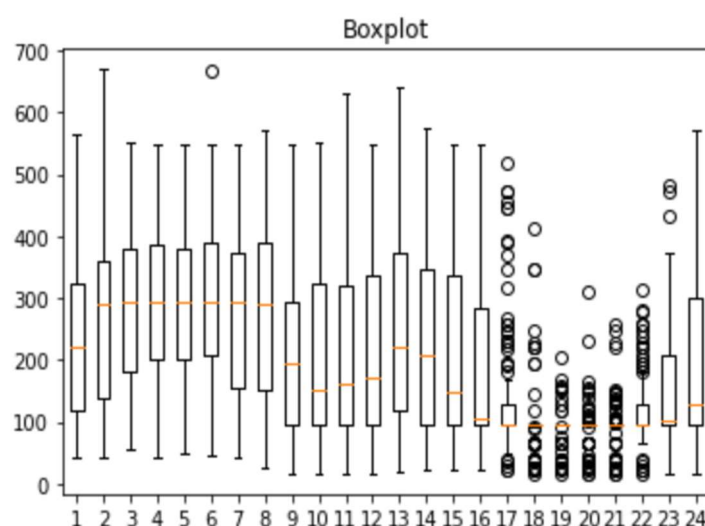


圖 9 24 小時內分別的總收購容量

(5) 若將資料分為平日與假日（圖 10 為平日，圖 11 為假日），Boxplot 將呈現出不一樣的結果，推測平假日可能為影響因素，故加入考量。並且因台灣產業用佔比明顯高於家庭用電，我們也考慮特殊日所帶來的影響，如春節假期等，作為變數加入模型中。

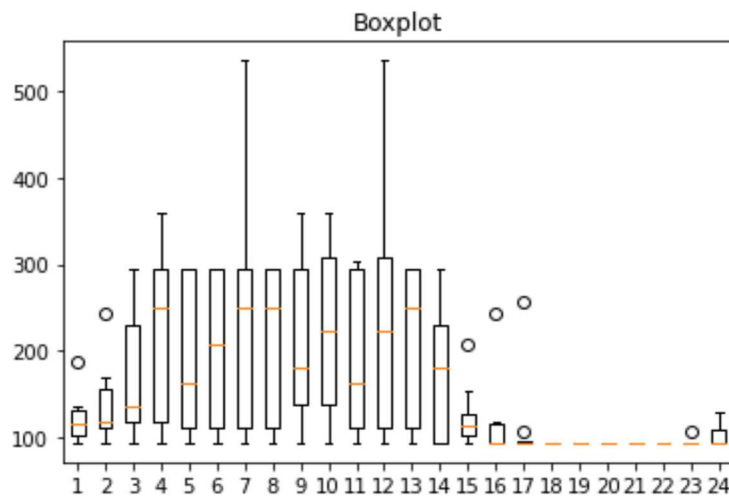


圖 10 平日資料

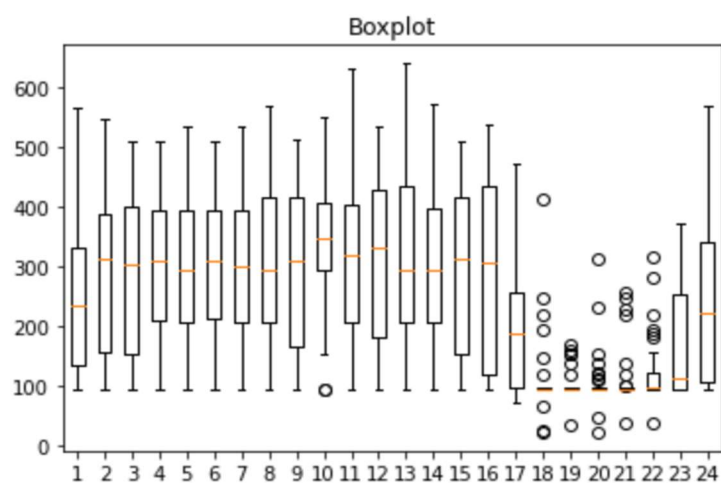


圖 11 假日資料

根據台電電力交易平台說明，調頻備轉係補充再生能源不穩定性的輔助作用，所以推測風力可能也為影響數據的變數。同時，文獻探討中提到氣溫也是對電力供需有顯著影響的變數，我們以緯度高低考慮，取台灣緯度中間值，以臺中農業改良場的氣溫數據作為變數加入模型中。

