

# 逢甲大學學生報告 ePaper

報告題名：

## 預測國內能源電力消費

Forecast For Domestic Electric Power Expenses

作者：吳坤益、王俊勛、呂冠霖、林欣霓、林強相

系級：統計學系三乙

學號：D9765656、D9765660、D9765554、D9765906、D9765851

開課老師：陳婉淑 教授

課程名稱：統計預測方法

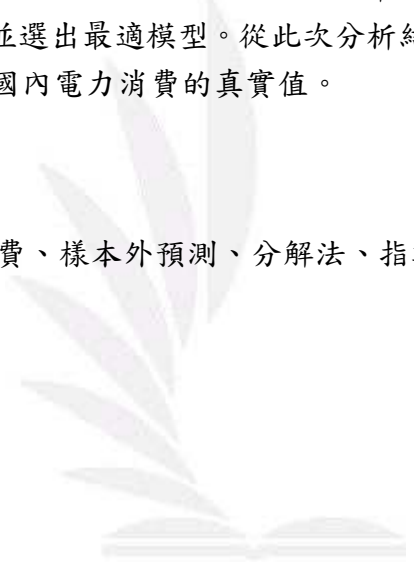
開課系所：統計學系

開課學年：99 學年度 第二 學期

## 中文摘要

近年全球皆面臨了能源的危機。不僅如此，能源的使用過度還使氣候有了異常的變遷，讓各國不得不開始重視節約能源與適當使用能源的重要性。其中電力也是各項能源中使用十分頻繁的，為防止全球暖化可從節能減碳，節約能源的方式立即落實。台灣是否也在討論這個議題及推廣許多活動之後能由國內每年每月消費電力來看出顯著成效？為了瞭解國內的電力消費，是否隨著政府的政策推廣與宣導而有效的減少用電量，我們使用 AREMOS 經濟統計資料庫中的「國內能源消費－電力」，時間為西元 1990 年 1 月至 2010 年 12 月的 252 筆月資料，並且保留最後 12 筆來做樣本外預測，以分析歷年來國內電力使用量的狀況。我們應用四種統計預測模型或方法，依序是時間序列迴歸法、指數平滑法、分解法、ARIMA 模型配適資料。最後利用 MSE、MAD、MPE、MAPE 四個準則評估樣本外預測，來評估這四種方法何者最佳，並選出最適模型。從此次分析結果能看出，分解法的配適模型最能有效的預測出國內電力消費的真實值。

**關鍵字：**ARIMA、電力消費、樣本外預測、分解法、指數平滑法



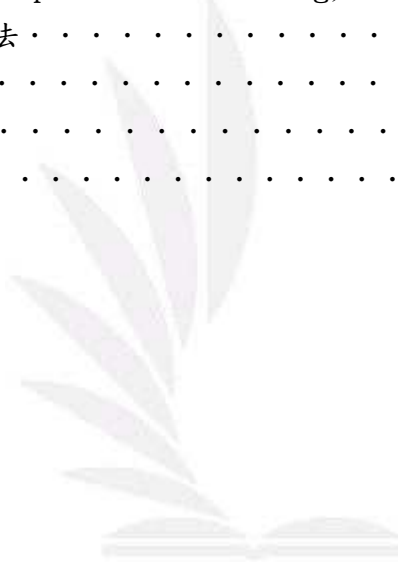
## Abstract

In order to better understand electric power expenses in Taiwan, we study “the domestic energy use - electric power” which was obtained from the AREMOS economic statistics information database. The time period is from January 1990 to December 2010, total 252 observations. We save the last 12 observations for out-of-sample forecasts. We apply four kinds of statistical methods/models for forecasting, which included the time series regression, exponential smoothing, decomposition method, and the ARIMA model to model this time series. In order to evaluate the forecast performance, we employ MSE, MAD, MPE, and MAPE four criteria to evaluate the performance of the four methods. Based on MAPE criterion, we found that the decomposition method with X11 and regression technique has the best performance (about 2.3%) among the four methods which is able to effectively forecast domestic electric power to expend in Taiwan for the future .

**Keywords** : ARIMA; Electric power expense; out-of-sample forecast;  
Decomposition Method; X11, Exponential Smoothing.

## 目次

第一章 緒論	5
第二章 研究方法	8
第一節、預測方法及模型	9
第二節、診斷分析方法	13
第三節、預測評估準則	13
第三章 資料分析	15
第一節、原始資料分析	15
第二節、時間序列迴歸法(Time Series Regression)	16
第三節、分解法(Decomposition Method)	20
第四節、指數平滑法(Exponential Smoothing)	24
第五節、ARIMA 分析法	25
第六節、最適模型	30
第四章 結論與建議	32
參考文獻	33



## 圖目錄

圖 2-1.1	研究流程圖	9
圖 3-1.1	原始時間序列圖	15
圖 3-2.1	原始資料取 log 後之圖	16
圖 3-2.2	時間序列分析法曲線圖	17
圖 3-2.3	實際值、估計值和 95%的上、下界預測區間曲線圖(時間序列法)	19
圖 3-3.1	原始時間序列圖	20
圖 3-3.2	去季節因子時間序列圖	20
圖 3-3.3	趨勢循環時間序列圖	21
圖 3-3.4	季節因子之時間序列圖	21
圖 3-3.5	不規則因子時間序列圖	21
圖 3-3.6	分解法曲線圖	22
圖 3-3.7	實際值、估計值和 95%的上、下界預測區間之曲線圖(分解法)	23
圖 3-4.1	原始資料的時間序列圖	24
圖 3-4.2	實際值、預測值、95%預測信賴上界、95%預測信賴下界之曲線圖	25
圖 3-5.1	原始資料時間序列圖	25
圖 3-5.2	經過取 log 後的時間序列圖	26
圖 3-5.3	一次差分圖	26
圖 3-5.4	一次差分後的 ACF 與 PACF 圖	27
圖 3-5.5	經過一次差分與季節差分後的 ACF 與 PACF 圖	27
圖 3-5.6	配適模型後的 ACF 與 PACF 圖	28
圖 3-5.7	白噪音與單根檢定檢定圖	29
圖 3-5.8	ARIMA 的估計值、實際值和 95%的上下界預測區間的預測圖(2010 年)	30

## 表目錄

表 1-1.1	取自台灣電力公司民國 99 年發電量表 . . . . .	7
表 2-3.1	MAPE 準則表 . . . . .	14
表 3-4.1	參數估計表(指數平滑法) . . . . .	24
表 3-5.1	ARIMA 模型的參數估計 . . . . .	29
表 3-6.1	模型比較表 . . . . .	31



## 第一章 緒論

在過去，科學家、工程師們已經知道利用許多原有的資源去產生出更多未知的資源。例如：知道煉石油可以得到汽油、塑膠類製品等；燃煤可發電；核能發電要使用鈾 235 及鈾 238 等發現。但使用這些資源的過程中，大家並沒有想到有一天這些資源會枯竭。為了防止惡化，科學家和政府無所不用其極的發展與推廣許多方法—石油代替品、綠建築、油電混合引擎、再生能源.....。

近年全球皆面臨了能源的危機。不僅如此，能源的使用過度還使氣候有了異常的變遷，讓各國不得不開始重視節約能源與適當使用能源的重要性。最近日本由於海嘯的強大衝擊，福島核電廠輻射外洩的問題，使得核電廠這個話題再次變成新聞焦點，其中電力也是各項能源中使用十分頻繁的，幾乎每一個人、每一個場所，都在使用電力。最近日本由於海嘯的強大衝擊，福島核電廠輻射外洩的問題，使得核電廠這個話題再次變成新聞焦點，但因核電廠的經濟效益較高且發電時不產生二氧化碳(化石燃料發電皆會產生大量的二氧化碳)。為了日益增加的用電需求量，政府須兼顧自然環境、經濟發展、不願意興建核電廠的民意，因此誕生許多節能減碳的推廣活動。

全球面臨了氣候暖化和能源危機，近年來氣候急遽的轉變和能源的即將耗盡，由於氣候暖化與能源使用有很大的關聯，使得各個國家不得不開始重視節約能源的重要。雖然電力的使用並不像石油或煤礦這類資源會用盡枯竭是因為電是「生產」出來的。電本身並不會對自然環境或人體有多大的影響，但是在產電供電或用電時卻可能對我們環境或健康有危害甚至是生命安全的疑慮。

電與氣候暖化的直接關係即是電器若耗電量為 1 度，發電時所排放的二氧化碳(溫室氣體)為 0.62 公斤。科學家提到除非我們減低導致全球暖化的溫室氣體的排放量，否則全球的平均氣溫在本世紀末將上升三到九度。依照台電的 2007 年年報中，各行業別的用電統計資料，家戶用電佔全台灣用電量的 2%；如果放寬範圍，將運輸、服務、娛樂、金融、通訊、電力業的用電，全部劃為民生用電，用電量也只佔全台的 24%。徐世輝(2008)加上科技的進步與以機械化取代人工的生產模式，未來對於電的需求勢必會日益增加。

台灣雖然面積小，但是核電廠的密度卻是佔全球第一，面對日本的前車之鑑，台灣應該如何自處？政府一方面降低工業用電的價格期望外來廠商的投資能增加，另一方面受民意影響卻又希望未來能夠避免建造新的核電廠。是否用電量在政府推廣的眾多節能活動依舊居高不下或有所改善？藉由這次的分析，我們將來討論這個議題。

環保議題，是近年來最熱門的議題，從政府高層單位到鄉間小公所上上下下都圍繞著節能的政策。根據環保署所推動的省電大作戰中提及的生活小動作，例如：空調—調製 28 度 C，平均一個月可省下 9.6 度；將傳統 60W 鎢絲燈泡改為 11W 省電燈泡，一個月可省下 36.8 度；吃飽飯後不要急著做下休息看電視，可到戶外散散步，不僅可消化更可省下 13.5 度。但是，是否政策一推行就都如風行草偃一般成效顯著？例如最近的日本大地震，因為重創了核電場，讓這幾天日本發佈省電政策預防供電量不均或不足。日本民眾上下其心一起遵守，連企業也放寬開例，上班族們不再被規定要穿西裝打領帶，而可以穿著”省電服裝” POLO 衫加牛仔褲，整體省電政策執行成效良好。台灣是否也在討論這個議題及推廣許多活動之後能由國內每年每月消費電力來看出顯著成效？我們利用歷史資料，試圖瞭解國內的電力消費是否隨著政府近年來的宣導而有效的減少用電量。

台灣用電可分為民生用電與產業用電，由徐世輝(2008)的報導，民生用電佔全台灣用電量的百分之二。若放寬標準，將除了工業用電以外的產業用電都視為民生用電，也只佔全台灣用電量的百分之二十四，所以我們可以知道台灣國內的消費用電是以工業用電為主。

國內電力的發購狀況，由表 1-1(取自台灣電力公司民國 99 年發電量表)可以看到國內的火力發電就佔了百分之七十二點一，核能發電佔百分之十九點三，也就是說對於不會產生二氧化碳的產電方式我們使用的並不頻繁，因此在最後我們也可以拿來討論是否能夠降低這種不環保的發電方式。

