

國立交通大學

運輸科技與管理學系

碩 士 論 文

高速公路旅行時間預測-以 k -NN法及分群方法探討

Freeway Travel Time Prediction by Using the k -NN Method and
Comparison of Different Data Classification

研 究 生：蔡繼光

指導教授：卓訓榮 教授

中 華 民 國 九 十 八 年 七 月

高速公路旅行時間預測-以k-NN法及分群方法探討

學生：蔡繼光

指導教授：卓訓榮 博士

國立交通大學運輸科技與管理學系碩士班

摘 要

旅行時間預測為先進旅行者資訊系統的一部分，其目的為利用即時的速度、流量等交通資訊，即時地、準確地預測旅行時間以供用路人、交通管理人員進行查詢或各種決策分析。對於用路人而言，大多以旅行時間最小為其決策目標進行路線規劃。然而用路人只能掌握部分的路況資訊，較難決定旅次之出發時間、並對路徑選擇進行最有效率的決策與預估，降低運輸行為之不確定性。因此旅行時間預測為智慧型運輸系統中一個重要的課題。

本研究利用電子收費系統之歷史資訊及高速公路偵測器所收集到的即時與歷史交通資訊(速度、流量)進行分析。在研究方法上，利用比對的方法，找出與即時交通資訊類似之歷史交通資訊，再利用電子收費系統之資訊推估該歷史時間點之旅行時間，進行旅行時間預測。

本研究在資料比對的方法上採用 k -NN 法(k -Nearest Neighbor Method)進行處理。在資訊比對上。為了進一步提高上述方法的準確率，由單時間點的比對擴大為，比對一段時間的交通資訊的變化，並考量在不同偵測器下交通特徵的差異，因此再加入偵測器的權重以進行分析。並討論在不同的資料分群下，是否可以提升預測準確率。

最後利用台灣國道三號高速公路為對象進行研究，驗證本研究方法是否有足夠能力得到準確的預測結果。

關鍵詞：k-NN 法，高速公路旅行時間預測，資料分群

Freeway Travel Time Prediction by Using the k -NN Method and Comparison of Different Data Classification

student : Chi-Kuang Tsai

Advisors : Dr. Hsun-Jung Cho

Department of Transportation Technology Management
National Chiao Tung University

ABSTRACT

The travel time prediction is a part of Advanced Traffic Information Systems, ATIS. It is using the instant speed and flow information to predict the travel time on certain the path. It helps the road user and traffic manager to survey or strategy analysis. Using the travel time information, it is useful for road user to do the trip planning by minimizing the travel time. Without travel time information, it is hard to gather enough information doing the efficient trip planning, nor estimating the travel time. Therefore, travel time prediction is the important issue in intelligent transportation system. It help trip planning and decrease the uncertainty of the trip.

The study uses the historical electronic toll collection, ETC, information and historical and real-time vehicle detector, VD, information, speed and flow, on the freeway. Using the comparison method find out the familiar traffic information between historical data and real-time data. Then, estimate the travel time of the historical time by ETC data. With the familiar period and estimated travel time predict the real-time travel time.

The study adopts the k -nearest neighbor method, k -NN method, in data comparison. For decreasing the rate of error, instead of comparing each time period compares the traffic information trend to find out the similar data. Besides, we consider the difference between different VDs, and weight the VD data to adjust. Finally, discussing different data classification increases the performance of prediction.

Finally, the study took the third freeway in Taiwan for example. Figure out if the method got enough predicting ability.

Keywords: k -NN method, freeway travel time prediction, data classification.

誌謝

首先，我要感謝這兩年給與我指導的卓訓榮老師，在這過程中，逐漸清楚了所謂學術論文的意義，以及做研究的方向，並且在待人處世上也給與了我相當大的影響及鼓勵，期望在未來的日子裡依然能夠與卓老師維持良好的互動。接下來要感謝系上的王晉元老師及資財系的周幼珍老師，時間與精力，幫我找出了論文上的瑕疵以及一些說明不清之處，也謝謝韓復華老師在期中報告時給了我後續可以研究的方向，讓我有能力將這篇論文順利完成。也要謝謝系上的老師在這兩年給我的指導與體諒，讓我增進不少見聞。

另外還要謝謝運研所的葉祖宏先生，在我進行程式的撰寫時，幫我檢查演算法是否得到正確的結果，並找出程式中的一些問題以助我完全我這個部分程式，給予我相當大的助力。另外也要謝謝彥佑學長、健綸學長、黃恆學長、佳儒、如君還有志霖，靠著大家的努力把這個計畫處理完畢，也讓我可以用來參考而得已做為論文的資料來源。

最後，要謝謝健綸學長、昱光學長、黃恆學長、隊長、帥哥錦、猴給、Issac、如君和亦晴陪伴我度過綜一 1106 室 730 天的日子(好像沒那麼多)，還要謝謝之音、宜珊還有小捲陪了我 365 天(因為你們走太快了)，還有志玲和怡婷陪我度過學生生涯的最後一年。這一切在壁虎的啾啾聲，還有鍵盤的答答聲中不知不覺就結束了，讓我在學生的最後兩年有點點滴滴的回憶。

也要謝謝我的酒肉朋友，帥哥號和蕭宅宅，當我在寂寞、飢餓的時候，助我一臂之力，帶我離開那種困境；也要謝謝我的好室友夸克和強翰(號?)，陪我度過無數個夜晚；還有像是老蔣、包子這些其他實驗室的好朋友，願意串串門子帶來一些新鮮空氣。謝謝大家的幫忙，讓我這兩年的生活沒有什麼空白。

最後，也是最重要的，要感謝我的父母，願意在我的求學生涯中，給我無邊無際的支援與支持，讓我能夠勇敢面對所有的問題，所有的難關，所有的所有，讓我知道我有後盾可以讓我依靠，謝謝你們，在此僅以此篇論文獻給在我身邊的你們。

蔡繼光 謹 誌于

國立交通大學運管所

中華民國九十八年七月

目 錄

中文摘要.....	I
英文摘要.....	II
誌謝.....	III
目錄.....	IV
表目錄.....	VI
圖目錄.....	VII
第一章 緒論.....	1
1.1 研究背景.....	1
1.2 研究目的.....	2
1.3 研究範圍.....	3
1.4 研究流程.....	4
第二章 文獻回顧.....	7
2.1 旅行時間預測相關文獻.....	7
2.2 k -NN 預測交通資訊相關文獻.....	13
2.3 小結.....	16
第三章 資料處理.....	17
3.1 偵測器資料與 ETC 資料的比較.....	17
3.2 高速公路偵測器資料.....	26
3.3 電子收費系統資料.....	28
3.4 ETC 與 VD 之間關係.....	31
第四章 旅行時間預測模式介紹.....	33
4.1 k -NN 介紹.....	33
4.2 k -NN 參數設定.....	33
4.3 k -NN 模式建立.....	37
4.4 本研究所設定之 k -NN 模型.....	38
4.5 模式預測流程.....	41

第五章 實例測試	42
5.1 研究範圍	42
5.2 旅行時間預測結果	45
5.2.1 龍潭收費站到樹林收費站測試結果	45
5.2.2 後龍收費站到樹林收費站測試結果	47
第六章 結論與建議	53
參考文獻	54
附錄一 各偵測器信賴區間	58
附錄二 其他 VD 與 ETC 比較結果	67
附錄三 其他預測結果	70
自傳	75



表 目 錄

表 2.1 類神經網路模式之構建類型整理	12
表 3.1 缺漏嚴重之偵測器代碼及缺漏率	26
表 3.2 研究範圍之偵測器代碼及缺漏率	27
表 5.1 臺灣地區國道三號高速公路各收費站通行車輛次數	44
表 5.2 k -NN 模式旅行時間預測整理(970330)	51



