

逢甲大學學生報告 ePaper

迴歸分析－環境變異因子對區域均溫的探討

Regression analysis - Environmental variability factor of regional average temperature

作者：李慶揚、楊子萱、林竑志、鐘偲珮、簡玉婷

系級：統計學系三年乙班

學號：D9963870、D9923935、D9923802、D0087795、D9963689

開課老師：陳婉淑 教授

課程名稱：迴歸分析

開課系所：統計學系

開課學年：101 學年度 第一 學期

中文摘要

為探討影響台灣本島各地溫度的各種環境因素，我們選定的樣本為2012年11月27日下午六點的大氣溫度，再考慮以下六種可能影響溫度的變數有相對溼度、懸浮粒子濃度、一氧化碳濃度、氮氧化物濃度、風速以及雨量。以敘述統計做為參考，觀察散佈圖的趨勢，使讀者對資料有初步認知。開始檢視變數是否有自我相關後，產生初步模型。為挑選重要變數，使用五大方法：向後消去法、前進選擇法、逐步選取法、調整 R^2 擷取法、 CP 選擇法，比較各方法的選取差異及共同點後，採用了相對溼度、懸浮粒子濃度及雨量作為篩選結果，進而建立最佳模式，變數中發現當相對溼度及雨量增加則氣溫會降低，懸浮粒子則反之。最後，觀察殘差分析中的誤差項是否符合基本假設後，再檢視資料中異常點或影響點的影響程度，經由分析過程發現，異常點並不存在，為了提升我們解釋能力，所以我們刪除了一點影響點為屏東縣恆春鄉，反映出此為最佳模式。



關鍵字：區域均溫、複迴歸分析、殘差分析、敘述統計、選取重要變數、
多重共線性、SAS、變異數分析

Abstract

We investigate several environmental factors which possibly affect the temperatures around the island of Taiwan. Six variables which might caused to Taiwan's temperature are considered in this project : relative humidity, concentration of suspended particulates, concentration of carbon monoxide, concentration of nitrogen oxides, wind speed, and rainfall. The data set was collected at 6:00 pm, November 27th, 2012 which is available at Environmental Protection Department's webpage. We fit a multiple linear regression using SAS statistical package. We use five model selection procedures to select important variables which affect the temperatures. There are backward elimination, forward selection, stepwise regression, adjusted R2 Selection, and CP Selection. Relative humidity, concentration of suspended particulates, and rainfall are the most important independent variables in our final model. Our results show when relative humidity and rainfall increase, temperatures would reduce, and when concentration of suspended particulates decline, temperatures would rise. Finally, we also work on diagnostic checking to check if there is any outlier or influential point. We remove two influential points Hengchun Township, Pingtung County in order to improve prediction power and to reflect the best model.

Keyword: Regional average temperature, Multiple Regression Analysis, Residual Analysis, Descriptive Statistics, Select important variables, Multicollinearity, SAS, ANOVA

目 次

第一章 緒論	7
第一節 研究背景、動機與目的	7
第二節 資料出處與介紹反應變數、解釋變數	8
第二章 資料分析	9
第一節 一般統計敘述、散佈圖	9
第二節 全部變數	12
第三節 選取重要變數	13
第四節 建立模式	16
第三章 殘差分析	19
第一節 檢查誤差項是否符合假設	19
第二節 異常點	20
第三節 影響點	24
第四章 結語	27
附錄一 原始資料	28
附錄二 參考文獻	29
附錄三 小組討論時間表	30

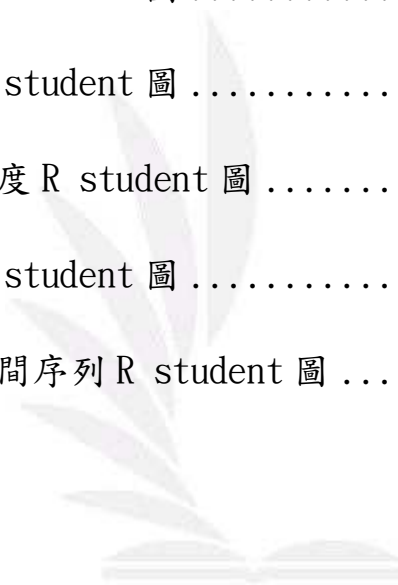
表 目 錄

表 2.1 敘述統計數據表.....	10
表 2.2 變異數膨脹因子分析表.....	13
表 2.3 全模式 ANOVA 表.....	13
表 2.4 向後消去法結果表.....	14
表 2.5 前進選擇法結果表.....	14
表 2.6 逐步迴歸法結果表.....	15
表 2.7 重要變數選取法整合表.....	16
表 2.8 選模後 ANOVA 表.....	17
表 3.1 誤差項平均檢定表.....	19
表 3.2 誤差項服從常態檢定表.....	20
表 3.3 D-W 值正負自我相關數值表.....	20
表 3.4 殘差分析數據表.....	21
表 3.5 DFFITS 影響點數值表.....	24
表 3.6 Hat value 影響點數值表.....	25
表 3.7 DFBETAS 影響點數值表.....	25
表 3.8 COVRATIO 影響點數值表.....	26
表 3.9 影響點統計歸納表.....	26

圖 目 錄

圖 2.1 台灣各地氣溫與相對濕度散佈圖.....	10
圖 2.2 台灣各地氣溫與懸浮粒子濃度散佈圖.....	11
圖 2.3 台灣各地氣溫與一氧化碳濃度散佈圖.....	11
圖 2.4 台灣各地氣溫與氮氧化物濃度散佈圖.....	11
圖 2.5 台灣各地氣溫與風速散佈圖.....	12
圖 2.6 台灣各地氣溫與雨量散佈圖.....	12
圖 2.7 C_p 選擇法變數圖.....	16
圖 2.8 台灣各地氣溫與相對濕度散佈圖.....	17
圖 2.9 台灣各地氣溫與懸浮粒子濃度散佈圖.....	17
圖 2.10 台灣各地氣溫與雨量散佈圖.....	18
圖 3.1 殘差圖.....	21
圖 3.1.1 各地氣溫殘差圖.....	21
圖 3.1.2 相對濕度殘差圖.....	21
圖 3.1.3 懸浮粒子濃度殘差圖.....	22
圖 3.1.4 各地雨量殘差圖.....	22
圖 3.1.5 各地氣溫時間序列殘差圖.....	22
圖 3.2 標準化殘差圖.....	22

圖 3.2.1 各地氣溫標準化殘差圖	22
圖 3.2.2 相對濕度標準化殘差圖	22
圖 3.2.3 懸浮粒子濃度標準化殘差圖	23
圖 3.2.4 各地雨量標準化殘差圖	23
圖 3.2.5 各地氣溫時間序列標準化殘差圖	23
圖 3.3 R student 圖	23
圖 3.3.1 各地氣溫 R student 圖	23
圖 3.3.2 相對濕度 R student 圖	23
圖 3.3.3 懸浮粒子濃度 R student 圖	24
圖 3.3.4 各地雨量 R student 圖	24
圖 3.3.5 各地氣溫時間序列 R student 圖	24



第一章 緒論

第一節 研究背景、動機與目的

一、研究背景

以環境類別為明確方向，進而從延伸、縮小專注於探討環境溫度變化。蒐集資料之初，因本組對資料來源可信度與原始資料需求的重視相當執著，為求謹慎，特別向氣象局研究人員諮詢測量溫度等許多變數的方法及限制條件，或有無其他忽略的環境變動因素存在。鑒因於觀察到這影響資料準確性的方面來說，將細節的處理更趨詳細，對於報告的呈現，希望能更臻準確。再者，資料內容，透過 SAS 分析、程式計算及五種分析方法的應用、交互比對，過程中也採用散佈圖的方式，來觀察其關聯性的趨勢變化。

二、研究動機與目的

環境，周圍所在的條件皆為其範圍。對不同對象和解釋角度來說，環境的內容也有所不同。例如：從生物領域來看，是指生物生活周圍的氣候、生態系統、周圍群體和其他物種群；從人文社會領域來看，是指具體的人生活周圍的情況和條件；對理化學科來說，可能是化學反應的溶液或受影響的物體等。由於環境涵蓋的範圍極廣，題材之豐富、種類之多元，與人類生活相關密不可分，甚至可以說只要存在於地球上任何物體，都必將受其影響、限制。加上近幾年來，全球環境變遷極大，甚至開始意識到，面對大自然的反撲，人類是有多麼脆弱。在生死存亡之際，在尚未發現其他能使人類生存的環境之前，只有一個地球，如何讓它能永續生存，牽動著有地球上所有的生命體的重大議題，值得我們好好省思。這也是為何，作為此次研究類別的主要原因。經由本組的討論發現，各位在對於環境的認知、偏好與其不同見解來看，大多以生物領域的研究方面興趣最濃厚，相較於其他領域，一來研究意願不高，二來考量到蒐集資料的方法與較缺乏，資料來源的依據和正確性較難掌握，難度較高。生物領域面，宏觀面不外乎大自然，岩石圈、水圈、大氣圈、生物圈等自然環境，雖然都有意願嘗試研究，礙於資料收集的難易與掌握性，決定將環境的研究範圍限制在台灣本島，範圍縮小，變數也隨之減少，且相較其他區域，對於台灣環境辨識度與熟悉度也較高，有助之後研