為了瞭解國內的電力消費，是否隨著政府的政策推廣與宣導而有效的減少用電量，我們使用 AREMOS 經濟統計資料庫中的「國內能源消費—電力」，時間為西元 1990 年 1 月至 2010 年 12 月的 252 筆月資料，並且保留最後 12 筆來做樣本外預測，以分析歷年來國內電力使用量的狀況。我們應用四種統計預測模型或方法，依序是時間序列迴歸法、指數平滑法、分解法、ARIMA 模型配適資料。這四種預測模型或方法都是非常普遍且廣為使用的預測方法。最後利用 MSE、MAD、MPE、MAPE 四個準則評估樣本外預測，來評估這四種方法何者最佳，並選出最適模型。



發電量			
項目	發電量 (億度)	構成比 (%)	較上年增加率 (%)
登購電量合計	2073.8	100.0	7.1
台電自有	1577.9	76.1	10.1
抽蓄水力	30.5	1.5	-7.4
火力	1107.6	53.4	14.5
核能	400.3	19.3	0.1
再生能源	39.5	1.9	18.7
購電	495.9	23.9	-1.3
民營火力	388.5	18.7	3.1
汽電共生	95.0	4.6	-16.1
再生能源	12.4	0.6	-0.4

表 1-1 取自台灣電力公司民國 99 年發電量表



## 第二章 研究方法

本報告的研究流程如圖 2-1.1。首先先擬定研究的議題：我們將議題的討論範圍鎖定在近年來大家持續關注的能源問題以及環保意識的興起與宣導，加上與時事配合(福島輻射外洩事件)，我們決定討論國內電力的消費預測以及期望藉由分析結果能了解目前國內發電量與發電方法是否未來有改善的必要。

接下來找尋合適的資料：我們利用 AREMOS 經濟統計資料庫找尋有關國內能源電力消費的部分，我們由選取出的數筆資料中再利用 Excel 做折線圖做基礎分析找出季節性最明顯的資料(因為用電量會隨不同季節有不同用量)為最合適資料—「國內能源消費—電力」，單位為公秉油當量(KLOE)。期間從西元 1990 年 1 月至 2010 年 12 月，總共 252 筆月資料。我們將總數 252 筆的月資料分為兩部分，第一部分為前 240 筆資料，即樣本內資料，主要用來建立模式分析與預測未來的走勢。第二部分為最後的 12 筆資料，即樣本外資料，主要用以與所預測出的結果做對照以了解模型的預測程度是否在一定的水準上。而樣本外的十二筆資料，即從西元 2010 年 1 月到 2010 年 12 月。

接著我們依序使用時間序列迴歸法、分解法、指數平滑法及 ARIMA 配適模型，並對 2010 年的消費量做預測。ARIMA 模型配適先判斷原始資料是否有趨勢，以及其平均數和變異數是否平穩：意指數據均不只要均值的上下波動且不可有明顯上升或下降的趨勢，若序列有此情況發生需對其進行差分使原始序列平穩化。最後用 MSE、MAD、MPE 及 MAPE 四個準則來做評估，並選出最佳的配適模型。

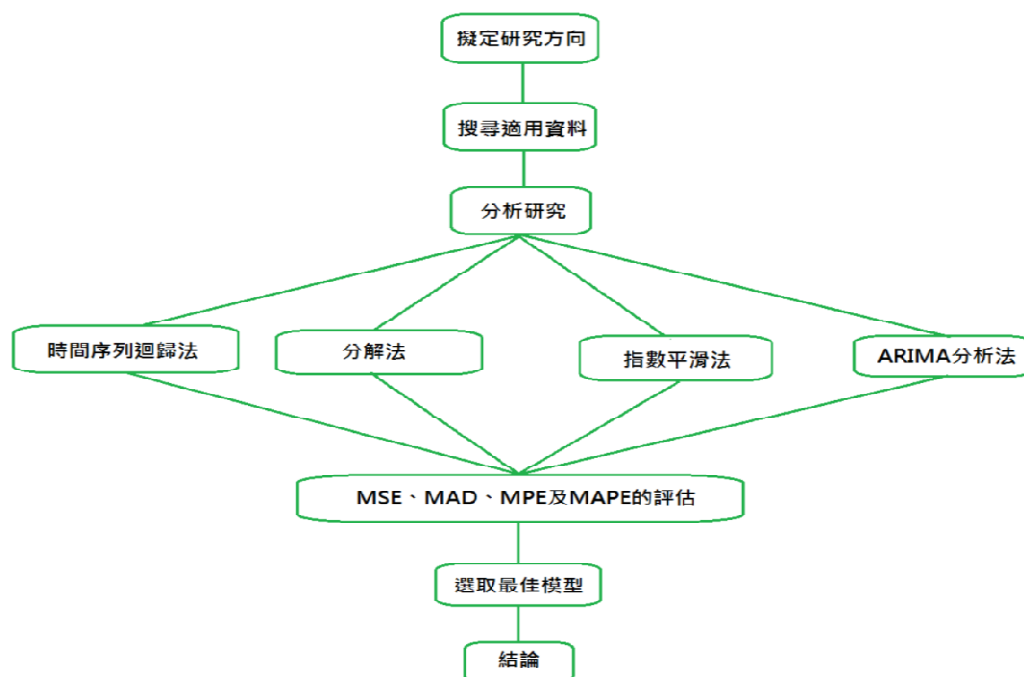


圖 2-1.1 研究流程圖

## 第一節、預測的方法/模型

本報告預計採用時間序列迴歸法、分解法、指數平滑法及 ARIMA 配適模型，並對 2010 年該年的國內電力消費量做一至十二期的月消費量預測(單位：公秉油當量)。

### 1. 時間序列迴歸法

時間序列是指把過去的觀察資料依時間順序排列而成的數據序列。時間序列分析即是應用數理統計方法對數據進行統計分析及推斷。它的基本原理：一是假設事物發展具有連續性。從過去的數據中，就能預測事物的發展趨勢。二是涉及到事物發展的隨機性。任何事物發展都可能受突發因素影響，因此要利用統計分析對歷史事件加以處理。時間序列迴歸分析是僅依照變數本身過去的資料所存有的變異型態來建立模型。時間序列迴歸法是將時間序列以趨勢

(Trend； $TR_t$ )、季節(Season； $SN_t$ )及殘差( $\varepsilon_t$ )來表示。

時間序列模型：

$$y_t = TR_t + SN_t + \varepsilon_t$$

## 2. 分解法(Decomposition Method)：

時間序列分解法是數年來一直非常有用的方法，這種方法包括譜分析、時間序列分析和傅立葉級數分析等。

而我們在統計預測中用到的分解法是將時間序列分解成四大因素：趨勢(Trend;  $TR_t$ )、季節(Season;  $SN_t$ )、循環(Cyclical;  $CL_t$ )及不規則(Irregular;  $IR_t$ )，將觀測值( $y_t$ )以這四大因素來表示，可用數學方程式表示為  $y_t = f(TR_t, SN_t, CL_t, IR_t)$ 。

再來我們利用 X11 的方法來分析我們所挑選的原始資料，對此資料的時間序列數據分別去做趨勢循環、不規則因子、季節因子、去季節因子的時間序列圖，接著再去觀察我們的資料是否有季節性存在，或呈現何種趨勢。

分解法又可分為兩種模型，若時間序列的變異數不會隨著時間的改變而改變時，我們採用加法模型；若反之，時間序列的變異數不平穩時，我們則使用乘法模型。

加法模型：

$$y_t = TR_t + SN_t + CL_t + IR_t$$

乘法模型：

$$y_t = TR_t \times SN_t \times CL_t \times IR_t$$

## 3. 指數平滑法 (Exponential Smoothing)：

假設時間序列的態勢具有穩定性或規則性，所以時間序列可以被合理的順勢推延。而指數平滑法將全期平均和移動平均的優點保留，不捨棄過去的數據，但是給予逐漸減弱的影響程度，即隨著數據資料的遠離，賦予逐漸收斂為零的權術。

指數平滑法是生產預測中常用的一種方法。當我們的資料呈現季節性且變異數平穩時，我們使用Winters Method-Additive做配適；若有季節性但變異數不平穩時，我們則使用Winters

Method-Multiplicative。而資料若是沒有呈現季節性，但有上升的趨勢且趨勢有減緩的現象時，我們使用 Damped Trend 做配適。

以下為所介紹之三種模型：

Winters Method-Additive模型：

$$L_t = \alpha(Y_t - S_{t-s}) + (1 - \alpha)(L_{t-1} + b_{t-1})$$

$$b_t = \gamma(L_t - L_{t-1}) + (1 - \gamma)b_{t-1}$$

$$S_t = \delta(Y_t - L_{t-1}) + (1 - \delta)S_{t-s}$$

$$F_{t+m} = (L_t + b_{tm})S_{t-s+m}$$

Winters Method-Multiplicative 模型：

$$L_t = \alpha(Y_t / S_{t-s}) + (1 - \alpha)(L_{t-1} + b_{t-1})$$

$$b_t = \gamma(L_t - L_{t-1}) + (1 - \gamma)b_{t-1}$$

$$S_t = \delta(Y_t / L_{t-1}) + (1 - \delta)S_{t-s}$$

$$F_{t+m} = (L_t + b_{tm})S_{t-s+m}$$

Damped Trend Exponential Smoothing的模型為：

$$L_t = \alpha Y_t + (1 - \alpha)(L_{t-1} + \Phi b_{t-1})$$

$$b_t = \gamma(L_t - L_{t-1}) + (1 - \gamma)\Phi b_{t-1}$$

#### 4. ARIMA 模型分析法

ARIMA 模型又稱 Box-Jenkins 模型，它的基本想法是：將資料隨時間前進而形成的數據序列視為一個隨機序列。用一定的數學模型來近似描述這個序列。ARIMA 模型可分為 3 種：(1)自回歸模型(簡稱 AR 模型)；(2)移動平均模型(簡稱 MA 模型)；(3)自回歸移動平均混合模型(簡稱 ARIMA 模型)。ARIMA 模型建模的基本條件是要求待預測的數列滿足平穩的條件，意指數據均不只要均值的上下波動且不可有明顯上升或下降的趨勢。如果不平穩則要對原始序列進行差分平穩化，在平穩的條件下再對序列作模型的配適。模型配適後則要對模型檢驗，通過檢驗的模型即為一個合適的模型。在做 ARIMA 模型分析法時，我們可分為四個步驟：(1)辨識模型：利用資料找出可適用的 ARIMA 模型；(2)估計：將資料放入模型中找出模型中的參數；(3)診斷分析：用不同的診斷分析法去了解模型是否對於研究資料為合適的，以獲得最終模型；(4)預測：當找出最終模型後，即可用於找出研究資料的時間序列預測值。以下為模型的介紹：

AR(P)：

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t$$

MA(q) :

$$Z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$$

ARMA(p, q) :

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$$

以上三種模型需假設白噪音來自常態分配，也就是平均數為零，變異數為常數，因此配適模式後，需要檢查殘差是否存在自我相關，當殘差彼此無自我相關，即代表配適模型是一個合適的模型。

所有使用到的檢定詳細敘述如下：

### 1. 白噪音檢定(White noise test) :

噪音檢定主要用以檢測時間數列模型當中的殘差之間是否存在相關性，當殘差符合白噪音，即殘差項為零，變異數為一常數，且殘差間彼此無自我相關者，即符合白噪音。當殘差符合白噪音代表模型唯一適合模型。以下為白噪音檢定之假設檢定與決策規則：