4.4 研究結果與分析

透過視覺化時間序列呈現資料後，我們發現資料有季節性規律，綜觀不同月份之間，每日午夜到清晨會出現得標容量高峰，具有以天為單位的週期性。且資料會受到日照時數、風速、溫度等外生變數影響。接著，我們使用 Dickey Fuller 檢定法檢測總交易量與價格資料是否平穩，證實兩者皆是平穩序列後，建立 SARIMAX 模型，加入季節性與外生變數組成 SARIMAX 模型，提高模型預測準確度。並開始訓練數據，將資料按比例隨機分為訓練集（75%）與測試集（25%），並將前項 SARIMAX 模型預測的殘差值作為變數加入 BPN 模型，以整合預測模型，完成 SARIMAX-BPN 的混合模型。最後進行準確度評估。

(一) SARIMAX 步驟

第一步：視覺化時間序列，觀察資料趨勢。

觀察 2021 年 11 月 1 日到 2022 年 2 月 28 日的資料，以下每月資料趨勢圖分為四個部分，由上到下分別為容量、趨勢、季節性和殘差，各月份資料具有以一天為單位的規律性及明顯的季節性，並且趨勢一欄雖有起伏但整體平穩，沒有明顯地持續上升或下降。

透過時間序列圖觀察台電電力交易平台調頻備轉資料得知，調頻備轉得標容量具有以一天為週期的規律性。請見圖 12，2021 年 11 月得標容量從 00:00 開始急速上升大約上午 08:00 之後趨近於零。請見圖 13，2021 年 12 月得標容量從前日半夜開始增加，隔一小時之後出現第二高峰，趨近中午 12:00 得標容量又開始上升。請見圖 14，2022 年 1 月得標容量從前日開始增加，大約下午 15:00 出現第二高峰，之後趨近於零。請見圖 15，2022 年 2 月得標容量高峰從前日一直持續到傍晚 17:00，之後急速下降趨近於零。綜觀上述不同月份之間，每日午夜到清晨會出現得標容量高峰，具有以一天為單位的季節性規律。