圖 目 錄

圖 1.1 研究流程圖	6
圖 2.1 SVR 示意圖	9
圖 3.1 970121(一)流量與旅行時間對應圖	18
圖 3.2 970122(二)流量與旅行時間對應圖	19
圖 3.3 970123(三)流量與旅行時間對應圖	20
圖 3.4 970124(四)流量與旅行時間對應圖	21
圖 3.5 970125(五)流量與旅行時間對應圖	22
圖 3.6 970126(六)流量與旅行時間對應圖	23
圖 3.7 970127(日)流量與旅行時間對應圖	24
圖 3.8 96 年歷史旅行時間信賴區間圖	28
圖 3.9 歷史旅行時間估計流程圖	30
圖 3.10 偵測器推估歷史旅行時間流程圖	31
圖 3.11 偵測器與 VD 歷史旅行時間推估圖	32
圖 4.1 旅行時間對應常態分配分佈圖	35
圖 4.2 不同 k 值對應誤差圖	36
圖 4.3 差異值說明圖	39
圖 4.4 預測流程架構圖	41
圖 5.1 研究範圍高速公路路線圖	43
圖 5.2 龍潭到樹林依單時間點 k -NN 法旅行時間預測圖	45
圖 5.3 龍潭到樹林依日期別分群旅行時間預測圖	46
圖 5.4 龍潭到樹林依星期別分群旅行時間預測圖	46
圖 5.6 k -NN 模式旅行時間預測圖(970315)	48
圖 5.7 k -NN 模式旅行時間預測圖(970316)	48
圖 5.8 k -NN 模式旅行時間預測圖(970303)	49
圖 5.9 k -NN 模式旅行時間預測圖(970304)	49
圖 5.10 k -NN 模式旅行時間預測圖(970305)	50

第一章 緒論

1.1 研究背景

由於運輸建設方案執行速度永遠不及交通需求之膨脹，當某些時段、路段交通需求超出道路容量負荷時，就有供不應求的狀況發生，而道路也就隨之產生了擁擠現象而造成車速下降，甚至造成塞車的發生，而行程的不確定性就隨之提高。因此為了改善現有的交通問題，在觀念與作法上已逐漸由增加道路容量的方式轉變成提升利用現有的資源與交通設施的使用效率為目標，在既有的資源下，達到最大的效率。

有鑑於以上的現象，各國開始著手研究應用電子、資訊、車輛、通訊、控制等技術配合交通管理與控制策略，以提升交通運輸系統的效率，發展運輸系統的智慧化。1980 中期以後，歐、美、日等先進國家更投入大量資源積極發展智慧型運輸系統(Intelligent Transport System, ITS)，希望不僅提高交通系統之運作效率，並且增加駕駛之安全性。我國也持續地針對智慧型運輸系統進行一連串的計劃，包括運輸政策白皮書(2002)、台灣地區 ITS 綱要計畫(2003)等。其中先進旅行者資訊系統(Advanced Traffic Information Systems, ATIS)為智慧型運輸系統中的一部分，其作用在於提供旅行者或是交通管理人員準確的且可信賴的交通資訊。而旅行時間預測為 ATIS 之中的一個方向，其目的為利用速度、流量等交通資訊，即時地、準確地預測某一段路徑之旅行時間供用路人、交通管理人員進行查詢以創造出附加價值，降低運輸行為之不確定性、並進行各種決策分析。

對用路人而言，在行駛至目的地的各種路線選擇下，如何運用有效方法完成旅次是用路人追求的目標。假設用路人在理性的決策下進行路線規劃，使用者在進行路線規劃時會考量個人的偏好、態度、知識以及對交通系統的了解，以追求旅行時間成本最小為目標進行規劃[1]。然而用路人只能掌握部分的路況資訊，無法針對旅次之出發時間、路徑選擇進行最有效率的決策與預估。因此 ATIS 的目的為提供可靠且準確的交通資訊，以增加用路人進行旅次規劃之效率，在效率提高下有以下兩個優點。

1. 節省道路使用者的旅行成本。

有助於道路使用者找尋最短旅行時間之路線選擇以降低旅行時間成本，並且讓道路使用者選擇其他路線以加速擁塞的抒解以降低因道路擁塞所帶來的社會成本，進而提高運輸系統之服務水準。

2. 有助於減少道路使用者減少不確定性，以增加道路使用者的舒適程度。

由於提供使用者較多資訊可以有助於減少道路使用者對於交通狀況的不確定性，因此道路使用者可以得到更多的交通資訊以增加舒適程度[2]。

對於交通管理人員，ATIS 有助於進行路網績效評估、控制策略研擬、駕駛人行為分析等決策應用分析。

我國也持續地對於 ATIS 方面進行發展，行政院於民國 95 年產業科技策略會議(SRB)[3]中提出「智慧型車輛產業的檢視與前瞻」議題，建議智慧型運輸系統(ITS)建置「北台灣科技走廊智慧型運輸系統示範區域」。利用先進感測設施，設計具人工智慧的交通分析系統進行交通控制並以多重管道資訊傳播系統發佈資訊，其目的在於協助用路人有足夠的資訊在高速公路及替代道路中進行路線選擇，以避開壅塞而快速到達目的地。因此提供快速、準確及充分的交通資訊給用路人是各國一直在研討的課題。

1.2 研究目的

路況資訊有很多種類型呈現，如速率、流量、佔有率、事故偵測等，然而對於用路人而言，如何讓用路人最有效的掌握路況，才是用路人最關心的交通資訊。用路人可以透過各種管道得到即時的交通資訊，再加上自己的經驗，自行推估在當時路況下需要多少時間完成目標的旅次。用路人在知道當時的交通資訊下，不一定有辦法迅速地且準確地預測在旅次進行過程中是否發生道路擁塞的現象，因此可能出發時交通順暢，但是隨著旅次的進行，愈來愈多車輛進入道路，而在旅次進行一半時或者在某些路段，發生道路擁塞的情形。甚至是用路人在沒有足夠的資訊及經驗下，忽略了其他更快到達目的地的路線。

由於即時的速度、流量等交通資訊只能反應當下的交通狀況，對於用路人的實際效用可能仍顯不足。因此本研究希望可以研究一種旅行時間預測的方法有能力即時地預測完成旅次所需要的旅行時間。而目前有相當多的研究，其旅行時間資料庫都是由歷史的速度、流量等資訊推估的結果，本研究希望可以建立一組比較有可靠性的歷史旅行時間資料庫，因此利用高速公路電子收費系統得到車輛通過各收費站的時間以得到實際的旅行時間資訊。

由即時的交通資訊對應歷史資訊預測出在相同的起迄點下所需要的旅行時間。在了解每一條路線的旅行時間後，可以鼓勵用路人於道路發生擁塞前即利用替代道路，除了給與用路人更低旅行成本的路線規劃，也可以減少道路的負荷，以降低發生擁擠的機會，或加速道路擁塞的抒解，進而減少塞車所帶來的社會成本。因此提供用路人旅行時間的資訊是有需要的。

目前在旅行時間預測上有相當多的研究是利用推估現在的交通資訊與未來的交通資訊之間的關係進行預測，研究當尖峰、離峰；平日、假日；甚至季節改變下，即時交通資訊與未來交通資訊之間的關係。然而如果當即時資訊發生不預期的變化時，像是連續假日、晴天、雨天，即時的交通資訊和未來交通資訊之間的關係都可能和平常的情形不同，而預測的能力就不一定能夠維持。

未來的資訊和即時的資訊有關聯性，因此在相同的交通情形下，相同的路段會對應類似的旅行時間長度。本研究希望找出與即時交通現象類似的歷史交通現象，因此會對應類似的旅行時間，而當有不預期的事件發生時，如果歷史有類似的交通現象下，就可以得到類似的旅行時間資訊。

本研究將試著收集即時交通資訊以及建立歷史旅行時間資料庫，並且利用 k -NN 方式做為本研究之模式進行旅行時間的預測，並試著提升其準確率。

1.3 研究範圍

本研究希望可以預測在某一組相同的起迄點下，各個時間點所可能的旅行時間，由於目前北台灣的運量較大，交通擁塞發生的機會較大，所以本研究希望可以有效的預測在交通擁塞下或是交通順暢下，該路段的旅行時間為何。由於目前我國偵測器密度最高的道路為高速公路，而且可以藉由電子收費系統(ETC)所收集的資訊，得到在各時間點下，不同收費站之間所需要的旅行時間為何。結合由偵測器與電子收費系統所收集的資訊，對交通資訊與旅行時間之間的關係為本研究的研究範圍。

因此本研究將以高速公路為主要的研究對象，並以國號三號上的樹林收費站到後龍收費站之間的旅行時間為分析對象，利用其間高速公路的偵測器及電子收費系統(ETC)所蒐集的交通資訊，預測旅行時間，並找出與實際值誤差最小的方法，以增加預測效率。