究分析，提升可信度。溫度，物體的冷熱程度，只要是在地球上，不論哪個位置、何種地形、任何狀態，都一定會有溫度的存在。在台灣這海島型國家，土地面積小，緯度的差異不大，除了日照，這影響溫度最主要的因素之外，降雨、地形及氣壓等。台灣面積雖然不大，但各地區天氣型態皆有所差異，影響氣溫變化幅度因素眾多，也因此台灣居民的生活習型態、產業的表現等皆受氣溫影響息息相關，希望藉此能了解台灣各地溫度變化而產生天氣現象。

第二節 資料出處與介紹反應變數、解釋變數

因應這次的主題，區域性溫度的變化，首先，以台灣各大縣市（包含其縣市內區域）為範圍，從北中南東四大區域中，平均挑選出三十個縣市區域，依據各區域所在的地理位置不同，進行調查區域內各項變數指標的數據顯示與差異性比較。

基於反應變數的選定為溫度的變化關係，所以環境中任何對溫度具有影響性的因子即為此次所研究的目標，從提及與溫度相關的書籍著手，其中不乏大氣科學、溫度的測量、溫度變異對氣候的影響性等書籍作為選定變數的參考依據，經過組員間的討論、議定後，在追求實驗結果的準確性與合理性之下，選取六項較為常見、出現次數頻繁且影響性較直接，受人為因素干擾少的變數，且將較為生僻，或者台灣不曾出現、缺乏的條件等變數因子，予以剔除。其中變數內容如下：台灣各地的相對濕度（ X_1 ），北部地區相較於南部，西部地區相較於東部，降雨次數均較為頻繁，加上地形差異皆影響溫度變化，因此選之。台灣各地的懸浮粒子濃度（ X_2 ）、一氧化碳濃度（ X_3 ）、氮氧化物濃度（ X_4 ），北中南東各地人口數、密集度不一，尤其是交通工具使用上的習慣、數量與種類，上述兩者皆會影響廢氣排放量進而造成空氣品質上的差異，加上紫外線照射下，反射程度均不同，受空氣中塵粒數量、成分、濃度及吸收紫外線程度影響，若滯留在大氣內的熱能增加，則致使溫度上升，因此選之。台灣各地的風速（ X_5 ），以地理位置來看，中央山脈為天然屏障，阻隔了夏季西南氣流吹拂，東部地區風速變化、雨量皆較小，溫度變化較小；冬季東北季風、西伯利亞高壓冷氣團影響，北部地區首當其衝，氣壓最高，風速相對較大，溫度變化較大，依序往南遞減，因此選之。台灣各地的雨量（ X_6 ），北部地區四季均有雨，分布較平均，濕度相對較高；中部地區，四季更迭較明顯，易於區分，雨量隨季節轉替而變化；南部地區，雨量偏集中至夏、秋兩季，春、冬則幾乎無雨，濕度偏向兩極化；東部地區，入秋、季節轉換之時，雨量增加幅度則最為明顯，加上夏季有焚風現象，濕度相對較低。有關資料方面

的彙整，為求得第一手資料的數據，決定以行政院環保署、交通部中央氣象局等政府相關單位內，具專業、直接性且未經分析的觀測資料作為參考出處。

第二章 資料分析

此章節我們先以敘述統計作為分析資料的始源，再者介紹並分析 6 個解釋變數，進而藉由系統中五種選取法來選出最重要的 3 個解釋變數，最後就可得到此次主題的最佳迴歸方程式。

第一節 一般敘述統計、散佈圖

此小節我們將詳述敘述統計的數據內容，並藉由反應變數與 6 個解釋變數的散佈圖來解釋他們的相關性。

一、一般敘述統計

我們蒐集 2012 年 11 月 27 日下午六點台灣本島中十一個縣及八個市，除了台東縣、花蓮縣、基隆市、新竹市、嘉義市這 5 個縣市只取一區外其他縣市皆各取兩個地區，所以共有 33 筆資料，初步分析由表 2.1 可得知平均氣溫 (y) 為 17.67°C ，最低氣溫出現在苗栗縣三義鄉為 14°C 而最高氣溫出現在高雄市小港區溫度為 23°C ，標準差為 2.06。

在解釋變數中，台灣各地的相對濕度 (x_1) 顯示出水蒸氣的飽和度有多高，從資料得知相對溼度最高為 95%，相對溼度最低為 70%，平均為 84.52% 顯示當天的濕度偏高；懸浮粒子濃度主要的來源是從地表揚起的塵土，及化石燃料（煤、石油等）和垃圾的燃燒造成的，在世界衛生組織提到懸浮粒子濃度大於 $20\mu\text{g}/\text{m}^3$ 其總死亡率、心肺疾病死亡率和肺癌的死亡率會增加 (95% 以上可信度)，則在歐盟限制標準懸浮粒子濃度不可大於 $50\mu\text{g}/\text{m}^3$ 而，當天平均懸浮粒子濃度 (x_2) 為 $50.42\mu\text{g}/\text{m}^3$ ，最大值為出現在屏東為 $114\mu\text{g}/\text{m}^3$ 都超過歐盟的標準，最小值為 $14\mu\text{g}/\text{m}^3$ 在世界衛生組織的安全標準內，而標準差為 21.94 是我們解釋變數中標準差最大的，可能是因為城鄉差距大的關係所導致；一氧化碳濃度規定容許量為 35ppm 而資料中平均一氧化碳濃度 (x_3) 為 0.8 ppm 及最大值為 1.52ppm 皆未超出容許量；氮氧化物是一種因高溫燃燒之後會產生的污染物，形成的原因大多為汽機車、火力發電廠，在空氣中的氮氧化物濃度如果超過 6ppm，即會對人體有不良影響，超過 1.0ppm 可使支氣管及呼吸道阻塞增加，超過 10ppm 則會喪失嗅知能力，而平均氮氧化物濃度 (x_4) 為 40.22 ppb，最大值為 94ppb 皆於標準內；平均風速 (x_5) 為 2.08 m/sec 以等級來分為 2 級的輕風，人面感覺有風，樹葉搖動，普通

之風標轉動，最大值出現在恆春為 9.3 m/sec 以等級來分為 5 級的清風，有葉枝小樹開始搖擺，資料表示當天風勢均不大；而以雨量來看要達大雨特報累積雨量需達 50 mm 以上，資料中平均雨量 (x_6) 為 0.22 mm，最大值為 1.4 mm 顯示當天雨量偏小；以上所有初步敘述統計資料均可由表 2.1 得知。

表 2.1 敘述統計數據表

變數	N	平均值	標準差	最小值	最大值
y：各地氣溫 (°C)	33	17.67	2.03	14	23
x_1 ：相對濕度 (%)	33	84.52	6.41	70	95
x_2 ：懸浮粒子濃度 ($\mu\text{g}/\text{m}^3$)	33	50.42	21.94	16	114
x_3 ：一氧化碳濃度 (ppm)	33	0.8	0.3	0.37	1.52
x_4 ：氮氧化物濃度 (ppb)	33	40.22	20.15	3.4	94
x_5 ：平均風速 (m/sec)	33	2.08	1.71	0.5	9.3
x_6 ：平均雨量 (mm)	33	0.22	0.41	0	1.4

二、散佈圖

由圖 2.1 得知台灣各地氣溫 (y) 與相對濕度 (x_1) 及圖 2.5 台灣各地風速 (x_5) 皆有明顯的負相關，在圖 2.6 中氣溫 (y) 與各地的雨量 (x_6) 有負相關但不明顯。由圖 2.2 可得知台灣各地氣溫 (y) 與懸浮粒子濃度 (x_2) 有明顯的正相關，在圖 2.3 中氣溫 (y) 一氧化碳濃度 (x_3) 及圖 2.4 中氣溫 (y) 與氮氧化物濃度 (x_4) 亦有正相關，但一氧化碳濃度 (x_3) 與氮氧化物濃度 (x_4) 比較之下並不明顯。