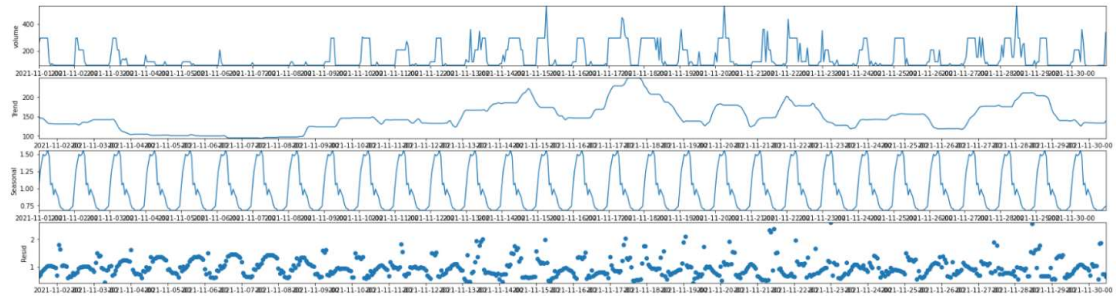


圖 12 2021 年 11 月資料趨勢

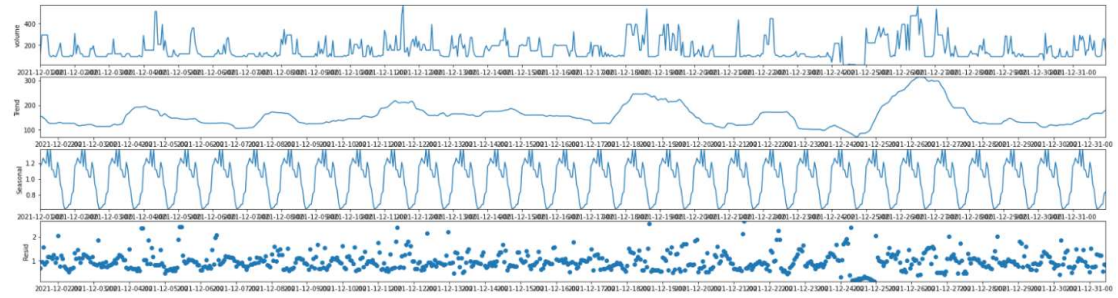


圖 13 2021 年 12 月資料趨勢

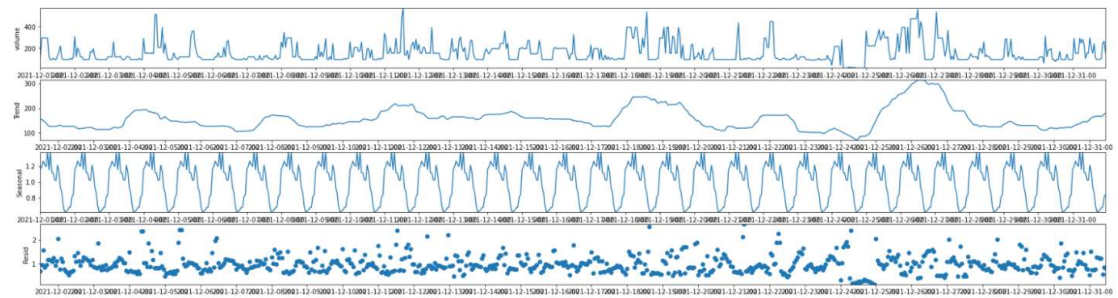


圖 14 2022 年 1 月資料趨勢

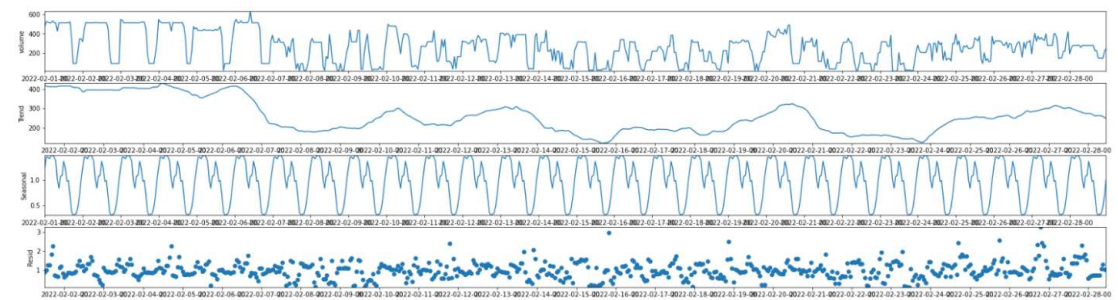


圖 15 2022 年 2 月資料趨勢

第二步：標準化總交易量與價格資料

使用 Dickey Fuller 檢定法檢測，得出檢驗統計量（Test Statistic = -3.79），小於臨界值（Critical Value = 1%），所以拒絕原假設，表示總交易量與價格皆為平穩序列。詳細數值如表 5、表 6。

表 5 使用 Dickey Fuller 檢定法之價格分析結果

Price Results of Dickey Fuller Test	
Test Statistic	-8.434437e+00
p-value	1.832843e-13
#lags used	2.800000e+01
Number of Observations Used	2.851000e+03
Critical Value(1%)	-3.432646e+00
Critical Value(5%)	-2.862554e+00
Critical Value(10%)	-2.567310e+00
The data is stationary. (Critical Value 1%)	

表 6 使用 Dickey Fuller 檢定法之得標容量分析結果

Volume Results of Dickey Fuller Test	
Test Statistic	-3.768875
p-value	0.003242
#lags used	27.000000
Number of Observations Used	2852.000000
Critical Value(1%)	-3.432645
Critical Value(5%)	-2.862554
Critical Value(10%)	-2.567310
The data is stationary. (Critical Value 1%)	

第三步：找出最佳參數

由於模型除了內生變數（例如：價格），還需考慮外生變數（例如：日照時數、風速等），故使用 SARIMAX（季節性時間序列分析）預測模型進行分析。因為此為平穩數據、d 固定為 0。

使用均方誤差（Mean-Square Error, MSE）來尋找最佳參數。MSE 度量的是預測值和實際觀測值間差的平方之均值，只考慮誤差之平均大小，且由於經過平方，與真實值偏離較多的預測值會受到更為嚴重的懲罰。MSE 為衡量資料變化程度的一種方法，MSE 的值越小，說明預測模型描述實驗資料具有更好的精確度。得出結論 SARIMAX(2,0,1)×(1,0,1,24)為最佳參數，如圖 16。