因此本研究的範圍為以下幾點：

1. 如何從歷史 ETC 的資料校估出實際的歷史旅行時間資訊。
2. 如何從即時的道路資訊與歷史的道路資訊中，預測旅行時間。
3. 如何提升旅行時間預測的績效。

1.4 研究流程

本研究的研究流程如圖 1.1 所示，茲將流程圖中各步驟詳細說明如下：

1. 研究範圍界定

決定實際案例的研究對象以及預測範圍起迄點，並根據研究背景與目的將問題進行清楚的描述與界定。

2. 文獻回顧

蒐集國內外研究高速公路旅行時間預測的相關文獻，接著回顧一些利用 k 最近鄰法(k nearest neighbor, k -NN)法預測交通資訊的相關文獻。利用既有的研究整理出本研究的研究方向。

3. 將資料進行分群(classification)

利用不同的分群將所蒐集到的歷史資料進行分群的動作，本研究將研究是否針對資料進行分群可以得到較佳的結果，若沒有經過分群動作，則可視為每一天為一組樣本進行分群；若以星期別分群，則將歷史與即時的資料依星期別分群，相同的星期別視為相同的群組。

4. 判斷輸入值為那一個集群(cluster)

若資料有經過分群動作，則即時資訊需要辨別位於那一個分群之中，因此需要進行集群的動作，依照不同星期別作直觀的分類。需要判斷該即時資訊位於何種分群下，再進行比對的動作，最後即時資訊與歷史資料相比對，利用 k 最近鄰法找出最類似的樣本值以預測旅行時間。

5. 利用 k -NN 進行旅行時間預測

本研究在進行 k 最近鄰法的比對前，需要找出比對的門檻值，將離群值刪除，接著決定比對之類似值的數量上限(k)，以及相關的參數，接著利用這些參數的設定以提升 k 最近鄰法的準確率。最後將實際案例利用分群後，得到分群或不分群的預測結果。

6. 比較這些分群下的績效

在各別預測以上三種分群的結果後，分析這些不同的分群下預測結果與實際值的誤差值，以判斷不同分群方法的優劣，並討論是否有其他改進的方法。

7. 結論與建議

對本研究過程與結果提出結論與建議，並提出後續可以進行研究的方向。



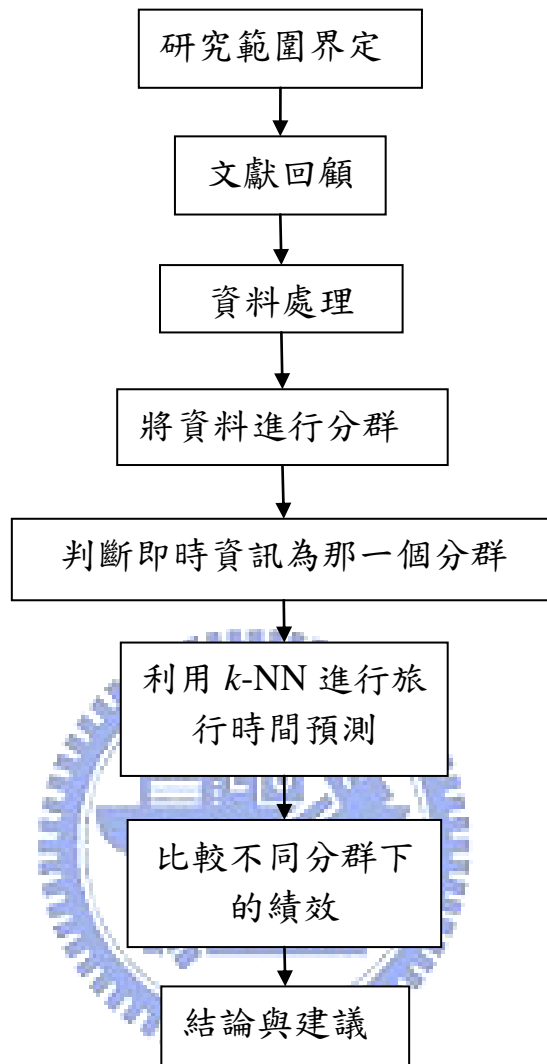


圖 1.1 研究流程圖

[資料來源：本研究整理]

第二章 文獻回顧

本研究目的為利用即時交通資訊發展一套旅行時間預測模式，並且利用 k -NN 法進行預測求解。以下將對於旅行時間預測、 k -NN 預測交通資訊方法，以上二種方向整理相關的研究且進行討論，歸納出本研究的預測方法。

2.1 旅行時間預測相關文獻

目前高速公路旅行時間預測文獻上，大致可分為有母數方法(parameter)及無母數(non-parameter)方法這兩種方法。其中有母數方法可以分為以下幾種：1.迴歸方法(regression)，2.時間序列方法(time-series)。以下將分別說明這兩種方法。

迴歸方法

迴歸分析是統計方法中廣為使用的分析方法，其運用涉足各個領域。目前迴歸方法運用於旅行時間的預測上，主要以兩個方向為主：1.交通資訊與旅行時間之間的關係，找出速度、流量、佔有率為輸入資訊，找出與旅行時間之間的關係。2.不同時間點旅行時間的關係，得到某個時間點的旅行時間與下一個時間點的旅行時間後，找出之間的關係。以下將分別說明：

1.交通資訊與旅行時間之間的關係

Kwon 等學者(2000) [4]利用美國 I880 資料庫來建立旅行時間預測模式，資料庫中包含單一線圈偵測器所偵測到的佔有率以及流量資訊，及探針車所得到的旅行時間，試著找出交通資訊與旅行時間之間的關係。模式中將偵測到的流量、佔有率、及當時的星期別及日期都視為一個集合 X_i ，對探針車得到的實際旅行時間做線性迴歸，最後以平均絕對百分誤差(Mean Absolute Percentage Error, MAPE) 來檢定預測績效。

2.不同時間點間旅行時間的關係

Sen 等人(1997)[5]引用美國芝加哥的 ADVANCE 計劃中，討論在某一些偵測器較少路段下，如何進行旅行時間預測。在這裡利用探針車以收集道路資訊，蒐集各車輛行駛時所在各路段下旅行時間為何。討論在不同的日期、不同的月份下，各時間點旅行時間之間的關係，並針對這些迴歸式的參數值是否具有顯著性進行研究。當某些偵測器充足的路段，則利用路段長度與平均速率之間的關係推估該路段旅行時間。該研究依照在某路段上有無佈設偵測器的兩種情況，採用探偵車或是偵測器兩種不同的資訊收

集方式來進行旅行時間預估。最後提出在都市幹道的網路上，受到號誌系統控制與環境因素之影響下，路段旅行時間之不確定性較高速公路系統來的複雜。

Rice 等學者(2004)[6]使用線性迴歸(linear regression)、主成份分析(principal component)以及最近鄰法(nearest neighborhood)三種方法，各別應用在 PeMS 計劃中預測旅行時間。其中迴歸方式利用單一線圈偵測器(single loop detector)收集資料，可以得到流量與佔有率的資料，接著利用流量與佔有率推估出速度，利用上下游兩偵測器之間的距離，及這兩個偵測器附近的的速度資訊，推估當時的旅行時間。利用這些旅行時間建立歷史資料庫，並推估一套前後時間點旅行時間的線性迴歸方程式，在已知兩偵測器之間的交通資訊，推估這兩個偵測器間的旅行時間後，預測從第二個偵測器出發到第三個偵測器間所需要的旅行時間。

此迴歸式分別以時間別及星期別進行分群，最後將以上三種預測方法所得結果以均方根誤差(Root Mean Square Error, RMSE)來檢定預測績效。

時間序列

由以上迴歸的方法，可以發現迴歸主要的目的是在找一組輸入參數和一組輸出參數之間的關係，然而在旅行時間上可以發現，各時間點的交通資訊之間都有關係，因此時間序列的方法主要是從利用多組的資料中，找出這些樣本之間的關係，再加入一些變異數以預測旅行時間，因此時間序列方法可說是利用多組輸入資訊以推估一組輸出值的方法，其中主要有兩個方法，1.自我迴歸平均整合(Autoregressive Integrated Moving Average, ARIMA)。2.支持向量機迴歸(Support Vector Machine Regression, SVR)。

1. 自我迴歸平均整合

Yang 等人(2006) [7]把時間序列模式應用在幹道的旅行時間預測研究上。透過實際採用裝有 GPS 系統的探針車於美國明尼蘇達州 194 號高速公路作實際研究。其方法是將蒐收所得資料視為時間序列，以 ARIMA 模式進行旅行時間預測。對明尼蘇達 194 號高速公路作實測，結果顯示此方法能夠有效預測短期內的旅行時間。

