

動態迴歸模型於短期區域電力負載預測之研究

李宜馨^{1*} 陳彥銘² 李明峯³

摘要

本研究分析臺灣本島負載與溫度變化的相關性，並將臺灣本島分成北部、中部、南部及東部等四個區域分別做負載預測，以兼顧區域的負載特性和氣象因素。本研究採用迴歸搭配時間序列的多方程方式進行各區域短期負載預測，依地區分別建立小時別負載預測模型，允許迴歸模型搭配動態誤差的架構，使其能夠從負載形狀的短暫變化中相對快速地恢復。以實際負載資料驗證模型的可行性，各區域的整體平均絕對誤差百分比(Mean Absolute Percentage Error, MAPE)都在3%以下，是屬於高準確的預測模型。透過電力負載預測，期望提供未來24小時的耗電量及用電趨勢，以確保合理電力調度與系統安全運轉。

關鍵詞：短期負載預測，時間序列，迴歸模型

1. 前言

隨著科技進步及生活品質提升，電力成為現代化工業不可或缺的能源，亦是普羅大眾日常賴以為生之中樞，越舒適便利的生活則越增加對電的依賴，有鑑於此，供電品質之穩定與否在國家、社會及經濟發展中扮演著關鍵角色。而電力屬於無形的商品，因此電業不同於一般製造業，是一種即時生產、即時消費的商品，沒有在製品、也幾乎沒有運輸時間，以現今的技術與經濟效益而言具有不能大量儲存之特性，電力業務有「同時同量」原則，在無存貨可備的情況下生產與消費是同時發生的，換言之，電力負載需隨著用戶需求而隨時變動，過多的生產會造成不必要的能源浪費及電力公司的供電成本增加。相反的，當負載需求大過於供電能力時則將造成跳電使得用戶及電力公

司都遭受巨大損失。因此，準確地預測電力需求量，是達成安全經濟運轉與電力系統供需平衡之首要工作，負載預測不論是在電力系統網路的運轉與安全分析、系統維護排程、電力經濟調度、緊急事故處理或燃料採購等決策中均扮演著舉足輕重的角色。根據不同的應用目的亦有不同的預測時間長度，預測時間為一年以上至數十年的長期負載預測可應用於籌備長期電源開發、電力政策研擬及規劃能源比重等。中期負載預測的時間長度為數月至數年，主要應用於年度機組維修、電價結構研擬及發電燃料採購等。而短期負載預測則為幾小時至幾週的預測時間範圍，主要用於備載容量的評估、電力潮流分析和尖離峰負載調配等以實現系統安全運轉與經濟調度。本文所探討的即為短期負載預測。

短期負載預測不但是電力負載系統調度

¹工業技術研究院綠能與環境研究所 副研究員

²工研院綠能所 研究員

³工研院綠能所 資深研究員

*通訊作者電話: 03-5917905, E-mail: YiHsinLi@itri.org.tw

收到日期: 2018年08月30日

修正日期: 2018年10月04日

接受日期: 2018年11月06日

的必要工具，尤其在未來面對電力自由化的趨勢，電力市場的價格調整，都將以短期負載預測作為主要的參考資訊。在許多過去文獻中(徐正威，2008；羅文毅，2009；林明河，2010)，氣溫是短期負載預測中廣泛使用且重要的氣象因素，原因為氣溫的變化大大影響了用戶使用空調設備等高耗電產品的情形與時間多寡。而臺灣本島雖然面積不大，但以通過嘉義等地的北回歸線(北緯23.5度)為界，以北為副熱帶季風氣候，以南為熱帶季風氣候，再加上多樣化的地理面貌影響使得氣溫分布差異甚大，造成南北氣候有明顯差距。故採取分區方式所進行的負載預測，比起直接預測全臺總體負載而言，可提高負載與氣象因素間的相關性，將有助於預測結果的準確度。本文在考慮區域與氣象對負載造成的影響後，將臺灣本島根據台電公司定義之北部、中部、南部以及東部等四個區域分別做短期負載預測，若想獲得全臺總體負載預測，可再經由區域負載之預測結果累計而成。

有鑑於電力負載預測的重要性，國內外學者提出的負載預測方法不勝枚舉，主要可區分為統計方法及人工智慧演算法兩大類。類神經網路近年來快速發展是人工智慧演算方面最廣泛應用的預測工具(Hippert *et al.*, 2001；Fan *et al.*, 2009；蘇仕博，2014)，可以處理複雜的非線性函數，具備強大的自我學習能力，本身沒有太多模型假設限制，主要的缺點是學習速度緩慢，通常需要運用大量的訓練資料來讓類神經網路學習，且在實務應用上常遇到的問題是模型變數難以直觀解釋，易造成使用者不易理解而產生後續維護上的困難。而統計方法中又以時間序列及迴歸分析為最大宗(Ramanathan *et al.*, 1997；Cui & Peng, 2015；Clements *et al.*, 2016)，此類方法應用於負載預測除了有方便且快速的特點外，還具備簡單的數學公式以及預測流程。本文即採用迴歸搭配時間序列的多方程方式，建立小時別負載預測模型，以進行各區域短期負載預測，期望能對電力負載預測提