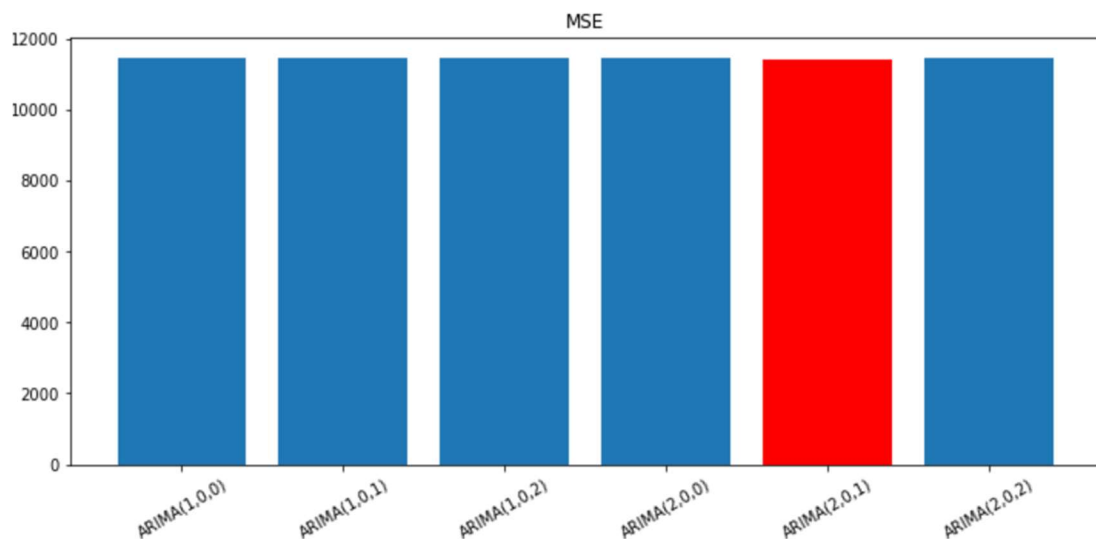


圖 16 運用 MSE 來尋找合適的 SARIMAX 參數

第四步：建立 SARIMAX 模型

將上述所提及的所有變數加入模型，變數為小時、風速、太陽日照時數、氣溫、平假日與特殊日。使用 Python Library: statsmodels 來進行 SARIMAX 預測。預測步驟如表 7 至表 9 所示。

表 7 各變數加入 ARIMA 模型後得出之 P 值

	coef	P> z
小時	-3.2261	0.000
風速	0.3792	0.795
太陽日照時數	49.4120	0.000
氣溫	-1.0797	0.475
平假日	-49.8763	0.000
特殊日	66.6448	0.000

表 8 刪除表 7 中 P 值 > 0.05 的變數（風速、氣溫）後再次運算之 P 值

	coef	P> z
小時	-2.4897	0.002
太陽日照時數	-0.2214	0.976
平假日	-45.4080	0.000
特殊日	63.6040	0.000

表 9 刪除表 8 中 P 值>0.05 的變數（太陽日照時數）後再次運算之 P 值

	coef	P> z
小時	-5.9831	0.000
平假日	-43.8785	0.000
特殊日	44.8830	0.000

第五步：分析預測結果

將 2021 年 11 月 1 日至 2022 年 2 月 28 日的數據資料，使用 One-step-ahead forecast 預測（紅色），並與原始資料（藍色）做比對，結果如圖 17。同時記錄預測結果與原始資料對比之殘差值，如圖 18。