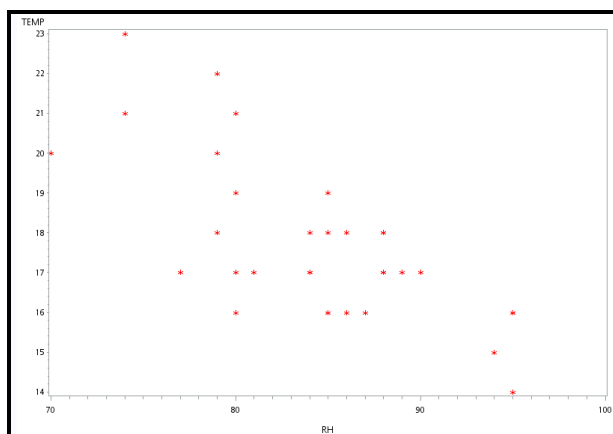


圖 2.1 台灣各地氣溫與相對濕度散佈圖

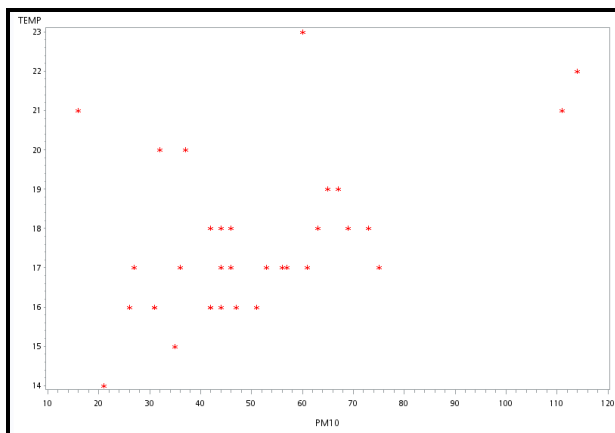


圖 2.2 台灣各地氣溫與懸浮粒子濃度散佈圖

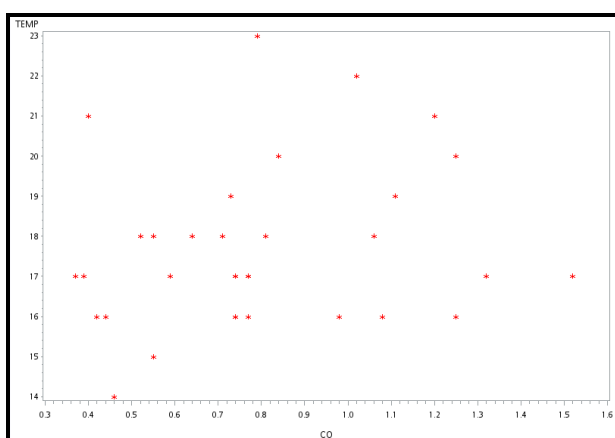


圖 2.3 台灣各地氣溫與一氧化碳濃度散佈圖

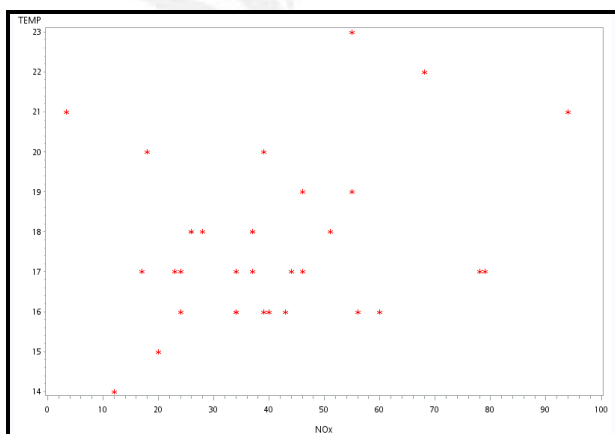


圖 2.4 台灣各地氣溫與氮氧化物濃度散佈圖

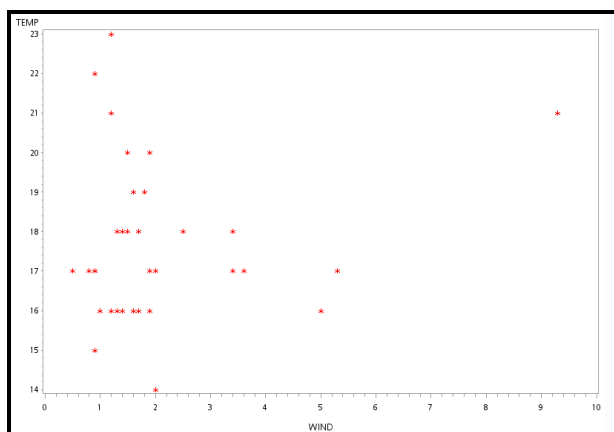


圖 2.5 台灣各地氣溫與風速散佈圖

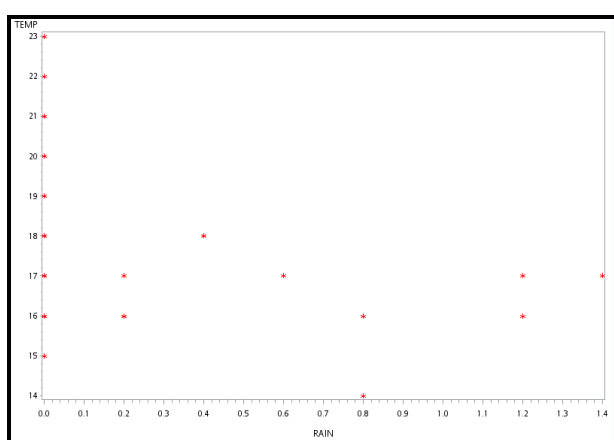


圖 2.6 台灣各地氣溫與雨量散佈圖

第二節 全部變數

所謂的多重共線性 (Multicollinearity) 是指解釋變數與解釋變數間存在高度相關性，造成分析時的干擾，因此可以選擇重新定義解釋變數，或者將高度相關性的變數與其他變數組合。假如發生多重共線性的情形，可以使用變異數膨脹因子 (Variance Inflation Factor; VIF) 大小作為判斷準則，VIF 越大表示解釋變數的容忍度越小，其中當 $VIF > 10$ ，表示解釋變數間存在高度相關性。

在我們所探討主題中從表 2.2 可得知各變數的 VIF，相對溼度 (1.52)、懸浮粒子濃度 (2.59)、一氧化碳濃度 (4.19)、氮氧化物濃度 (6.12)、風速 (2.11) 和雨量 (1.28) 的變異數膨脹因子皆沒有大於 10，故我們所要探討的 6 個解釋變數間無多重相關性。

表 2.2 變異數膨脹因子分析表

Variable	Parameter	Pr > t	Variance Inflation
----------	-----------	---------	--------------------

	Estimate		
Intercept	33.18715	<.0001	0
X ₁	-0.19935	<.0001	1.52232
X ₂	0.02263	0.1736	2.58615
X ₃	0.23988	0.8746	4.19282
X ₄	0.00900	0.7425	6.12228
X ₅	-0.03043	0.8721	2.11257
X ₆	-1.35020	0.0373	1.27690

以下是使用 6 個變數組成的複迴歸模型：

$$\hat{y} = 33.187 - 0.199 X_1 + 0.023 X_2 + 0.24 X_3 + 0.009 X_4 - 0.0304 X_5 - 1.35 X_6,$$

(<.0001) (<.0001) (0.1736) (0.8746) (0.7425) (0.8721)

(0.0373)

此處括號內的數值為 P 值。

從下表 2.3 可得知全模式變異數分析的 P 值小於 0.0001，而其中此複迴歸方程式標準差為 1.25，也可得到相對濕度 (X₁) 及雨量 (X₆) 與溫度 (y) 有相關性；則懸浮粒子濃度 (X₂)、一氧化碳濃度 (X₃)、氮氧化物濃度 (X₄) 及風速 (X₅) 與溫度 (y) 無相關性。

表 2.3 全模式 ANOVA 表

Source	DF	Sum of Squares	Mean Square	F-value	Pr>F
Model	6	90.81400	15.13567	9.71	<.0001
Error	26	40.51934	1.55844	-	-
Corrected Total	32	131.33333	-	-	-

第三節 選取重要變數

此小節我們使用五種選取法來選出 3 個對反應變數最有影響力的解釋變數，也就作為最終模型的重要變數。

一、向後消去法 (Backward Elimination)

向後消去法指將所有指定變數輸入至模型，在每一次分析會逐步移除對該模型中最不顯著的變數，直到所有剩餘的變數在統計分析上都對該模型具有顯著貢獻。

首先把所有變數選入模型，接著做模型分析，分析結果中 P 值最大且 P 值大於 0.1 中的變數移除，所以先移除 P 值為 0.8746 的台灣各地的一氧化碳濃度變數 (x_3) 後再做分析，接著再移除 P 值為 0.8148 的台灣各地的風速變數 (x_5) 再做分析，然後再移除 P 值 0.3863 台灣各地的氮氧化物濃度變數 (x_4)，最後如表 2.4 所示留下的變數為台灣各地的相對濕度 (x_1) 其 P 值 ($<.0001$)、台灣各地的懸浮粒子濃度 (x_2) 其 P 值 (0.0040) 及台灣各地的雨量 (x_6) 其 P 值 (0.0409)。故我們使用向後消去法所選取的重要變數是台灣各地的相對濕度 (x_1)、懸浮粒子濃度 (x_2) 以及雨量 (x_6)。

表 2.4 向後消去法結果表

Variable	Parameter Estimate	Standard Error	Type II SS	F-Value	Pr > F
Intercept	32.82377	3.02184	169.93072	117.99	<.0001
x_1	-0.19480	0.03412	46.93299	32.59	<.0001
x_2	0.03094	0.00991	14.03712	9.75	0.0040
x_6	-1.13215	0.52914	6.59319	4.58	0.0409

二、前進選擇法 (Forward Selection)

前進選擇法是把最顯著的變數逐一選入模型中，直到剩餘變數不再對模型具有顯著貢獻。

首先，輸入相對濕度 (x_1)，分析後 P 值 ($<.0001$) 小於檢定規則的 0.1，則選入模型如表 2.5，接著輸入懸浮粒子濃度 (x_2) 分析 P 值 (0.0044) 小於檢定規則 0.1，則選入模型，再輸入變數一氧化碳濃度 (x_3)、氮氧化物濃度 (x_4) 及台灣各地風速 (x_5) 做分析之後的 P 值皆大於檢定規則 0.1 則不選入模型，最後，輸入台灣各地雨量 (x_6) 做分析結果 P 值 (0.0409) 小於檢定規則的 0.1 則選入模型。故我們使用前進選擇法所選取的重要變數是台灣各地的相對濕度 (x_1)、懸浮粒子濃度 (x_2) 以及雨量 (x_6)。