$H_0$  : 殘差符合白噪音

$H_1$  : 殘差不符合白噪音

決策規則：

P-value > 0.05，則不拒絕 $H_0$ ，符合白噪音，表模型配適合適。

P-value < 0.05，則拒絕 $H_0$ ，不符合白噪音，表模型配適不合適。

因此我們要让模型不拒絕 $H_0$ ，即 P-value > 0.05，才合適。

### 2. 單根檢定(Unit Root Test) :

單根檢定乃是檢測時間序列是否為平穩，平穩即代表平均數不隨時間改變而改變，亦表示殘差的期望值等於零。在 ARIMA 的分析過程中，序列是否平穩是首要工作，當序列平穩後才可再做進一步的分析，因此，單根檢定的目的在於決定要以原始資料或差分調整後之資料進行分析研究。以下為單根檢定之假設檢定與決策規則：

$H_0$ ：有單根（時間序列不平穩）

$H_1$ ：沒有單根（時間序列平穩）

決策規則：

$P\text{-value} > 0.05$ ，則不拒絕 $H_0$ ，表示有單根，模型配適不合適。

$P\text{-value} < 0.05$ ，則拒絕 $H_0$ ，表示沒有單根，模型配適合適。

因此我們要讓模型拒絕 $H_0$ ，即 $P\text{-value} < 0.05$ ，才合適。

### 3. 介入分析法

當資料在某個時間點或者某個時段出現異於之前的走向，則對異常的情況作介入分析，在做介入分析前，須先討論資料產生變異的原因與影響點，並對之作介入分析，使之後的模型預測能力更佳，減少因異常的情形而影響之後的預測結果。介入分析可分為以下不同的型態：(1) 只有對某一發生點產生影響；(2) 除了發生點有影響外，對後續也產生影響；(3) 除了發生點以外，對後續也產生影響但影響漸漸減少。不同的資料型態會用不同的介入分析法使模型調整成為具有較佳的預測能力。

## 第二節、診斷分析方法

DW 檢定法：

主要用以檢定迴歸分析中的殘差是否存在自我相關，在一個合適的模型中，我們假設殘差彼此無相關，表示一個合適的模型中，殘差彼此的相關性為零。

\*檢定規則：

當  $Pr < DW$  的值小於顯著水準  $\alpha = 0.05$  時，表示顯著存在正自我相關。

當  $Pr > DW$  的值小於顯著水準  $\alpha = 0.05$  時，表示顯著存在負自我相關。

## 第三節、預測評估準則

預測能力的好壞可用準則來衡量，我們將12筆樣本外預測資料與各種預測方法的預測值用不同的準則做比較，以獲得最適模型。以下是我們用來衡量不同預測方法的評估準則。假設 $n$ 為總樣本數， $y_t$ 表示第 $i$ 個時間點的實際值， $\hat{y}_t$ 為第 $t$ 個時間點的預測值，其評估準則如下：



1.平均絕對離差(Mean Absolute Deviation, MAD)：用以評估評價模型與資料之誤差程度，越接近0表其誤差程度越小。公式如下：

$$MAD = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t|$$

2.均方差(Mean Squared Error, MSE)：計算實際值與預測值之預測誤差取平方加總後並加以平均，通常數字太大則會將之取根號，值越大表示預測越不精準。計算公式如下：

$$MSE = \frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2 \quad RMSE = \sqrt{MSE}$$

3.平均誤差率(Mean Percentage Error, MPE)：對誤差加以懲罰，且每一個誤差率均採用相同的權重。計算公式如下：

$$MPE = \left[ \frac{1}{n} \sum_{t=1}^n \frac{y_t - \hat{y}_t}{y_t} \right] \times 100\%$$

4.平均絕對誤差百分比(Mean Absolute Percentage Error, MAPE)：用以作為預測模式好壞之評估指標，計算公式如下：

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|y_t - \hat{y}_t|}{y_t} \times 100\%$$

依 MAPE 大小將模式預測能力分為三個等級，如下圖所示：

MAPE(%)	評估
<5	高準確的預測
6-10	尚可的預測
>10	不準確的預測

表 2-3.1 MAPE 準則表



## 第三章 資料分析

首先我們判斷原始資料是否有趨勢，以及其平均數和變異數是否平穩，當序列平穩代表接下來的預測才有意義。接著我們依序使用時間序列回歸法、分解法、指數平滑法及 ARIMA 配適模型，並對 2010 年的電力消費量做樣本外預測預測。最後將保留的十二筆真實值與預測值用 MSE、MAD、MPE 及 MAPE 四個準則來做評估，並選出最佳的配適模型。

### 第一節 原始資料分析

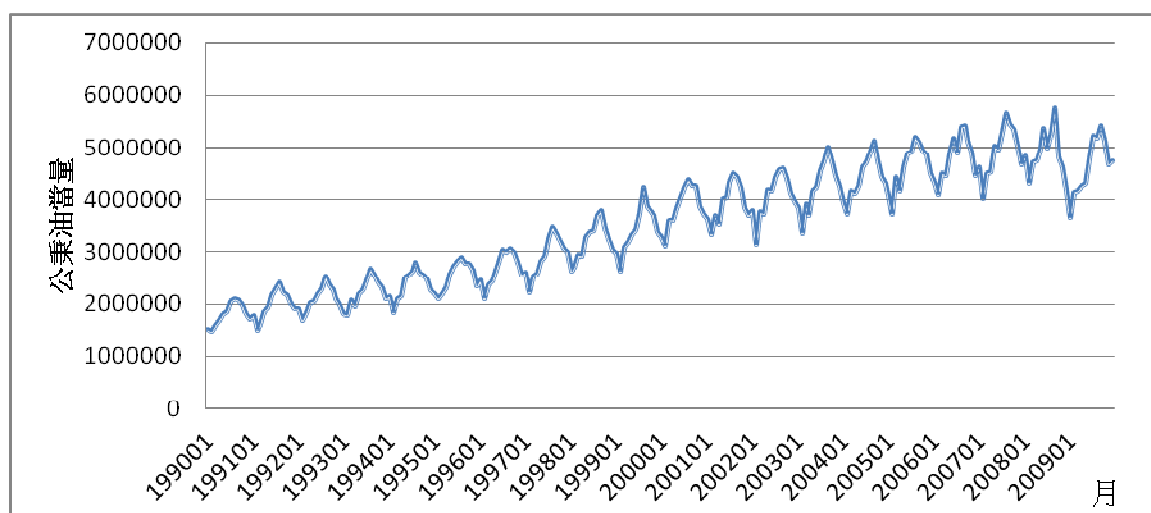


圖 3-1.1 原始時間序列圖

由圖 3-1.1 可看出，電力消費有明顯增加的趨勢代表平均數有上升的現象發生，且變異的範圍也隨著時間的增加而擴大，但在 2008 年底到 2009 年初增加的趨勢有漸緩的現象，由上述可得知此筆資料的平均數與變異數皆不平穩。由圖 3-1.1 可發現電力消費的使用量上在每年 7 月到 9 月為使用量的高峰，而 12 月到 2 月為用電量相對較少的月分，因此可瞭解此筆資料是具有季節性的。

由圖 3-1.1 發現在 2008 年 9 月開始降低的幅度特別大，我們認為是受到 2008 年七月起至年底，陸續調漲電費共百分之二十，並且在 2009 年 1 月時，又調漲電費百分之二十，且當時金融海嘯使得經濟在短時間內蕭條，造成許多工廠周轉不靈而倒閉，間接讓電力的消費量大減。

## 第二節 時間序列迴歸法

### 模型的配適

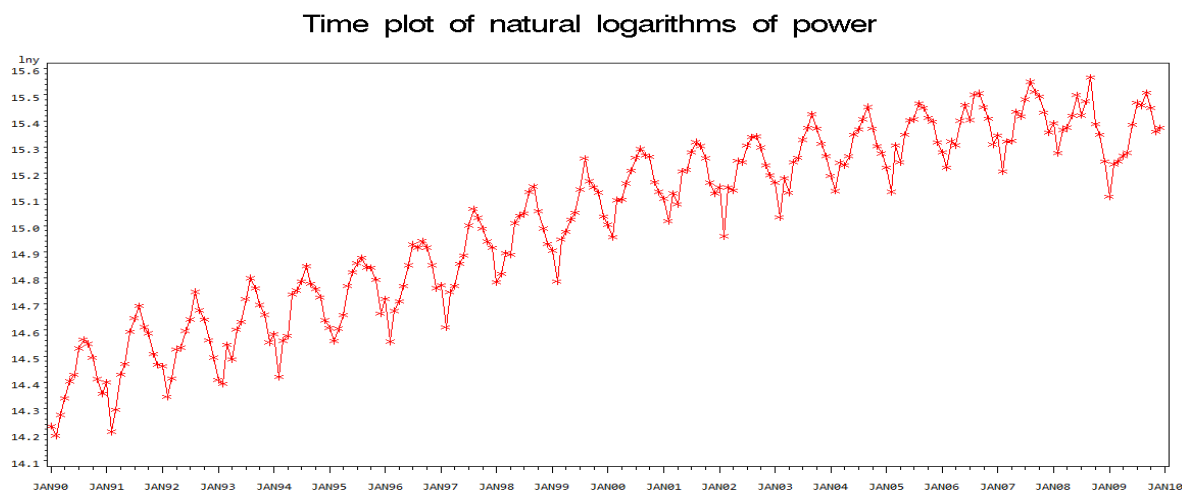


圖 3-2.1 原始資料取 log 後之圖

首先，我們要知道資料的變異是否平穩，從 3-1.1 中發現變異並不平穩，因此我們取 log 做轉換，從圖 3-2.1 我們可發現 log 轉換的資料變異數有趨於一致。因此我們可判斷此筆資料的序列已為平穩。在序列平穩後，我們可得其配適之模型，為了瞭解配適的模型是否為一合適的模型，因此我們要對模型的殘差做診斷分析。當模型的殘差不存在自我相關性時，才是一個好的配適模型。而我們所使用的檢定方法是「DW 檢定法」。接下來我們對配適模型的殘差做檢定後，發現其  $Pr < DW$  的值小於 0.05，表示殘差顯著存在正自我相關性。由於模型存在自我相關，表示此模型不是一個好的配適模型，因此我們需加入自我相關項，用來消除自我相關性的存在，使模型的配適更合適。在我們加入三階自我相關項後，可發現殘差已無自我相關存在，代表原模型加入三階自我相關項後即為合適模型。

將原模型加入三階自我相關項後的所有參數估計值代入模型中，得到時間序列迴歸法的預測方程式：

$$y_t^* = 14.3744 + 0.004557t - 0.0305M_1 - 0.01232M_2 - 0.0146M_3 - 0.00722M_4 \\ + 0.0699M_5 + 0.0112M_6 + 0.1533M_7 + 0.2015M_8 + 0.1853M_9 + 0.1327M_{10} \\ + 0.0658M_{11} + \varepsilon_t$$

$$y_t^* = \log(y_t)$$

$$\varepsilon_t = 0.358025\varepsilon_{t-1} + 0.291779\varepsilon_{t-2} + 0.275252\varepsilon_{t-3} + a_t$$

$$a_t \stackrel{i.i.d.}{\sim} N(\mu, \sigma^2) \quad \hat{\sigma} = 0.003742$$

Mt為季節虛擬解釋變數，其中：

$$M_1 = \begin{cases} 1 & \text{when Jan.} \\ 0 & \text{Otherwise} \end{cases} \quad M_2 = \begin{cases} 1 & \text{when Feb.} \\ 0 & \text{Otherwise} \end{cases}$$

$$M_3 = \begin{cases} 1 & \text{when Mar.} \\ 0 & \text{Otherwise} \end{cases} \quad M_4 = \begin{cases} 1 & \text{when Apr.} \\ 0 & \text{Otherwise} \end{cases}$$

$$M_5 = \begin{cases} 1 & \text{when May.} \\ 0 & \text{Otherwise} \end{cases} \quad M_6 = \begin{cases} 1 & \text{when June} \\ 0 & \text{Otherwise} \end{cases}$$

$$M_7 = \begin{cases} 1 & \text{when July} \\ 0 & \text{Otherwise} \end{cases} \quad M_8 = \begin{cases} 1 & \text{when Aug.} \\ 0 & \text{Otherwise} \end{cases}$$

$$M_9 = \begin{cases} 1 & \text{when Sep.} \\ 0 & \text{Otherwise} \end{cases} \quad M_{10} = \begin{cases} 1 & \text{when Oct.} \\ 0 & \text{Otherwise} \end{cases}$$

$$M_{11} = \begin{cases} 1 & \text{when Nov.} \\ 0 & \text{Otherwise} \end{cases}$$