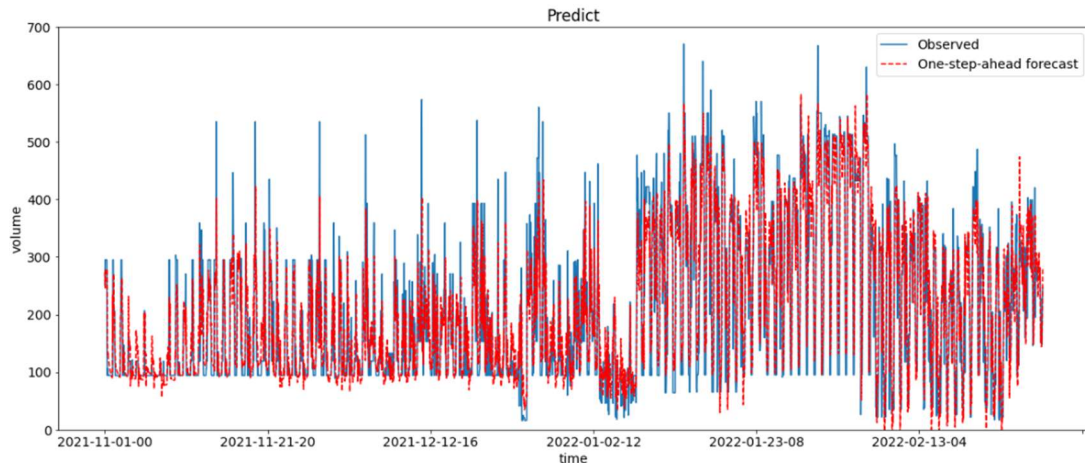


圖 17 One-step-ahead forecast 預測結果與原始數據之比對圖

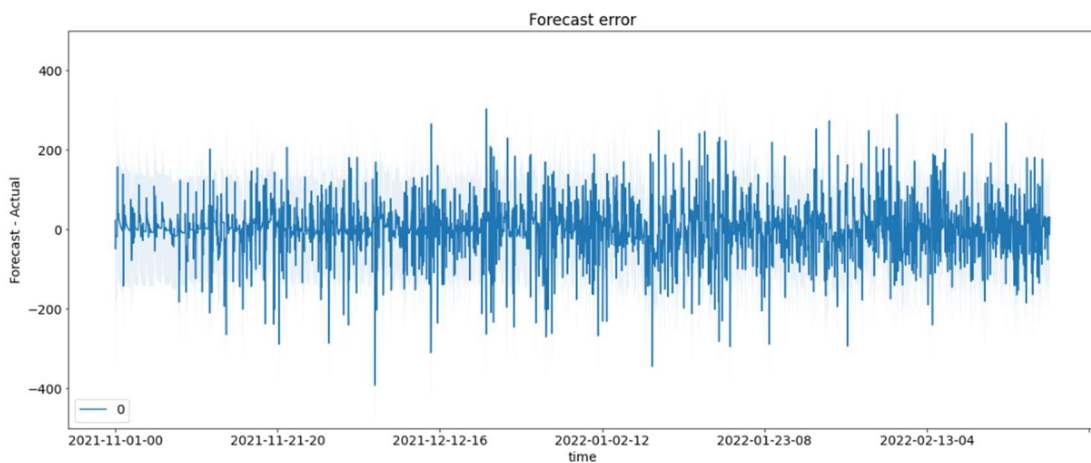


圖 18 2021 年 11 月 1 日至 2022 年 2 月 28 日之預測殘差

使用 RMSE 與 MAE 計算誤差值

$$\text{RMSE} = 68.9703458790658$$

$$\text{MAE} = 46.58854718117474$$

(二) BPN 步驟

第一步：選擇數據

將小時、平假日、特殊日、風速、太陽日照時數、氣溫、作為輸入項，建立 X array，而將分類類別放入 Y array 當作參考值。並將資料按比例隨機分為訓練集（75%）與測試集（25%）。特徵處理方面，將現有特徵標準化使其縮放成同一規模。利用 scikit-learn 標準化 X_train 與 X_test 裡的變數。

第二步：訓練數據

預設數值為訓練次數為 200 次，學習率 0.001，L2 為 0.01，隱藏單元為 30，隱藏層則固定為 1。經過測試，選擇準確率高且測試集準確率與訓練集準確率相差小者，選定學習率 0.0001，L2 為 0.01，隱藏單元為 30，結果如表 10。

表 10 訓練參數篩選

n_hidden	l2	eta	accuracy-train	accuracy-valid	accuracy
30	0.01	0.001	70.4	68.12	60.97
30	0.01	0.0005	67.2	65.62	61.94
30	0.1	0.0005	65.85	64.38	62.78
30	0.01	0.0001	61.95	63.75	62.36
20	0.01	0.0001	61.6	63.75	62.22

第三步：分析預測結果

隨著訓練次數的增加，cost（Cost Function，代價函數：計算整個訓練集所有損失之和的平均值）與準確率都有所進步，如圖 19 和圖 20 所示，也沒有太明顯的過度適配現象。此時準確率為 62.36%，雖然超過半數，但仍然偏低，因此接下來試著將前一 SARIMAX 預測的殘差值作為變數加入 BPN 模型裡，以融合整合數列的特性。