供簡單易解釋且配適度良好的模型。

2. 研究方法

2.1 時間序列ARIMA模型

時間序列就是按照時間順序所取得的一系列觀測資料。Box與Jenkins於1970年代提出整合自迴歸移動平均(Autoregressive Integrated Moving Average; ARIMA)模型，又可稱為Box-Jenkins模型，利用時間序列變數本身去分析其隨機性，亦即資料為自己說話(Box & Jenkins, 1970)。ARIMA模型包含了三個部分，第一部分為差分(Integrated)，是將非平穩序列轉換成平穩(stationary)序列的步驟，若本就為平穩序列則不需經過差分步驟，其模型可簡化為ARMA (Autoregressive Moving Average; ARMA)，此剩餘的兩部分由自迴歸(Autoregressive; AR)模型與移動平均(Moving Average; MA)模型為基礎「混合」構成。根據參考相關文獻，將其內容分述如下(Shumway & Stoffer, 2006；Hyndman & Athanasopoulos, 2018；葉淑媚等，2007)。

(1) 自迴歸模型(AR)

假設存在時間序列 y_t ，使用落後 p 期的資料作為自變數預測未來資料，則定義為 p 階自迴歸模型，簡稱AR(p)如下所示：

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t \quad (1)$$

其中 p 為階次(order)， $\phi_1, \phi_2, \dots, \phi_p$ 為自迴歸係數，隨機誤差項 ε_t 為白噪聲(White noise)序列，其平均值為零，標準差為 σ_t 。式(1)可透過後移運算子B(backshift operators)來表示，後移運算子代表意義為 $B^i y_t = y_{t-i}$ ，則AR(p)模型可改寫成

$$\phi_p(B) y_t = \varepsilon_t \quad (2)$$

式中 $\phi_p(\cdot)$ 為 p 次多項式，稱其為 p 階自迴歸係數多項式，定義如下

$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p \quad (3)$$

(2) 移動平均模型(MA)

假設存在時間序列 y_t ，可以表達成下式，則稱此時間序列為移動平均模型，記為MA(q):

$$y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} \quad (4)$$

其中 q 為階次， $\theta_1, \theta_2, \dots, \theta_q$ 為移動平均係數，隨機誤差項 ε_t 為白噪聲序列，平均值為零，標準差為 σ_ε 。亦可透過後移運算子 B 表示為

$$y_t = \theta_q(B) \varepsilon_t \quad (5)$$

式中 $\theta_q(\cdot)$ 為 q 次多項式，稱其為 q 階移動平均係數多項式，定義為

$$\theta_q(B) = 1 + \theta_1 B + \theta_1 B^2 + \cdots + \theta_q B^q \quad (6)$$

將自迴歸模型AR(p)與移動平均模型MA(q)混合以建立自迴歸移動平均模型ARMA(p,q)，透過後移運算子 B 、 p 階自迴歸係數多項式(式3)及 q 階移動平均係數多項式(式6)簡記為

$$\phi_p(B)y_t = \theta_q(B)\varepsilon_t \quad (7)$$

其中 p 及 q 分別為自迴歸及移動平均項次，隨機誤差項 ε_t 為白噪聲序列，其平均值為零，標準差為 σ_ε 。式中的 y_t 平均值為零，若 y_t 的平均值為 μ 而非為零，則以 $(y_t - \mu)$ 替代式中的 y_t (Shumway & Stoffer, 2006)。如前所述，當 $q = 0$ ，ARMA(p,q)模型即退化為 p 階自迴歸模型AR(p)。反之，若 $p = 0$ ，則ARMA(p,q)模型即退化為 q 階移動平均模型MA(q)，所以AR模型和MA模型實際上是ARMA模型的特例。

以上所探討的AR、MA以及ARMA模型皆需要求時間序列為平穩序列，若時間序列 y_t 為非平穩序列，可經由差分轉換成平穩序列，其作法如下所示：

$$\nabla y_t = y_t - y_{t-1} = (1-B)y_t \quad (8)$$

一般稱 ∇y_t 為一階差分，故 $\nabla^d y_t$ 為 d 階差分。將非平穩序列經過 d 階差分的ARMA(p,q)模型記成ARIMA(p,d,q)，可表達成下式：

$$\phi_p(B)(1-B)^d y_t = \theta_q(B)\varepsilon_t \quad (9)$$

其中 p 是自迴歸項數， d 是時間序列在成為平穩序列前需差分的次數， q 是移動平均項數。時間序列方法著重在根據已有的歷史數據對未來進行預測，而不考慮任何外在變數，在負載的變化較複雜時或在短時間內變化較大時，其準確度將有可能大為降低。

2.2 迴歸模型

迴歸模型是了解變數間是否相關與相關方向和強度的分析方式，在一般實證分析中被廣泛的應用。將研究資料分成因變數(dependent variable)與自變數(independent variable)，並建立因變數為自變數的函數模型，以便觀察與預測所感興趣的變數。換句話說，迴歸分析的目的在于找出一條最能代表所有觀測樣本的函數，並藉由此函數來表示所感興趣的變數與其它一個或多個變數之間的關係。

迴歸分析主要用於預測或判斷變數間關係，假設隨機變數 y 與一變數 x 存在關係，可表達成：

$$y = \beta_0 + \beta_1 x + \varepsilon \quad (10)$$

稱其為簡單線性迴歸模型，其中 y 為因變數，亦稱為反應變數(Response variable)， x 為自變數，亦稱為解釋變數(explanatory variable)， β_0 及 β_1 為迴歸係數， ε 為隨機誤差。而實務上影響因變數的因素通常不只一個，多元線性迴歸(Multiple Linear Regression)模型即探討兩個或兩個以上自變數來解釋一個因變數的迴歸分析，其模型可以寫成下列矩陣形式：

$$y = X\beta + \varepsilon \quad (11)$$

$$\text{其中, } y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix},$$

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}$$

通常 y 為 $n \times 1$ 的觀測值向量， β 為 $(k+1) \times 1$ 的迴

歸係數向量， ε 為 $n \times 1$ 的隨機誤差向量， \mathbf{X} 為 $n \times (k+1)$ 的自變數矩陣，亦可稱為設計矩陣 (design matrix)。線性迴歸模型一般包含下列的基本假設：