為了瞭解模型之估計是否適當，我們保留了最後的12筆真實數據，和未來一年預測的比較用。我們將這12筆數據的實際值、估計值和95%的上、下界預測區間畫成預測曲線圖：

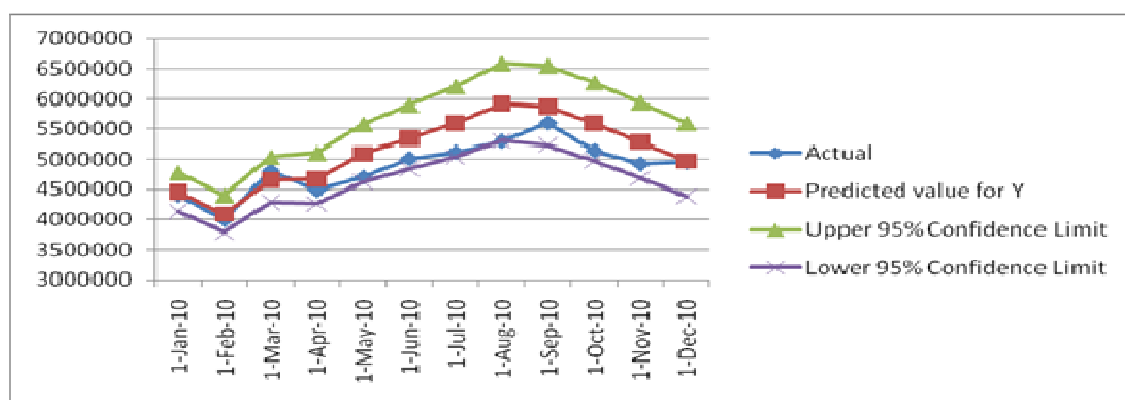


圖3-2.2 時間序列分析法曲線圖

從圖3-2.2可看出真實值在前面幾月都很貼近預測值但在後面幾個月卻發現漸漸偏離預測值而貼近95%下界直到最後才又回到預測值附近因此認為此模型的預測能力較弱，因為有高估的現象發生，所以我們試著用介入分析來改善高估實際值的狀況。

當資料在某個時間點或者某個時段出現異於之前的走向，則對異常的情況作介入分析，使之後的模型預測能力更佳，減少因異常的情形而影響之後的預測結果。從原始資料圖我們明顯的看出在2009年1月的用電量有明顯異於偏離原本的走勢，於是我們決定進行介入分析，而在介入分析之前，我們須先找尋當時使異常點走勢有所改變的主要因素與發生時間點，發現影響此點異常的原因有其三：

1. 2008年7月到9月期間行政院致力於節能減碳的政策與實施。
2. 2008年7月底公布，夏季用電調漲百分之二十，一般用電在該年十一月至隔年一月也將陸續調漲百分之二十。
3. 2008年9月時，金融海嘯使得經濟在短時間內蕭條，造成許多工廠周轉不靈而倒閉。

綜合以上三點以及透過原始時間序列圖的觀察，我們推估，讓2009年1月的走在向呈現異常的事件以及時間點為2008年9月。因此，我們從2008年9月(第227筆資料)進行介入分析。

以下為我們對介入分析所設定的參數：

$$S_t = \begin{cases} 0 & \text{if } t < 227 \\ 1 & \text{if } t \geq 227 \end{cases}$$

在介入分析後我們可獲得新的配適模型，接著對模型的殘差做診斷分析。當殘差不存在自我相關性時，才是一個好的配適模型。而我們的檢定方法是「DW檢定統計量檢定」。在對介入後的配適模型做殘差分析後發現，其 $Pr < DW$ 之機率值小於顯著水準 $\alpha = 0.05$ ，表示顯著存在正自我相關。因此我們要配適自我相關項來解決這個問題。在配適二階自我相關項後，可觀察出殘差已無自我相關存在，表示在配適二階自我相關項之後，殘差的自我相關性已消除，此即為合適模型。同時，可觀察出原本需要加入三階自我相關項才能消除的自我相關性，在做完介入分析後，只需要配適到二階自我相關項時即可達到合適模型。因此可了解，介入分析在此模型中是必要存在的。

以下是我們的时间序列迴歸法的預測方程式：

$$y_t^* = 14.3679 + 0.004839t - 0.0315M_1 - 0.1237M_2 - 0.0165M_3 - 0.009055M_4 + 0.0675M_5 + 0.1082M_6 + 0.1498M_7 + 0.1975M_8 + 0.1806M_9 + 0.1277M_{10} + 0.0621M_{11} - 0.1271S_t + \varepsilon_t$$

$$Y_t^* = \log(y_t)$$

$$\varepsilon_t = 0.387173\varepsilon_{t-1} + 0.432154\varepsilon_{t-2} + \alpha_t, \alpha_t \sim N(0, \sigma^2)$$

$$\hat{\sigma} = 41.08065$$

$$S_t = \begin{cases} 0 & \text{if } t < 227 \\ 1 & \text{if } t \geq 227 \end{cases}$$

為了瞭解在做完介入分析後，模型是否能表達出預測好壞，我們將保留的最後的12筆真實數據，和對未來一年的預測數據做比較。我們將這12筆數據的實際值、估計值和95%的上、下界預測區間畫成預測曲線圖：