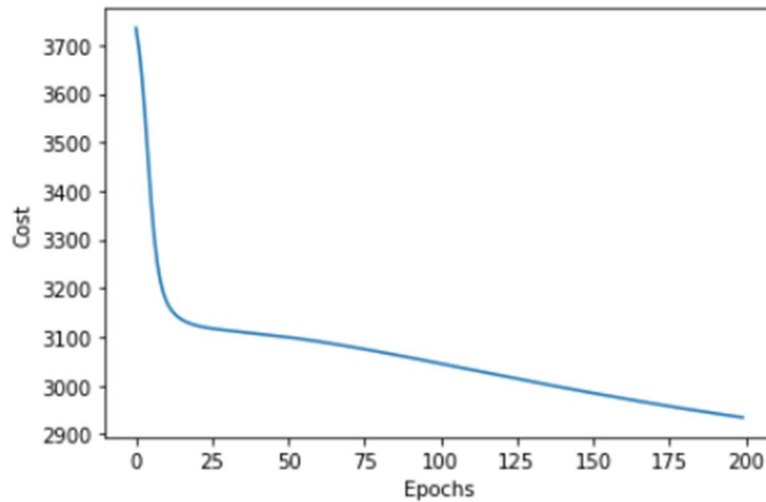


圖 19 SARIMAX 之 Cost 圖

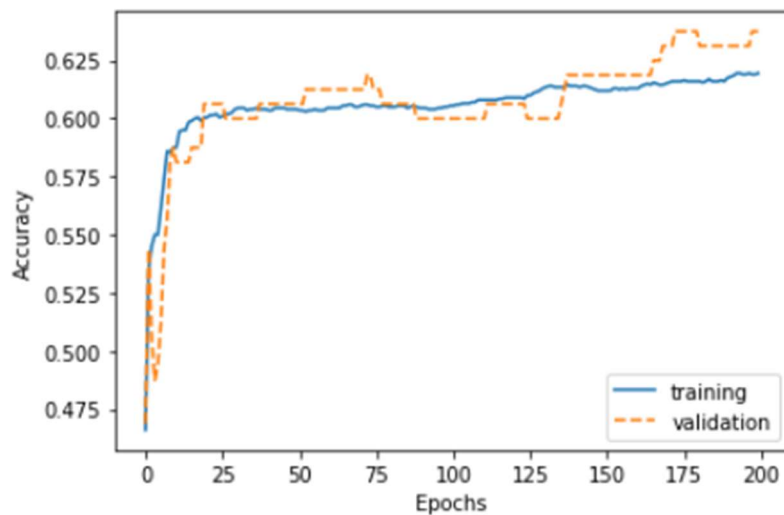


圖 20 SARIMAX 之準確率圖

(三) SARIMAX-BPN 模型

第一步：加入殘差值。

將先前 SARIMAX 模型預測得到的殘差作為變數加入 BPN 模型裡。

第二步：分析預測結果

加入殘差值，利用先前 BPN 得到的設定參數，學習率 0.0001，L2 為 0.01，隱藏單元為 30。重新預測後，有了很大的進步，cost 從 2934.24 下降至 2520.56，預測準確率也從 62.36% 提升至 69.44%，提升 7.08%，如圖 21 和圖 22，可見殘差值對預測結果有顯著影響。

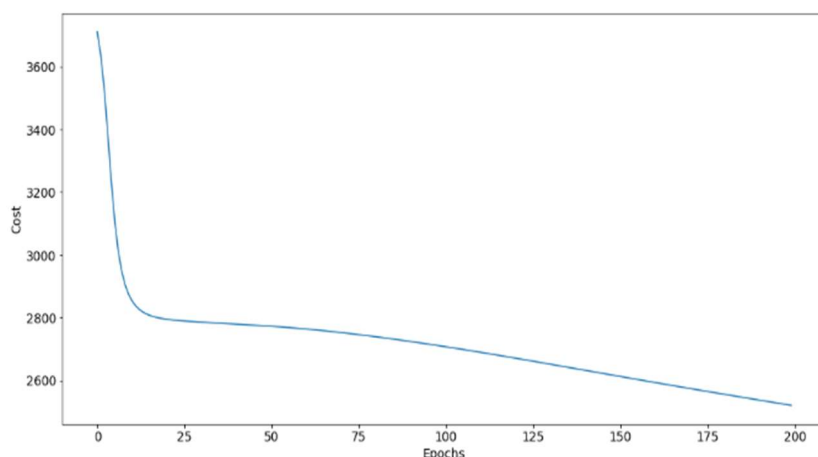


圖 21 SARIMAX-BPN 模型的 Cost 圖

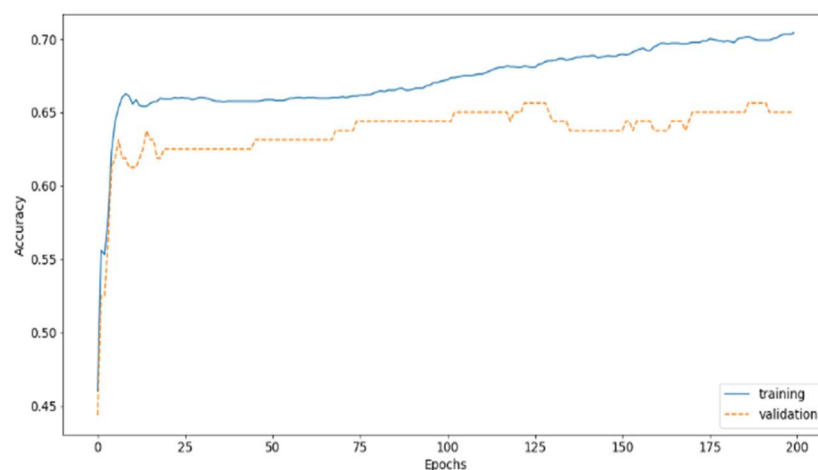


圖 22 SARIMAX-BPN 模型的準確率圖

4.5 預測結果績效評比

相比其單 SARIMAX 模型與單 BPN 模型的預測結果，將 SARIMAX 殘差值加入為變數的 SARIMAX-BPN 模型準確率更佳，符合文獻探討中的結論，將 SARIMAX 的整合特性與 BPN 的非線性特性結合利用確實能更有效的預測電力資訊。而本專題預測之準確率僅有 69.44%，不到 70%，推測可能原因為還有關鍵變數尚未考慮。另一方面可能因台電電力平台收購量也為台電預測值數據的體現，因此與實際氣候變化有些為區別，也可能參雜其他人為決策與因素。