- (i) 誤差項 ε_i 之平均數 $E(\varepsilon_i) = 0$ 且變異數 $\text{Var}(\varepsilon_i) = \sigma^2$
- (ii) 誤差項之間獨立，即表示其共變異數 $\text{Cov}(\varepsilon_i, \varepsilon_j) = 0, i \neq j$

多元線性迴歸模型可以很好地展現因變數與各個自變數間的相關性，然而卻無法直接地描述一組隨機觀測數據內部之間的相互依賴關係。

2.3 動態迴歸模型

前兩節所述之 ARIMA 及迴歸模型要不是只考慮了觀測數據內部之間的相互依賴關係，就是只考慮了其它自變數的影響，然而各個自變數和時間的綜合影響並未被考慮到。將外在自變數整合 ARIMA 模型有以下兩種方式，為了方便解釋在此假設資料為平穩序列且只考慮一個自變數，其結果也可以很容易的擴展到多個自變數或需要差分的模型，根據參考相關文獻 (Hyndman, 2010)，將其內容分述如下。

(1) ARMAX 模型 (Autoregressive Moving Average with Exogenous Variables; ARMAX)

ARMAX 模型為直接在 ARMA(p,q) 模型等號右側加入自變數，如下所示：

$$y_t = \beta x_t + \varepsilon_t + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} \quad (12)$$

其中 x_t 為時刻 t 的自變數而 β 為其係數。雖然這個公式看起來很簡單，但缺點是自變數係數 β 難以解釋，不像在迴歸模型中， β 可以解釋為每增加一單位 x_t 時對 y_t 的影響。上述等式的右邊存在因變數的滯後值，即意味著 β 只能以因變數的先前值為條件來解釋，這非常的不直觀。上述式子亦可透過後移運算子 B 表示為

$$y_t = \frac{\beta}{\phi_p(B)} x_t + \frac{\theta_q(B)}{\phi_p(B)} \varepsilon_t \quad (13)$$

可以清楚看出自迴歸係數與自變數及誤差項混合在一起。

(2) 動態迴歸模型

動態迴歸模型 (dynamic regression models) 為迴歸模型與 ARMA 模型的組合 (Regression with ARMA errors)，不同於一般迴歸模型假設誤差項獨立，在此模型中允許迴歸誤差項自相關 (Hyndman & Athanasopoulos, 2018)，其組合方式為對迴歸模型中的誤差項配適時間序列 ARMA 模型，如下所示：

$$y_t = \beta x_t + n_t \quad (14)$$

$$n_t = \varepsilon_t + \sum_{i=1}^p \phi_i n_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} \quad (15)$$

其中 n_t 為迴歸模型的誤差項， ε_t 為白噪聲序列。在此模型下，係數 β 與一般迴歸模型中的解釋方式沒有不同。上述式子也可透過後移運算子 B 表示為

$$y_t = \beta x_t + \frac{\theta_q(B)}{\phi_p(B)} \varepsilon_t \quad (16)$$

在預測能力方面，ARMAX 模型與動態迴歸模型差異不大，但動態迴歸模型的額外易解性使其具有吸引力 (Hyndman, 2010)，因此本研究採用動態迴歸模型架構，並以實際資料進行後續建模與驗證。

3. 區域電力負載預測結果

建模之電力負載資料來源為台電公司網站公告之資料，台電公司將臺灣本島劃分為北中南東等四個區域，北部區域是指龍潭超高壓變電所以北，地理位置為新竹縣鳳山溪以北。南部地區是指嘉民超高壓變電所以南，地理位置為雲林縣濁水溪以南。中部區域是指新竹縣鳳山溪以南，雲林濁水溪以北，介於北、南部區域之間。東部地區是指花蓮及臺東 (台灣電力公司, 2013)。氣象資料來源為中央氣象局網站公告之資料，選擇板橋、臺中、高雄及花蓮等四個測站記錄做為四區負載預測氣象影響因子。本文將根據台電公司定義之北中南東等四個區域分別建模做短期負載預測，若想獲得全臺總

體負載預測，可再經由區域負載之預測結果累計而成。

3.1 模型建構

資料內容為2016年7月28日到2017年11月30日期間的四區每個小時用電負載及氣象紀錄。影響用電負載變化的因素包羅萬象，且隨著預測時間長度不同，所考慮的影響變數便不相同，要建立一個優良的預測模型則要藉由適當的影響變數挑選來達成。用電負載紀錄是一個隨時間推移而變化的過程，整體而言用電負載存在一種規律性，不論是一天週期還是一週週期都有顯示出電力負載的特徵性，如圖1及圖2所示。圖1為全臺灣一天內平均每小時的用電趨勢

勢，可以發現在06:00之後從夜間低點開始非常快速地加載，並在白天保持高位，但中午12:00到13:00之間出現明顯降載，而在18:00會出現第二個高峰後才慢慢拖尾降載。圖2則為全臺灣一週內平均每小時的用電趨勢，每週的負載模式也非常明顯，週一至週五負載較高，其中週一及週二的負載比起週三到週五稍低，而週末因較多用電單位處於休息狀態使得負載明顯低於週間，且同是週末週六又與週日存在著極大的差異。由於一天24小時不同時段的負載情況大不相同，一週7日也是類似的狀況，若僅使用單一模型對所有時間點預測必須在模型內建立此用電行為週期特徵之變數，然而要同時處理日週期跟星期週期變化以及其可能的交互作用