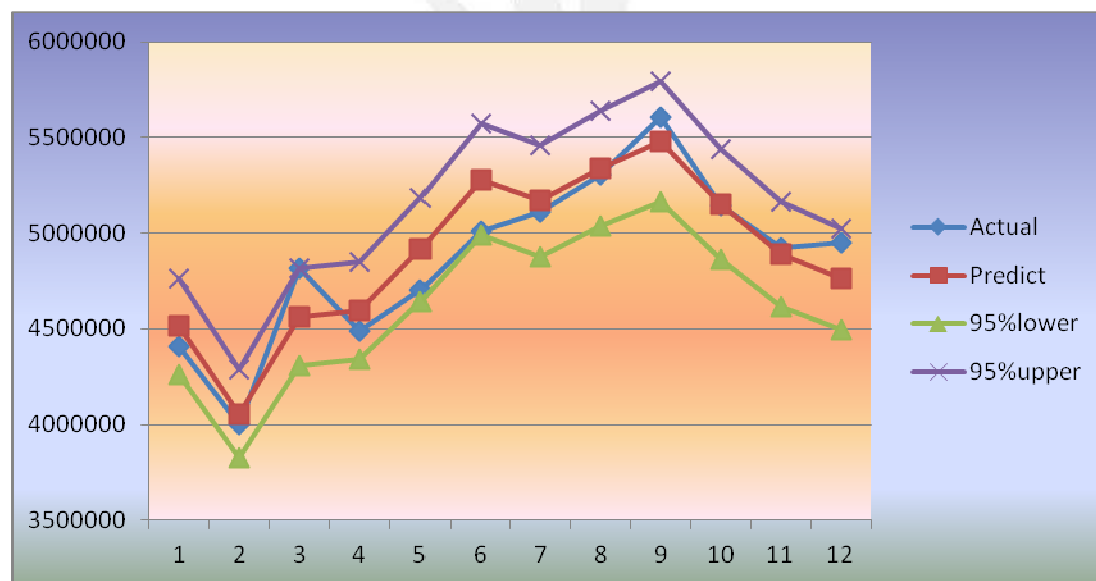


圖3-2.3 實際值、估計值和95%的上、下界預測區間之曲線圖（時間序列法）

由圖3-2.3可得知，在做完介入分析後，已無高估的現象，所以我們認為此預測方法優於無介入分析之曲線圖。藉此可認定此模型為較合適且較具預測能力之模型。

### 第三節 分解法(Decomposition Method)

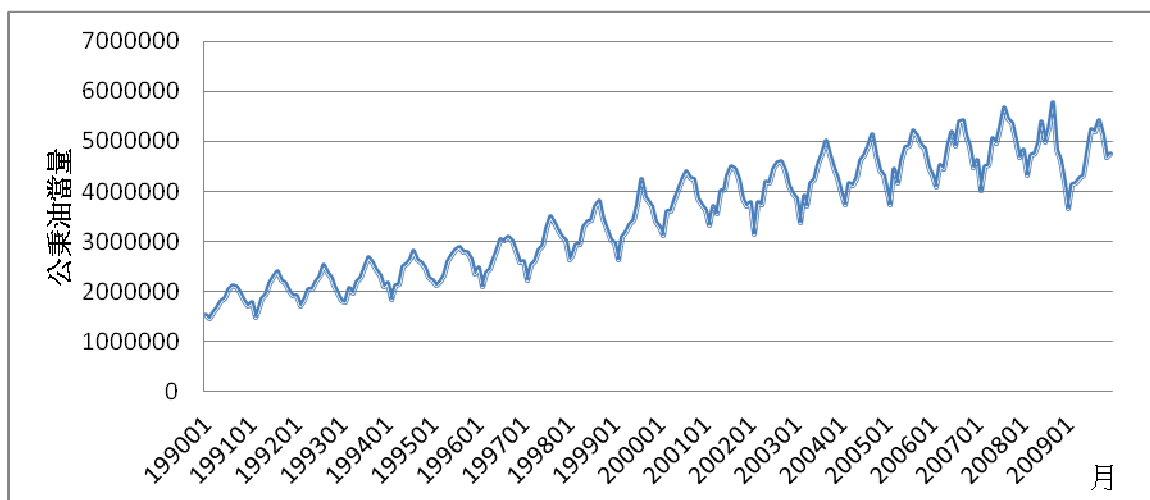


圖 3-3.1 原始時間序列圖

首先我們選擇適當的配適模型。由圖 3-3.1 可看出時間序列的平均數上升不一致，變異程度隨時間的增加而擴大。即變異數不平穩。因此我們選用乘法模型來做預測。

#### Time plot of deseasonalized power

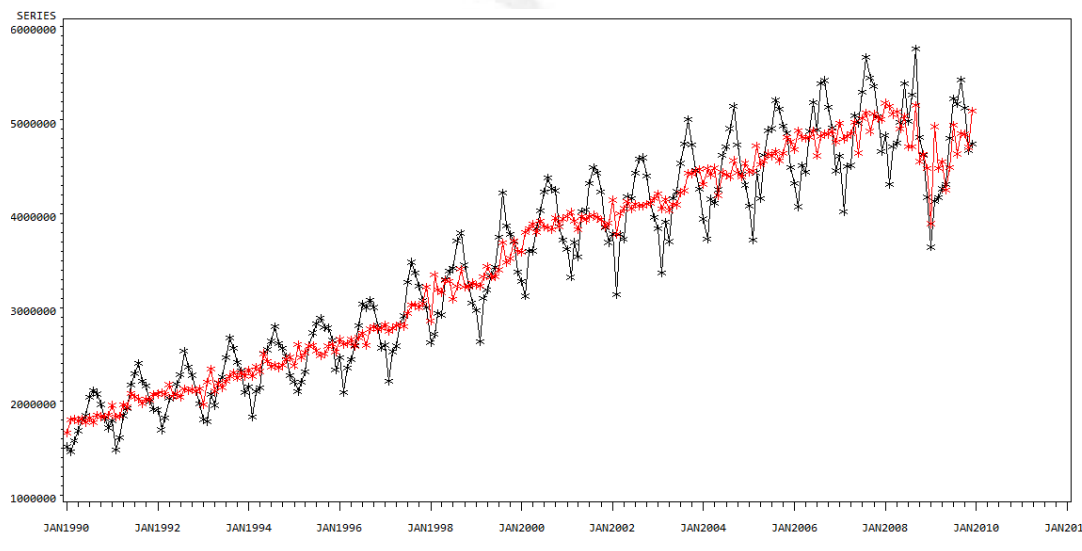


圖 3-3.2 去季節因子時間序列圖

Trend—cycle component

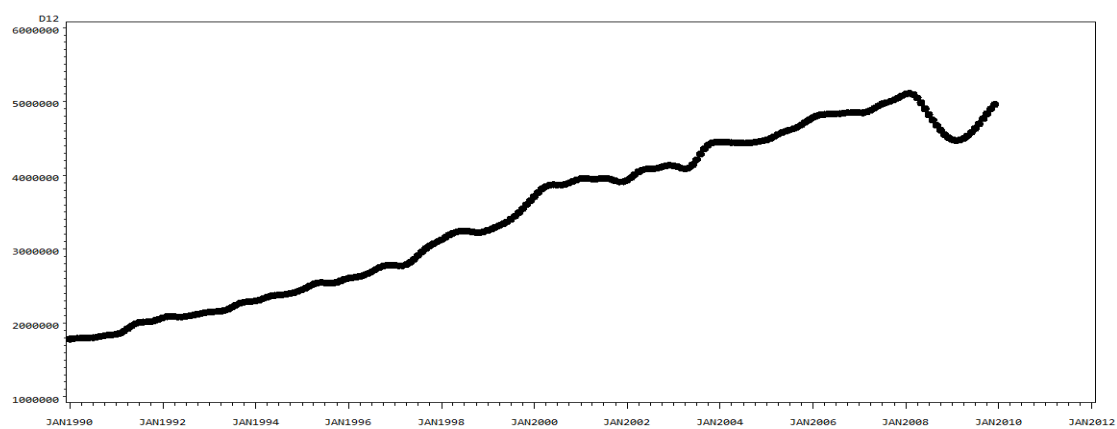


圖 3-3.3 趨勢循環時間序列圖

Seasonal component

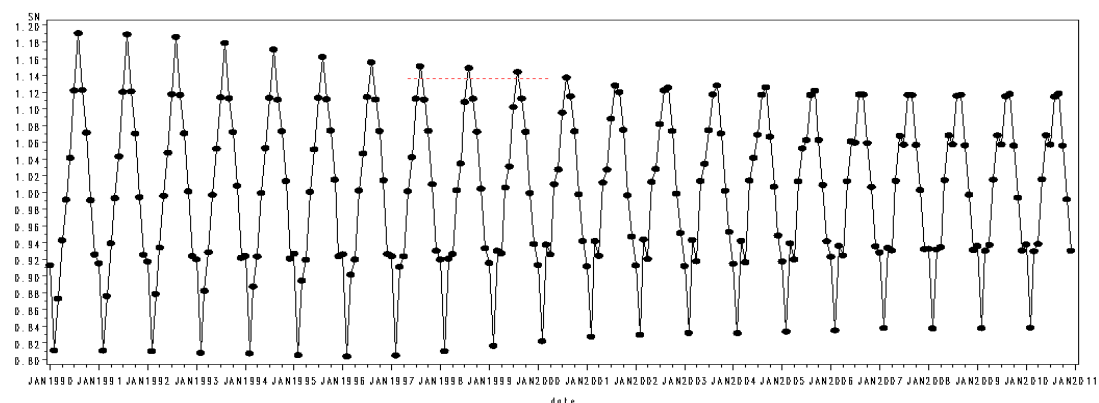


圖 3-3.4 季節因子之時間序列圖

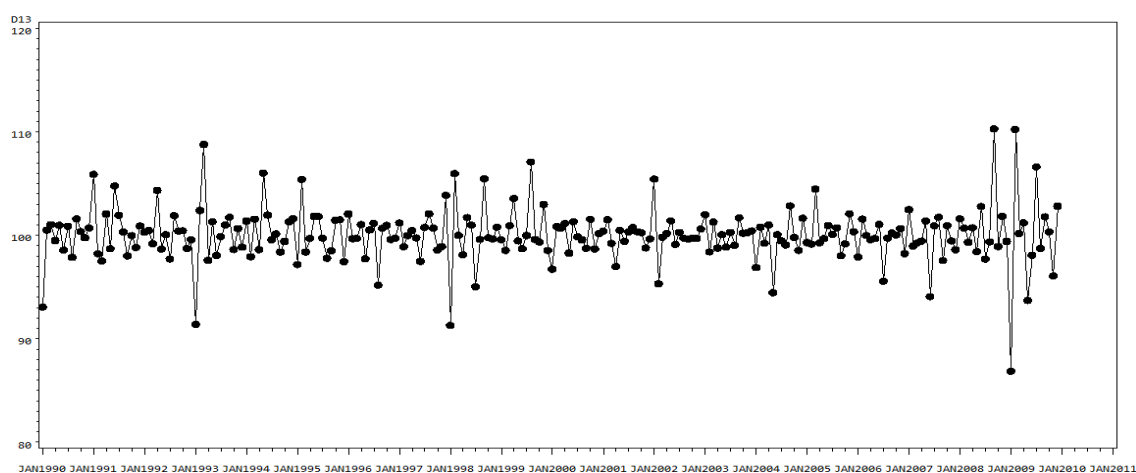


圖 3-3.5 不規則因子時間序列圖

從圖 3-3.2 中可看出在沒有季節因素的情況下，電力消費量也會隨著時間的增加而上升。從圖 3-3.3 可看出整體的趨勢有往上的現象。從圖 3-3.4 可發現此



資料有明顯的季節變化，但有越來越小的趨勢。在圖 3-3.5 中可看出平均數平穩。

接著對模型的殘差做診斷分析。當殘差不存在自我相關性時，才是個好的配適模型。而我們的檢定方法是「DW 檢定法」，在我們對模型的殘差做檢定後發現，其  $Pr < DW < 0.05$ ，表示殘差顯著存在正自我相關。由於模型存在自我相關性時，意謂著此模型不是一個好的配適模型，因此我們需加入自我相關項，以消除自我相關性使模型的配適更合適。在配適三階自我相關項後，其  $Pr < DW$  與  $Pr > DW$  的值皆大於 0.05，表示殘差已無自我相關存在，因此代表此一模型為合適模型。

為了瞭解模型之估計是否合適，我們將保留的最後的 12 筆真實數據，和對未來一年的預測數據做比較。我們將這 12 筆數據的實際值、估計值和 95% 的上、下界預測區間畫成預測曲線圖：

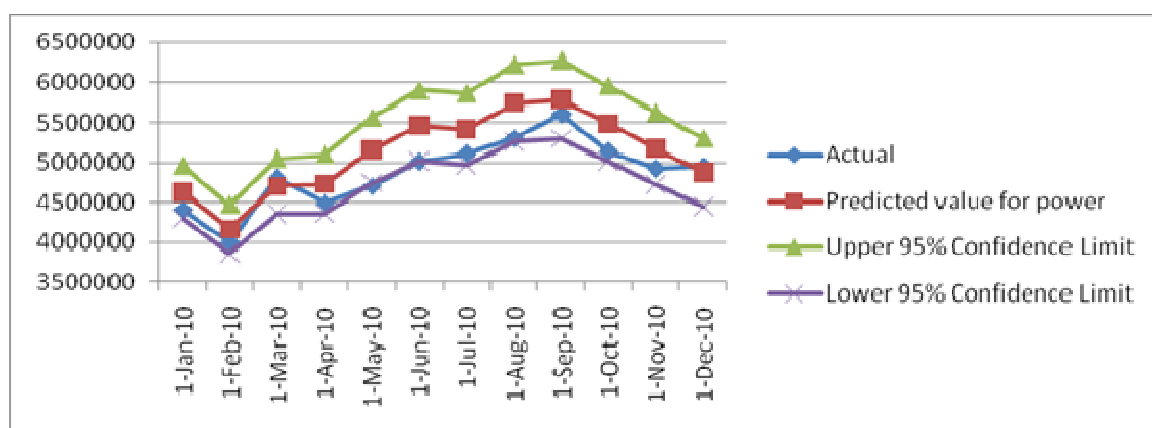


圖 3-3.6 分解法曲線圖

從圖中發現大部分值落在信賴區間 95% 的下界，因此我們可知道此模型有高估預測值的現象發生，所以我們選擇做介入分析。同時時間序列法，我們在 2008 年 9 月做介入分析，去了解模型在做完介入分析後是否有所改善。在獲得我們做完介入分析後的原模型後，需對模型的殘差做診斷分析，以了解模型是否合適。我們用「DW 檢定統計量檢測」，在我們對模型的殘差做檢定後發現，其  $Pr < DW$  之機率值小於顯著水準  $\alpha = 0.05$ ，表示顯著存在正自我相關。因此我們要配適自我相關項來解決這個問題。當我們配適三階自我相關項。在配適完三階自我相關項後，殘差已不具有自我相關存在，表示此為合適模型。



以下是我們的分解法的預測方程式：

$$dy_t = 1643157 + 15506t - 500343S_t + \varepsilon_t$$

$$\hat{y}_t = TR_t \times SN_t$$

$$\varepsilon_t = 0.149\varepsilon_{t-1} + 0.300\varepsilon_{t-2} + 0.336\varepsilon_{t-3} + \alpha_t$$

$$S_t = \begin{cases} 0 & \text{if } t < 227 \\ 1 & \text{if } t \geq 227 \end{cases}$$

為了瞭解在做完介入分析後的模型之估計是否適當，我們將保留的最後的12筆真實數據，和對未來一年的預測數據做比較。我們將這12筆數據的實際值、估計值和95%的上、下界預測區間畫成預測曲線圖：