2. 支持向量機迴歸

支持向量機迴歸(support vector machine regression, SVR)方法是由 Vapnik's(1997)[8]提出，主要應用於時間序列的預測上。

Wu 等人(2004)[9]嘗試將其運用在旅行時間預測上，利用中研院的 ITWS 計劃在國道一號高速公路上選取一公里的範圍作 9 為研究對象，以迴圈偵測器(loop detector)，將所收集到的道路速度資訊，建立每三分鐘一筆的歷史資料庫。最後使用以下三種方法：支持向量機迴歸(support vector machine regression)法、當前旅行時間預測法(current travel time prediction method)以及歷史平均旅行時間預測法(historical mean prediction method)來預測旅行時間。並比較三種旅行時間的推估方法：

A. Support vector regression(支持向量機)

利用所收集到交通資料得到在不同時間下旅行者所花的旅行時間，先將有這些樣本資料經過轉換得到 $\Phi(x)$ ，再試著找出利用一條分隔線，將具有相似特徵資料視為同一分群，最後找出該分群中交通資料與旅行時間之間的關係。因此在不同的分群中，需要建立不同的旅行時間預測推估模式。

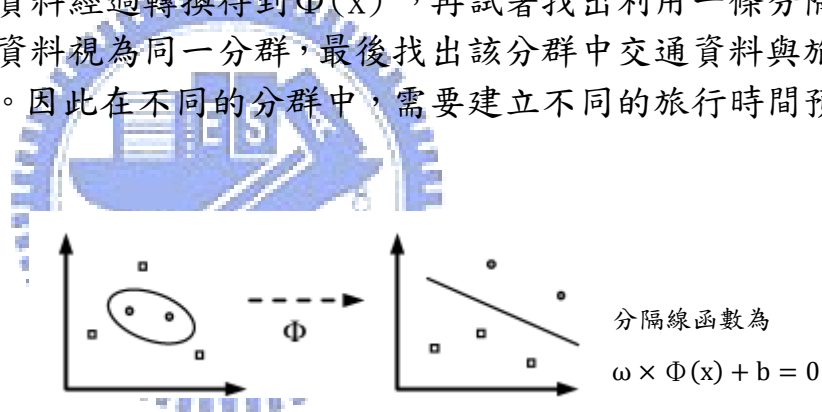


圖 2.1 SVR 示意圖

[資料來源：Wu 等人(2004)[9]]

B. current travel time prediction method

收集即時的交通資料，以及各偵測器間的距離，並考慮資料傳遞的延誤後，預測所需要的旅行時間，預測函式如下式所示。利用該旅行時間預測函數，直接由速率預測旅行時間。

$$T(t, \Delta) = \sum_{i=0}^{L-1} \frac{x_{i+1} - x_i}{\gamma(x_i, t - \Delta)}$$

Δ ：資訊延誤

$(x_i, t - \Delta)$ ：實驗的路段長度

$\gamma(x_i, t - \Delta)$ ：進入路段的速率

C. historical mean prediction method

收集歷史各時間下旅行者所花的旅行時間資料，代入下式後，得到一組在不同時段下的旅行時間估計函數，利用平均的歷史旅行時間預測未來的旅行時間。

$$\tilde{T}(t) = \frac{1}{w} \sum_{i=1}^w T(i, t)$$

w ：資料收集的星期數； $T(i, t)$ ：不同星期所收集的旅行時間

以上的方法主要是根據偵測器所收集的旅行時間樣本資料做分析，由於偵測器每三分鐘更新一次，所以造成資料有中斷或是錯誤的現象發生，因此將有效的資料收集後，再根據不同的估計方法預測旅行時間的長度。

而在績效評估上，以均根差(root mean error, RME)和均方根差(root mean square error, RMSE)來評估績效。比較後認為 SVR 有較佳的績效。以上的方法主要是根據偵測器所收集的旅行時間樣本資料做分析，由於偵測器每三分鐘更新一次，所以造成資料有中斷或是錯誤的現象發生，因此將有效的資料收集後，再根據不同的估計方法預測旅行時間的長度。

除了以上的母數方法外，另外還有無母數方法以預測旅行時間，而無母數方法主要可以分為以下三種：1.類神經網路(ANN)方法，2.模擬方法，3.k-最近鄰(k-NN)方法。以下將分別介紹以上三種方法。

類神經網路法

Yoshikazu 等人(1998)[10] 利用架設在路段的 AVI 系統蒐集車輛資料，應用混合式類神經網路方法解釋每個路段的旅行時間與整個路徑的旅行時間之間的關係。而國內相關文獻有李季森(2001)[11]探討國內高速公路駕駛人變換車道行為與變換車道時間，進而研究於不同預測時間、流量、探針車混合比例與區段長度等相關參數之實驗組合，並透過類神經網路進行旅行時間之預測。張修榕(2001)[12]透過類神經網路模式來進行雙階段高速公路旅行時間之預測。針對感應線圈偵測器可蒐集車流速度及流量的特性，利用模擬的方式產生所需之交通資料並作驗證；接著是預測部分，採用倒傳遞(feed-forward back propagation)類神經網路模式來建立不同車流型態下之旅行時間預測模式。黃裕文(2003)[13]以微觀的角度探討國內高速公路施工路段的車流變化，同樣以上述的方法建立旅行時間預測模式。溫志元(2002)[14]係針對高速公路進口匝道匯流路段之變換車道行為與加速車道變換車道匯入主線行為動機與條件進行界定，透過類神經網路進行旅行時間預測。此外，路段線形(Road Profile)亦可能造成旅行時間之推算誤差，林士傑(2001)[15]以中華顧問工程司交通千里眼(e-traffic)所提供之即時交通播報資訊，再加上高速公路幾何、交通量調查與客運車輛 GPS 等資料，運用類神經網路準確預測高速公路旅行時間，來供用路人參考以降低不確定性。鑒於國內偵測器普遍設置不足，吳佳峰等人(2001)[16]希望透過 GPS 車輛歷史旅行資料預估車輛旅行時間，為了能夠正確預估車輛旅行時間，該研究設定了車輛運行路線分段以及車輛歷史旅行資料劃分時段之準則。近年來有許多研究利用多項偵測單元進行資料融合，藉以提升旅行時間推算之準確率，如李穎(2002)[17]融合國道客運班車 GPS 資料、車輛偵測器資料、事件資料等真實資料，以類神經網路法尋找各項資料來源其參數與旅行時間之關係。張慶麟等人(2002)[18]以 AVI 方式針對高速公路平常日之車流情形，先行應用車流模擬方式考量不同資料輸出時距、偵測器佈設間距及 AVI 辨識率等產生相關資料，配合簡單指數平滑法、Holt's 指數平滑法、ARIMA 模式及倒傳遞神經網路構建四種旅行時間預測模式，分別進行預測績效分析。魏健宏等人(2003)[19]收集探針車、道路偵測器以及事件等資訊，利用類神經網路中的倒傳遞神經網路進行實證評析。

整理相關之研究後，類神經網路模式之構建類型可以分成六類如下表說明。

表 2.1 類神經網路模式之構建類型整理

模式名稱	模式特性	模式構建概念	輸入變數類型
ANN1 時間序列模式	時間序列模式	利用歷史旅行時間資料預測旅行時間	不同時期之旅行時間
ANN2 時間序列+ 交通特性模式	時間序列+ 交通特性模式	由於歷史旅行時間資料間隔過大，較不能即時、準確地預測目前的旅行時間，因此加入一交通特性之變數，以增加預測績效	不同時間之旅行時間、交通特性變數(流量、速度、時間)
ANN3 車輛偵測器	路段與速度模式	由於速度與路段長度已知，則可計算出各偵測器間之旅行時間，若計算頻率愈高，準確率愈高	車道變化路段數、各路段之長度、車道數與車輛偵測器速度
ANN4 客運車輛 GPS 資料	利用客運車輛 GPS 資料及道路幾何特性	由客運車輛之 GPS 資料進行旅行時間預測。	客運車輛 GPS 速度、車道變化路段數、各路段之長度、車道數與 GPS 位置
ANN5 總體資料融合 時間考量模式	1. 考慮當時之交通資訊 2. 考慮當時及推估一段時間推移後之交通資訊	1. 考慮客運 GPS、車輛偵測器、事件、道路幾何等因素以構建預測模式 2. 利用以上資料推估衍生之樣本進行模式訓練、測試	偵測器資料、事件資料、客運車輛 GPS 資料、交通量資料(主線與匝道)、幾何特性資料、匝道型態、日期時間、車輛出發時間等資料
ANN6 總體資料融合 空間考量模式	ANN5 資訊+ 上游路段交通 狀況建立模式	利用客運 GPS 資料、車輛偵測器、事件、道路幾何特性，再加上上游路段交通狀況一併考量以構建模式	ANN5+上游路段群組之客運及車輛偵測器資訊