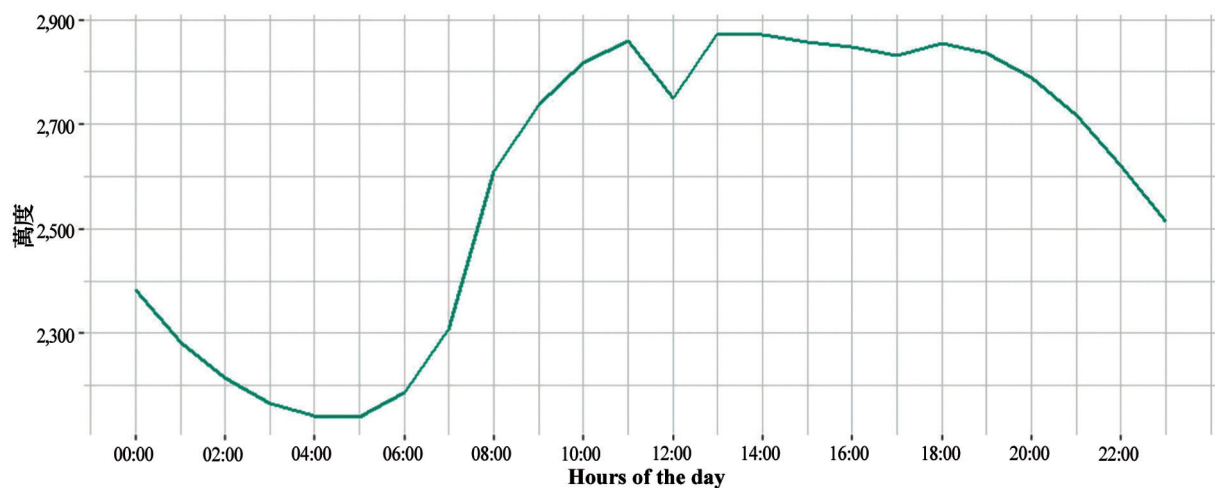


圖1 全臺灣一天內平均每小時的用電趨勢示例(本研究整理繪製)

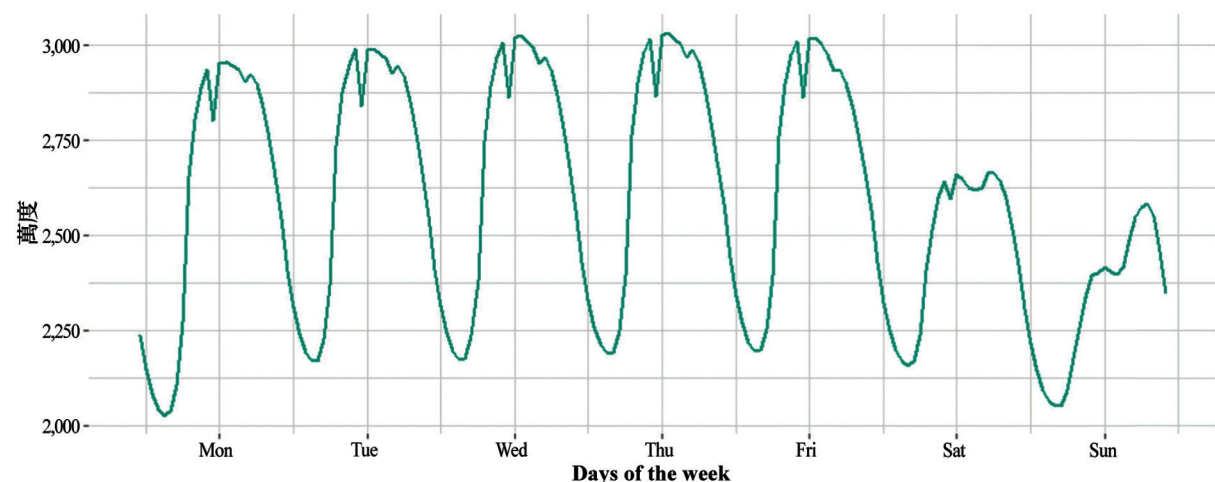


圖2 全臺灣一週內平均每小時的用電趨勢(本研究整理繪製)

將大幅增加模型複雜度，而另一方法則是直接對每個小時個別建模，避免需在模型內處理日週期變數，除減少模型複雜度外更增加了模型對其他影響變數的靈敏度。因此本研究在建立各個區域負載預測模型時，採用個別小時建模方式，並在模型中設定代表星期型態的虛擬變數(Dummy Variable)以反映出星期別間之差異。此外，工作日與假日的負載差異明顯，因此，日期資訊中除了星期型態的改變會影響用電負載變化外，是否為國定假日的影響也是相當明顯的，在模型中亦將建立虛擬變數來表示為是否為非週末的國定假日。

氣溫對於短期負載需求的波動非常重要，因為氣溫的變化影響了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。圖3為全臺小時用電負載與氣溫散布圖，可以看出氣溫在每個小時對負載的影響是有差異的，因此也顯示了依小時分別建模的必要性。由圖亦可顯而易見氣溫對負載的關係本質上是非線性的，可以大致以 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不至於太低，使用暖氣空調的比例不高，而 23°C 以上負載則是與溫度明顯呈現正相關，負載隨著溫度升高而增加，此現象則跟民眾普遍使用冷氣來達到冷房效果有關。因此，本研究將依此設定氣溫對負載的影響為發生在大於 23°C 時。