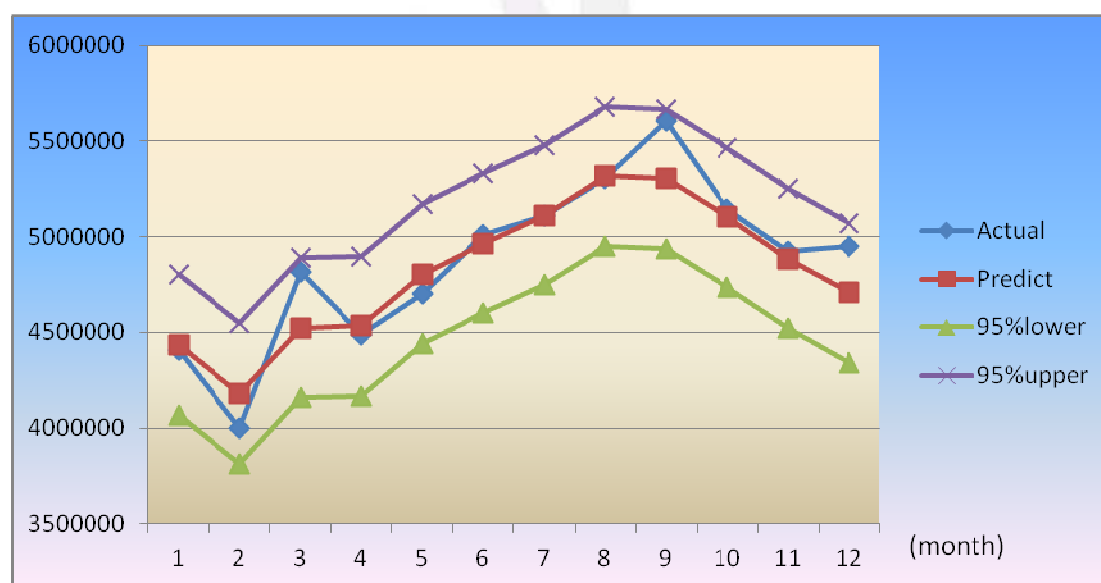


圖3-3.7實際值、估計值和95%的上、下界預測區間之曲線圖（分解法）

由圖3-3.7可得知，在做完介入分析後，已無高估的現象，而且大部分的值都在預測值附近，因此我們認為此模型的預測能力優於無介入分析之模型。

## 第四節 指數平滑法

### 模型的配適

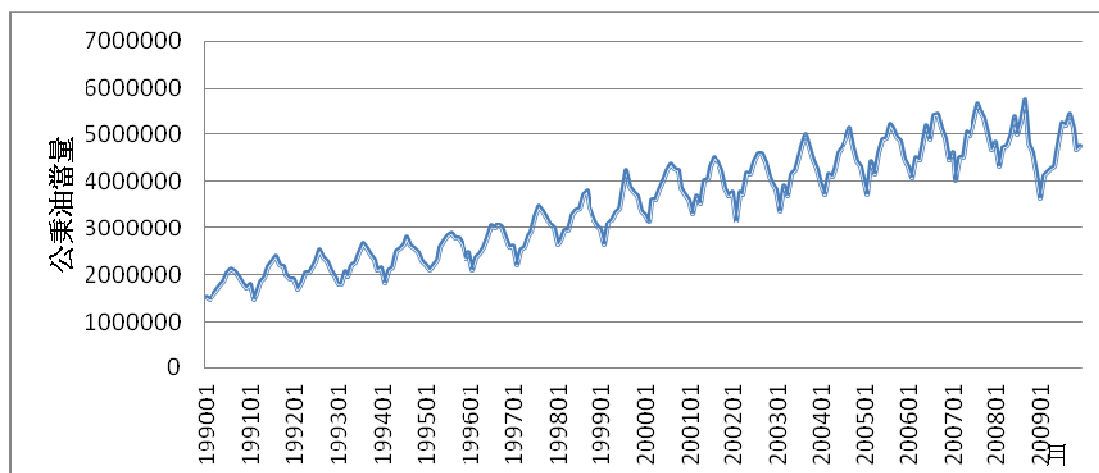


圖 3-4.1 原始資料的時間序列圖

首先我們找出適當的配適模型。由圖 3.4-1 可看出原始資料具有季節性，其變異程度隨時間的增加而擴大。即變異數不平穩。因此我們選用 Winters Method-Multiplicative 模型做配適。

Model Parameter	Estimate	Std. Error	T	Prob> T
LEVEL Smoothing Weight	0.32300	0.0338	9.5683	<.0001
TREND Smoothing Weight	0.00100	0.0063	0.1594	0.8735
SEASONAL Smoothing Weight	0.23971	0.0394	6.0788	<.0001
Residual Variance (sigma squared)	2.02345E10	.	.	.

表 3-4.1 參數估計表 (指數平滑法)

由表 3-4.1 可獲得各個參數的估計值，可知趨勢項的估計較不顯著，代表這個變數對此模型的影響較小。接下來將各個參數估計值帶入，得到我們的指數平滑法預測方程式：

$$L_t = 0.323 \left( \frac{Y_t}{S_{t-s}} \right) + (1 - 0.323)(L_{t-1} + b_{t-1})$$

$$b_t = 0.001(L_t - L_{t-1}) + (1 - 0.001)b_{t-1}$$

$$S_t = 0.23791 \left( \frac{Y_t}{L_{t-1}} \right) + (1 - 0.23971)S_{t-s}$$

為了瞭解我們所做的模型之估計是否適當，我們將保留的最後的 12 筆真實數據，和對未來一年的預測數據做比較。我們將 12 筆的實際值、估計值和 95%

的上、下界預測區間做成表格，再將這四種值畫出預測曲線圖：

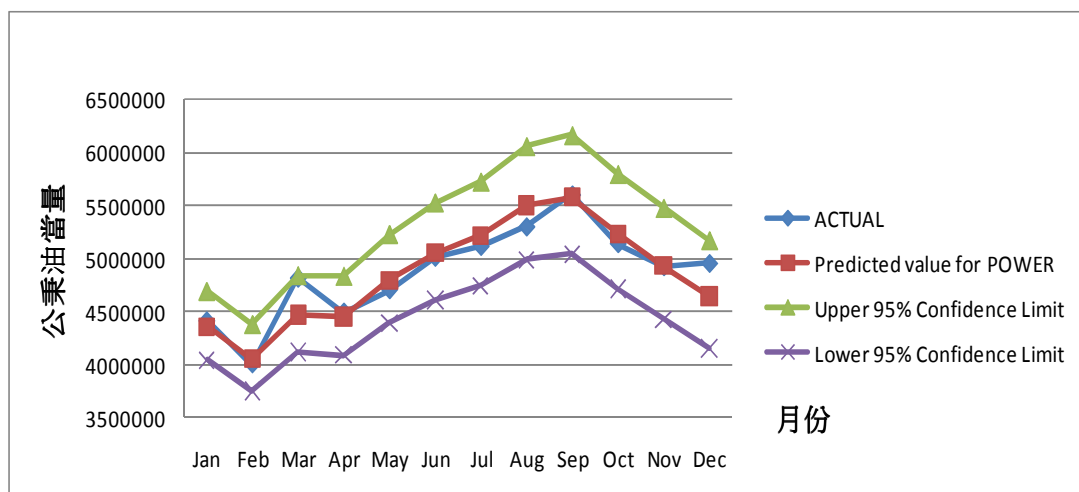


圖 3-4.2 實際值、預測值、95%預測信賴上界、95%預測信賴下界之曲線圖

由圖 3-3.2 的預測結果可看出，大部分的實際值都很貼近預測值，只有一筆接近 95%信賴區間上界，因此我們認為指數平滑法所做出的預測模型，適合用來分析我們的資料。