[資料來源：魏健宏等人(2003)[19]]

模擬方法

模擬方法為利用軟體模擬實際交通的狀況，主要可分為巨觀粒子模擬及微觀粒子模擬。

1. 巨觀粒子

Chang 等人(1994)[20]建構一套巨觀粒子模擬(MPSM)系統，分別採用 MPSM，修正的 MPSM(M-MPSM)，和微觀(micro)三種車流模擬模型，用於先進交通管理系統的應用，並以 MPSM 模擬高速公路上的車流，M-MPSM 模擬都市街道車流系統，而微觀模型則用來針對已產生擁塞的街道模擬，則可以模擬結果推導出包含旅行時間等各項車流系統資訊。Johnston 等人(1999)[21]建構一套巨觀的車流模擬系統，以平行計算的方式來運作，並將明尼阿波理斯市(Minneapolis)的公路路網的車流資訊代入以驗證此模型的準確度與計算模擬速度。Kachani 等人(2001)[22]利用巨觀模式將車流視為液體流動，使二階多項式旅行時間(PTT)和指數旅行時間(ETT)等模式來模擬出駕駛對上游的擁塞與否所產生的反應行為與路段密度或鄰近路段密度所造成的車流效應這兩種現象，藉以估計旅行時間，可增加等候線來將此類模型用於使用者動態均衡的問題上。

2. 微觀粒子

Kiesling 等人(2005)[23]提出微觀的車流模擬系統平行時間(time-parallel)模擬的方式將各節點時間區隔成好幾個區段(interval)，此模擬方法有兩個優點，其一為即使不存在適當的變數仍可以使用此模擬方法來模擬，其二為每節點可以單獨模擬出其時間間距，而不需要考慮其它節點。經過驗證此模型後得到，在低密度時使用有較佳的適用性。

而本篇文章主要針對 k -NN 方法進行研究，因此將該文獻回顧獨立於 2.2 小節進行討論。

2.2 k -NN 預測交通資訊相關文獻

最近鄰分群法(Nearest Neighbor classification)是一個廣為使用的一套分群方法，在已知分群的數量下，將類似的樣本分類至同一分群中。

k 最近鄰法(k -Nearest Neighbor Method, k -NN Method)並非為一套分群

方法，而是一種利用比較歷史資料與即時資料，找出擁有相同特性的資料的方法。最初是由 Benedetti(1977)[24]、Stone(1977)[25]及 Tukey(1977)[26]這些學者提出了近鄰法(nearest neighbor method)的概念發展而來。這些學者利用一元位置估計(Univariate Location Estimators)，建立了無母數迴歸式的模型，進而引申出最近鄰法的概念，可以利用輸入的資訊找尋出與資料庫中最相似的資料。Altman(1992)[27]將以上的研究進行整理，將一元位置估計引申至多元位置估計(Multivariate Location Estimators)並且提出了 k -NN 法，該方法將歷史資料區分群，再將輸入資料與之前的歷史資料相比對，比對出該資料與那一個分群較為類似，即可用歷史資料推估資料。

而 Gary 等人(1991)[28]利用迴歸模型的概念，延伸出無母數迴歸的模型以預測資訊。首先利用一階自迴歸模型(first-order autoregressive system)整理出如何在利用單變數建立模型 $x'(t) = b \cdot x(t) + e(t)$ ，並利用該公式以推估資料。

其中 b 代表自迴歸之相關係數，而 $e(t)$ 代表獨立、常態分佈且平均為零之隨機變數。在利用最小均方差(minimum mean-squared error, MMSE)的方法求出自變數與應變數間最接近的解以預測資料，發現當 $e(t) = \text{平均數} = 0$ 時，可以得到最小均方差的模型 $\hat{x}'(t) = b \cdot x(t)$ 可以利用自變數而得到最佳的預測。然而在多變數下，由於 $e(t)$ 的平均值受多變數的影響而不再為零，因此將由最小均方差的方法為找出與實際值最接近的歷史資料做為預測值，公式如下：Minimize $Q = (x'(t) - x(t))^2 = (x'(t) - x(t)) \cdot (x'(t) - x(t))'$

因此 k -NN 的方法均試著找出最小的平方差做為預測結果。

Smith、Demetsky(1997)[29]對 k -NN 法進行績效評估，分析比較以下四種交通流量的預測方法：歷史平均法，時間序列法、類神經網路法與 k -NN 法。評估的方法為利用歷史的流量資料預測未來的流量再和實際值做比對，判斷那一個方法的績效較差，結果發現當歷史資料大時， k -NN 法所預測的誤差結果比其他三種方法所得到的結果誤差來得小，因此利用 k -NN 針對流量進行預測是可行的，比較誤差率的結果為迴歸 9.57%，時間序列為 9.03%，類神為經 8.93 而 k -NN 為 7.54%。因此 k -NN 法有較佳的預測能力。Smith、Williams 以及 Oswald(2002)[30]將權重(Weight)的想法加入 k -NN 的模型，考慮目前的流量可能會與先前的交通流量有關係，但是隨著時間距離的增加，其影響的程度愈小，所以利用權重的方法進行修正，將不同時間所造成的影響納入考慮。

Clark(2003)[31]除了流量外，試著利用 k -NN 法試著針對其他可蒐集到的交通資訊進行預測，如速度和佔有率。接著進行一一分析和交叉分析，結果發現當同時利用流量、佔有率和速率這三項變數進行分析時，所得到的預測值相對於個別比對而言，誤差有下降的現象，所以愈多變數進行討論可以得到更精確的結論。接著 Rice、Zwet(2004)[10]希望可以預測出旅行時間，利用流量、佔有率推估該路段之速率，並由該速率資訊得到旅行時間資料庫。因此可以由估計得之旅行時間應用 k -NN 法以推估下一時段旅行時間。

Robinson、Polak(2005)[32]提出了四點建立 k -NN 模型時應該先設定的條件。

1. 決定模式中的特徵向量。
2. 利用加權法減少 k -NN 模式的誤差。
3. 設定距離量度(distance metric)。
4. 決定判斷類似樣本的數量。

Chang 等人(2006)[33]提出 k -NN 法需要加入門檻以提升預測績效，加入的門檻為以下兩種類型。

1. 交通資訊門檻

設定交通資訊的上下界門檻，當交通擁擠或流量很少下，所對應的交通情形相差不大，因此當交通資訊超過上下界時，即以上下界資訊為當時之資訊。

2. 比對時間門檻

當道路開始擁塞或擁塞逐漸解除期間，可能會對應著相同的交通資訊，然而其代表的意義並不相同，因此這兩種期間所對應的旅行時間也不會相同。故為了避免進行比對時，將這兩種情形都同時考慮，因此加入了時間的門檻，避免在比對時發現早上的交通資訊與某一歷史傍晚的交通資訊相類似而提高預測的誤差。

2.3 小結

本研究整理旅行時間預測的相關研究，大致可以分為以下五種方法：迴歸分析法、時間序列法、類神經網路法、模擬法及 k -NN 方法。

其中迴歸及時間序列法根據不同的特徵建立不同的模型後，再利用校估所得的參數進行預測，然而這些參數隨著不同的特徵而改變，因此在外在環境的改變下，就需要重新校估參數以維持其準確率。

類神經網路法收集各種交通資訊，再依照不同資料來源、各輸入變數之組合以調整模式架構，進行訓練與測試，以得到變數與旅行時間之間的關係，因此在模式的訓練與測試上，隨著變數的增加，需要較長的運算時間。而模擬法，則是利用當時的交通狀況以及車流模式，模擬未來的交通狀況，以推得旅行時間，因此隨著路網複雜度的增加，會需要較長的計算時間。

而根據 k -NN 相關研究發現，除模擬方法外，在比較 ANN、迴歸方法、時間序列方法及 k -NN 法後， k -NN 方法在預測交通資訊上(速度、流量)下，可以有相當好的績效。因此本研究將試著研究是否在旅行時間預測上也有足夠的準確率。而目前旅行時間預測上，大部分的文獻均以模擬方法、ANN、迴歸方法、時間序列方法及 k -NN 法進行預測。本篇研究將試著利用 k -NN 方法進行旅行時間的預測，並且將該方法應用於實際的案例進行分析，最後試著利用不同的分群以找出較佳的預測結果。