但若僅探討預測為高或低需求且實際也為高或低需求，將預測重點聚焦在供電穩定與不穩定的時段，準確率可提升至 72.98%，顯示本專題之預測模型雖不能有效的區分出中段需求，但仍能預測出高需求或低需求時段。因此也能將其運用在預測高需求時段上，使廠商能選擇高需求時段加入市場。

5、結論

5.1 研究結論

過去的研究已指出集成學習方法能夠結合不同運算模型的優點，提升預測的準確度，故本專題研究以結合 SARIMAX 和 BPN 的混合模型，加入交易量波動的內生變數，以及日照時數、風速、溫度等外生變數，使預測模型更為準確。

本專題透過初步的圖形分析與資料拆解得知時間序列數據的季節性波動與平穩化特性後，首先使用 SARIMAX 進行時間序列預測，除考量資料之時間相依性外，並將小時、平假日與特殊日等外生變數加入模型，透過驗證資料得出模型預測之 RMSE 為 68.97，顯示單一 ARIMA 模型預測準確率略低。本研究亦使用 BPN 進行預測，標準化變數後加入模型，並將資料分為三個區間（類別），得到分類預測之準確率 62.36%。

本研究最後將兩個模型合併，將 SARIMAX 模型求得之殘差做為新變數加入 BPN 模型中，並對比先前建立之模型。經由 SARIMAX 結合 BPN 模型的預測過程與結果，將準確率提升至 69.44%，顯示線性與非線性模型相結合能互補出更加準確的預測值。其中，若聚焦於供電狀況穩定與不穩定的高與低需求，其準確率可提升至 72.98%，即變動量較大的需求有更高的可預測性。

綜上所述，集成學習比之單一預測模型擁有更高的準確率。透過結合不同類型的模型以互補缺失的因素，並完善模型，測試出的準確率明顯大幅提升。雖電力輔助市場之調頻備轉之供需可能因為市場尚未成熟，亦尚未衍生出穩定的規律，導致預測準確率略低。加上價格部分從一開始的大範圍變動到一直居高不下，也讓其失去可預測性，降低變動與競爭性。不過本研究透過將 SARIMAX 數列殘差值做為輸入值加入 BPN 模型，整合線性與非線性的特點，並確認組合後的模型擁有更好的預測能力。

本研究之模型所提出的預測能提供民間更高的參考價值，以鼓勵更多人投入此電力輔助市場，增加調頻備轉的容量，進而達到穩定供電與預防停電的目的，提升整體電力市場的效率與穩定性。

5.2 管理意涵

根據本專題建立的 SARIMAX-BPN 預測模型所提供的分析結果，以及觀察台電電力輔助市場的交易現況，可以給出以下三個重要的管理意涵：

1. 建議台電可以依照目前的交易量劃分為三種競標價格上限，如：550、600、650。在需求少時降低競標價格上限，需求高時利用高競標價格吸引中小企業來參與，使其有更自由的時段彈性，降低進入門檻並提供更合理的誘因機制。
2. 中小企業或未來想要參與市場的家庭可以參考本專題的預測模型，在台電可能為較缺電(交易量大)時，調度自己工廠或家庭的用電量，在缺電時以較佳的價格賣給台電，而其他時候自用。亦即可以挑時段參與市場。
3. 用戶代表群帶領零散用戶。透過通過台電電力交易平台合格交易者的用戶群代表，集結更多零散業者，一個用戶群代表手下有很多業者，首先，當台電電力出現短缺時，用戶群代表可以呼叫底下的業者們降低用電力使用量，緩解電力短缺。其次，也可以提供電力輔助服務。用戶群代表若將零散業者的多個發電機組和儲能設備聚集在一起，將集成可觀的電量，擁有自發電設備，或儲能設備的業者不但能間接參與台電輔助市場的電力交易，也能在用戶群代表供電不穩定時，補足其電力短缺的問題，提供更穩定的電力生態。最終，將現階段電力輔助市場台電對民間業者上下游的合作關係，擴展成為上中下游、一對多的合作關係，讓更多業者進入輔助市場，更有效利用閒置電力。

5.3 研究限制與未來研究建議

本專題預測準確率接近 70%，仍有待提升，推測原因為我們雖然以日前輔助市場之數據進行預測，但該數據（即台電的交易量）也僅為預測數值，電力傳輸當天才會再做調整。而特殊日的定義等也具主觀性，容易出現差異。

再者，模型中作為變數的天氣因子等是以大致範圍取得的，選取位於台灣中部的臺中農改場天氣作為變數，建議未來研究可根據發電廠所在地設定權重以計算更合宜的天氣因素變數，將有助於提高預測之準確率。

最後，台電輔助市場為一新興市場，現階段民間業者於其中所占比率並不

高，除了還沒太多人關注外，還有些正在觀望的廠商尚未加入。若能提高價格誘因，以鼓勵更多民間業者投入這個新興市場，將更能達到最終穩定供電之目的，未來的研究可嘗試聚焦於市場尚未到達足夠的競爭強度時的價格誘因機制建置。