負載與時間及歷史資訊具有很強的關聯性，若前一天的總用電量很高，預計接續的該天負載應該也會很高，故本研究將前一日之總用電量放入模型當中，也將考慮是否為工作日轉變為假日以及是否為假日轉變為工作日的虛擬變數，原因為假日(包含週末及國定假日)的用電負載通常都比工作日低，舉例來說以週日的總用電量來估計週一的負載通常會低估，反之亦然，以週五的總用電量來估計週六的負載則通常會高估，故以工作日與假日間轉換的虛擬變數來做適當的調整。

考慮圖1中的負載曲線特徵，為捕獲小時別的負載變化，在建立各別區域負載預測模型時，每個地區都採用每小時單獨建模的模型結構，而本研究選用動態迴歸模型，故先根據上

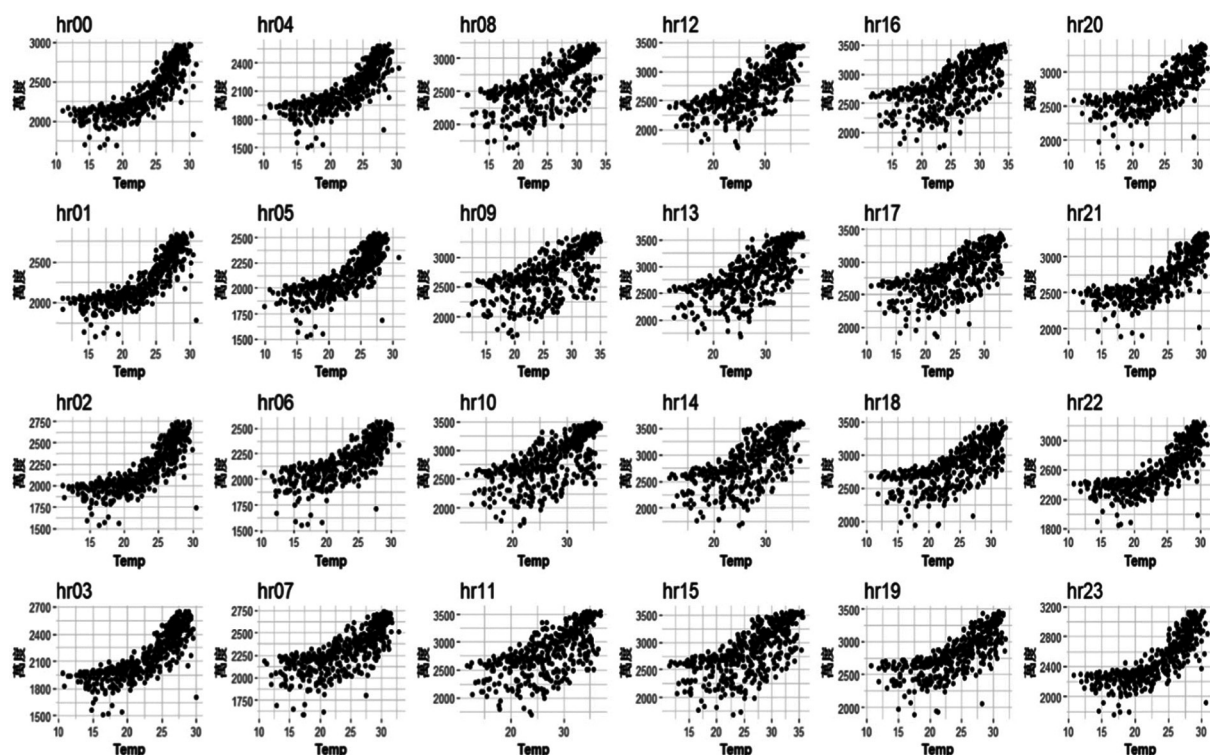


圖3 為全臺小時用電負載與氣溫散布圖(本研究整理繪製)

述所挑選之影響變數作為解釋變數配適迴歸模型，再將迴歸模型中之誤差項配適ARMA(1,1)模型。以 a 表示所屬區域， h 表示所屬小時， d 表示所屬的天，則反應變數 L_{ahd} 表示為區域 a 第 d 天在 h 小時的用電負載，其模型架構如下所示

$$L_{ahd} = \beta_{0ah} + \beta_{1ah} C_{ahd} + \beta_{2ah} DL_{ad-1} + \beta_{3ah} H_d + \beta_{4ah} HW_d + \beta_{5ah} WH_d + \sum_{p=1}^6 \alpha_{pah} W_{pd} + n_{ahd} \quad (17)$$

$$n_{ahd} = \varepsilon_{ahd} + \phi_{ah} n_{ahd-1} + \theta_{ah} \varepsilon_{ahd-1} \quad (18)$$

其中 $a=1,2,3,4$ ， $h=1,2,\dots,24$ ， n_{ahd} 為迴歸模型誤差， n_{ahd-1} 表示為同區域同一小時前一天的迴歸模型誤差， ε_{ahd} 為隨機誤差， ε_{ahd-1} 表示為同區域同一小時前一天的隨機誤差。自變數依序如下詳述，以 $Temp_{ahd}$ 表示為區域 a 第 d 天在 h 小時的氣溫，而為表示大於 23°C 的氣溫影響定義

$$C_{ahd} = \begin{cases} 0 & Temp_{ahd} \leq 23 \\ Temp_{ahd} - 23 & Temp_{ahd} > 23 \end{cases}$$