表 2.5 前進選擇法結果表

Step	Variable Entered	Label	Number Vars In	Partial R-Square	Model R-Square	C_p	F-Value	Pr > F
1	x_1	RH	1	0.5156	0.5156	11.8176	33.00	<.0001
2	x_2	PM10	2	0.1161	0.6318	4.0314	9.46	0.0044
3	x_6	Rain	3	0.0502	0.6820	1.8007	4.58	0.0409

三、逐步迴歸法 (Stepwise Regression)

在使用逐步迴歸法時，首先不用在模式中放入任何的變數，接著採取前進選擇法把最重要的變數選進迴歸模式，然後再使用向後消去法將選進模型中的不重要變數剔除，重複以上步驟多次，直到沒有變數被選取或剔除為止，在模型中的變數即為我們所以選取的重要變數。

把我們所探討主題的解釋變數使用逐步迴歸法選取重要變數，首先，輸入台灣各地的相對濕度 (x_1) 分析後 P 值 (<.0001) 小於檢定規則的 0.1，則選入模型如表 2.6，接著沒有變數需要剔除，所以輸入台灣各地的懸浮粒子濃度 (x_2) 分析 P 值 (0.0044) 小於檢定規則的 0.1，則選入模型，然後台灣各地的相對濕度 (x_1) 和懸浮粒子濃度 (x_2) 皆有符合向後消去法的檢定規則，所以無須剔除，最後，輸入台灣各地的雨量 (x_6) 做分析結果 P 值 (0.0409) 小於檢定規則的 0.1 則選入模型。故我們使用逐步迴歸法所選取的重要變數是台灣各地的相對濕度 (x_1)、懸浮粒子濃度 (x_2) 以及雨量 (x_6)。

表 2.6 逐步迴歸法結果表

Step	Variable Entered	Variable Removed	Label	Number Vars In	Partial R-Square	Model R-Square	C_p	F-value	Pr > F
1	x_1	-	RH	1	0.5156	0.5156	11.8176	33.00	<.0001
2	x_2	-	PM10	2	0.1161	0.6318	4.0314	9.46	0.0044
3	x_6	-	Rain	3	0.0502	0.6820	1.8007	4.58	0.0409

四、調整 R^2 擷取法 (R^2_{adj} Selection)

調整後判定係數法是利用全部變數迴歸模式最大的 R^2_{adj} 值去判定，且選取最佳迴歸模式的準則，已達成選取重要變數的模型準確性。

當調整 R^2 (0.6491) 是多種選擇方式中的最大值，且選取模型中的變數數目最少，所以依調整 R^2 擷取法所選取重要變數為台灣各地相對濕度 (x_1)、台灣各地懸浮粒子濃度 (x_2)、台灣各地雨量 (x_6) 為最終決定的解釋變數。

五、 C_P 選擇法

C_P 選擇法應用於已知的迴歸模型中，選取的迴歸模式必須使得 C_P 值夠小且滿足 C_P 值接近參數個數 (P) 的條件。

當 C_P 值 (1.8007) 是多種選擇方式中的最小值，且選取模型中的變數數目最少，從圖 2.7 顯示出 C_P 選擇法所選取重要變數為台灣各地的相對濕度 (X_1)、台灣各地的懸浮粒子濃度 (X_2)、台灣各地的雨量 (X_6) 為最終決定的解釋變數。

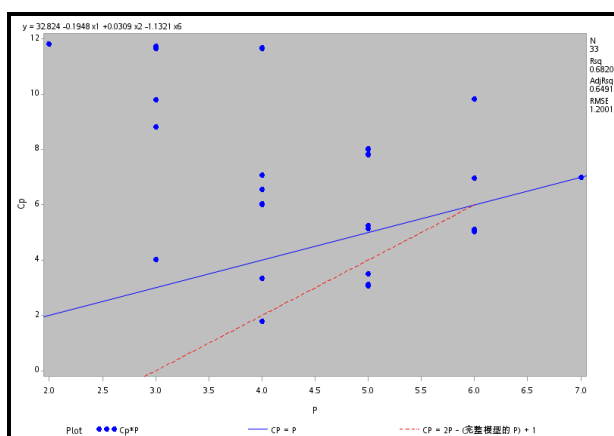


圖 2.7 C_P 選擇法變數圖

綜合以上所有選取法的最佳結果如表 2.7 可得知，最後我們選取的重要變數為台灣各地的相對濕度 (X_1)、台灣各地的懸浮粒子濃度 (X_2)、台灣各地的雨量 (X_6)。

表 2.7 重要變數選取法整合表

重要變數選取法	選取變數
向後消去法	X_1 相對濕度 (%)、 X_2 懸浮粒子濃度 ($\mu\text{g}/\text{m}^3$)、 X_6 各地的雨量 (mm)
前進選擇法	X_1 相對濕度 (%)、 X_2 懸浮粒子濃度 ($\mu\text{g}/\text{m}^3$)、 X_6 各地的雨量 (mm)
逐步選取法	X_1 相對濕度 (%)、 X_2 懸浮粒子濃度 ($\mu\text{g}/\text{m}^3$)、 X_6 各地的雨量 (mm)
調整 R^2 擷取法	X_1 相對濕度 (%)、 X_2 懸浮粒子濃度 ($\mu\text{g}/\text{m}^3$)、 X_6 各地的雨量 (mm)
C_P 選擇法	X_1 相對濕度 (%)、 X_2 懸浮粒子濃度 ($\mu\text{g}/\text{m}^3$)、 X_6 各地的雨量 (mm)

第四節 建立模式

以下是由上節的重要變數選取法，比較後選出的變數組成的最佳模式：

$$\hat{y} = 32.824 - 0.1948 x_1 + 0.0309 x_2 - 1.35 x_6,$$

(<.0001) (<.0001) (0.0044) (0.0409)

此處括號內的數值為 P 值。

根據表 2.8 分析選模後的 P 值小於 0.0001，所以表示三個解釋變數具有線性相關。從以上的迴歸模型，我們可以知道各地的相對溼度 (x_1) 只要上升一單位 (%)，台灣各地氣溫 (y) 就會下降 0.1948 度，再依據圖 2.8 可看出相對濕度與台灣各地氣溫的關係呈現負相關；空氣中懸浮粒子的濃度 (x_2) 只要上升一單位 ($\mu\text{g}/\text{m}^3$)，台灣各地氣溫就會上升 0.0309 度，再依據圖 2.9 可看出懸浮粒子的濃度與台灣各地氣溫的關係呈現正相關；各地的雨量 (x_6) 只要增加一單位 (mm)，台灣各地氣溫就會下降 1.35 度，再依據圖 2.10 可看出各地的雨量與台灣各地氣溫的關係呈現負相關。

表 2.8 選模後 ANOVA 表

Source	DF	Sum of Squares	Mean Square	F-value	Pr > F
Model	3	89.56615	29.85538	20.73	<.0001
Error	29	41.76719	1.44025	-	-
Corrected Total	32	131.33333	-	-	-

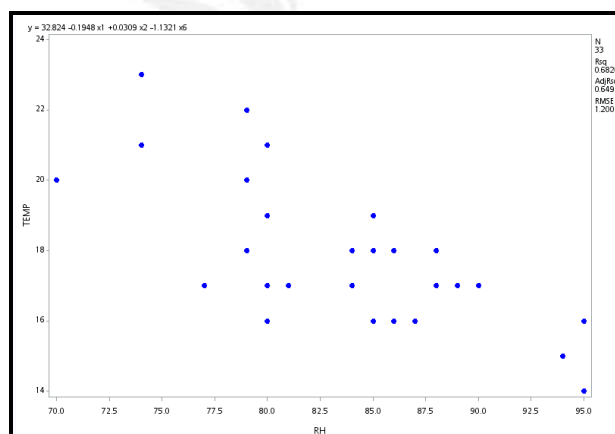


圖 2.8 台灣各地氣溫與相對濕度散佈圖

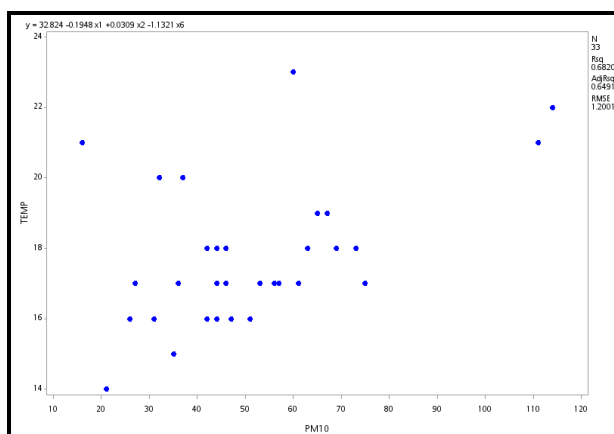


圖 2.9 台灣各地氣溫與懸浮粒子濃度散佈圖

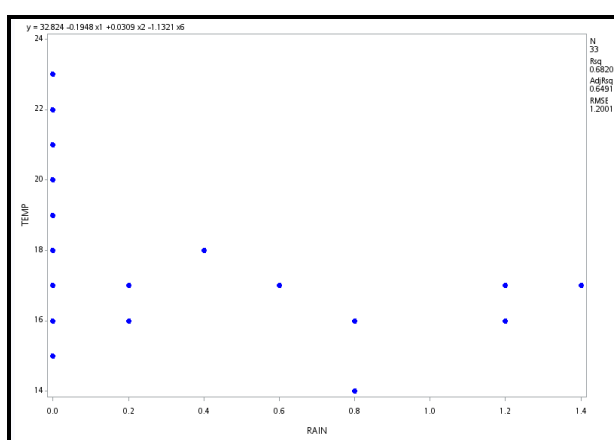


圖 2.10 台灣各地氣溫與雨量散佈圖

第三章 殘差分析

此章節我們先檢測殘差分析誤差項是否符合基本假設，否則，當資料違反殘差分析的基本假設時，會導致分析偏誤的發生。所以在符合基本假設情況下，在第二、三節我們使用幾個檢定方法檢視是否有無異常點與影響點。

