



逢甲大學學生報告 ePaper

國內能源消費-能源用

作者：林貞好、翁瑞嬪、楊程皓、李銘峻

系級：統計三乙

學號：D9639335、D9639174、D9639424、D9659982

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： 98 學年度 第 2 學期

中文摘要

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。如何節約能源及提升能源效率是我們現階段該認真思考的問題。

為了瞭解國內的能源消費是否隨著政府的宣導而有效的減少能源用量，我們利用 AREMOS 資料庫下載了「國內能源消費-能源用」從 1992 年到 2009 年的 216 筆月資料，應用統計預測方法來分析歷年來國內能源使用量的狀況。並且保留最後的 12 筆數據，用以和預測分析做比較。

我們先判斷原始資料是否具有趨勢，以及其平均數和變異數是否為常數。接著利用 SAS 軟體，依序使用時間序列迴歸法、分解法、指數平滑法及 ARIMA (Autoregressive Integrated Moving Average Model) 配適模型，並做預測分析。最後利用 MSE、MAD、MPE 及 MAPE 四個準則來評估預測表現，選出配適度最佳的模型為時間序列迴歸法所配適的模型。

我們預測未來幾年的能源消費量依舊會年年持續增加，面對近來環保議題，政府如何有效宣導及人民應當如何節約能源是急需探討的議題。

關鍵字：能源消費、時間序列迴歸法、分解法、指數平滑法、Autoregressive Integrated Moving Average (ARIMA) Model

目 次

第一章	緒論.....	5
第一節	研究背景.....	5
第二節	研究動機.....	5
第三節	研究目的.....	6
第四節	資料來源.....	6
第五節	資料分析.....	7
第六節	研究流程(流程圖).....	8
第二章	研究方法.....	9
第一節	時間序列迴歸法(Time Series Regression).....	9
一、	何謂時間序列迴歸法(Time Series Regression)：.....	9
二、	模型的配適.....	10
一、	何謂分解法(Decomposition Method)：.....	16
二、	模型的配適.....	16
第三節	指數平滑法(Exponential Smoothing).....	23
一、	何謂指數平滑法(Exponential Smoothing).....	23
二、	模型的配適：.....	24
第四節	ARIMA 分析法.....	27
一、	何謂 ARIMA.....	27
二、	模型的配適.....	28
第五節	最佳模型的選擇.....	35
第三章	結論與建議.....	36
參考文獻	38

圖 目 錄

圖 1-1	原始時間序列圖.....	7
圖 2-1.1	原始時間序列圖.....	10
圖 2-1.2	轉換後的時間序列圖.....	10
圖 2-1.3	DW 檢定統計量	12
圖 2-1.4	DW 檢定.....	12
圖 2-1.5	加入一階自我相關項後 DW 檢定.....	13
圖 2-1.6	實際值、預測值、95%預測信賴上界、95%預測區信賴下界之圖...	15
圖 2-2.1	原始時間序列圖.....	16
圖 2-2.2	去季節因子時間序列圖.....	17
圖 2-2.3	趨勢循環時間序列圖.....	17
圖 2-2.4	季節因子時間序列圖.....	18
圖 2-2.5	不規則因子時間序列圖.....	18
圖 2-2.6	DW 檢定統計量.....	20
圖 2-2.7	殘差的自我相關性之檢定.....	20
圖 2-2.8	加入一階自我相關後殘差的檢定.....	20
圖 2-2.9	實際值、預測值、95%信賴區間上界、95%信賴區間下界之圖.....	22
圖 2-3.1	原始資料的時間序列圖.....	24
圖 2-3.2	指數平滑的估計值、實際值和 95%的上下界預測區間的預測圖.....	26
圖 2-4.1	原始時間序列圖.....	28
圖 2-4.2	原始時間序列的 ACF 及 PACF.....	28
圖 2-4.3	square root 轉換且加入一次差分的 ACF 及 PACF.....	29
圖 2-4.4	轉換後且加入非季節差分及季節差分的 ACF 及 PACF.....	30
圖 2-4.5	我們的模型的 ACF 及 PACF.....	30
圖 2-4.6	我們的模型的殘差檢定	32
圖 2-4.7	我們的模型的好壞判斷	33
圖 2-4.8	ARIMA 的估計值、實際值和 95%的上下界預測區間的預測.....	34

表 目 錄

表 2-1.1	參數估計.....	13
表 2-1.2	時間序列迴歸法之真實值及預測值.....	15
表 2-2.1	參數估計.....	21
表 2-2.2	分解法之真實值及預測值.....	22
表 2-3.1	參數估計.....	25
表 2-3.2	指數平滑法真實值及預測值.....	26
表 2-4.1	我們的模型的參數估計.....	33
表 2-4.2	ARIMA 的實際值及預測值.....	34
表 2-5.1	模型評估.....	35



第一章 緒論

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。如何節約能源及提升能源效率是我們現階段該認真思考的問題。

第一節 研究背景

對於現代人而言，能源就像空氣一樣重要，全世界有 80% 的能源供應來自於煤、石油、天然氣……等石化燃料。根據專家估計，石化燃料大約再 40 年即將面臨危機。能源是一切活動的原動力，與經濟、政治、國防、科技、交通、社會及衛生環境等各方面均有極密切的關係。因此，能源的重要性是無庸置疑的。

近年來在國際市場炒作、政治問題、氣候等因素的影響，國際能源價格波動非常劇烈，不管是產業或者是民生都受到極大的影響，整個地球環境也因過度的能源使用，面臨了種種警訊，例如：全球氣候極端異常、全球溫室效應變化、空氣污染嚴重、海平面上升、農作物生產受影響、動物大遷移……等等。人們對於能源的依賴性不斷增加，但能源卻是有限的，因此我們必須透過能源科技的不斷研發，使能源的取得與供應方式更多元化，尋找另一項具有乾淨、環保又永不耗盡的新能源，以供應人們迫切的需求。

第二節 研究動機

近年來隨著時代與科技的進步，人類的需求越來越大，消耗能源的速度也越來越快，但是人們只想著要如何使用它，卻忽略了該如何善用它，導致地球上的資源不斷的減少。目前，人類消耗的一些重要能源大都是不可再生的，例如：石油、天然氣、電力、煤炭……等等。而且隨著人類物質需求不斷增長，能源的消耗量越來越高，使得地球上有限資源的再生速度遠遠不及人類所使用的速度。能源是人類生活的必需品，更是國家建設的原動力，節約能源、有效的

運用能源必須從全體國民做起，加強宣導全民正確使用能源的觀念。雖然近年來環保意識抬頭，提倡永續經營的概念，但為了提高生活品質以及人與人之間的互通便利，使得各國在經濟發長與能源的使用兩者之間出現了矛盾，要在兩者間做出選擇，達到最佳的平衡點並不簡單，但若不再加以重視，繼續不節制的使用能源，只會加快地球死亡的速度，人類也因此滅亡。

第三節 研究目的

近年來由於氣候暖化的緣故，對環境產生重大影響，造成世界各地天災頻傳，各國科學家一致認為是使用能源所產生的二氧化碳導致的，因此開始宣導節能減碳。不只是使用化石燃料會排放溫室氣體，就連電的使用也會造成溫室氣體的增加，根據研究指出，每產生 1 度電，所消耗的化石燃料約釋放 0.67 公斤的二氧化碳。所以要減少溫室氣體排放量，最基本的方法就是節約能源。因此我們使用統計預測方法，以 1992 年到 2008 年的數據來預測未來一年(2009 年)的能源使用量。若估計出來的預測值還是呈現上升趨勢，代表國人對於氣候暖化的危機意識還是不夠，政府對於節能減碳的宣導效果不佳，需更進一步探討重視此問題。

第四節 資料來源

為了瞭解歷年來能源使用量的狀況，我們利用 AREMOS 資料庫下載了「國內能源消費-能源用」的資料，單位為公秉油當量(KLOE)。國內能源用消費是我國工業部門、運輸部門、農業部門、住宅部門、商業部門和其他部門的每個月使用有效能源的使用量。有效能源也就是已經轉換為可以直接使用的能源，在能源消費裡面占大部分的是石油產品和電力。我們利用西元 1992 年 01 月到 2009 年 12 月的 216 筆月資料做預測分析，並且保留最後的 12 筆真實資料，做為和預測值比較用。