第三章 資料處理

本研究希望可以利用即時的交通資料找出歷史資料和即時資料之間的關係以預測旅行時間，因此將以 k -NN 做為資料分析的方法先找出那一些歷史資料與即時資料相類似。目前本研究的資料來源，主要來由於高速公路的偵測器資料及電子收費系統(ETC)的通過車輛資訊，以下將說明這兩項資料之型態及資料處理過程。

3.1 偵測器資料與 ETC 資料的比較

已知偵測器資料屬於速率及流量之道路資訊，而 ETC 資料為旅行時間之資訊，其中 ETC 資料為經過比對過濾而得之旅行時間資料，而流量資料為原始資料，以下將從民國 97 年 1 月中選出星期一到星期天各一天，從龍潭收費站到樹林收費站間北上的流量資訊及旅行時間資訊進行比較。



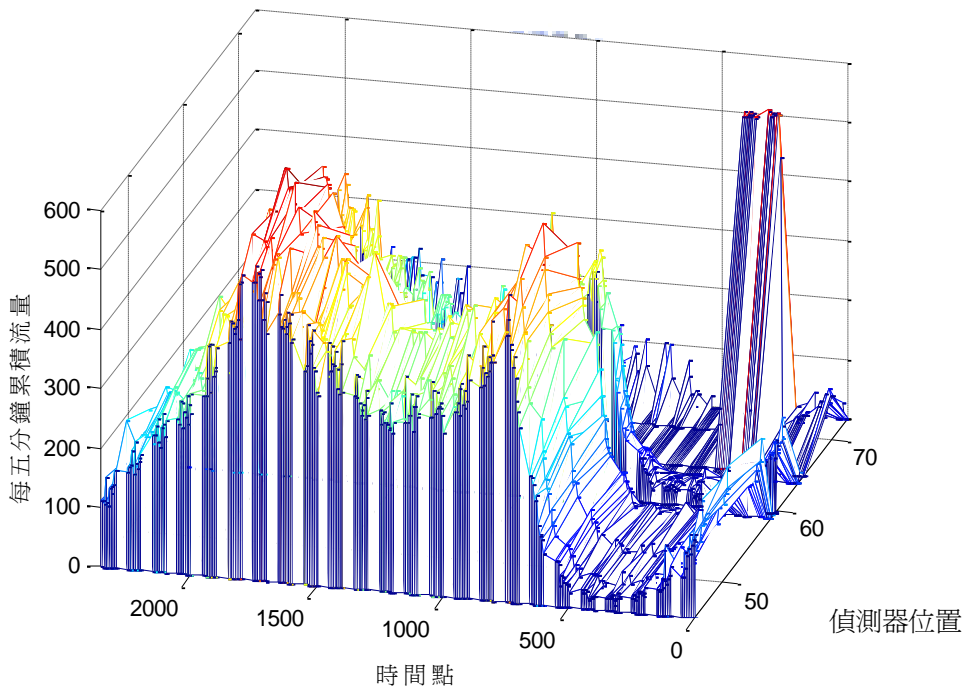
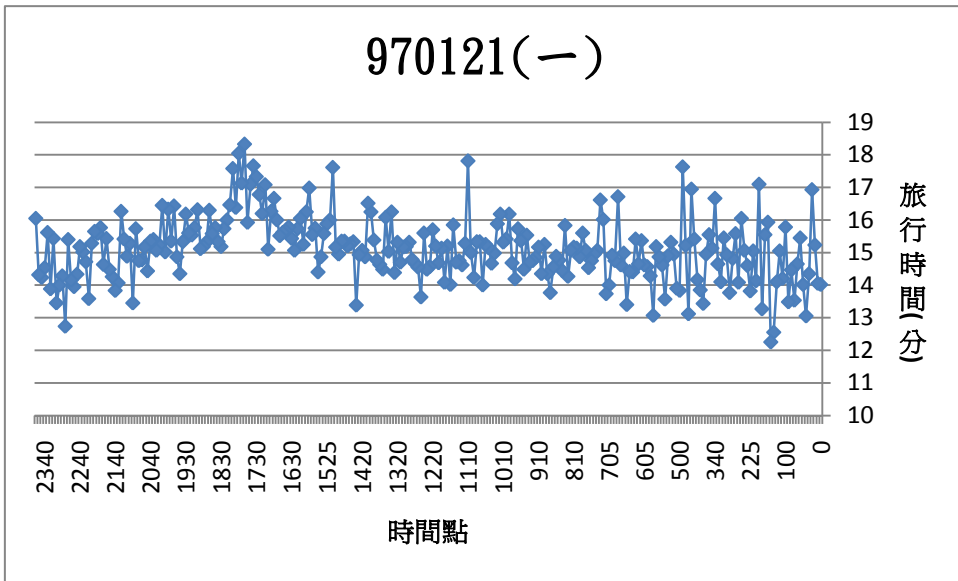


圖 3.1 970121(一)流量與旅行時間對應圖

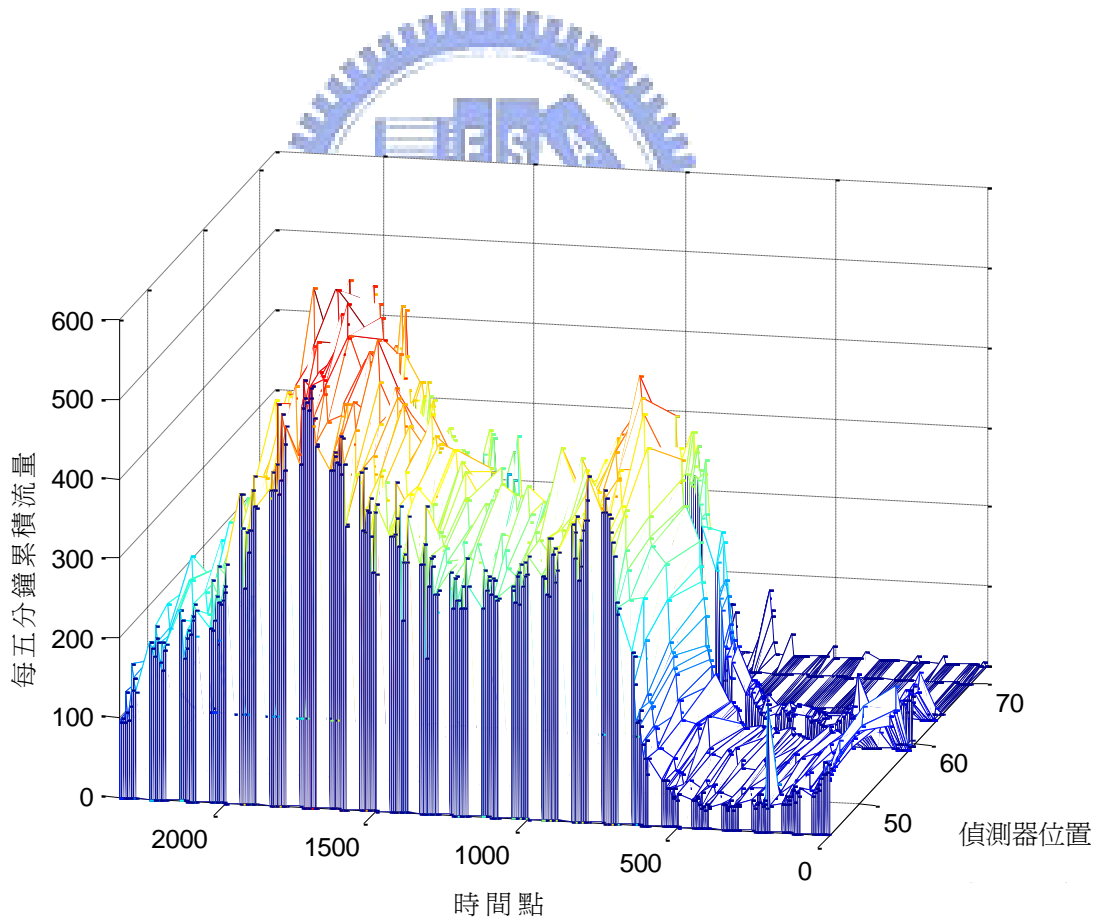
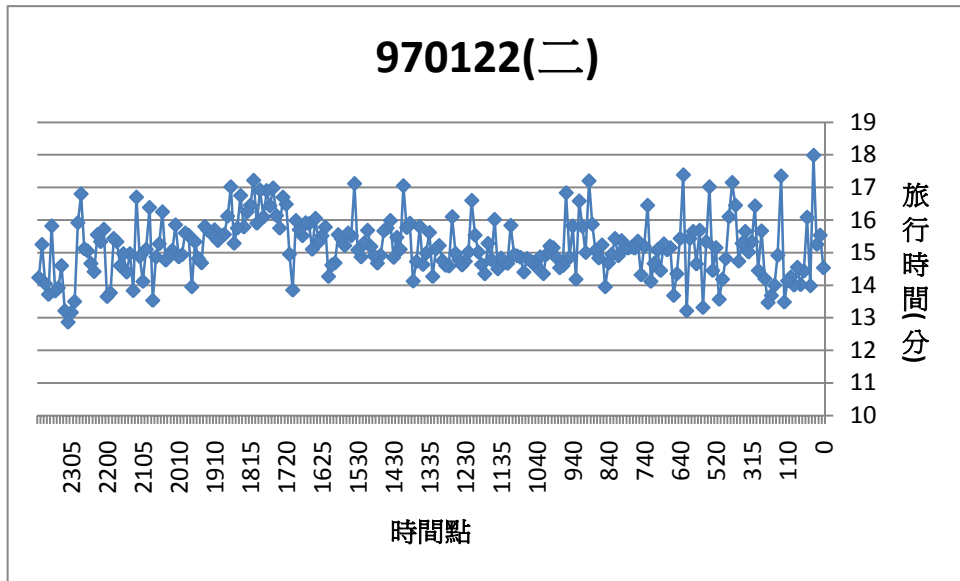