第一節 檢測誤差項是否符合假設

此小節我們須分析誤差項是否符合基本假設，此假設的內容有檢測誤差項平均是否為 0、誤差項變異數是否為常數、誤差項是否符合常態分配以及誤差項是否獨立。

一、檢測誤差項平均是否為 0， $E(\varepsilon_i)=0$

我們由表 3.1 得知，要檢測誤差項平均是否為 0，我們得利用 Student' s t、Sign 和 Signed Rank 三種檢定方法得到 P 值皆大於顯著水準 (0.05)，不拒絕虛無假設，故誤差項平均為 0。

表 3.1 誤差項平均檢定表

Test for Location: $\mu_0=0$				
檢定方法	統計量		P 值	
Student' s t	T	0.085	$\text{Pr} > t $	0.9328
Sign	M	0.5	$\text{Pr} \geq M $	1
Signed Rank	S	14.5	$\text{Pr} \geq S $	0.8001

二、檢測誤差項變異數是否為常數， $\text{Var}(\varepsilon_i)=\sigma^2$

利用卡方值判斷，檢定 H_0 ：誤差項的變異數是常數，因為P值為0.4125，所以不拒絕虛無假設，即誤差項的變異數是常數。

三、誤差項是否服從常態分配， $\varepsilon_i \sim N(0, \sigma^2)$

我們由表 3.2 得知，要檢測誤差項是否服從常態分配，我們可利用 Shapiro-Wilk、Kolmogorov-Smirnov、Cramer-von Mises 和 Anderson-Darling 四種檢定方法得到 P 值皆大於顯著水準 (0.05) 不拒絕虛無假設，故由誤差項是服從常態分配的。

表 3.2 誤差項服從常態檢定表

Tests for Normality				
檢定方法	統計量		P 值	
Shapiro-Wilk	W	0.9784	Pr < W	0.738
Kolmogorov-Smirnov	D	0.0831	Pr < D	> 0.15
Cramer-von Mises	W-Sq	0.0446	Pr > W-Sq	> 0.25
Anderson-Darling	A-Sq	0.3186	Pr > A-Sq	> 0.25

四、誤差項是否獨立， $Cov(\varepsilon_i, \varepsilon_j) = 0, \forall i \neq j$

Durbin-Watson test 是檢定誤差項的假設是否獨立，而 $d = 2(1 - \rho)$ ，d 值會介於 0 到 4 之間，若 d 值越接近 4，則誤差項呈自我負相關；反之，若 d 值越接近 0，則誤差項呈自我正相關；最後，當 d 值等於 2 時，則誤差項沒有自我相關，而我們所探討主題的 d 值為 1.8421，非常接近 2，所以較偏向誤差項沒有自我相關。

我們更進一步檢定誤差項是否有正自我相關，首先做假設檢定：

$$H_0: \rho = 0 \text{ v. s } H_1: \rho > 0,$$

由表 3.3 得知 P 值 (0.264) 大於顯著水準 (0.05)，所以可證明誤差項沒有正自我相關。最後我們檢定誤差項是否有負自我相關，以下做假設檢定：

$$H_0: \rho = 0 \text{ v. s } H_1: \rho < 0,$$

檢定誤差項是否有負自我相關由表 3.3 得知 P 值 (0.736) 大於顯著水準 (0.05)，所以綜合以上亦可證明誤差項獨立。

表 3.3 D-W 值正負自我相關數值表

Durbin-Watson	1.8421
Pr < DW	0.2640
Pr > DW	0.7360

第二節 異常點

此小節我們使用 Residual 檢定、Studentized Residuals 檢定和 Studentized Deleted Residuals 檢定來檢測是否存在異常點。

一、Residual 檢定

根據 Residuals 檢定規則為 $Residual > 3\sqrt{MSE}$ 為異常點，我們探討主題的 $|Residual| > 3.6002$ ，由表 3.4 顯示沒有資料符合此檢定標準，所以此檢定規則沒有異常點。且由圖 3.1.1 到圖 3.1.5 可知殘差皆落在以 0 為中心的

區塊內，圖 3.1.1 呈喇叭狀，圖 3.1.4 有逐漸縮小的趨勢，但依上述並無異常點的出現。

表 3.4 殘差分析數據表

bs	Dependent Variable	Predicted Value	Std Error Mean Predict	Residual	Std Error Residual	Student Residual	-2 -1 0 1 2	Cook's D	RStudent	Hat Diag H	Cov Ratio	DFFITS	DFBETAS			
													Intercept	x1	x2	x6
1	17.0000	18.7778	0.2612	-1.7778	1.171	-1.518	***	0.029	-1.5544	0.0474	0.8675	-0.3466	-0.1386	0.1251	-0.0318	0.1337
2	17.0000	17.8840	0.2457	-0.8840	1.175	-0.753	*	0.006	-0.7468	0.0419	1.1099	-0.1562	-0.0251	0.0076	0.0346	0.0757
3	23.0000	20.2651	0.4123	2.7349	1.127	2.427	****	0.197	2.6709	0.1180	0.5263	0.9771	0.7701	-0.7706	0.0363	-0.1642
4	21.0000	20.6743	0.6257	0.3257	1.024	0.318		0.009	0.3131	0.2719	1.5584	0.1913	-0.0179	-0.0056	0.1719	-0.0197
5	16.0000	15.2772	0.4561	0.7228	1.110	0.651	*	0.018	0.6446	0.1444	1.2680	0.2648	-0.1602	0.1935	-0.0738	-0.1008
6	18.0000	18.2814	0.3231	-0.2814	1.156	-0.243		0.001	-0.2395	0.0725	1.2305	-0.0669	-0.0489	0.0448	0.0240	-0.0230
7	17.0000	17.2948	0.6919	-0.2948	0.981	-0.301		0.011	-0.2958	0.3324	1.7023	-0.2088	-0.0599	0.0678	-0.0094	-0.1938
8	17.0000	16.0155	0.5816	0.9845	1.050	0.938	*	0.067	0.9358	0.2349	1.3296	0.5184	-0.1240	0.0992	0.1482	0.4491
9	18.0000	17.8164	0.3398	0.1836	1.151	0.159		0.001	0.1568	0.0802	1.2466	0.0463	-0.0241	0.0232	0.0274	-0.0167
10	19.0000	18.3389	0.2903	0.6611	1.164	0.568	*	0.005	0.5609	0.0585	1.1688	0.1398	-0.0333	0.0309	0.0774	-0.0535
11	19.0000	19.2510	0.2947	-0.2510	1.163	-0.216		0.001	-0.2122	0.0603	1.2167	-0.0538	-0.0193	0.0201	-0.0194	0.0167
12	16.0000	16.7216	0.3699	-0.7216	1.142	-0.632	*	0.010	-0.6254	0.0950	1.2029	-0.2026	-0.0228	0.0183	0.0232	-0.1637
13	16.0000	15.3911	0.4098	0.6089	1.128	0.540	*	0.010	0.5332	0.1166	1.2510	0.1937	-0.1456	0.1619	-0.0056	-0.0290
14	16.0000	15.1225	0.4723	0.8775	1.103	0.795	*	0.029	0.7902	0.1549	1.2466	0.3383	-0.1848	0.2316	-0.1266	-0.1265
15	15.0000	15.5957	0.4211	-0.5957	1.124	-0.530	*	0.010	-0.5234	0.1231	1.2622	-0.1961	0.1195	-0.1431	0.0440	0.0779
16	17.0000	18.9382	0.3916	-1.9382	1.134	-1.709	***	0.087	-1.7703	0.1065	0.8425	-0.6111	-0.4839	0.4227	0.3150	0.1496
17	16.0000	17.6491	0.2486	-1.6491	1.174	-1.405	**	0.022	-1.4297	0.0429	0.9070	-0.3028	0.0553	-0.0807	-0.0109	0.1516
18	17.0000	16.1274	0.3711	0.8726	1.141	0.765	*	0.015	0.7590	0.0956	1.1728	0.2468	-0.0624	0.1029	-0.1339	-0.1032
19	16.0000	18.5394	0.3006	-2.5394	1.162	-2.186	****	0.080	-2.3499	0.0627	0.5976	-0.6079	-0.3803	0.3165	0.2461	0.2120
20	14.0000	14.0620	0.5198	-0.0620	1.082	-0.0574		0.000	-0.0564	0.1876	1.4158	-0.0271	0.0110	-0.0136	0.0105	-0.0127
21	16.0000	16.6091	0.2864	-0.6091	1.165	-0.523	*	0.004	-0.5160	0.0569	1.1748	-0.1268	0.0001	-0.0210	0.0780	0.0153
22	18.0000	17.4312	0.2340	0.5688	1.177	0.483		0.002	0.4768	0.0380	1.1581	0.0948	0.0218	-0.0153	-0.0170	0.0375
23	16.0000	16.3616	0.5543	-0.3616	1.064	-0.340		0.008	-0.3345	0.2133	1.4396	-0.1742	-0.0142	0.0167	0.0000	-0.1610
24	17.0000	16.8165	0.2437	0.1835	1.175	0.156		0.000	0.1535	0.0412	1.1962	0.0318	-0.0105	0.0141	-0.0053	-0.0041
25	17.0000	16.7659	0.3170	0.2341	1.157	0.202		0.001	0.1989	0.0698	1.2300	0.0545	-0.0193	0.0185	0.0174	0.0323
26	18.0000	17.6308	0.3084	0.3692	1.160	0.318		0.002	0.3133	0.0660	1.2149	0.0833	-0.0420	0.0428	0.0383	-0.0340
27	18.0000	17.4325	0.2557	0.5675	1.173	0.484		0.003	0.4775	0.0454	1.1670	0.1041	-0.0083	0.0213	-0.0246	-0.0524
28	18.0000	18.5246	0.3269	-0.5246	1.155	-0.454		0.004	-0.4480	0.0742	1.2079	-0.1268	0.0351	-0.0296	-0.0854	0.0417
29	17.0000	18.7812	0.3342	-1.7812	1.153	-1.545	***	0.050	-1.5852	0.0776	0.8846	-0.4596	0.0882	-0.0635	-0.3204	0.1404
30	22.0000	20.9619	0.6541	1.0381	1.006	1.032	**	0.112	1.0330	0.2971	1.4096	0.6715	-0.0368	-0.0475	0.6006	-0.0600
31	21.0000	18.9038	0.5889	2.0962	1.046	2.005	****	0.319	2.1223	0.2408	0.8346	1.1952	0.9788	-0.8389	-0.8539	-0.1930
32	20.0000	20.1780	0.6010	-0.1780	1.039	-0.171		0.002	-0.1685	0.2508	1.5297	-0.0975	-0.0907	0.0840	0.0470	0.0112
33	20.0000	18.5795	0.3426	1.4205	1.150	1.235	**	0.034	1.2468	0.0815	1.0093	0.3714	0.2596	-0.2175	-0.1931	-0.1123