第五節 資料分析

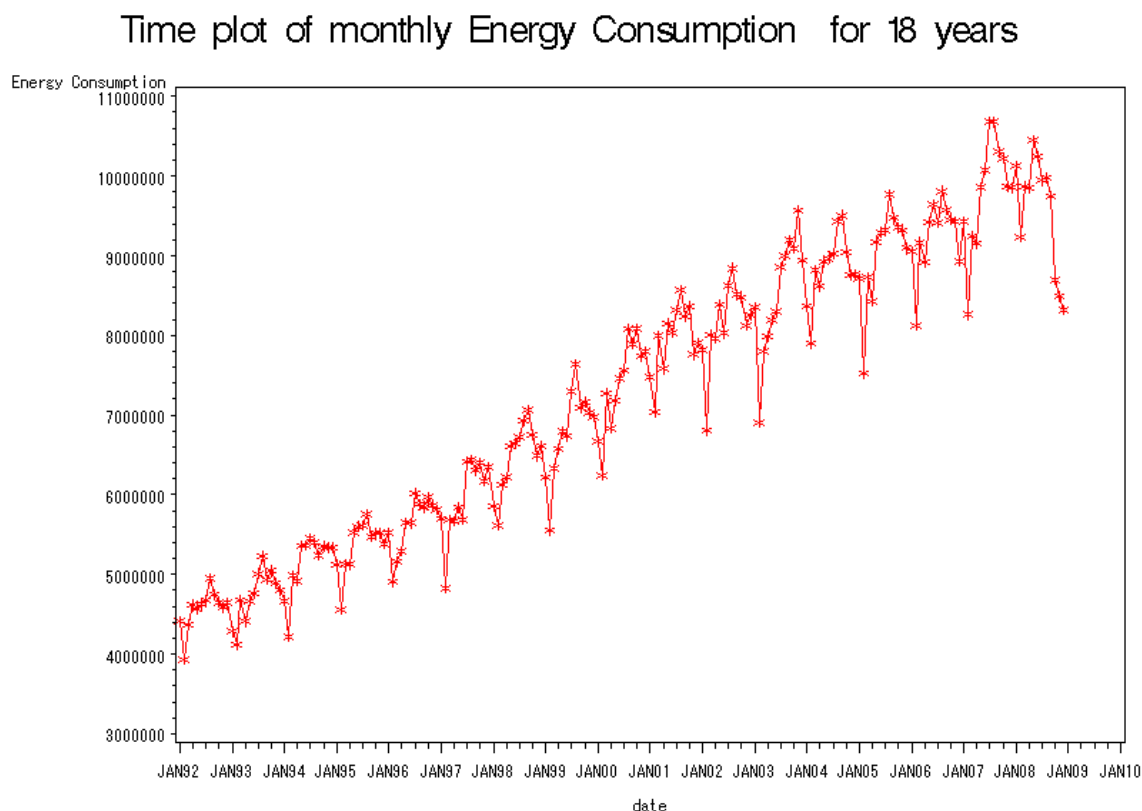
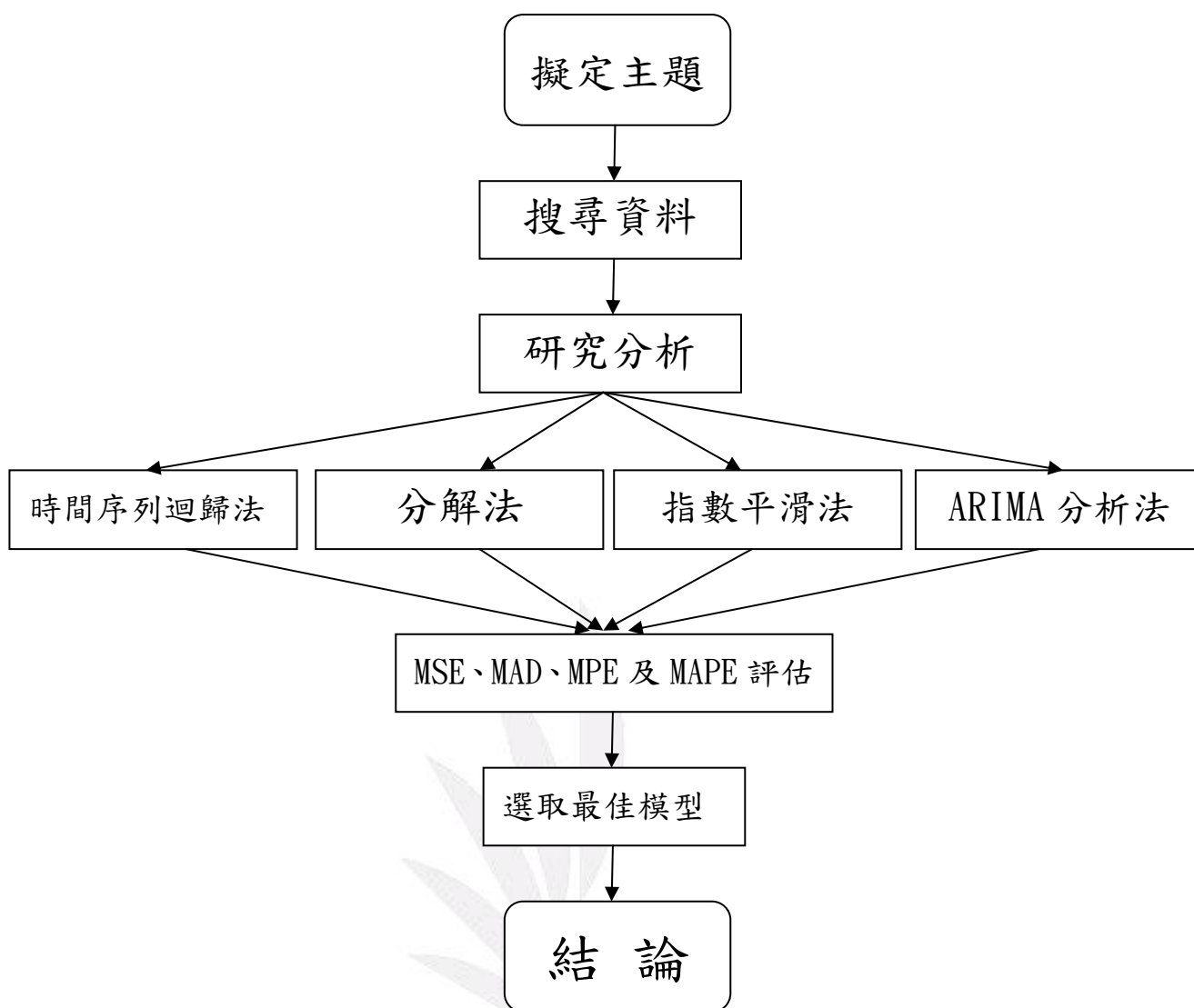


圖 1-1 原始時間序列圖

由原始時間序列圖可看出能源消費使用量有歷年增加的趨勢，且變異隨著時間的增加而增加，但在 2008 年左右，能源消費量增加的趨勢有減緩的現象，由上述可得知此筆資料的平均數及變異數皆不平穩。由圖 1-1 亦可發現能源消費使用量在每年年中(6~8 月)都是較多的，而年初及年底的用量相對較少，由此得知此筆資料具季節性。

從原始資料圖發現在 2008 年中時，能源消費量降低的趨勢特別多，我們認為是受到原油價格飆漲的影響，連帶影響國內油品價格，再加上 2008 年 9 月的金融海嘯，導致國內的消費意願降低，能源消費量減少。

第六節 研究流程(流程圖)



為了瞭解國內人民使用消費能源是否有減緩的趨勢，我們利用 AREMOS 資料庫下載了「國內能源消費-能源用」的資料，單位為公秉油當量。時間為西元 1992 年 01 月到 2009 年 12 月的 216 筆月資料做預測分析。我們利用時間序列迴歸法、分解法、指數平滑法及 ARIMA 分析法，四種預測方法做預測模型，並且對 2009 年做預測。接著利用 MAPE、MPE、MSE、MAD 這四種評估指標，作為模式精準判斷的依據，選出最佳的模型。最後進行討論並做出結論。

第二章 研究方法

首先判斷原始資料是否具有趨勢，以及其平均數和變異數是否平穩。接著我們利用 SAS 軟體，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 2009 年的使用量做預測。最後利用保留的真實值與預測值以 MSE、MAD、MPE 及 MAPE 四個準則來做評估，選出配適度最佳的模型。

第一節 時間序列迴歸法(Time Series Regression)

一、何謂時間序列迴歸法(Time Series Regression)：

時間序列是指將被觀測到的數據依時間次序排列而成的數據序列。從經濟、金融到工程技術，從天文、地理到氣象，從醫學到生物，幾乎在各個領域中都涉及到時間序列。時間序列分析即是對時間序列數據進行統計分析及推斷。時間序列的資料往往不能以迴歸分析的方法來建立模型加以分析，因為迴歸分析想要建立的是因果模型。而時間序列中之殘差與各觀測值間通常都存在相關性，這種序列相關性已經違背了各觀測值為獨立的必要假設。因此，時間序列迴歸分析是僅依照變數本身過去的資料所存在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(ε_t)來表示。

時間序列模型：

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

二、模型配適

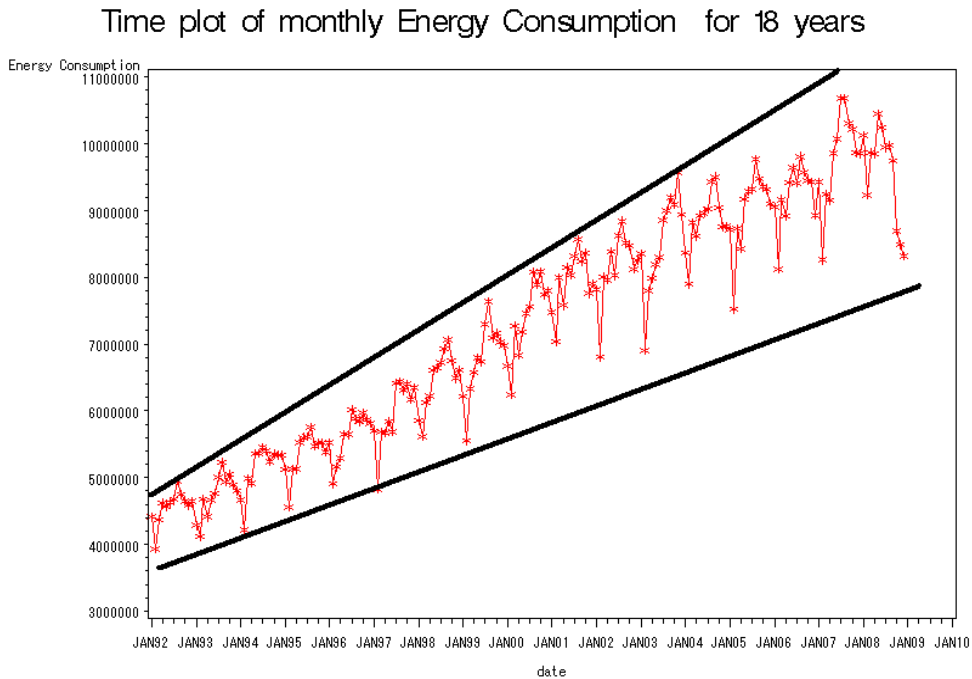


圖 2-1.1 原始時間序列圖

首先判斷此筆資料的序列是否平穩。由圖 2-1.1 可看出時間序列的波動不一致，變異數隨著時間的增加而增加，且季節變異波動中，隨著時間呈現波動越來越大的情形，表示變異數不為常數也就是不平穩。為使變異數趨近於平穩，我們透過 square roots 對原始數據加以轉換。

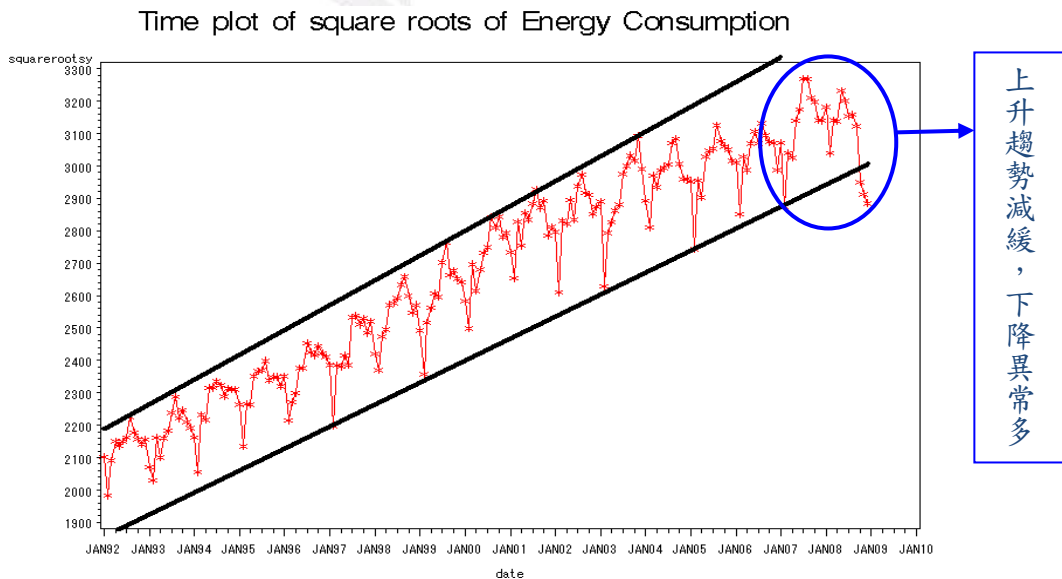


圖 2-1.2 轉換後的時間序列圖

由圖 2-1.2 可發現經過轉換後的時間序列圖之變異大致為一常數，即變異數趨近於平穩。但由圖 2-1.2 可看出 2008 年上升的趨勢有減緩的現象，且最後幾個月的能源消費量持續的降低，低於歷年來上升趨勢應有的值許多。

我們對 2008 年進行調查，發現 2008 年初全球股市因憂慮美國經濟衰退，而紛紛慘跌，且美國與日本在當時公布的經濟數據都略優於市場預期等因素，導致國際油價開始上漲。在 2008 年 5 月奈及利亞情勢再度傳出動盪，關切全球能源供應吃緊，而引發投機交易，油價首度飆破 120 美元。油價、物價的飆漲加上當時的經濟不景氣，使得人民多搭乘大眾運輸交通工具，減少自行開車就減少石油能源的消費量，以及節約能源減少用電量，以對應油價高漲及經濟不景氣。又受 2008 年九月金融風暴的影響，使得 2008 年中後之能源的消費量遞減的趨勢異常的多。