## 第五節 ARIMA 模型分析法

### 模型的配適

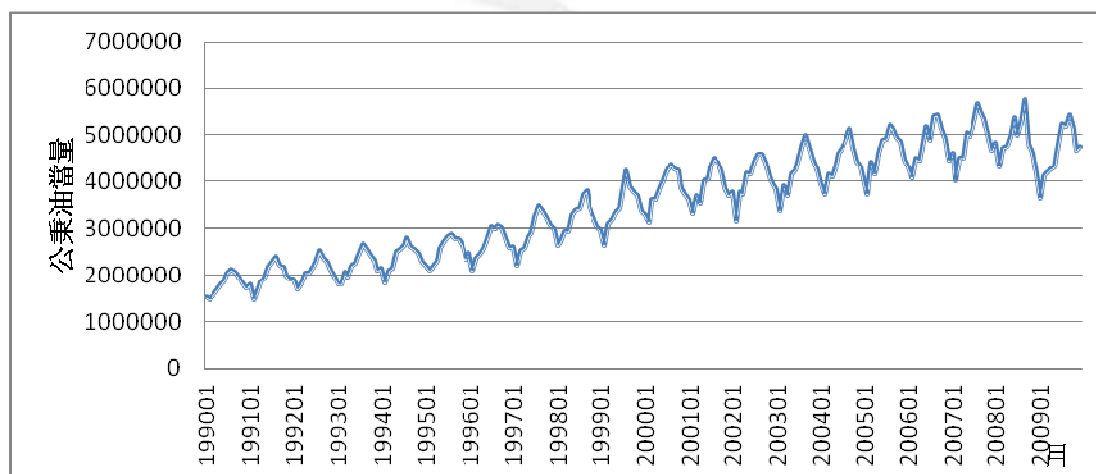


圖 3-5.1 原始資料時間序列圖

從圖 3-5.1 可看出時間序列波動不一致，變異會隨著時間的增加而增加，表示此變異數不為常數，即變異數不平穩。且資料的平均數有上升的趨

勢，代表平均數不平穩，因此必須做轉換和差分，讓模型合適。

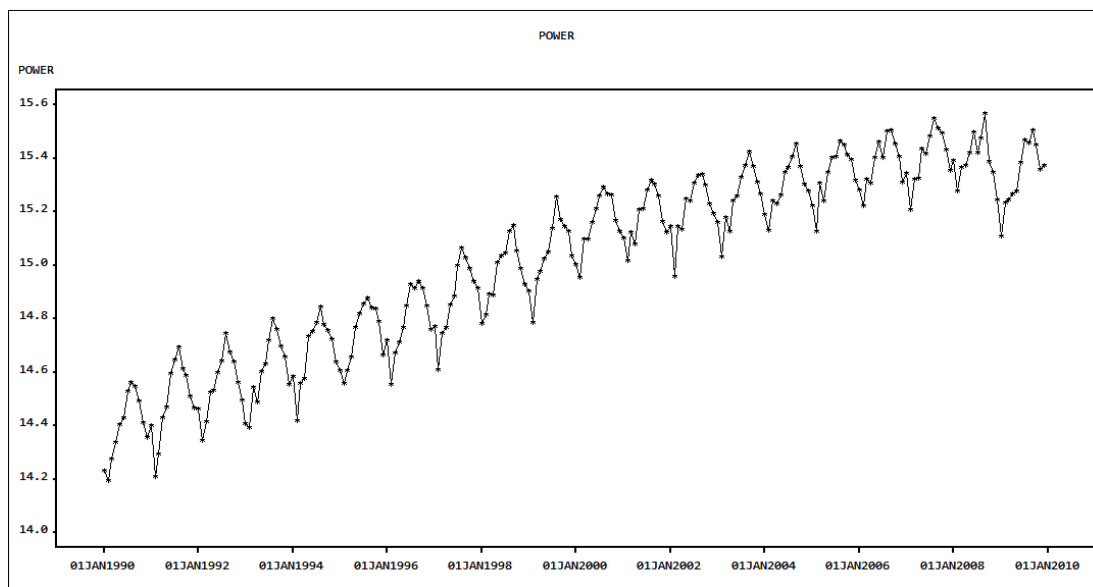


圖 3-5.2 經過取 log 後的时间序列圖

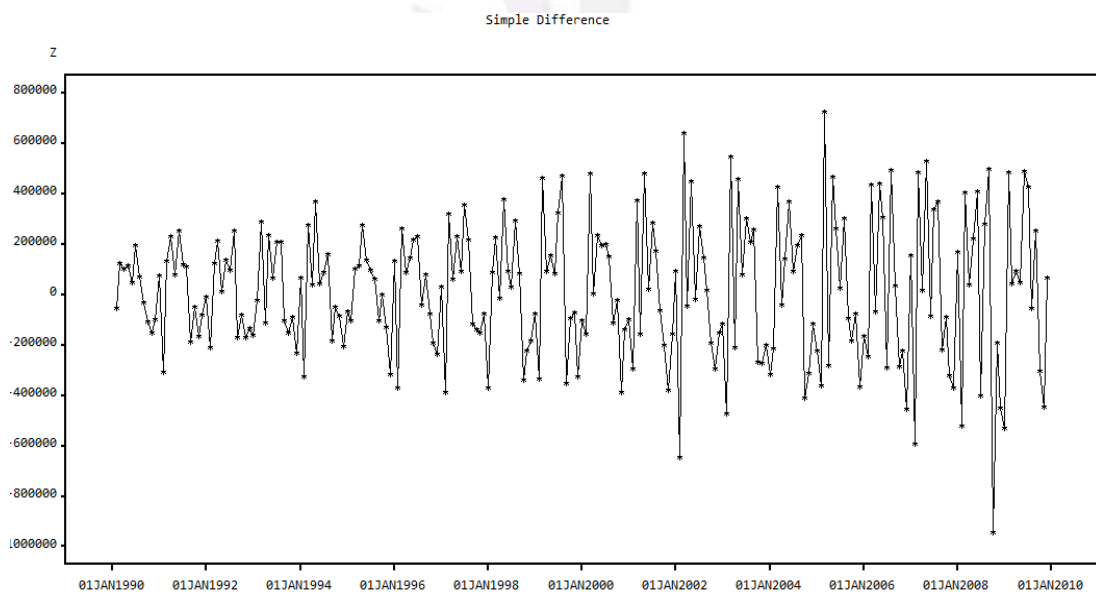


圖 3-5.3 一次差分圖

從圖 3-5.2 可得知，經過轉換後變異數有趨於一致的現象。但從圖中能看出平均數不為一常數，因此我們做一次差分來改進。從圖 3-5.3 能看出平均數在做完一次差分後有平均數平穩。

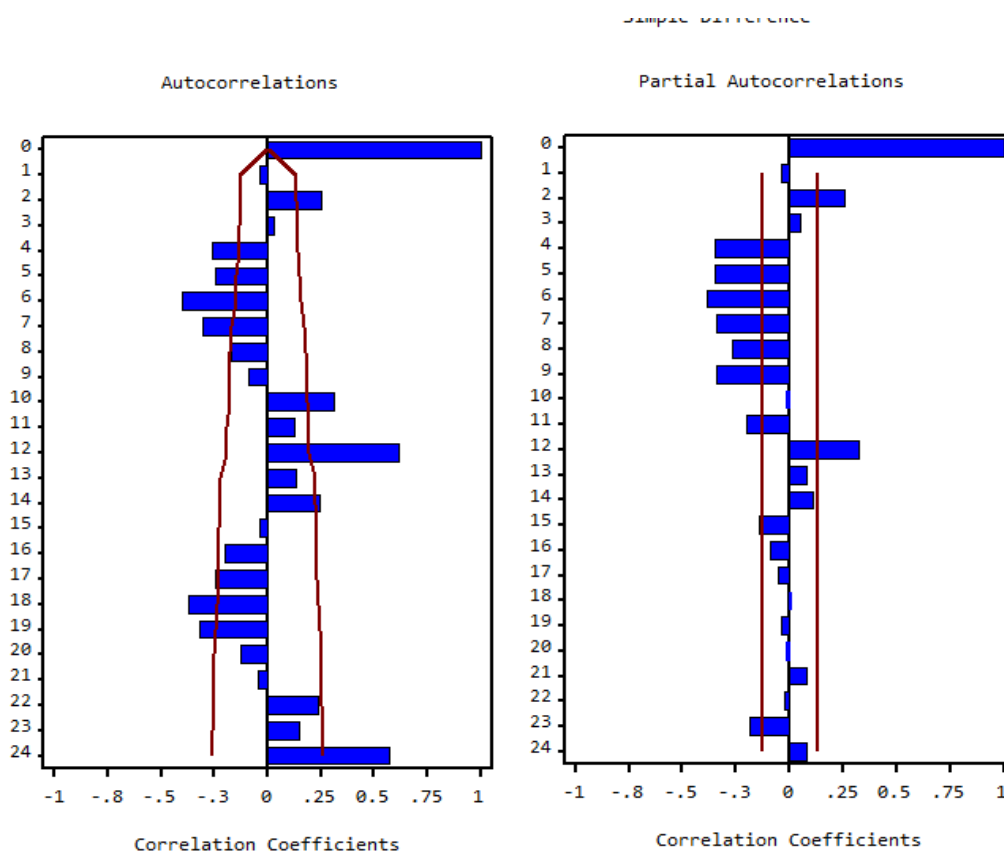


圖 3-5.4 一次差分後的 ACF 與 PACF 圖

從圖 3-5.4 看出在 ACF 圖(左邊)中的 lag12 與 lag24 有明顯的突出因此我們對此再做季節差分。

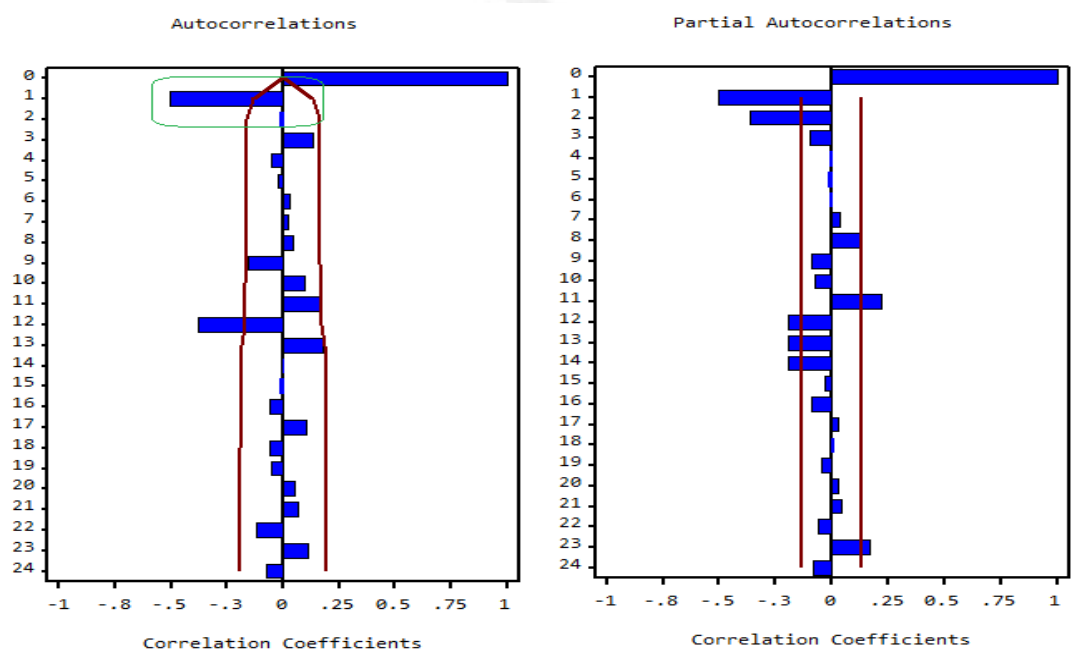


圖 3-5.5 經過一次差分與季節差分後的 ACF 與 PACF 圖

由圖 3-5.5 我們判斷 ACF 有突出兩倍標準差的現象，因此我們決定配適模型 MA(q)。又由圖的 ACF 可得知 lag1 及 lag12 有顯著的現象發生，因此我們加以配適，(q=(1)(12))，又稱為 Airline model。

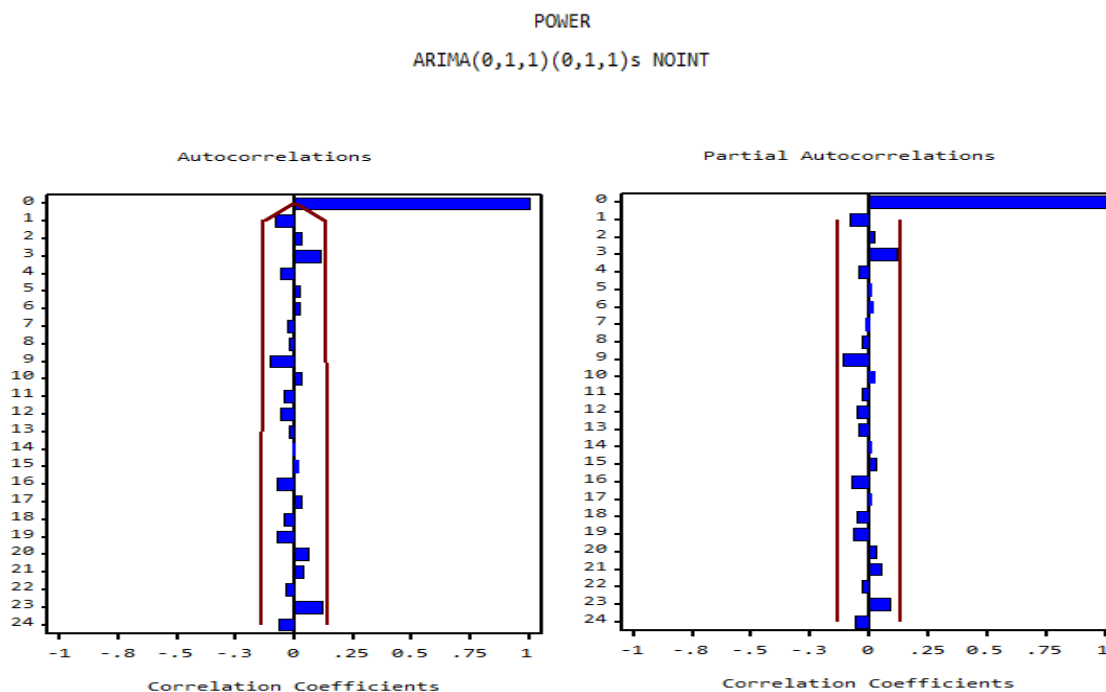


圖 3-5.6 配適模型後的 ACF 與 PACF 圖

經過取 log 並且配適 ARIMA(0, 1, 1)(0, 1, 1)s NOINT 的模型後，由圖 3-5.6 的 ACF 與 PACF 可看出，殘差都在兩倍標準差之內，表示此模型的配適是合適的。

配適出模型後，接著我們配適的模型之殘差做診斷分析。若殘差具有單根，即表示此模型的時間序列不平穩，會隨著時間的改變而有所改變，這樣的模型並非一個好的配適。而一個好的配適模型殘差其平均數為 0，變異數為常數且殘差彼此不相關，即符合白噪音，並且不存在單根，才會是一個好的配適。因此，我們對殘差做白噪音及單根檢定，診斷此模型殘差是否符合配適模型的假設。