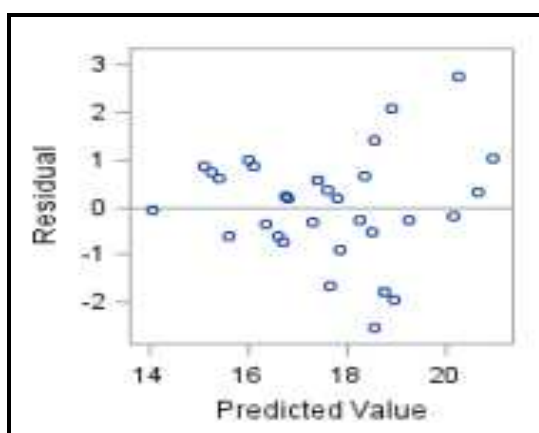


圖 3.1.1 各地氣溫殘差圖

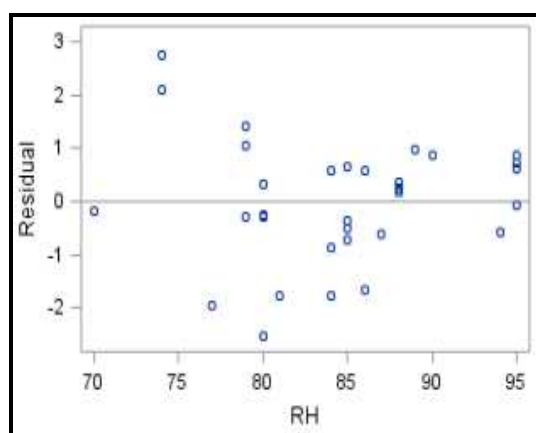


圖 3.1.2 相對溼度殘差圖

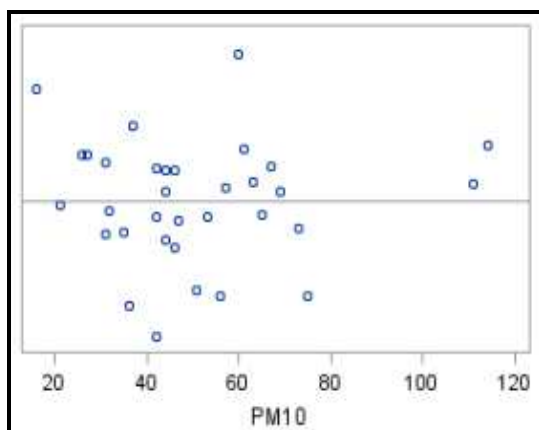


圖 3.1.3 懸浮粒子濃度殘差圖

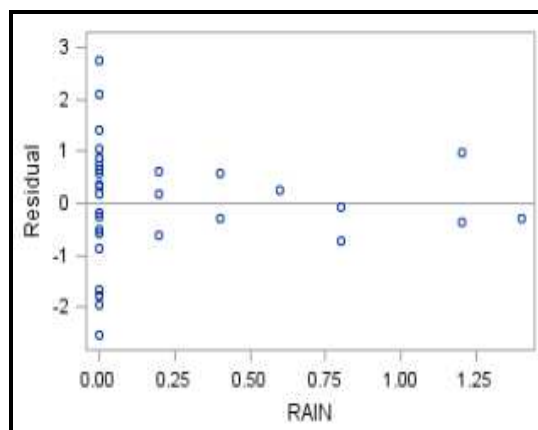


圖 3.1.4 各地雨量殘差圖

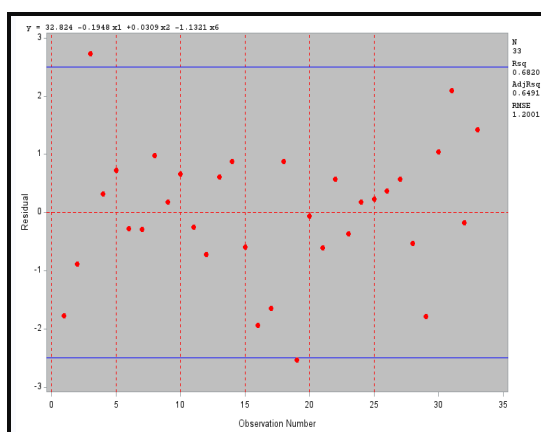


圖 3.1.5 各地氣溫時間序列殘差圖

二、Studentized Residuals 檢定

根據 Studentized Residuals 檢定標準為 $|r_i| > 3$ 則為異常點，但由上表 3.4 統計顯示出沒有一筆資料符合此檢定標準，所以此檢定規則下沒有異常點。且由圖 3.2.1 到圖 3.2.5 可知殘差皆落在以 0 為中心的區塊內，圖 3.2.1 呈現喇叭狀，圖 3.2.4 有逐漸縮小的趨勢，但依上述並沒有異常點的出現。

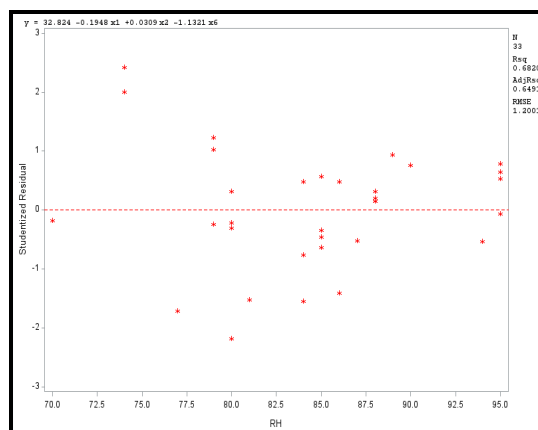
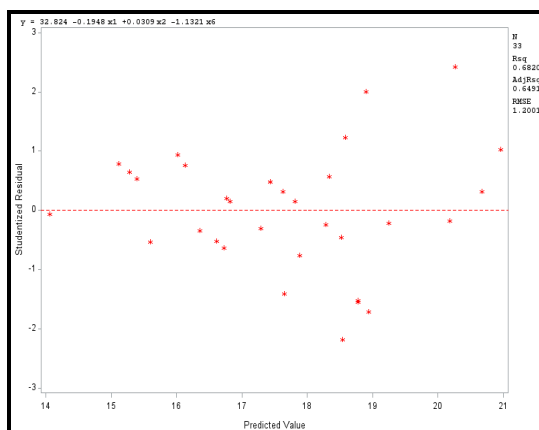