DL_{ad-1} 表示為區域 a 第 d 天的前一天之日總用電量， H_d 表示為第 d 天是否為非週末的國定假日之虛擬變數， HW_d 表示為第 d 天是否假日轉變為工作日的虛擬變數， WH_d 表示為第 d 天是否工作日轉變為假日的虛擬變數， W_{pd} 表示為第 d 天星期型態的虛擬變數， $p=1,\dots,6$ 。因此，對於特定區域(a)特定小時(h)的模型而言，其模型參數僅基於包含該區域該小時的觀測數據子集來估計之。此種迴歸模型搭配動態誤差的架構，使得昨天錯過的因素將預期可在今天部分重現，該模型的這一特性使其能夠從負載形狀

的短暫變化中相對快速地恢復。

3.2 用電負載預測驗證

本研究將各個區域各個小時之用電資料依上一小節所述模型分別建模，即表示每一個區域都有其所有共24個小時用電模型，每一個模型預測一天當中的一個小時，整合起來即可預測一日24小時之負載。為驗證模型的可行性本研究將所得之觀測資料切分為訓練資料(training data)與測試資料(testing data)，將2016年7月28日到2017年11月30日期間的用電紀錄切割出完整的一年2016-07-28至2017-07-27 (365天)做為訓練資料，剩餘的2017-07-28至2017-11-30 (126天)做為測試資料進行預測測試，並以平均絕對誤差百分比(Mean Absolute Percentage Error, MAPE)表示預測模型的好壞，MAPE為相對數值，若越接近零表示預測效果越好，且不受觀測值單位與大小之影響，能夠客觀取得實際值與預測值間之差異程度，其計算公式如下所示

$$MAPE = \frac{1}{N} \sum_{n=1}^N \left| \frac{Y_n - \hat{Y}_n}{Y_n} \right| \quad (19)$$

其中 Y_n 表實際值， \hat{Y}_n 表預測值， N 為評估時間內的預測值數目。

3.2.1 北部用電預測

表1為北部區域的MAPE(%)驗證分析，提供了詳細的預測性能細分，整體的負載預測MAPE為1.88%，最大的絕對誤差百分比為23.91%出現於颱風天，而絕對誤差百分比

表1 北部區域的MAPE(%)驗證分析(本研究整理)

	Overall MAPE	Maximum APE	No. APE >5%	No. APE >10%	No. APE >15%	No. APE >25%	Obs. (個)
Overall	1.88	23.91	146	27	10	0	2,856
Normal	1.54	21.68	48	4	2	0	1,800
Weekend	1.89	7.42	22	0	0	0	696
Holiday	4.18	16.61	19	11	2	0	72
Typhoon	3.35	23.91	57	12	6	0	288

註：表中obs.個數已剔除缺失值個數。No. APE > 5表示為APE大於5%的個數，以此類推。

(Absolute Percentage Error, APE)大於5%的個數為146個，表示絕大多數預測誤差都低於5%。週末的MAPE為1.89%，正常工作日的MAPE為1.54%，其中最大的絕對誤差百分比為21.68%出現於2017-08-15 17:00，原因為該天發生臺灣中油對台電大潭發電廠的天然氣供應管線突然停止運作使得機組跳脫，在備載容量不足的情況下為避免全面崩潰，供電系統啟動保護措施而緊急分區停電，繼而引發2017-08-15 16:51起臺灣本島各地發生大規模停電事件。總體而言，在正常情況下不論是週間的正常工作日還是週末的MAPE都良好的維持在2%以下，而國定假日的MAPE為4.18%雖小於5%，但對於一般情況下的工作日及週末而言算是相對的大，原因可能為一年當中國定假日出現的次數相對少，也就是說其訓練樣本數相對不足，且負載變化又大，因此相對地不易預測，另一特殊情況則為颱風所造成，颱風天氣象變化多端又是大風又是大雨，且可能致災放假，故颱風天的MAPE為3.35%亦屬相對的大，上述兩種特殊情況都較容易出現重大預測誤差，誠如表中所顯示的大於10%或15%的絕對誤差百分比大多都落在國定假日及颱風天的情況下。

3.2.2 中部用電預測

表2為中部區域的MAPE(%)驗證分析，根據不同情況條件下詳細的分析預測性能，整體的負載預測MAPE為2.07%，而最大的絕對誤差百分比為15%出現於國定假日的情況下，絕

對誤差百分比大於5%的個數為160個，亦表示絕大多數預測誤差都低於5%。週末的MAPE為1.97%，正常工作日的MAPE亦為1.97%，而正常日當中的重大誤差也多是由2017-08-15臺灣本島各地發生大規模無預警停電事件所引起的。總體而言，在正常情況下不論是週間的正常工作日還是週末的MAPE都良好的維持在2%以下。與北部區域負載預測一樣，特殊情況下的國定假日與颱風天的MAPE都比正常情況下相對的大，分別為3.49%及2.59%但都小於5%，而大於10%的重大誤差亦大多都落在國定假日及颱風天的情況下，但跟北部區域負載預測不同的是中部區域沒有大於15%以上的負載預測誤差。

3.2.3 南部用電預測

表3為南部區域的MAPE(%)驗證分析，依據不同情況條件下詳細的分析預測性能，南部區域負載預測不論是整體的MAPE還是大於5%的絕對誤差百分比個數都是四個區域中最低的，整體的負載預測MAPE為1.67%，而最大的絕對誤差百分比為15.33%出現於颱風天的情況下，絕對誤差百分比大於5%的個數為99個，表示絕大多數預測誤差都低於5%。週末的MAPE為1.56%，正常工作日的MAPE為1.53%，總體而言，在正常情況下不論是週間的正常工作日還是週末的MAPE都維持在1.5%上下相當良好，而正常日當中大於10%的重大誤差出現於連續兩個颱風尼莎及海棠入侵後的2017-08-