圖 3.2 970122(二)流量與旅行時間對應圖

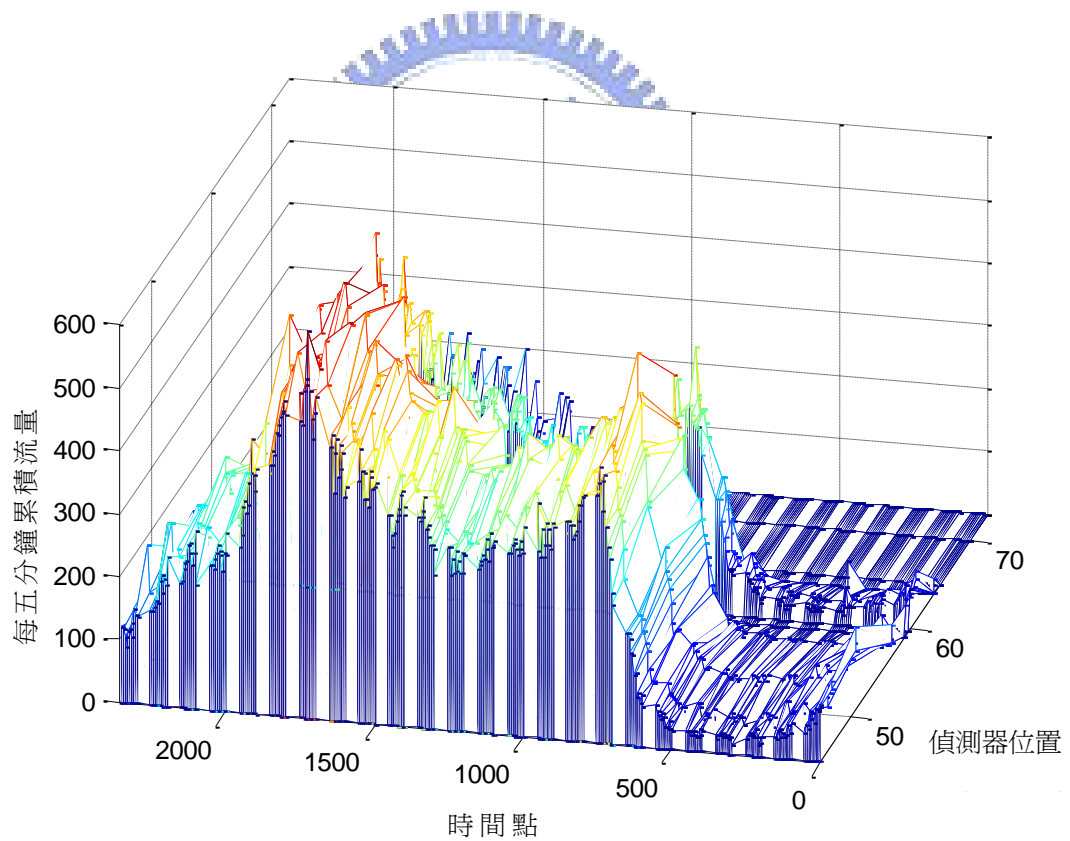
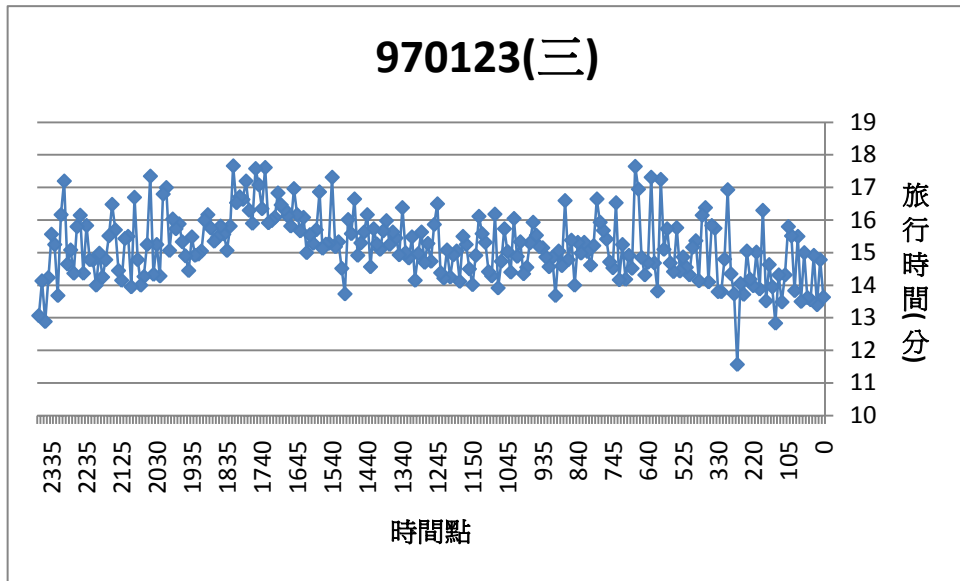