圖 3.2.1 各地氣溫標準化殘差圖

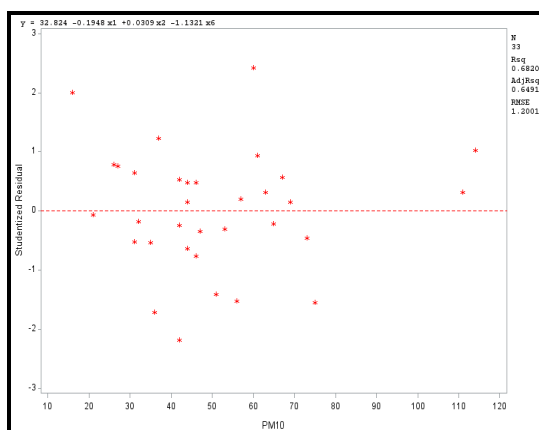


圖 3.2.2 相對溼度標準化殘差圖

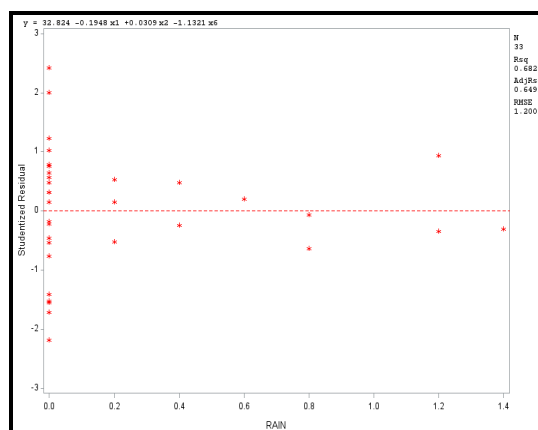


圖 3.2.3 懸浮粒子濃度標準化殘差圖

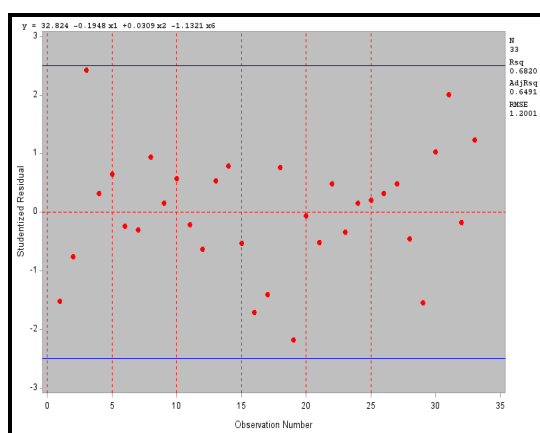


圖 3.2.4 各地雨量標準化殘差圖

圖 3.2.5 各地氣溫時間序列標準化殘差圖

三、Studentized Deleted Residuals 檢定

根據 Studentized deleted residuals 檢定標準 $|t_i| > 3$ 為異常點，由表 3.4 顯示 33 筆資料中均沒有異常點。且由圖 3.3.1 到圖 3.3.5 可知殘差皆落在以 0 為中心的區塊內，圖 3.3.1 呈現喇叭狀，圖 3.3.4 有逐漸縮小的趨勢，但依上述並沒有異常點的出現。

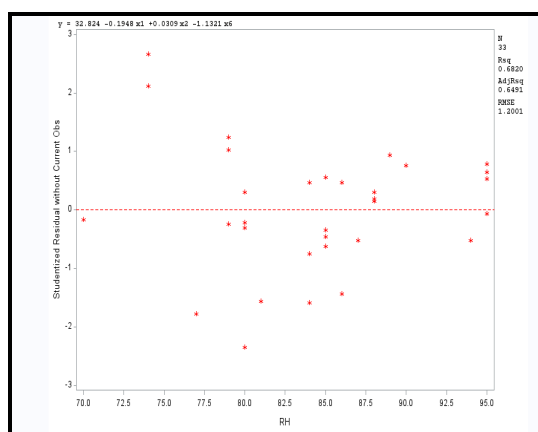
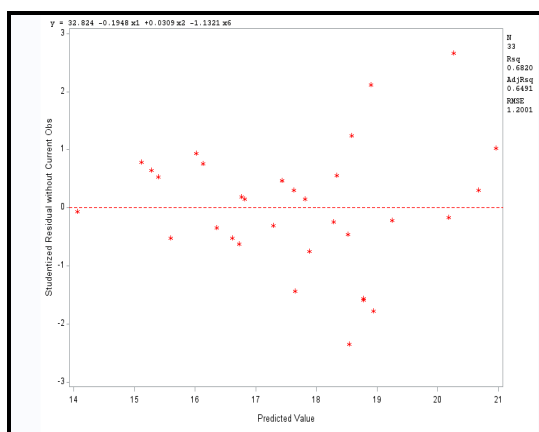


圖 3.3.1 各地氣溫 R student 圖

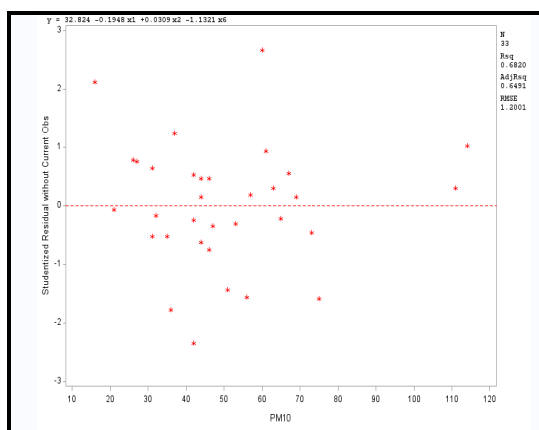


圖 3.3.3 懸浮粒子濃度 R student 圖

圖 3.3.2 相對溼度 R student 圖

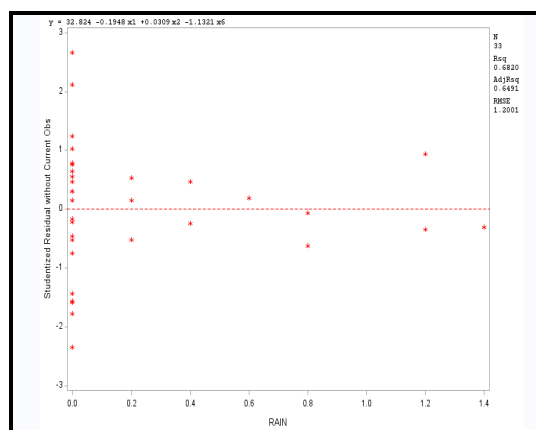


圖 3.3.4 各地雨量 R student 圖

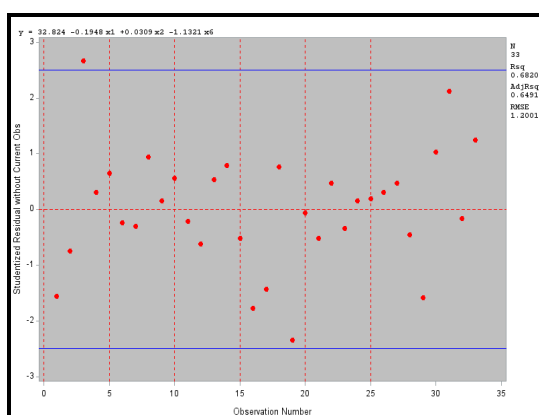


圖 3.3.5 各地氣溫時間序列 R student 圖

第三節 影響點

此小節我們使用 DFFITS 檢定、The hat matrix elements 檢定、Cook' s distance statistic 檢定、DFBETAS 檢定和 COVRATIO 檢定來檢測是否有影響點。

一、DFFITS 檢定

根據 DFFITS 判斷準則，若大於 $2\sqrt{\frac{p}{n}}$ 則為影響點，其中 n 為樣本數、p 為參數個數， $|DFFITS| > 0.6963$ ，由表 3.4 統計出下表 3.5 這幾筆資料為影響點，差異最大的地區為屏東縣恆春鄉。

表 3.5 DFFITS 影響點數值表

地區代號、名稱	DFFITS
3，高雄市-小港	0.9771
31，屏東縣-恆春	1.1952

二、The hat matrix elements 檢定

根據 The hat matrix elements 判斷準則，若大於 $2 \frac{p}{n}$ 則為影響點， $|\text{Hat value}| > 0.2424$ ，由表 3.4 統計出下表 3.6 這幾筆資料為影響點，差異最大的地區為台中市大里區。

表 3.6 Hat value 影響點數值表

地區代號、名稱	\text{Hat value}
4, 高雄市-仁武	0.2719
7, 台中市-大里	0.3324
30, 屏東縣-屏東	0.2971
32, 台東縣-台東	0.2508

三、Cook' s distance statiatic 檢定

根據 distance 準則，若大於 1 但由表 3.1 沒有任何一在此檢定法是沒有影響點。

Cook' s statiatic 判斷 1 則為影響點，統計出來的資料筆大於 1，所以的標準之下結果

四、DFBETAS 檢定

根據 DFBETAS 判斷準則，若 $|\text{DFBETA}| > \frac{2}{\sqrt{n}}$ 則為影響點， $|\text{DFBETAS}| > 0.3482$ ，由表 3.4 統計出下表 3.7 這幾筆資料為影響點，差異最大的地區為屏東縣恆春鄉。