表2 中部區域的MAPE(%)驗證分析(本研究整理)

	Overall MAPE	Maximum APE	No. APE >5%	No. APE >10%	No. APE >15%	No. APE >25%	Obs. (個)
Overall	2.07	15.00	160	16	0	0	2,856
Normal	1.97	14.43	74	7	0	0	1,800
Weekend	1.97	9.04	42	0	0	0	696
Holiday	3.49	15.00	15	7	0	0	72
Typhoon	2.59	10.25	29	2	0	0	288

註：表中obs.個數已剔除缺失值個數。No. APE > 5表示為APE大於5%的個數，以此類推。

表3 南部區域的MAPE(%)驗證分析(本研究整理)

	Overall MAPE	Maximum APE	No. APE >5%	No. APE >10%	No. APE >15%	No. APE >25%	Obs. (個)
Overall	1.67	15.33	99	15	1	0	2,856
Normal	1.53	12.82	41	4	0	0	1,800
Weekend	1.56	6.56	11	0	0	0	696
Holiday	2.29	5.99	3	0	0	0	72
Typhoon	2.60	15.33	44	11	1	0	288

註：表中obs.個數已剔除缺失值個數。No. APE > 5表示為APE大於5%的個數，以此類推。

01，而大於5%的預測誤差中也有由2017-08-15全臺大停電事件所引起的。與其他區域負載預測一樣，特殊情況下的國定假日與颱風天的MAPE都比正常情況下相對的大，分別為2.29%及2.6%，而大於10%及15%的重大誤差亦大多都落在國定假日及颱風天的情況下。

3.2.4 東部用電預測

表4為東部區域的MAPE(%)驗證分析，由於東部區域負載相對於其他三個區域而言相對的小，平均小時負載才為37.12萬度(北部955.28萬度，中部706.49萬度，南部891.97萬度)，而MAPE為預測值與實際值間的差距相對實際值而言的比例，因此在計算MAPE時會比其他是個區域稍微大一些。整體的MAPE為2.88%，而正常工作日及週末的MAPE分別為2.79%及2.80%，都小於3%預測表現依然是相當不錯。

3.3 結果分析與討論

據研究成果推測預測誤差除了隨機因素

外，主要可能受以下幾種原因影響：

(一) 氣象資料代表性

研究中所採用之氣象站資料皆選擇在代表性的主要人口及用電稠密區，例如北部區域選擇以板橋氣象站為代表，而氣候條件上會與代表氣象站出現較大差異地區多是在相對較為偏遠的山區或海區，此類地區其用電量占比較小，對整體預測誤差不至於產生巨大影響，但仍有可能在某些狀況下即使在用電稠密地區內也有較大的氣候差異，如夏日的對流雲或地形雨，冬天的季風影響等，當同一區域內產生較大氣象差異時就可能造成預測模型較大之誤差。

(二) 特殊日或事件

由表1至表4的誤差統計可以看出最大誤差多出現在特殊事件，如2017-08-15大停電，或者出現在特殊日，例如北部、南部為颱風日，中部為中秋節假日，東部為突發之高用電日，大體來說特殊日的預測誤差皆較平常日(週間日)高，而因為特殊日的重要用電行為影響因子

表4 東部區域的MAPE(%)驗證分析(本研究整理)

	Overall MAPE	Maximum APE	No. APE >5%	No. APE >10%	No. APE >15%	No. APE >25%	Obs. (個)
Overall	2.88	78.52	440	53	14	4	2,856
Normal	2.79	74.54	257	22	8	1	1,800
Weekend	2.80	78.52	96	12	2	2	696
Holiday	3.10	12.75	13	5	0	0	72
Typhoon	3.64	27.41	74	14	4	1	288

註：表中obs.個數已剔除缺失值個數。No. APE > 5表示為APE大於5%的個數，以此類推。

可能與平常日有很大差異，因此以平常日為主體之資料所訓練出來的模型無法準確預測特殊日的用電狀況，且因為特殊日資料累積的速度慢，即使獨立建模資料量亦為影響準確度因素之一。

4. 結 論

負載預測是電力公司研擬各種營運計畫的基礎。因此，準確的負載預測是相當重要的課題。本研究透過電力負載預測，期望提供未來 24 小時的耗電量估計及用電變化趨勢。將臺灣本島依北部、中部、南部及東部等四個區域分別作負載預測，期望透過分區的方式避免因區域太大可能造成的多種區域負載特性與氣候類型因稀釋而影響預測的準確度，且研究成果也顯示區域越小負載與氣象因素的關係性則會愈高，進而可以得到更佳的預測結果。此外，個別區域的預測結果也可提供電業對特定區域做配電層級之運作規劃，該結果亦能顯示該區域的負載特性，亦可提供更細緻的遠期輸配電系統規劃以及需求面管理措施制定的參考。

本研究採用迴歸搭配時間序列的多方程方式進行各區域短期負載預測，依地區分別建立小時別負載預測模型，期望能對電力負載預測提供簡單易解釋且配適度良好的模型，有別於人工智慧模型有學習速度緩慢之缺點且模型架構難以解釋。而相對於文獻中的其他模型而言，所提出的模型中最重要的區別因素可能是著重在同一小時的負載影響，並允許迴歸模型搭配動態誤差架構的靈活性，使得昨天錯過的因素將預期可在今天重現，這一特性使其能夠從負載形狀的短暫變化中相對快速地恢復，可以產生短期預測績效的重要改進。以實際負載資料，透過平均絕對誤差百分比(MAPE)的驗證，達到各區域的整體平均絕對誤差百分比(MAPE)都在 3% 以下，是屬於高準確的預測模型，可對確保合理且經濟的電力調度與系統安全運轉做出有效貢獻。