參考文獻

1. 王舜葳、劉智淵、陳歆怡、紅藍創意團隊（2020）。再生能源發電預測，解讀風與光。台電月刊，第 691 期，6-8。
2. 台電電力交易平台（2022）。電力交易平台，<https://etp.taipower.com.tw/>
3. 李宜馨、陳彥銘、李明峯（2018）。動態迴歸模型於短期區域電力負載預測之研究。臺灣能源期刊，5（4），395-406。
4. 吳進忠（2021）。電力市場的權力遊戲六：預測電力交易試營運的命運。關鍵評論網，<https://www.inside.com.tw/article/23374-game-of-power-trading>
5. 姚文隆、尹居才（2002）。以自迴歸式建模傳遞網路為基礎之即時用電需量預測研究。國立高雄第一科技大學機械與自動化工程系碩士論文，高雄市。
6. 許志義、楊宗霖、蔡志祥、葉法明（2020），論市場導向之綠電發展策略。臺灣能源期刊，7（4），365-487。
7. 許志義、林振玄（2021），智慧家庭能源管理系統營運模式及其經濟分析：電能產消者 vs. 產消儲電者。臺灣能源期刊，8（3），255-276。
8. 電力市場訓練教材（2021）。台電電力交易平台，<https://reurl.cc/Gxpnmx>
9. 農二代（2021）。ARIMA 時間序列模型 python 應用-銅價格預測(一)，<https://reurl.cc/6ZdKlr>
10. 葉怡成、楊耀華、張萬鈞（2009）。ARIMA-BPN 時間數列神經網路。臺灣技術學刊，24（1），77-86。
11. 經濟部能源局（2020）。智慧電網總體規劃方案（核定本），經濟部能源局。
12. 楊豐碩、卓金和、陸臺根、于席正、許立翰、黃從葆（2018）。國外輔助服務制度之探討。台電工程月刊，第 841 期，17-46。
13. 管嫫媛（2022）。一座廠出包就讓全台停電，台灣「大電廠病」怎麼治。商業週刊，第 1791 期，34-36。
14. 管嫫媛、黃靖萱（2021）。公投、假球、台積電。商業週刊，第 1779 期，90-99。
15. 劉志文（2018）。輸電級再生能源併網情境建立及供電穩定之影響評

- 估。財團法人中技社專題報告：確保我國供電穩定之再生能源管理與運轉策略，33-57。
16. 劉志文、黃鳳翎 (2021)。以集成學習方法做極短期太陽能輸出預測。國立台灣大學電機工程學系碩士論文，台北市。
 17. 盧展南、許元禹 (2018)。電力系統短期負載預測之研究。國立中山大學電機工程學系碩士論文，高雄市。
 18. Asrari, T.X. Wu, B. Ramos (2017). A Hybrid Algorithm for Short-Term Solar Power Prediction–Sunshine State Case Study. *IEEE Transactions on Sustainable Energy*, 8 (2), 582-591.
 19. Box, G.E.P., Jenkins, G.M. (1970). Time Series Analysis: Forecasting and Control. San Francisco: Holden Day.
 20. Chen, Q., Guo, H., Zheng, K., Wang, Y. (2021). Data Analytics in Power Markets. Springer.
 21. Dickey, D.A., W.A. Fuller (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association*, 74, 427–431.
 22. J. Lee, W. Wang, F. Harrou, Y. Sun (2020). Wind Power Prediction Using Ensemble Learning-Based Models. *IEEE Access*, 8, 61517-61527.
 23. Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, E., Winkler, R. (1982). The accuracy of extrapolation (time series) methods: Results of a forecasting competition. *Journal of Forecasting*, 1 (2), 111-153.
 24. Makridakis, S., Spiliotis, E., Assimakopoulos, V. (2018). The M4 Competition: Results, findings, conclusion and way forward. *International Journal of Forecasting*, 34 (4), 802-808.
 25. McCulloch, W., Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5 (4), 115-133.
 26. Raschka, Sebastian, Vahid Mirjalili (2019). *Python Machine Learning, 3rd Ed.* Birmingham, UK: Packt Publishing.
 27. RE100 2021 Report (2021), <https://www.there100.org/stepping-re100-gathers-speed-challenging-markets>
 28. RE100 Taiwan (2022), <https://www.re100.org.tw/>

29. Rumelhart, D., Hinton, G., Williams R. (1986). Learning representations by back-propagating errors. *Nature*, 323, 533-536.
30. Seabold, Skipper, and Josef Perktold (2010). Statsmodels: Econometric and Statistical Modeling with Python. *Proceedings of the 9th Python in Science Conference*, 92-96.