當資料在某個時間點或者某個時段出現了異常的表現，我們就會針對此狀況進行介入分析，使預測能力不會受到此異常狀況而影響。讓預測結果可以更佳。因此我們針對 2008 年 5 月(197 筆資料)進行介入分析：

$$S_t = \begin{cases} 0 & \text{if } t < 197 \\ 1 & \text{if } t \geq 197 \end{cases}$$

接著對此模型的殘差做診斷。當殘差不存在自我相關性時，才是一個好的配適模型。因此我們對殘差做以下的檢定：

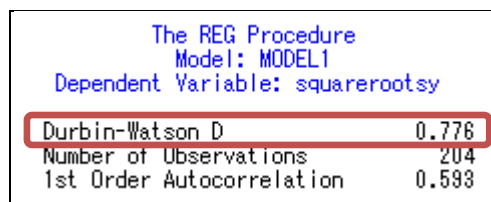
DW 檢定統計量檢測規則：

當 DW 值為 2 時，表示此模型的殘差不具自我相關性，若 DW 值介於 0 到 2 之間則表示殘差存在正自我相關，而 DW 值介於 2 到 4 之間則表示殘差存在負自我相關。

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。

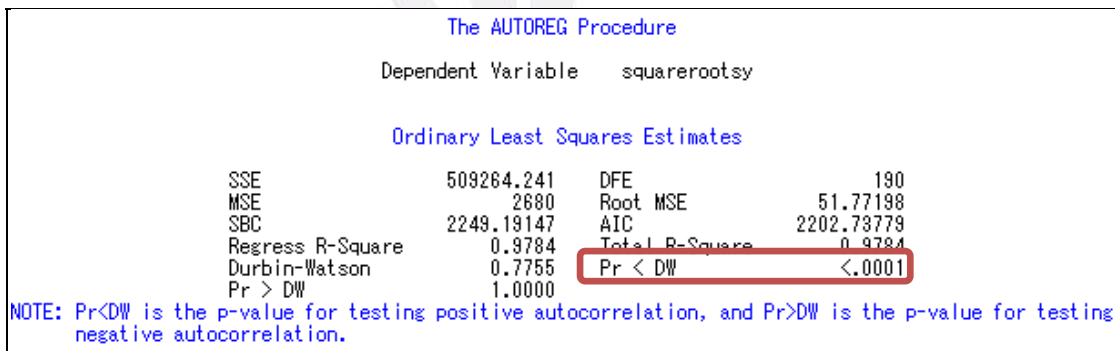
當 $Pr > DW$ 的值小於顯著水準 $\alpha = 0.05$ 時，表示顯著
→ 存在負自我相關。



```
The REG Procedure
Model: MODEL1
Dependent Variable: squarerootsy

Durbin-Watson D      0.776
Number of Observations      204
1st Order Autocorrelation    0.593
```

圖 2-1.3 DW 檢定統計量



```
The AUTOREG Procedure
Dependent Variable      squarerootsy

Ordinary Least Squares Estimates

SSE      509264.241      DFE      190
MSE      2680          Root MSE      51.77198
SBC      2249.19147      AIC      2202.78779
Regress R-Square      0.9784      Total R-Square      0.9784
Durbin-Watson      0.7755      Pr < DW      <.0001
Pr > DW      1.0000

NOTE: Pr<DW is the p-value for testing positive autocorrelation, and Pr>DW is the p-value for testing
negative autocorrelation.
```

圖 2-1.4 DW 檢定

由圖 2-1.3 可知 DW 值為 0.776，屬於正自我相關。我們進一步做自我相關的檢定，由圖 2-1.4 的殘差檢定，發現 $Pr < DW$ 的值小於顯著水準，由此得知此模型存在正自我相關。

因為模型存在自我相關性時不是一個好的配適，因此我們加入一階自我相關項，消除自我相關性使模型的配適更合適。

The AUTOREG Procedure

Estimates of Autoregressive Parameters

Lag	Coefficient	Standard Error	t Value
1	-0.592922	0.058574	-10.12

Yule-Walker Estimates

SSE	318960.138	DFE	189
MSE	1688	Root MSE	41.08065
SBC	2159.49099	AIC	2109.71919
Regress R-Square	0.9329	Total R-Square	0.9885
Durbin-Watson	2.1919	Pr < DW	0.8058
Pr > DW	0.1942		

NOTE: Pr<DW is the p-value for testing positive autocorrelation, and Pr>DW is the p-value for testing negative autocorrelation.

圖 2-1.5 加入一階自我相關項後 DW 檢定

由圖 2-1.5 得知 Pr>DW 的值及 Pr<DW 的值皆大於顯著水準 $\alpha=0.05$ ，表示不具有自我相關性，此模型配適佳。

表 2-1.1 參數估計

Variable	DF	Estimate	Standard error	t value	Approx Pr> t
ϕ	1	0.592922	0.058574	-10.12	<.0001
β_0	1	2082	17.471	119.19	<.0001
β_1	1	5.662	0.1277	46.14	<.0001
β_2	1	-27.6784	11.3844	-2.43	0.016
β_3	1	-166.7925	14.1905	-11.75	<.0001
β_4	1	-16.32	15.5846	-1.05	0.2964
β_5	1	-33.1447	16.3147	-2.03	0.0436
β_6	1	32.8079	16.6707	1.97	0.0505
β_7	1	42.4961	16.6959	2.55	0.0117
β_8	1	79.7619	16.578	4.81	<.0001
β_9	1	110.1505	16.207	6.8	<.0001
β_{10}	1	69.4164	15.4547	4.49	<.0001
β_{11}	1	50.6532	14.0226	3.61	0.0004
β_{12}	1	15.9213	11.1353	1.43	0.1544
β_{13}	1	-183.2318	31.2839	-5.86	<.0001

由表 2-1.1 及圖 2-1.5 得到此模型的參數估計值及標準差。將所有的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：

$$y_t^* = 2082 + 5.662t - 27.6784M_1 - 166.7925M_2 - 16.32M_3 - 33.1447M_4 + 32.8079M_5 + 42.4961M_6 + 79.7619M_7 + 110.1505M_8 + 69.4164M_9 + 50.6532M_{10} + 15.9213M_{11} - 183.2318S_t + \varepsilon_t$$

$$Y_t^* = \sqrt{y_t}$$

$$\varepsilon_t = 0.592922\varepsilon_{t-1} + a_t, \quad a_t \sim N(0, \sigma^2)$$

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

$$S_t = \begin{cases} 0 & \text{if } t < 197 \\ 1 & \text{if } t \geq 197 \end{cases}$$

M_t 為季節虛擬解釋變數，其中：

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

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

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

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

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

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

為了瞭解我們所做的模型之估計是否適當，我們保留了最後的 12 筆真實數據，做為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、下界預測區間做成表格(表 2-1.2)，再將這四種值畫成預測曲線圖：

表 2-1.2 時間序列迴歸法之真實值及預測值

日期	實際值	預測值	95%預測信賴下界	95%預測信賴上界
2009/1/1	7715603	8597148.39	8006976.47	9208303.83
2009/2/1	7930365	8061982.81	7429879.57	8719888.03
2009/3/1	8501703	9118254.55	8423312.85	9840738.46
2009/4/1	8808304	9137316.39	8433766.79	9869047.56
2009/5/1	9135677	9628007.91	8902414.06	10382024.67
2009/6/1	9541611	9754919.72	9032885.39	10504712.81
2009/7/1	9936410	10043821.53	9310459.39	10804981.81
2009/8/1	9697328	10284970.93	9542446.27	11055312.45
2009/9/1	9771000	10067903.61	9333260.78	10830374.34
2009/10/1	9779946	9988866.88	9257055.44	10748514.96
2009/11/1	9026592	9808270.7	9083126.11	10561260.17
2009/12/1	9698809	9745480.59	9022631.56	10496180.29

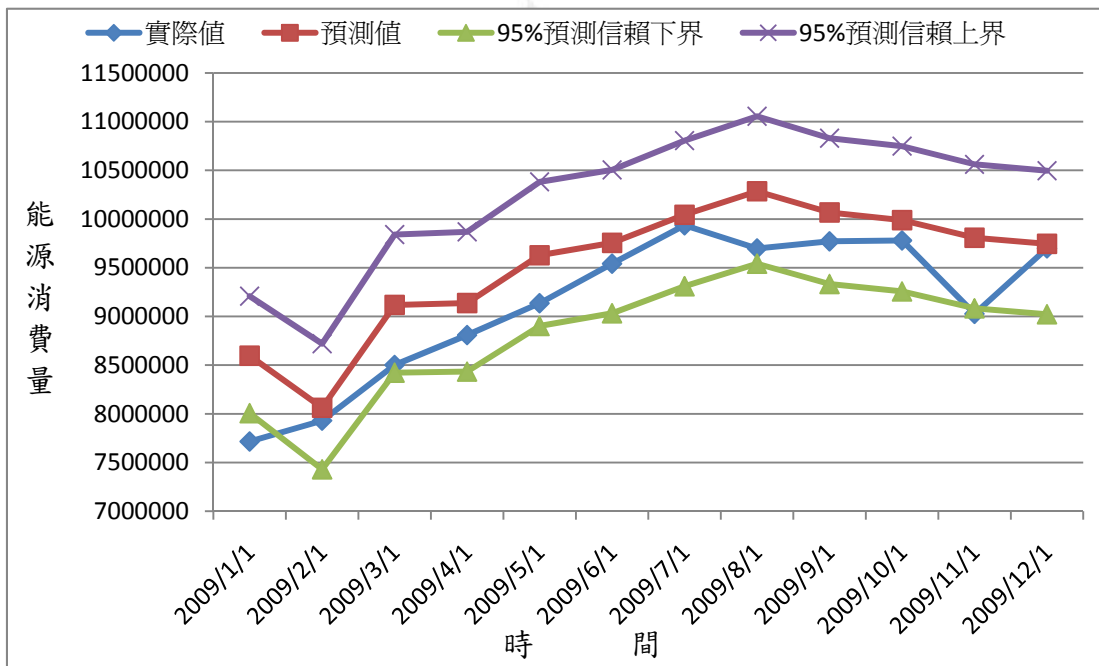


圖 2-1.6 實際值、預測值、95%預測信賴上界、95%預測信賴下界之圖

由圖 2-1.6 的預測結果看出大部分的實際值都落在 95%信賴區間之內，因此我們認為以此方法做出的預測模型是適合拿來分析我們的資料。

第二節 分解法(Decomposition Method)

一、何謂分解法(Decomposition Method)：

分解法是將時間序列分解成趨勢(Trend； TR_t)、季節(Season； SN_t)、循環(Cyclical； CL_t)及不規則(Irregular； IR_t)四大因素，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 $y_t = f(T_t, S_t, C_t, I_t)$ 。我們利用 X11 的方法來分解原始的資料，對原始時間序列資料分別做去季節因子、趨勢循環、季節因子、不規則因子的時間序列圖，可看出資料是否存在季節性，或者有何趨勢。

分解法又可分為加法模型及乘法模型兩種。當時間序列的變異數為一常數，也就是時間序列的變異數不會隨著時間的改變而改變時，適用加法模型；若時間序列的變異數不平穩，表示序列的變異數會隨著時間的改變而改變，則適用乘法模型。