對於電力公司來說必須先能夠預測短中長期負載的變化，才能夠做好相應的準備以有效地調度供電機組，維持電網可靠運轉裕度，以及更經濟地配置電力設施投資。本研究成果產出屬於日前逐時負載預測，應用上可作為電力公司在進行機組併離網排程之重要決策參考，因發電機組依不同燃料價格及發電量可區分為基載(如核能、燃煤)，中載(如燃氣)，尖載(如氣渦輪、燃油、水力)，各機組除發電量和發電成本差異甚大外，起動停機所需的前置時間亦不同，實務上必須根據逐時負載預測進行各機組的併離網時間排程，才在能夠滿足負載的條件下減少高價機組運轉或不必要的機組待機以降低整體發電成本，此外電力公司另需針對當日最高負載預測值進行備轉容量的準備以應付機組故障、預測誤差或者其他突發狀況。個別區域的負載預測結果也可提供電業對特定區域做配電層級之運作規劃，在配電操作時時常面對電力負載移轉最佳化問題，須解決分配變電所多餘的電力給其它不足的變電所，然而仍需依靠負載預測才能更精準掌握移轉量及時機，在維持電網穩定地前提條件下，使得流失的能源最少或是成本最低獲致分配最佳決策。

誌 謝

本文承蒙經濟部能源局之計畫經費補助，謹此致謝。

參考文獻

- 葉淑媚、李佳樺及許天維，2007。ARIMA 模式分析與預測—以鴻海股票市場日收盤價與報酬率為例，臺中教育大學學報：數理科技類，21(2)，第 51-69 頁。
- 徐正威，2008。考慮氣象因素之分區短期負載預測，大同大學電機工程研究所碩士論文，大同大學。
- 羅文毅，2009。基於氣象資料的配電饋線負載

- 預測之研究，國立臺灣科技大學電機工程學系碩士論文，國立臺灣科技大學。
- 林明河，2010。短期饋線負載預測模型之研究，國立臺灣海洋大學電機工程學系碩士論文，國立臺灣海洋大學。
- 蘇仕博，2014。類神經網路於短期電力負載預測之研究，國立臺灣海洋大學航運管理學系研究所碩士論文，國立臺灣海洋大學。
- 台灣電力公司，2013。台灣電力股份有限公司輸電系統規劃準則。
- Box, G.E.P. and G.M. Jenkins, 1970. Time series analysis: Forecasting and control, San Francisco: Holden-Day.
- Cui, H. and X. Peng, 2015. Short-Term City Electric Load Forecasting with Considering Temperature Effects: An Improved ARIMAX Model. *Mathematical Problems in Engineering*, 2015, Article ID 589374, 10 pages.
- Clements, A.E., A.S. Hurn and Z. Li, 2016. Forecasting day-ahead electricity load using a multiple equation time series approach. *European Journal of Operational Research*, 251, 522-530.
- Fan, S., L. Chen and W.J. Lee, 2009. Short-term Load Forecasting Using Comprehensive Combination Based on Multimeteorological Information. *IEEE Trans. On Industry Applications*, 45, 1460-1466.
- Hippert, H.S., C.E. Pedreira and R.C. Souza, 2001. Neural Networks for Short-Term Load Forecasting: A Review and Evaluation. *IEEE Trans. on Power Systems*, 16, 44-55.
- Hyndman, R.J., 2010. The ARIMAX modelmuddle, Retrieved August, 2018, from <https://robjhyndman.com/hyndsight/arimax/>.
- Hyndman, R.J. and G. Athanasopoulos, 2018,. Forecasting: principles and practice, Retrieved August, 2018, from <https://otexts.org/fpp2/>.
- Ramanathan, R., R. Engle, C.W.J. Granger, F. Vahid-Araghi and C. Brace, 1997. Short-run forecasts of electricity loads and peaks. *International Journal of Forecasting*, 13, 161-174.
- Shumway, R.H. and D.S. Stoffer, 2006. Time Series Analysis and Its Applications: With R Examples. Springer, New York.

A Study of Short-term Regional Power Load Forecasting By Dynamic Regression Models

Yi-Hsin Li^{1*} Yen-Ming Chen² Ming-Feng Lee³

ABSTRACT

The object of this research is to analyze correlations between electric load and temperature in Taiwan. In order to adapt to local weather and usage characteristics, regional load forecasting models of norther, middle, southern and eastern Taiwan are separately built and explained in this study. These short-term forecasting models integrate both regression and time series methods to predict regional hourly load profiles, and dynamic error adjustments are also applied for the purpose of quick recovery when sudden load changes happen. The models are tested by real data and verified as high accuracy with MAPE less than 3%. Through this study of electric load forecasting, the total usage and profile changes in future 24 hours could be obtained to improve grid security and reasonable unit scheduling.

Keywords: short-term load forecasting, time series, regression model

¹ Associate Researcher, Green Energy & Environment Research Laboratories, Industrial Technology Research Institute.

² Researcher, GEL, ITRI.

³ Senior Researcher, GEL, ITRI.

*Corresponding Author, Phone: +886-5917905, E-mail: YiHsinLi@itri.org.tw

Received Date: August 30, 2018

Revised Date: October 4, 2018

Accepted Date: November 6, 2018