表 3.7 DFBETAS 影響點數值表

地區代號、名稱	Y	X ₁	X ₂	X ₆
---------	---	----------------	----------------	----------------

3, 高雄市-小港	<u>0.7701</u>	<u>0.7706</u>	0.0363	0.1642
8, 台中市-西屯	0.1240	0.0992	0.1482	<u>0.4491</u>
16, 桃園縣-大園	<u>0.4839</u>	<u>0.4227</u>	0.3150	0.1496
19, 新竹縣-湖口	<u>0.3803</u>	0.3165	0.2461	0.2120
30, 屏東縣-屏東	0.0368	0.0475	<u>0.6006</u>	0.0600
31, 屏東縣-恆春	<u>0.9788</u>	<u>0.8389</u>	<u>0.8539</u>	0.1930

五、COVRATIO 檢定

根據 COVRATIO 的判斷準則，若大於 $1+3\frac{\sigma}{\mu}$ 或者小於 $1-3\frac{\sigma}{\mu}$ 則為影響點，

$|\text{COVRATIO}| > 1.3636$ 或 $|\text{COVRATIO}| < 0.6364$ ，由表 3.4 統計出下表 3.8 這幾筆資料為影響點，差異最大的地區為台中市大里區。

表 3.8 COVRATIO 影響點數值表

地區代號	COVRATIO
3, 高雄市-小港	0.5263
4, 高雄市-仁武	1.5584
7, 台中市-大里	1.7023
19, 新竹縣-湖口	0.5976
20, 苗栗縣-三義	1.4158
23, 彰化縣-彰化	1.4396
30, 屏東縣-屏東	1.4096
32, 台東縣-台東	1.5297

綜合以上五種檢測影響點的方法，由下表 3.9 我們統計出可能為影響點的地區，第 7 筆資料台中大里區與第 31 筆資料屏東縣恆春鄉都在這五種檢測方法差異最大。所以我們原先有 69.15% 的解釋能力，在刪除第 7 筆資料台中大里區後，我們的解釋能力下降為 68.19%，故決定不刪除此影響點；接著在第 31 筆資料屏東縣恆春鄉後，解釋能力提高為 69.99%，因此決定刪除此影響點以增加解釋能力。

表 3.9 影響點統計歸納表

檢測方法	地區代號							
	3	<u>31</u>	-	-	-	-	-	-
DFFITS	3	<u>31</u>	-	-	-	-	-	-
Hat value	4	<u>7</u>	30	32	-	-	-	-
Cook' s D	-	-	-	-	-	-	-	-
DFBETAS	3	8	16	19	30	<u>31</u>	-	-
COVRATIO	3	4	<u>7</u>	19	20	23	30	32



第四章 結語

本次研究主題蒐集的資料為影響 2012 年 11 月 27 日下午六點台灣本島各地溫度的各種變數來分析，依據選取重要變數方法，得到相對濕度、懸浮粒子濃度及雨量為可能影響溫度的最佳模式變數。

此外，使用檢測異常點和影響點的方法，發現並沒有異常點存在，但是有影響點，所以為了提升解釋能力選擇刪除第 31 筆資料，使解釋能力提升至 69.99%。

在日常生活中發現，空氣中的濕度和降雨的雨量會使溫度降低，所以我們更進一步分析得到，以在台灣當日的相對濕度而言，相對濕度增加 1% 時，則溫度會減少 0.1678°C ；降雨量只要增加 1 mm，台灣各地的溫度就會下降 1.0367°C ，且明顯地看出降雨量比濕度更能直接影響台灣當時溫度的變化。接著，由於現在化石燃料、焚化廠及交通工具使用上的數量比起以往有很大不同，上述三者皆會影響空氣中懸浮粒子濃度，進而造成空氣品質上的差異，在紫外線照射下，受空氣中懸浮粒子濃度吸收紫外線程度不同所影響，使滯留在大氣內的熱能增加，則導致溫度上升，根據分析結果發現每當懸浮粒子濃度增加 $1\ \mu\text{g}/\text{m}^3$ ，溫度則會增加 0.0389°C 。由此更能確切知道，台灣本島的降雨量影響溫度的變化最為明顯，其次為相對濕度和懸浮粒子濃度。

附錄一、原始資料

本附錄的資料數據是從行政院環境保護署—空氣品質監測網提供。

	時間	反應變數 Y	解釋變數 X					
	2012/11/27 18:00	大氣溫度 (°C)	相對溼度 (%)	懸浮粒子濃度 ($\mu\text{g}/\text{m}^3$)	一氧化碳濃度 (ppm)	氮氧化物濃度 (ppb)	風速 (m/sec)	雨量 (mm)
1	台北市-士林	17.00	81	56	0.39	17	3.4	0
2	台北市-中山	17.00	84	46	0.74	46	3.6	0
3	高雄市-小港	23.00	74	60	0.79	55	1.2	0
4	高雄市-仁武	21.00	80	111	1.2	94	1.2	0
5	基隆市-基隆	16.00	95	31	0.74	39	1.2	0
6	新竹市-新竹	18.00	79	42	0.64	37	1.7	0.4
7	台中市-大里	17.00	80	53	1.32	79	0.5	1.4
8	台中市-西屯	17.00	89	61	0.74	44	2	1.2
9	嘉義市-嘉義	18.00	88	69	1.06	51	1.5	0
10	台南市-台南	19.00	85	67	1.11	55	1.8	0
11	台南市-安南	19.00	80	65	0.73	46	1.6	0
12	新北市-板橋	16.00	85	44	0.74	40	1.7	0.8
13	新北市-汐止	16.00	95	42	0.42	34	1.9	0.2
14	宜蘭縣-冬山	16.00	95	26	0.77	34	1.4	0
15	宜蘭縣-宜蘭	15.00	94	35	0.55	20	0.9	0
16	桃園縣-大園	17.00	77	36	0.37	24	5.3	0
17	桃園縣-中壢	16.00	86	51	1.25	60	1.3	0
18	新竹縣-竹東	17.00	90	27	0.59	23	0.9	0
19	新竹縣-湖口	16.00	80	42	0.44	24	5	0
20	苗栗縣-三義	14.00	95	21	0.46	12	2	0.8
21	苗栗縣-苗栗	16.00	87	31	0.98	43	1.6	0.2
22	彰化縣-二林	18.00	84	46	0.52	26	2.5	0.4
23	彰化縣-彰化	16.00	85	47	1.08	56	1	1.2
24	南投縣-竹山	17.00	88	44	0.77	37	0.8	0.2
25	南投縣-南投	17.00	88	57	1.52	78	0.9	0.6
26	雲林縣-斗六	18.00	88	63	0.81	37	1.3	0
27	雲林縣-台西	18.00	86	44	0.55	26	3.4	0
28	嘉義縣-朴子	18.00	85	73	0.71	28	1.4	0
29	嘉義縣-新港	17.00	84	75	0.77	34	1.9	0
30	屏東縣-屏東	22.00	79	114	1.02	68	0.9	0
31	屏東縣-恆春	21.00	74	16	0.4	3.4	9.3	0
32	台東縣-台東	20.00	70	32	0.84	18	1.9	0
33	花蓮縣-花蓮	20.00	79	37	1.25	39	1.5	0

附錄二、參考文獻

Abraham, B. and Ledolter, J. (2006) Introduction to regression modeling. Belmont, CA:Duxbury Press.



附錄三、小組討論時間表

時間	地點	討論內容
2012 年 11 月 26 日下午兩點	逢甲大學創意知識工坊	決定主題、蒐集資料。
2012 年 11 月 27 日下午一點	逢甲大學創意知識工坊	蒐集資料、初步資料分析。
2012 年 11 月 28 日下午六點	逢甲大學圖書館討論室	使用 SAS 軟體、資料分析。
2012 年 12 月 3 日下午一點	逢甲大學創意知識工坊	重新蒐集資料、使用 SAS 軟體、資料分析。
2012 年 12 月 5 日下午六點	逢甲大學圖書館討論室	使用 SAS 軟體、緒論、資料分析。
2012 年 12 月 10 日下午六點	逢甲大學圖書館討論室	使用 SAS 軟體、書面初步排版。
2012 年 12 月 12 日下午六點	逢甲大學圖書館討論室	殘差分析、結論。
2012 年 12 月 18 日下午兩點	逢甲大學創意知識工坊	修改中文摘要、殘差分析。
2012 年 12 月 19 日下午六點	逢甲大學圖書館討論室	製作書面報告、投影片。
2012 年 12 月 21 日下午四點	逢甲大學圖書館討論室	書面報告修改、製作投影片。
2012 年 12 月 22 日下午一點	逢甲大學圖書館討論室	製作投影片。
2012 年 12 月 23 日上午十點	逢甲大學圖書館討論室	製作投影片、口頭報告演練。
2012 年 12 月 24 日下午一點	逢甲大學創意知識工坊	口頭報告演練。
2012 年 12 月 26 日下午六點	逢甲大學忠勤樓 304 教室	最後確認書面報告、投影片、口頭報告演練。