加法模型：

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

乘法模型：

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

二、模型的配適

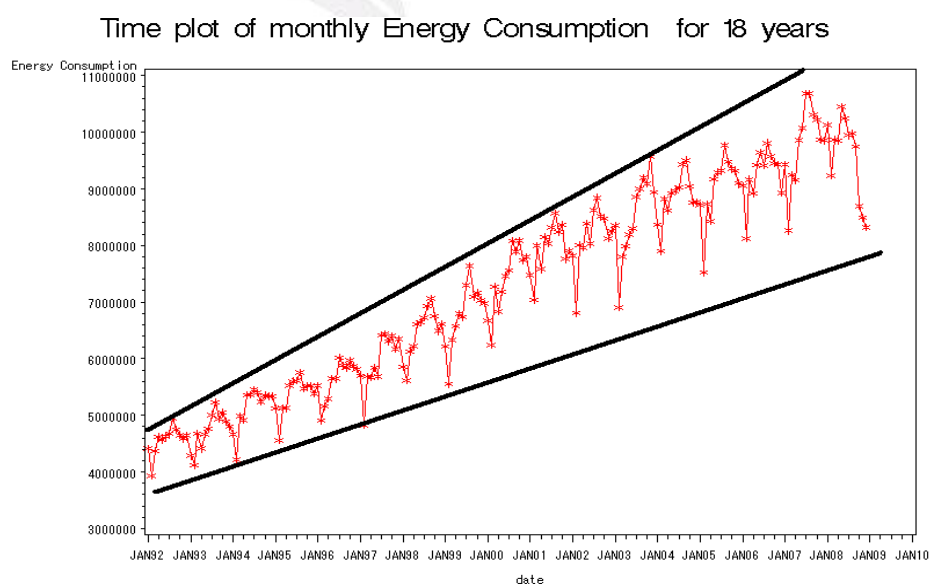


圖 2-2.1 原始時間序列圖

首先我們選擇適當的配適模型。由圖 2-2.1 可看出時間序列的波動不一致，變異隨時間的增加而增加，即變異數不平穩。因此我們決定選用乘法模型來做估計預測。

Time plot of deseasonalized Energy Consumption

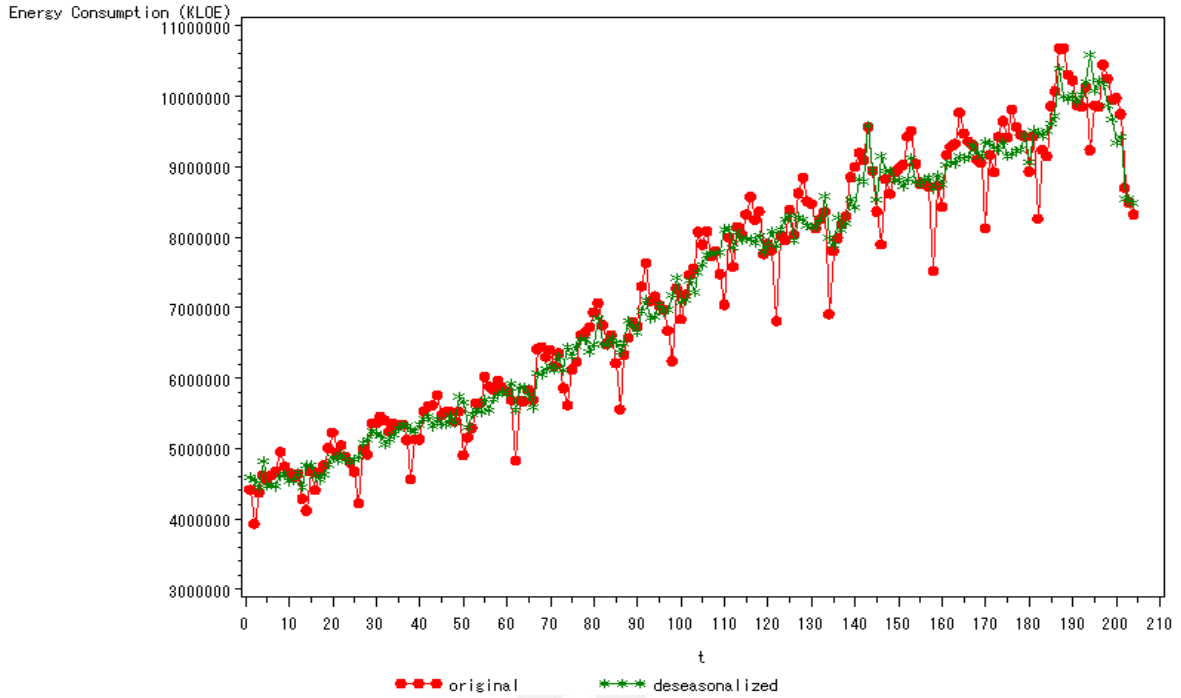


圖 2-2.2 去季節因子時間序列圖

Trend-cycle component

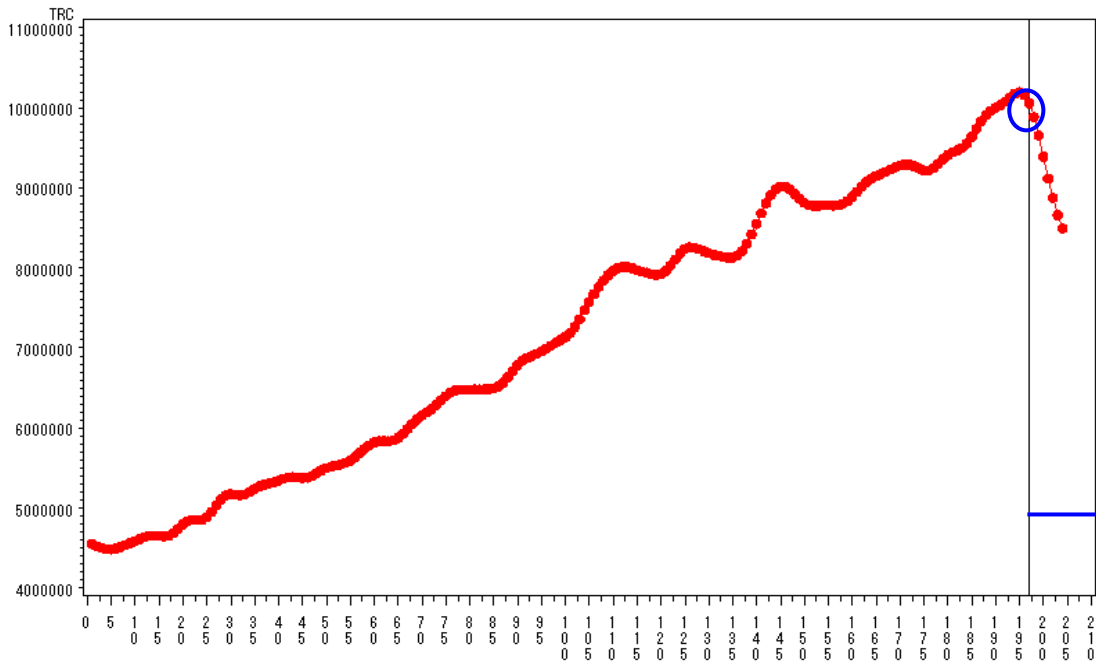


圖 2-2.3 趨勢循環時間序列圖

第 197 筆資料，2008 年 5 月

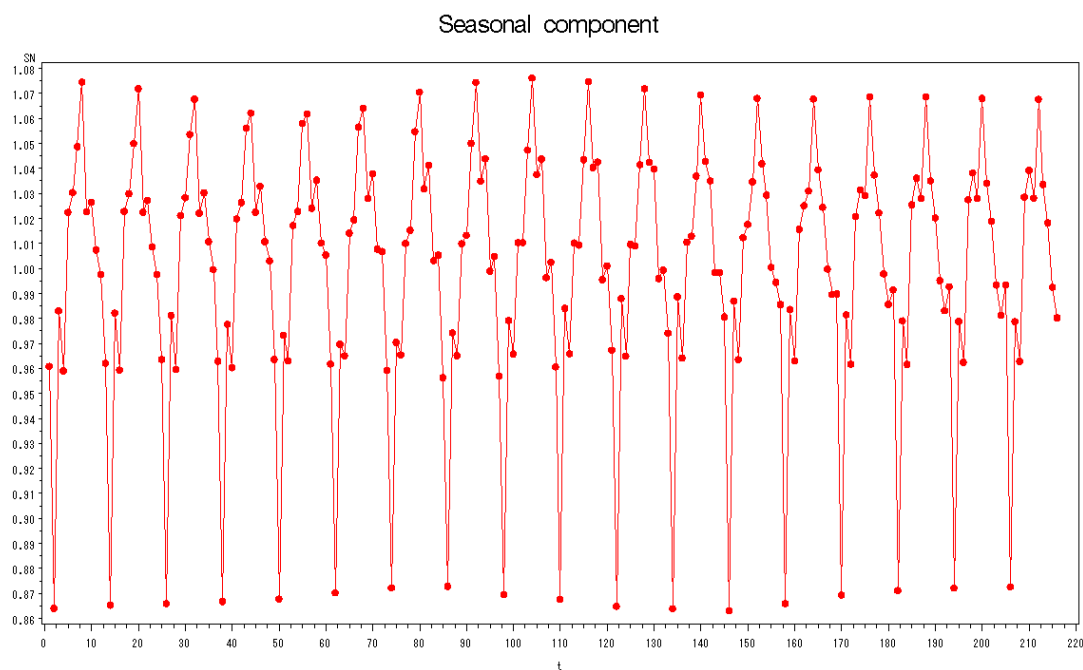


圖 2-2.4 季節因子時間序列圖

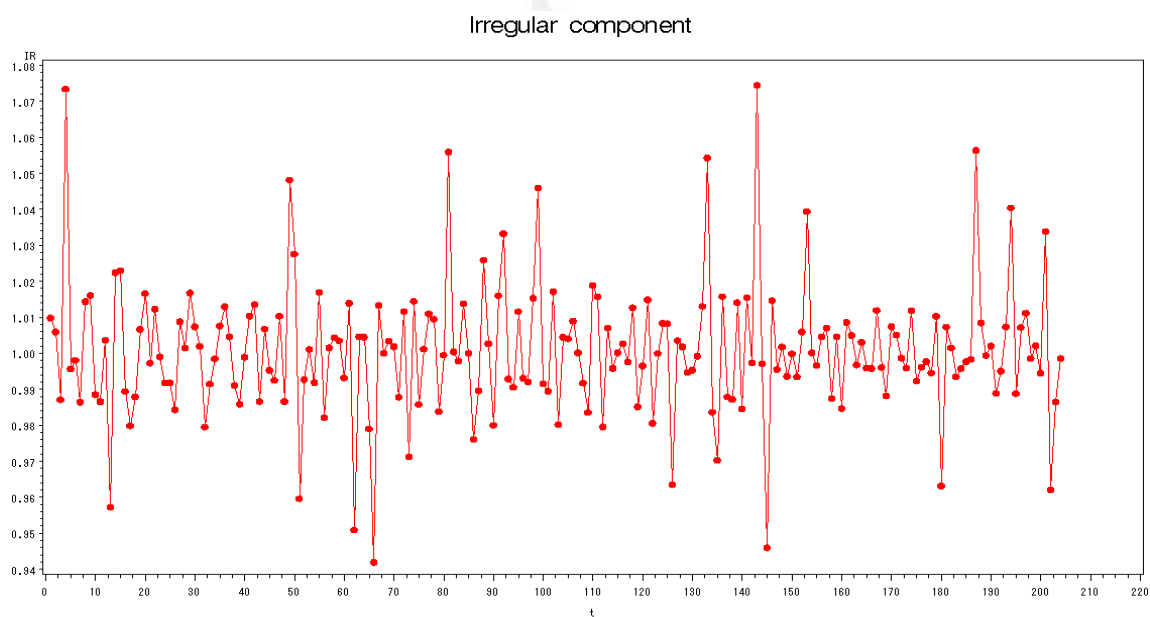


圖 2-2.5 不規則因子時間序列圖

接著觀察資料的趨勢季節性。由圖 2-2.2 可看出去季節因素後的時間序列仍有上升的趨勢，意味著即使沒有季節因素的影響，能源消費量也會隨著時間的增加而增加。圖 2-2.4 可看出能源消費量有明顯的季節變化。