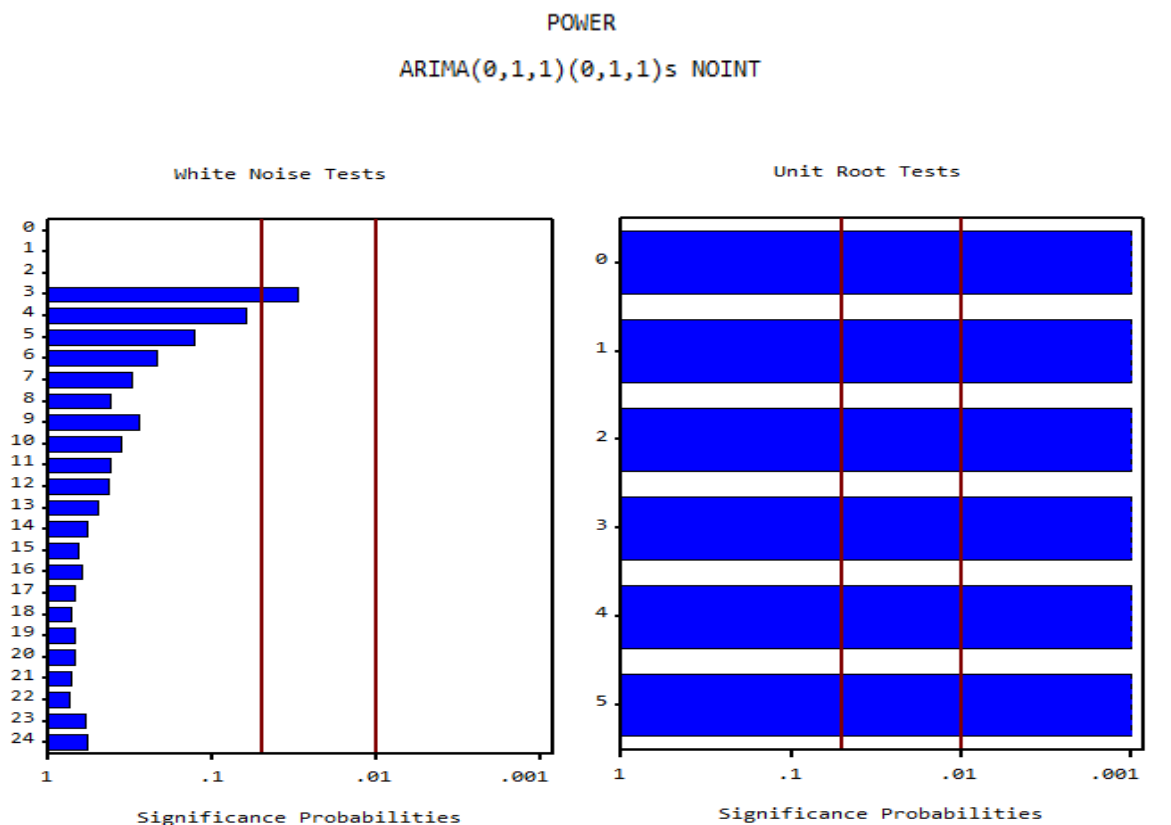


圖 3-5.7 白噪音與單根檢定檢定圖

由圖 3-5.7 的白噪音檢定(左圖)可知，所有的 lag 都大於顯著水準  $\alpha=0.05$ ，因此不拒絕  $H_0$ ，表示殘差有白噪音現象。又從單根檢定得知所有 lag 皆小於顯著水準  $\alpha=0.05$ ，意謂著拒絕  $H_0$ ，表示時間序列已經平穩。由以上檢定可得，我們的配適模型 ARIMA(0, 1, 1)(0, 1, 1)s NOINT 是合適的。

POWER  
ARIMA(0,1,1)(0,1,1)s NOINT

Model Parameter	Estimate	Std. Error	T	Prob> T
Moving Average, Lag 1	0.68274	0.0485	14.0656	<.0001
Seasonal Moving Average, Lag 12	0.80104	0.0542	14.7705	<.0001
Model Variance (sigma squared)	0.00149	.	.	.

表 3-5.1 ARIMA 模型的參數估計

由表 3-5.1 可得各個參數值，且從兩個參數值得 P-value 皆小於顯著水準  $\alpha=0.05$ ，表示參數顯著，都可留在模型中做估計。總結上述結果，我們使用 ARIMA(0, 1, 1)(0, 1, 1)s NOINT 來預測模型是合適的。將表中的參數估計值帶入模型中，得到我們的預測方程是如下：

$$(1-B)(1-B^{12}) \log(Y_t) = (1 - 0.68274B)(1 -$$

$$0.80104B^{12}) a_t, a_t \stackrel{i.i.d.}{\sim} N(\mu, \sigma^2) \hat{\sigma} = 0.00149$$

為了瞭解我們所做的模型之估計是否適當，我們將保留的最後的12 筆真實數據，和對未來一年的預測數據做比較。我們將這12筆數據的實際值、估計值和95%的上、下界預測區間上表，再將這四種值畫成預測曲線圖：

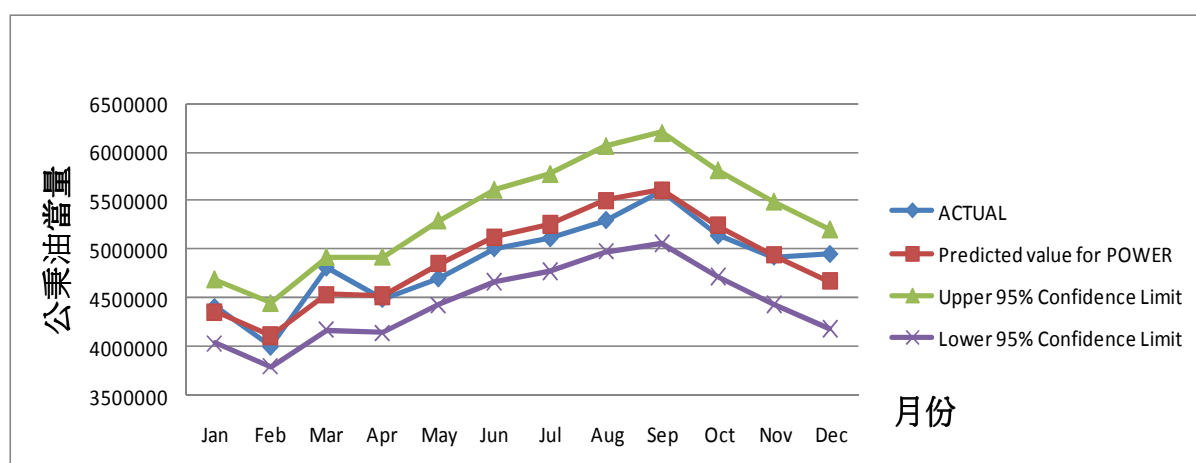


圖 3-5.8 ARIMA 的估計值、實際值和 95%的上下界預測區間的曲線圖(2010 年)

由圖 3-5.8 中，經由我們最後估計式計算出來的預測值與實際值相當接近，並且實際值幾乎都落在 95%信賴區間上、下界之內，表示此配適模型的預測能力不錯，適合拿來分析此資料。由於前面方法有做過介入分析且效果都比未作前加，因此我們也對 ARIMA 做介入分析。在做完 ARIMA 介入分析後發現，與無介入分析的所做出的估計值、實際值與 95%信賴區間預測值並無太大差別，因此我們決定不採用 ARIMA 介入分析後的模型。

## 第六節 最適模型

模型的準確度需要靠指標來衡量，而衡量的方法為樣本外預測，即除了在樣本以外，選取的十二筆真實值與各模型所預測出的十二比預測值與95%信賴區間值做比較，用以比較各模型預測能力的好壞。在用以比較樣本外的實際值與預測值之方法，我們使用下列四種評估指標：均方誤差（Root Mean Square Error, RMSE）、平均誤差百分比（Mean Percentage Error, MPE）、平均絕對誤差百分



比 (Mean Absolute Percentage Error, MAPE)、平均絕對誤差 (Mean Absolute Deviation, MAD)。當RMSE、MPE、MAPE、MAD這四種評估指標越接近零，表示預測值含實際值的總誤差越小，即代表解釋能力越好，而可得知何種方法較為適合預測本資料。接著再從各個評估指標中找出各指標中解釋能力最佳的模型，最後獲得我們所選定的最終模型。

欄 1	MAD	RMSE	MPE	MAPE
分解法	110906.4167	2.38E+10	0.846	2.2855
指數平滑法	113347.8333	24358458830	0.3266	2.3146
Time Series	120988.0283	22036390046	-0.4955	2.5006
ARIMA	124012.9167	23197502466	-0.4501	2.5524

表3-6.1 模型比較表

由表3-6.1可發現，以上四種方法在MAPE的準則下皆在2%~3%之間，因此可得知這四種方法的預測能力都算良好，而時間序列迴歸法在MSE的評估準則下表現最佳，而指數平滑法在MPE的準則下為最佳，分解法在MAD與MAPE的評估準則下表現最佳。在以上比較之下我們可得知最佳的預測模型為分解法。



## 第四章 結論與建議

從我們的研究發現，分解法的實際值幾乎落在預測區間之內，因此我們選擇使用此模型來預測。用電消費量雖然在 2008 年 8-9 月之後受到台電連續調漲電費的影響，再加上政府開始呼籲節能減碳以及推廣相關政策而大幅下降。因此我們認為除政府政策之外，電費的價格會直接影響到用電的消費量。

我們發現政府在電費的調節對於用電量的減少有成效，在來是由原始序列圖來看可以得到一個推論，就是目前工業在台廠房的數量與 2007 年前後的數量相仿。因為國內的用電大部分都是供給給工業用電，所以我們有此推論。

再來是火力發電的原料的價格會間接影響消費者的用電量，因為原物料的價格提高，政府購入原料的成本就會提高，因此不是將成本轉嫁給消費者就是提高用電的收費量

這樣一來也會影響用電量，雖然火力發電的成本較高並且對環境污染也較大，但是因為輻射對人體的危害使大家不禁未雨綢繆，即使核電廠產電過程中不產生溫室氣體，成本也較低，大家還是對於興建核電廠意願不高，因此這些還有待相關單位去取得經濟效益與自然環境維護的平衡。

從預測圖中我們可以發現對用電的需求量依舊會上升，在這個倡導節能省碳的時代中使用煤氣、石油等原物料的火力發電，必定漸漸地減少生產量，而在此情況下生產量居次位的核能發電必定會提高產電量：也就是短時間內核能發電是必要存在的。有鑒於日本的核輻射事件，政府現在必須同時進行在核電廠建造的結構安全上與尋找再生能源並將其普及化為當前最重要的兩件事。

### 參考文獻

1. 陳婉淑 (2011) , 統計預測方法講義。
2. 歷屆逢甲大學優質學生報告
3. AREMOS 經濟統計資料庫網站  
[http://cache.moe.edu.tw/aremos\\_ly/search.html](http://cache.moe.edu.tw/aremos_ly/search.html)
4. 統計預測評估準則參考網站：  
<http://acqy.csai.cn/user2/51384/archives/2009/41819.html>
5. 計算用電量網站  
<http://www.peopo.org/sukusan/post/23202>
6. 台灣電力公司網站  
<http://www.taipower.com.tw/>
7. Bowerman, B., O'Connell, R., and Koehler, A. (2005) Forecasting, Time Series, and Regression, 4th edition, Duxbury Press.