由圖 2-2.3 可明顯看出能源消費量隨時間的增加有明顯上升的趨勢循環，但在 2008 年 5 月後卻異常持續的下降，我們調查此時間點後發現當時奈及利亞情勢再度傳出動盪，關切全球能源供應吃緊，引發投機交易，導致油價不段的飆漲，首度飆破 120 美元。加上當時經濟不景氣，使得人民多搭乘大眾運輸交通工具，減少自行開車減少石油能源消費量，以對應油價高漲及經濟不景氣。再加上 2008 年九月金融風暴的影響，使得 2008 年中後能源消費量的趨勢循環呈現遞減的現象。

當資料在某個時間點或者某個時段出現了異常的表現，我們就會針對此狀況進行介入分析，使預測能力不會受到此異常狀況而影響。讓預測結果可以更佳。因此我們針對 2008 年 5 月(197 筆資料)進行介入分析：

$$S_t = \begin{cases} 0 & \text{if } t < 197 \\ 1 & \text{if } t \geq 197 \end{cases}$$

接著對此模型的殘差做診斷。當殘差不存在自我相關性時，才是一個好的配適模型。因此我們對殘差做以下的檢定：

DW 檢定統計量檢測規則：

當 DW 值為 2 時，表示此模型的殘差不具自我相關性，若 DW 值介於 0 到 2 之間則表示殘差存在正自我相關。而 DW 值介於 2 到 4 之間則表示殘差存在負自我相關。

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。

The REG Procedure	
Model: MODEL1	
Dependent Variable: DY	
Durbin-Watson D	0.753
Number of Observations	204
1st Order Autocorrelation	0.597

圖 2-2.6 DW 檢定統計量

The AUTOREG Procedure			
Dependent Variable DY			
Ordinary Least Squares Estimates			
SSE	1.32342E13	DFE	201
MSE	6.5842E10	Root MSE	256597
SBC	5673.60578	AIC	5663.65142
Regress R-Square	0.9790	Total R-Square	0.9790
Durbin-Watson	0.7529	Pr < DW	<.0001
Pr > DW	1.0000		

NOTE: Pr<DW is the p-value for testing positive autocorrelation, and Pr>DW is the p-value for testing negative autocorrelation.

圖 2-2.7 殘差的自我相關性之檢定

由圖 2-2.6 得到 DW 的值為 0.753，屬於正自我相關。又由圖 2-2.7 的殘差檢定，發現 Pr<DW 的值小於顯著水準，由此得知此模型存在正自我相關。

因為模型存在自我相關性時，不是一個好的配適，因此我們加入一階自我相關項，消除自我相關性使模型的配適更合適。

Estimates of Autoregressive Parameters			
Lag	Coefficient	Standard Error	t Value
1	-0.596971	0.056729	-10.52

Yule-Walker Estimates			
SSE	8.18875E12	DFE	200
MSE	4.09437E10	Root MSE	202346
SBC	5581.43504	AIC	5568.16256
Regress R-Square	0.9256	Total R-Square	0.9270
Durbin-Watson	2.0995	Pr < DW	0.7224
Pr > DW	0.2776		

NOTE: Pr<DW is the p-value for testing positive autocorrelation, and Pr>DW is the p-value for testing negative autocorrelation.

圖 2-2.8 加入一階自我相關後殘差的檢定

由圖 2-2.8 得知 Pr>DW 的值及 Pr<DW 的值皆大於顯著水準 $\alpha=0.05$ ，表示殘差不具有自我相關性，得知此模型配適合適。

表 2-2.1 參數估計

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > t
ϕ_1		0.596971	0.056729	-10.52	<.0001
β_0	1	4199405	70266	59.76	<.0001
β_1	1	30008	609.527	49.23	<.0001
β_2	1	-959069	152451	-6.29	<.0001

由表 2-2.1 得到各個參數估計值。且由此表可知此模型的參數估計之 P-value 都小於顯著水準 $\alpha = 0.05$ ，故參數估計皆顯著，表示此模型的配適是合適的。將所有參數估計的值帶入模型，得到分解法的預測方程式：

$$dy_t = 4199405 + 30008t - 959069S_t + \varepsilon_t$$

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

$$\varepsilon_t = 0.596971\varepsilon_{t-1} + a_t$$

$$S_t = \begin{cases} 0 & \text{if } t < 197 \\ 1 & \text{if } t \geq 197 \end{cases}$$

為了瞭解我們所做的模型之估計是否適當，我們保留了最後的 12 筆真實數據，做為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、下界預測區間做成表格(表 2-2.2)，再將這四種值畫成預測曲線圖。

表 2-2.2 分解法之真實值及預測值

日期	實際值	預測值	95%預測信賴下界	95%預測信賴上界
2009/1/1	7715603	8809655.39	8321026.46	9298284.32
2009/2/1	7930365	7949688.02	7472665.14	8426710.9
2009/3/1	8501703	9068747.34	8515843.11	9621651.58
2009/4/1	8808304	9022798.67	8472677.84	9572919.5
2009/5/1	9135677	9715415.81	9125360.04	10305471.58
2009/6/1	9541611	9875811.68	9278671.75	10472951.61
2009/7/1	9936410	9817298.53	9226163.29	10408433.76
2009/8/1	9697328	10237136.85	9623048.57	10851225.12
2009/9/1	9771000	9947226.68	9352598.36	10541855.01
2009/10/1	9779946	9833877.3	9247937.04	10419817.57
2009/11/1	9026592	9616664.3	9045448.35	10187880.24
2009/12/1	9698809	9528938.55	8964640.31	10093236.79

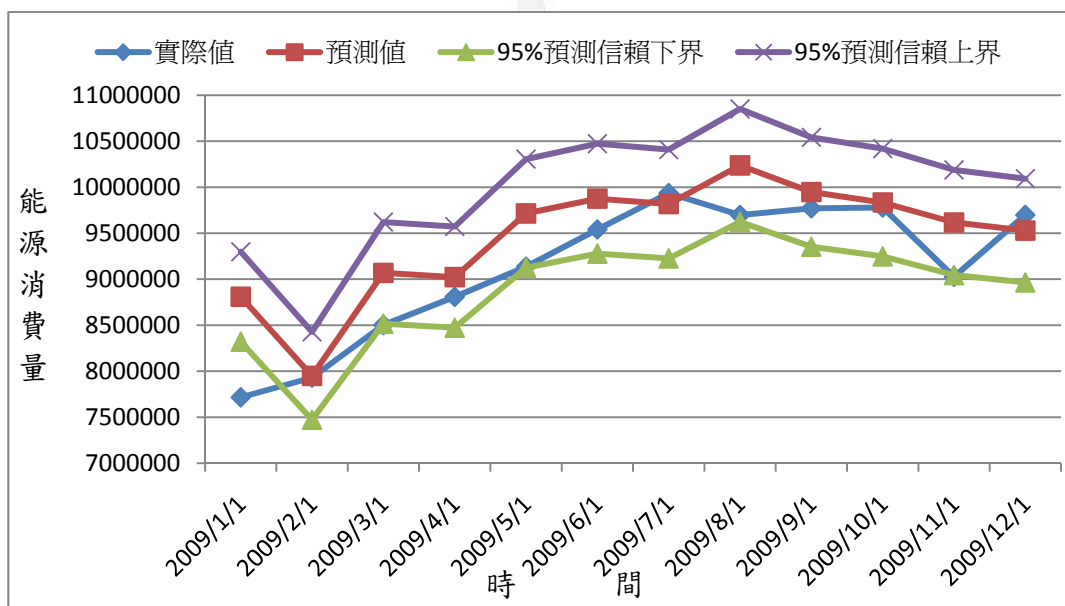


圖 2-2.9 實際值、預測值、95%信賴區間上界、95%信賴區間下界之圖

由圖 2-2.9 的預測結果可以看出大部分的實際值都落在 95%信賴區間之內，因此我們認為以此方法做出的預測模型是適合拿來分析我們的資料。

第三節 指數平滑法(Exponential Smoothing)

一、何謂指數平滑法(Exponential Smoothing)

指數平滑法又稱指數加權平均法，實際是加權的移動平均法，它是選取各時期權重數值為遞減指數數列的均值方法。相容了全期平均和移動平均所長，不捨棄過去的資料，但是僅給予逐漸減弱的影響程度，即隨著資料的遠離，賦予逐漸收斂為零的權數。指數平滑法解決了移動平均法需要n個觀測值和不考慮 t_n 前時期數據的缺點，通過某種平均方式，消除歷史統計序列中的隨機波動，找出其中的主要發展趨勢，對未來進行預測。

「指數平滑法」是生產預測中常用的一種方法，也用於中短期經濟發展趨勢之預測。當資料有季節性且變異數平穩，不隨時間的改變而改變時，我們使用Winters Method-Additive做配適；若資料具季節性但變異數不平穩，會隨著時間的改變而改變時，我們則使用Winters Method-Multiplicative。而若資料不具季節性，有上升的趨勢，但上升趨勢有減緩的現象，則使用Damped Trend 做配適。

Winters Method-Additive模型：

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

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

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

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

Winters Method-Multiplicative 模型：

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

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

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

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

Damped Trend Exponential Smoothing的模型為：

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

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

二、模型的配適：

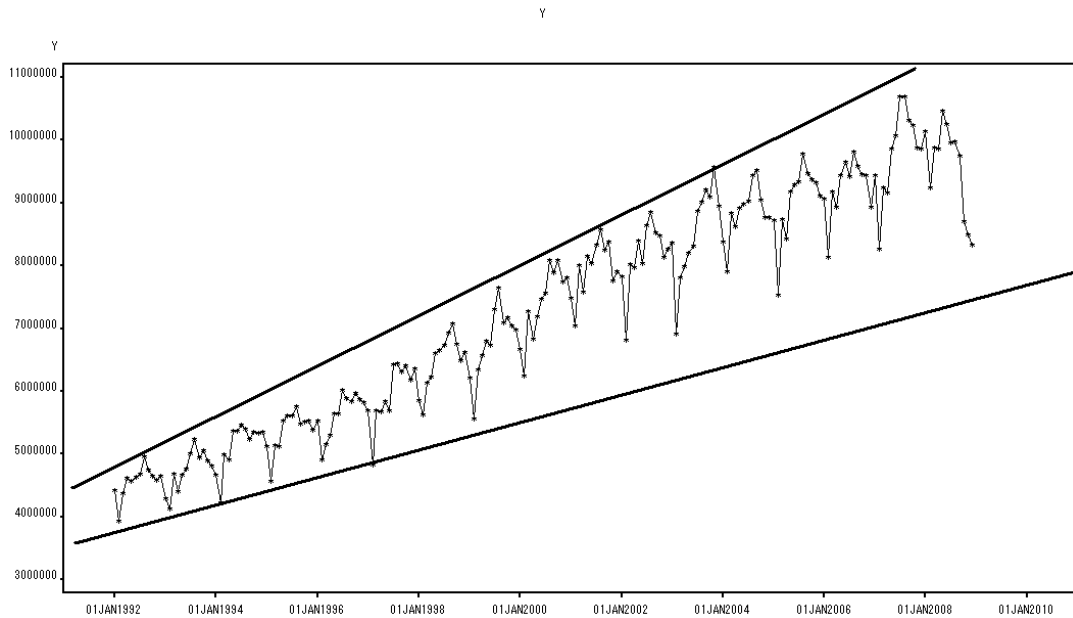


圖2-3.1 原始資料的時間序列圖

首先我們選擇適當的配適模型。由圖2-3.1可以看出資料具季節性，且變異隨著時間的增加而增加，表示變異數不為常數，也就是不平穩。因此我們選取Winters Method-Multiplicative模型做配適。

由表2-3.1可得到各個參數的估計值，又可知 $\hat{\gamma}$ 的估計並不顯著，代表這個變數對此模型的影響並不大。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：

$$L_t = 0.59704(Y_t/S_{t-s}) + (1-0.59704)(L_{t-1} + b_{t-1})$$

$$b_t = 0.00208(L_t - L_{t-1}) + (1-0.00208)b_{t-1}$$

$$S_t = 0.30737(Y_t/L_t) + (1-0.30737)S_{t-s}$$

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

表2-3.1 參數估計

Model Parameter	Estimate	Std. Error	T	Prob> T
$\hat{\alpha}$	0.59704	0.0445	13.4149	<.0001
$\hat{\gamma}$	0.00208	0.0075	0.2757	0.7831
$\hat{\delta}$	0.30737	0.0695	4.4243	<.0001
σ^2	5.34E+10	.	.	.
Smoothed Level	8463910	.	.	.
Smoothed Trend	23059	.	.	.
Smoothed Seasonal Factor 1	0.98004	.	.	.
Smoothed Seasonal Factor 2	0.86491	.	.	.
Smoothed Seasonal Factor 3	0.96667	.	.	.
Smoothed Seasonal Factor 4	0.9557	.	.	.
Smoothed Seasonal Factor 5	1.01078	.	.	.
Smoothed Seasonal Factor 6	1.01143	.	.	.
Smoothed Seasonal Factor 7	1.03518	.	.	.
Smoothed Seasonal Factor 8	1.06779	.	.	.
Smoothed Seasonal Factor 9	1.04485	.	.	.
Smoothed Seasonal Factor 10	1.02083	.	.	.
Smoothed Seasonal Factor 11	1.00473	.	.	.
Smoothed Seasonal Factor 12	0.99098	.	.	.

為了瞭解我們所做的模型之估計是否適當，我們保留了最後的12筆真實數據，做為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、下界預測區間做成表格(表2-3.2)，再將這四種值畫成預測曲線圖。

表2-3.2 指數平滑法真實值及預測值

日期	實際值	預測值	95%信賴區間上限	95%信賴區間下限
2009/1/1	7715603	8317590	8770499	7864682
2009/2/1	7930365	7360369	7872532	6848206
2009/3/1	8501703	8248707	8855810	7641603
2009/4/1	8808304	8177147	8838445	7515849
2009/5/1	9135677	8671675	9411980	7931370
2009/6/1	9541611	8700590	9490168	7911013
2009/7/1	9936410	8928744	9778466	8079022
2009/8/1	9697328	9234648	10148705	8320590
2009/9/1	9771000	9060355	9999236	8121474
2009/10/1	9779946	8875583	9836570	7914595
2009/11/1	9026592	8758791	9746216	7771366
2009/12/1	9698809	8661748	9675937	7647559

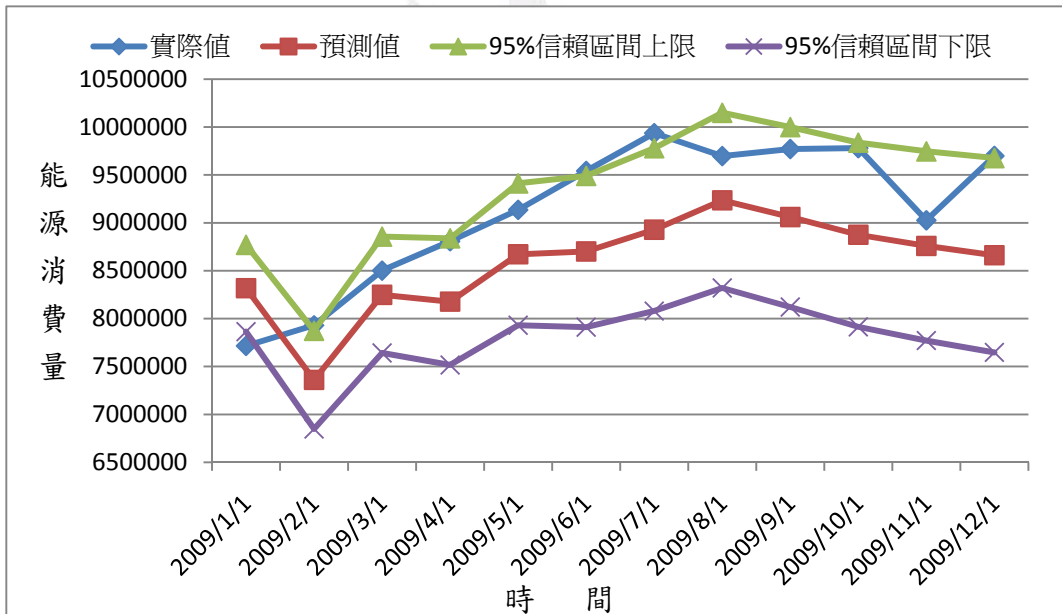


圖 2-3.2 指數平滑的估計值、實際值和 95%的上下界預測區間的預測圖

由圖 2-3.2 可看出實際值大致都落在 95%上下區間內，綜合以上結果，我們認為此模型預測可以拿來分析我們的資料。

第四節 ARIMA 分析法

一、何謂 ARIMA

ARIMA模型全稱為自迴歸移動平均模型(Autoregressive Integrated Moving Average Model)，是由博克思(Box)和詹金斯(Jenkins)於70年代初提出的時間序列預測法，所以又稱為博克思-詹金斯法或box-jenkins模型。

ARIMA模型的基本思想，是將預測對象隨時間推移而形成的數據序列，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。此模型一旦被識別後就可以從時間序列的過去值及現在值來預測未來值。

其中ARIMA(p, d, q)稱為差分自迴歸移動平均模型，AR(p)是自迴歸模型(autoregressive model)，其意義簡單來說就是現在的某一變數值，和同一變數過去p期的變數值有關，p為自迴歸項；MA(q)為移動平均(moving average)，其意義簡單來說就是現在的和過去幾個q期的隨機項，q為移動平均項數有關；d則是時間序列成為平穩時所做的差分次數。

ARIMA(p, q)模式敘述如下：

$$Z_t = \delta + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \cdots + \phi_n Z_{t-n} + a_t \\ - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_m a_{t-m}$$

二、模型的配適

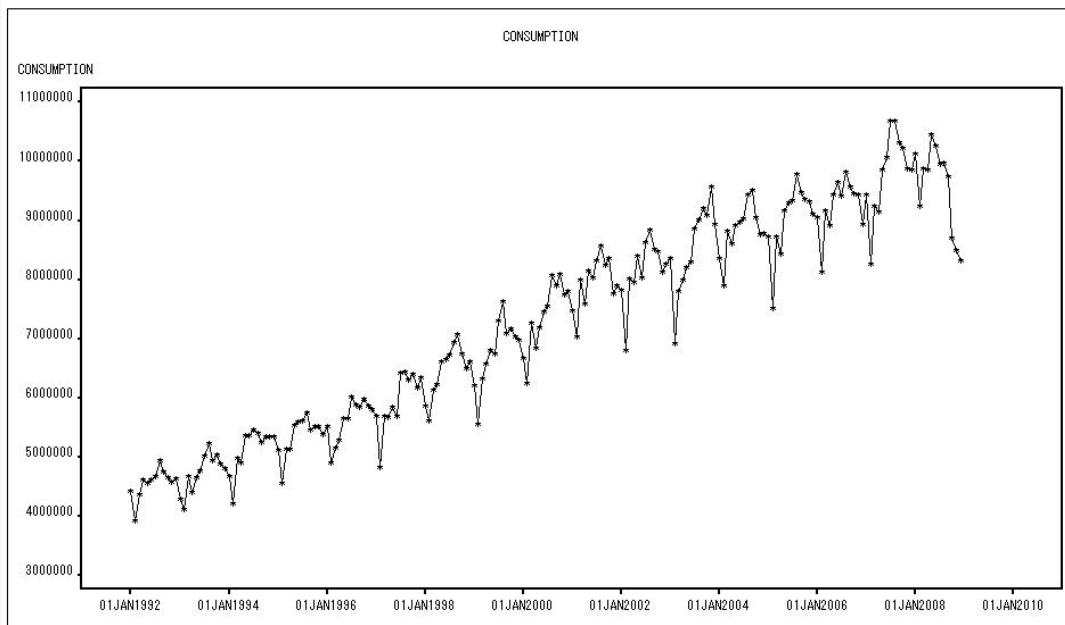


圖 2-4.1 原始時間序列圖

首先用肉眼觀察原始資料的時間序列圖，判斷資料是否平穩，由圖 2-4.1 可以得知，時間序列的波動不一致，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。且資料呈現逐漸上升的趨勢，代表平均數不平穩。因此，必須透過轉換和差分加以改善，使其平穩。

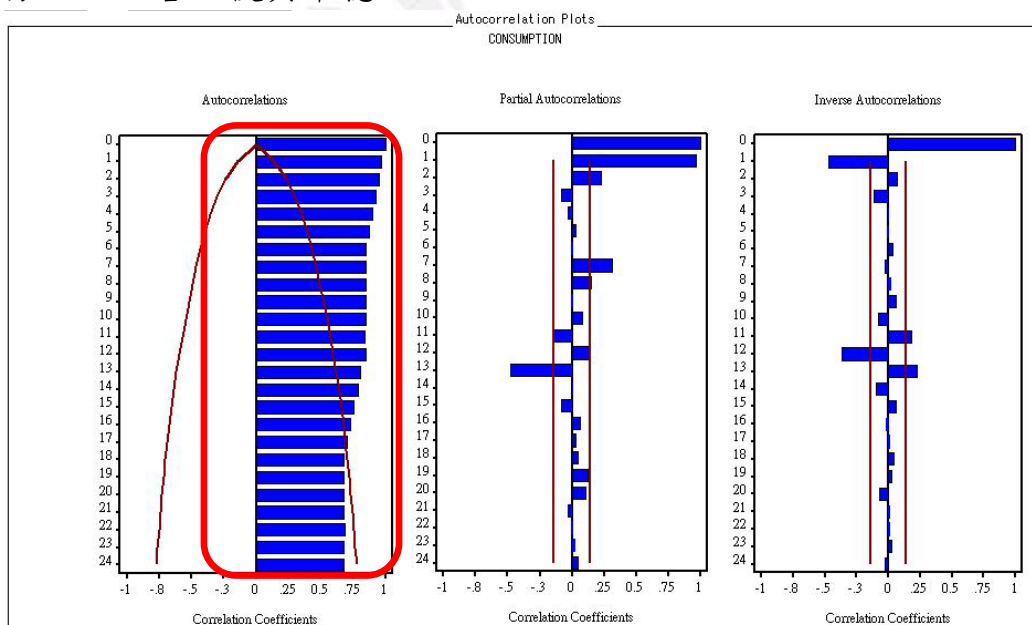


圖 2-4.2 原始時間序列的 ACF 及 PACF

原始資料的 ACF 及 PACF 圖亦可用來判別資料是否平穩。當 ACF 圖呈現下降非常緩慢(dies down slowly)的狀態時，表示此序列平均數不平穩，需做差分加以改善。

由圖 2-4.2 可得知 ACF 圖呈現下降緩慢(dies down slowly)的狀態，得知資料不平穩，因此我們做一次差分加以調整。

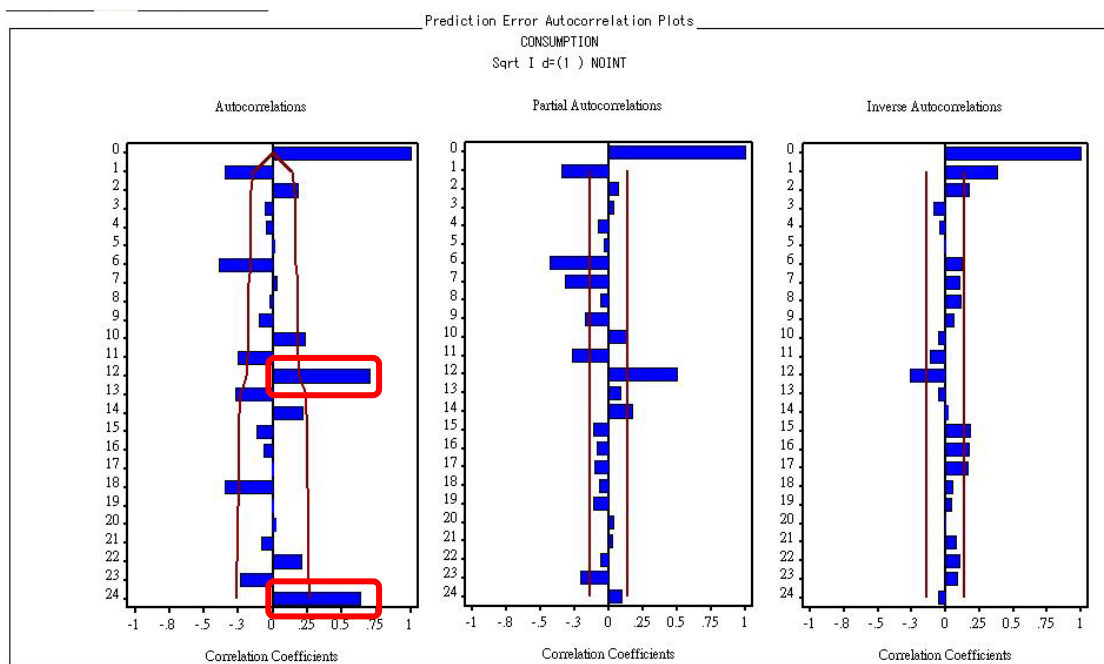


圖 2-4.3 square root 轉換且加入一次差分的 ACF 及 PACF

經由圖 2-4.3 的 ACF 可看出，殘差在 lag12 和 lag24 這兩處特別凸出，超過兩倍標準差的範圍許多，表示此資料呈現季節性波動，必須對此資料做季節差分來解決此不平穩因素。

解決所有不平穩因素後，接著配適適當的模型，ARIMA 預測方法的配適規則如下：

當 ACF 為 dies down 且 PACF 呈現 cuts off 時，配適 AR(p)。

當 ACF 為 cuts off 且 PACF 呈現 dies down 時，配適 MA(q)。

當 ACF 以及 PACF 皆為 dies down 時，則配適 ARMA(p, q)。

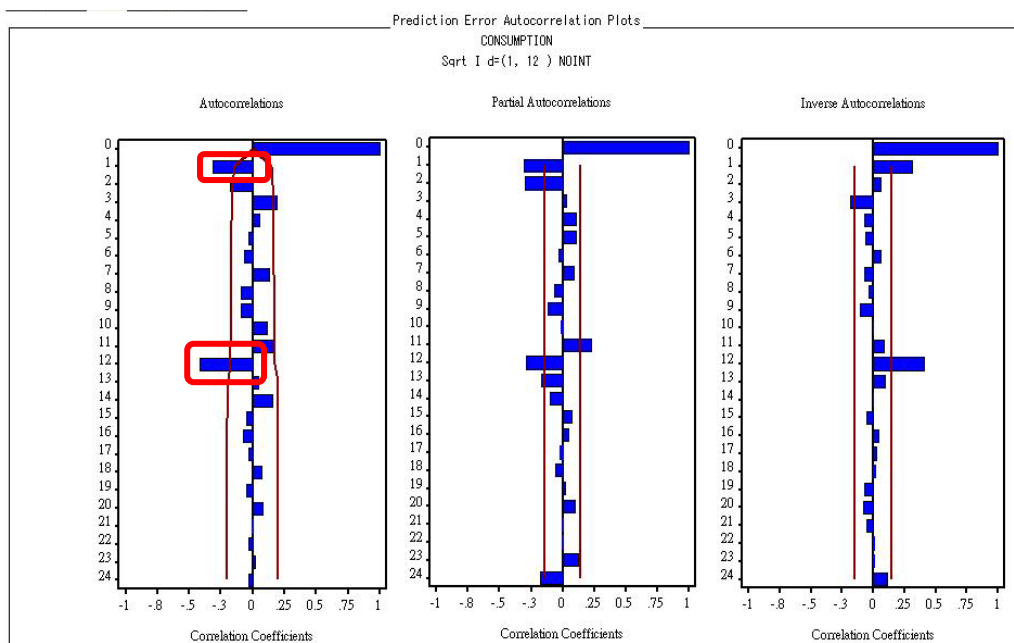


圖 2-4.4 轉換後且加入非季節差分及季節差分的 ACF 及 PACF

由圖 2-4.4 我們判斷 ACF 是 cuts off 而 PACF 為 dies down 所以我們配適模型 $MA(q)$ 。又由圖 2-4.4 的 ACF 得知 lag1 及 lag12 凸出兩倍標準差特別的多，因此我們加入這兩項加以配適($q=(1)(12)$)。

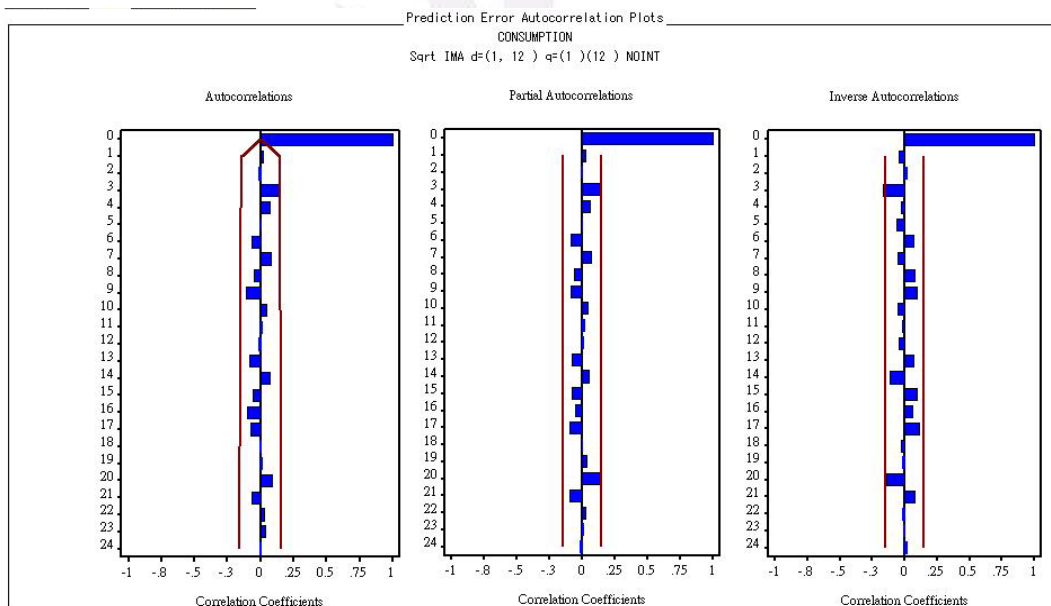


圖 2-4.5 我們的模型的 ACF 及 PACF

經過開根號轉換並且配適 $ARIMA(0,1,1)(0,1,1)_s$ NOINT 的模型後，由圖 2-4.5 的 ACF 及 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。

接著診斷配適的模型之殘差。若殘差具有單根，即表示此模型不平穩，會隨著時間的改變而改變，這樣的模型並不是一個好的配適。一個好的配適模型，其殘差值會落在兩倍標準差以內。且殘差的平均為零，變異數為常數，彼此之間不具有相互關係，也就是符合白噪音，並且不存在單根，這才是一個好的配式。因此我們對此模型的殘差做白噪音及單根的檢定，診斷此模型的殘差，確定所配適的模型是否恰當。

⊙白噪音之檢定：

假設檢定：
$$\begin{cases} H_0 : \text{white noise} & (\text{殘差符合白噪音}) \\ H_1 : \text{no white noise} & (\text{殘差不符合白噪音}) \end{cases}$$

決策規則：

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

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

⊙單根之檢定：

假設檢定：
$$\begin{cases} H_0 : \text{有單根} & (\text{時間序列不平穩}) \\ H_1 : \text{沒有單根} & (\text{時間序列平穩}) \end{cases}$$

決策規則：

P-value大於0.05，則不拒絕 H_0 ，表示有單根，模型配適不合適。

P-value小於0.05，則拒絕 H_0 ，表示沒有單根，模型配適合適。

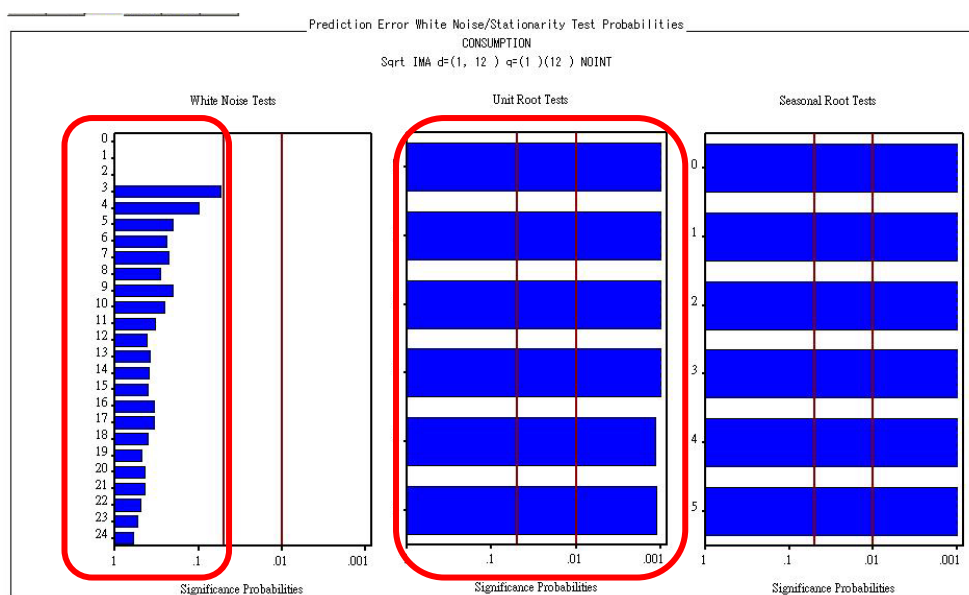


圖 2-4.6 我們的模型的殘差檢定

由圖 2-4.6 的 White Noise Test 知，所有的 lag 都大於顯著水準 $\alpha = 0.05$ ，不拒絕 H_0 ，表示殘差項具有 White Noise 現象。又由圖 2-4.6 的 Unit Root Test 得知，所有的 lag 皆小於顯著水準 $\alpha = 0.05$ ，拒絕 H_0 ，表示時間序列已為平穩狀態。由上之檢定得知，我們所配適的模型 Sqrt ARIMA(0,1,1)(0,1,1)_s NOINT 是合適的。

由於我們利用時間序列迴歸法及分解法進行預測時，預測結果總是高估，大部分的實際值都不在預測區間內。因此我們對 2008 年 5 月進行介入分析，加入介入分析後使兩種預測方法的預測表現都變得更好。而我們對此筆資料利用 ARIMA 預測方法預測的結果，實際值大部分就都落在預測區間內，進一步進行介入分析的預測結果並沒有特別好，表現不佳，因此我們不對 ARIMA 做介入分析。

表 2-4.1 我們的模型的參數估計

Model parameter	Estimate	Std. Error	T	Prob> T
MA factor 1 lag 1	0.4074	0.0657	6.2018	<.0001
MA factor 2 lag 12	0.792	0.0711	11.1346	<.0001
Model Variance (sigma squared)	1764	.	.	.

由表 2-4.1 可得到各個參數估計值，且知兩個參數估計值的 P-value 皆小於顯著水準 $\alpha=0.05$ ，表示參數顯著，都應被留在模型中。總括上述結果可得知，我們所配適的 Sqrt ARIMA(0,1,1)(0,1,1)s NOINT 模型是合適的。將表 2-4.1 的參數估計值帶入模型中得到我們的 ARIMA 預測方程式如下：

$$(1-B)(1-B^{12}) \sqrt{y_t} = (1-0.40740B)(1-0.79200B^{12})a_t$$

$$\hat{\sigma}^2 = 1764$$

Statistic of Fit	Value
Mean Square Error	5.67984E10
Root Mean Square Error	238324.1
Mean Absolute Percent Error	2.45605
Mean Absolute Error	179376.9
R-Square	0.981

圖 2-4.7 我們的模型的好壞判斷

由圖 2-4.7 得到 R-square 等於 0.981，可知此模型對此筆資料具有 98.1%的解釋能力。

為了瞭解我們所做的模型之估計是否適當，我們保留了最後的 12 筆真實數據，做為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、下界預測區間做成表格(表 2-4.2)，再將這四種值畫成預測曲線圖。

表 2-4.2 ARIMA 的實際值及預測值

日期	實際值	預測值	95%預測信賴上界	95%預測信賴下界
2009/1/1	7715603	8429039	8911983	7956118
2009/2/1	7930365	7535787	8067818	7017300
2009/3/1	8501703	8433194	9065411	7818042
2009/4/1	8808304	8324023	9014840	7653792
2009/5/1	9135677	8828554	9598996	8082218
2009/6/1	9541611	8860019	9687129	8060536
2009/7/1	9936410	9003797	9889884	8148856
2009/8/1	9697328	9208308	10154380	8296903
2009/9/1	9771000	8997576	9980164	8053174
2009/10/1	9779946	8713556	9725269	7743550
2009/11/1	9026592	8539571	9583615	7540754
2009/12/1	9698809	8398532	9474503	7371309

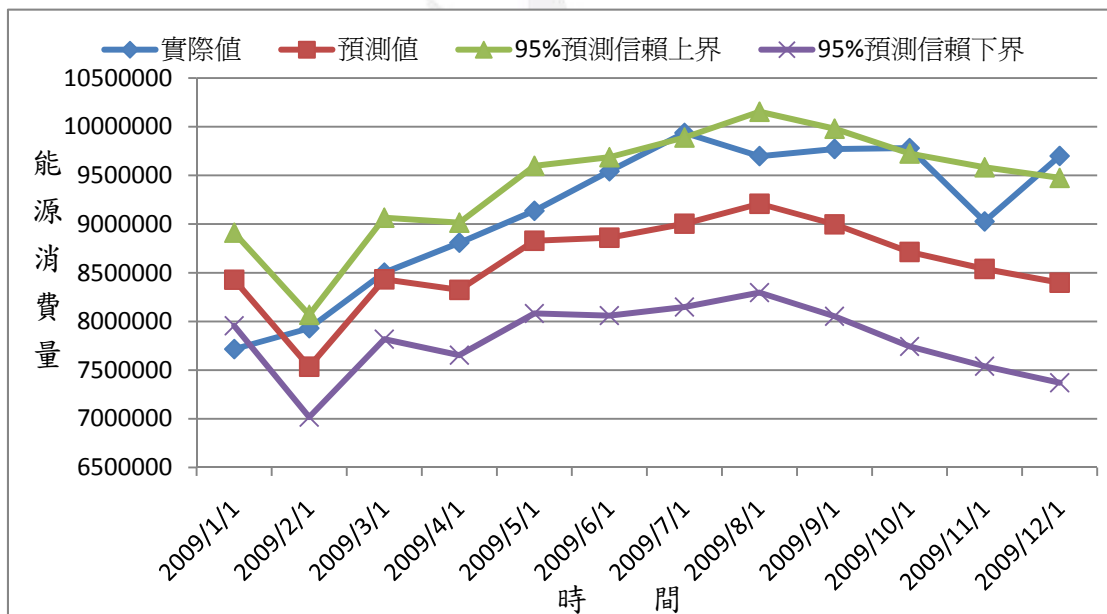


圖 2-4.8 ARIMA 的估計值、實際值和 95%的上下界預測區間的預測圖

藉由圖 2-4.8 可看出，經由我們最後預測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還不錯，適合拿來分析此筆資料。

第五節 最佳模型的選擇

模型是否準確，也就是精確度，需要靠指標來做衡量，我們選擇使用均方誤差(Mean Square Error, MSE)、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) 作為模式精準判斷的依據，選出最佳的模型。

當MAPE、MPE、MSE、MAD這四種評估指標的值越接近0，表示預測值和實際值的總誤差越小，由以得知預測方法越適當。

其中：

$$MAD = \frac{\sum |y_t - \hat{y}_t|}{n}$$

$$MSE = \frac{\sum (y_t - \hat{y}_t)^2}{n}$$

$$PE_t = \frac{y_t - \hat{y}_t}{y_t} \times 100 \quad \Rightarrow MPE = \frac{\sum PE_t}{n}$$

$$APE_t = \frac{|y_t - \hat{y}_t|}{y_t} \times 100 \quad \Rightarrow MAPE = \frac{\sum APE_t}{n}$$

表 2-5.1 模型評估

	MAD	MSE	MPE	MAPE
時間序列迴歸法	391133	223000000000.00	-4.42948	4.429482
分解法	371789.6	227000000000.00	-3.74043	4.232132
指數平滑法	641522	520000000000.00	5.382319	6.92343
ARIMA	645947.9	481000000000.00	5.709318	7.009685

由時間序列迴歸法的實際值和 95%預測區間的預測圖可以發現，在此模型下的預測值和實際值相近。又由表 2-5.1 可以發現，時間序列迴歸法在 MAD 和 MSE 的評估準則下表現最佳，且在 MPE 或 MAPE 的準則下也為次佳的。相較於其他的模型預測，時間序列迴歸法的預測最為良好，由此我們得到最佳的模型為時間序列迴歸法。

第三章 結論與建議

從我們的研究發現，時間序列迴歸法的實際值幾乎落在預測區間之內，所以我們選擇使用此模型來預測。能源消費量雖然在 2008 年受到美國原油庫存量下滑、美元急遽走貶，又產油國組織決定維持原油產量配額不變，再加上非洲最大產油國奈及利亞油管傳出遭攻擊，使得國際油價不斷的飆漲之影響，又在 2008 年 9 月華爾街投資銀行雷曼兄弟公司破產，金融情勢令人擔心美國經濟體質，使全球陷入經濟恐慌，發生金融海嘯，這些因素使得能源的消費量減少，所以我們覺得石油價格與金融狀況會影響能源的消費量。

從資料中發現在 2008 年中的時候，能源消費量有降低的趨勢，我們認為是受到原油價格飆漲的影響，連帶影響國內油品價格導致國內的消費意願降低。而 2008 年 9 月的金融海嘯也使得能源消費量持續呈現負成長的現象。雖然受到金融風暴影響，但是依照近幾個月來景氣逐漸回昇的趨勢，我們預測認為，於未來幾年的能源消費量也會年年持續增加。

如今台灣面臨了自產能源的不足，大部分的能源依賴進口，因此節約能源與尋找替代能源在未來資源短缺的時代，就越來越重要。多一份新能源開發利用，就能少一份傳統能源進口。我們必須徹底減少能源的浪費，從「節約能源」做起，讓生活及經濟得以永續發展。面對以上的問題，政府如何有效宣導及人民應當如何節約能源是急需探討的議題。

在 2007 年到 2008 年之間能源消費呈現負成長，符合了政府的節能減碳政策，政府當局也希望能一直保持下去。但假如未來汽油價格降低且經濟復甦，能源消費量也隨之恢復了上升的趨勢，是不是就此說明政府的政策效果不彰，必須再加強宣導與執行！

能源對我們來說是生活的必需品，無論做什麼都需要運用到能源，原以為取之不盡用之不竭，卻發現它正一點一滴的加速減少當中，因此，應當如何解決此問題，最主要要從我們自身著手，政府與產業也該積極地正視，而不只是紙上談兵。

為了使能源可以永續留存，應當在日常生活中養成良好的節能減碳習慣，例如：多搭乘大眾交通運輸工具、家庭與企業節約用水用電、積極尋找環保的替代能源……等等，必須以實際的行動徹底實行宣導的標語，如此一來才能確實地降低能源的使用量，讓能源不至於消耗殆盡。



參考文獻

1. 陳婉淑 (2010), 統計預測方法講義。
2. 羅新衡, 高紹惠 (2007) 減少二氧化碳排放—由日常生活著手, 科學發展, 413 期, 13-17 頁。
3. AREMOS 經濟統計資料庫網站：
http://cache.moe.edu.tw/aremos_ly/search.html
4. 經濟部能源統計網站：
<http://proj.moeaidb.gov.tw/ghg/files/Law/%E8%83%BD%E6%BA%90%E7%B5%B1%E8%A8%88%E6%89%8B%E5%86%8A-93.pdf>
<http://tw.group.knowledge.yahoo.com/oldboy-jay/article/view?aid=3113>
5. 統計預測評估準則參考網站：
<http://acqy.csai.cn/user2/51384/archives/2009/41819.html>
6. 世界年鑑參考網站：
<http://www7.www.gov.tw/todaytw/2008/intestine/ch06/2-6-12-0.html>