圖 3.3 970123(三)流量與旅行時間對應圖

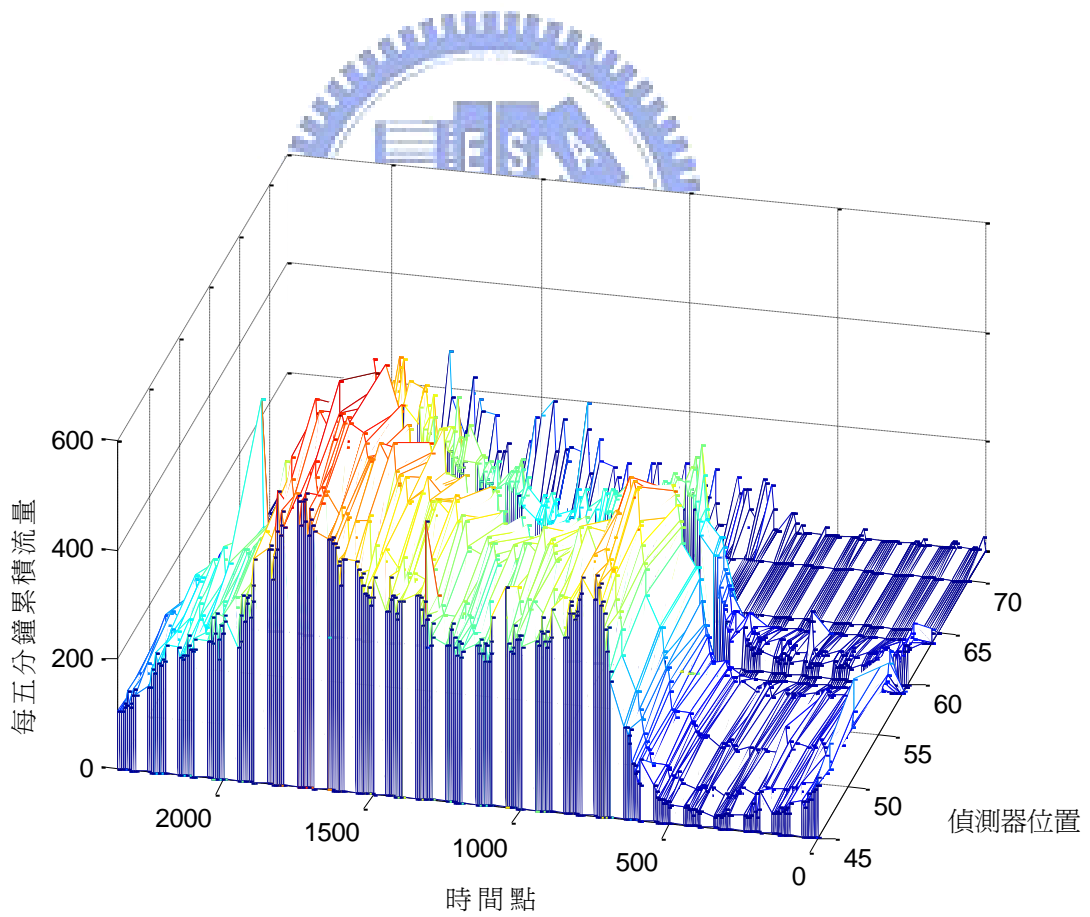
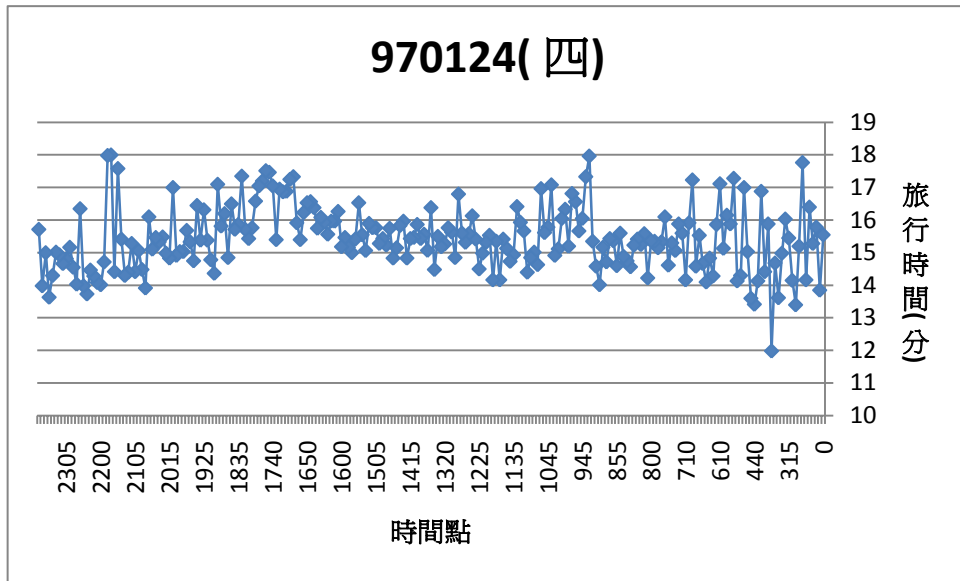


圖 3.4 970124(四)流量與旅行時間對應圖

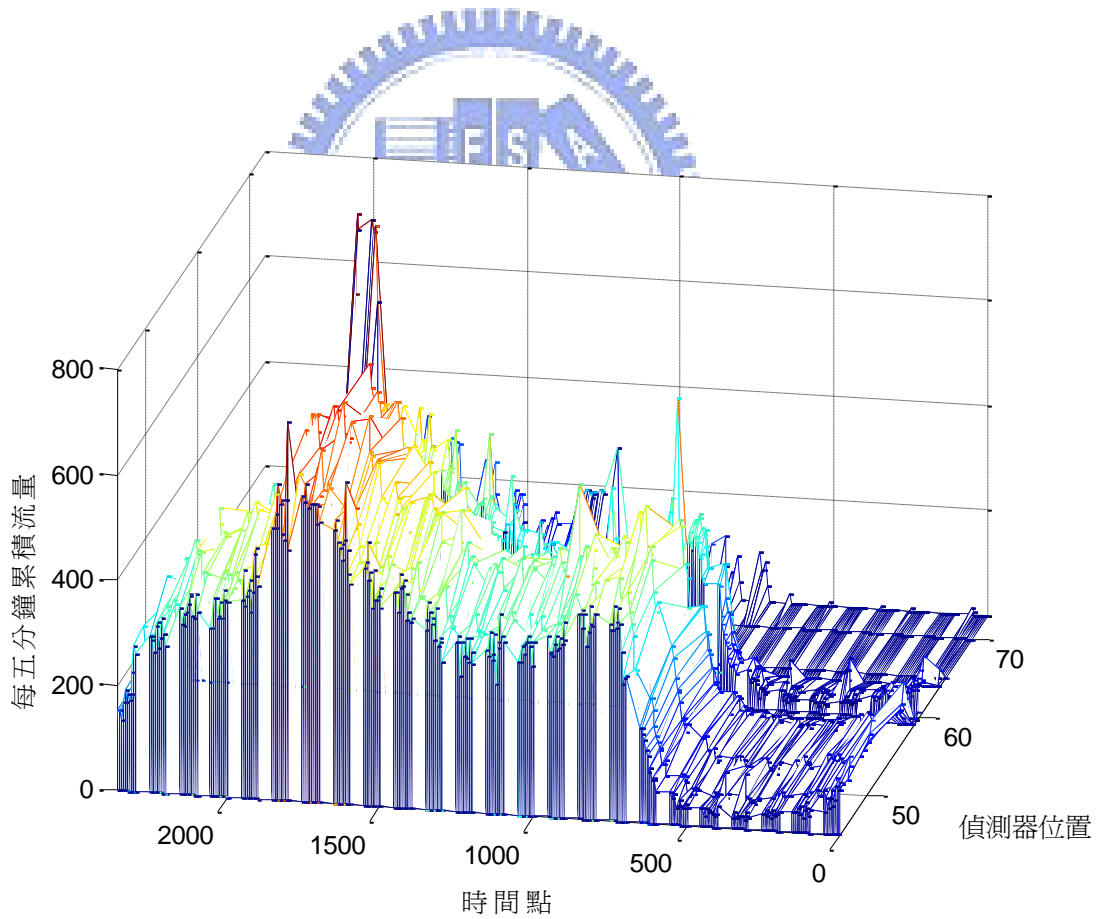
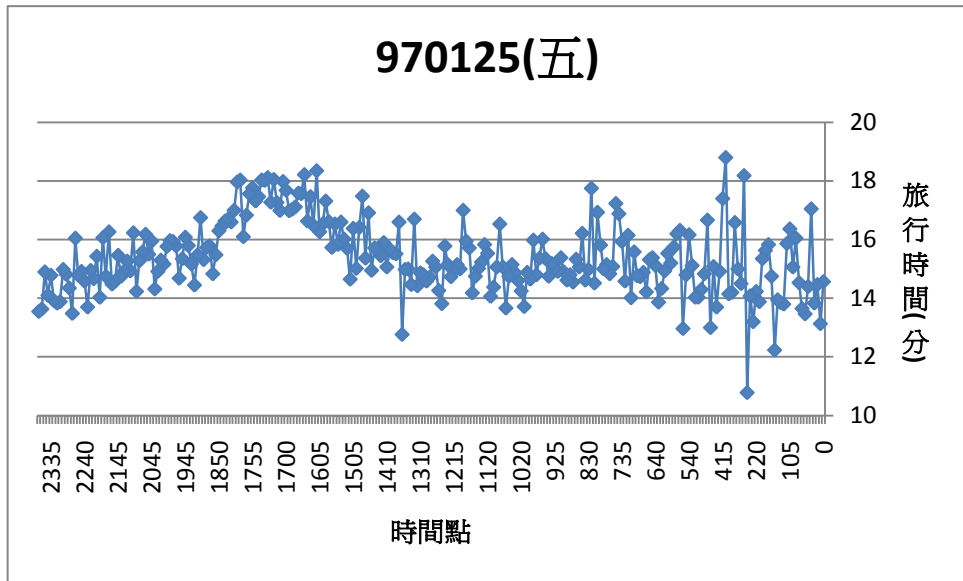


圖 3.5 970125(五)流量與旅行時間對應圖

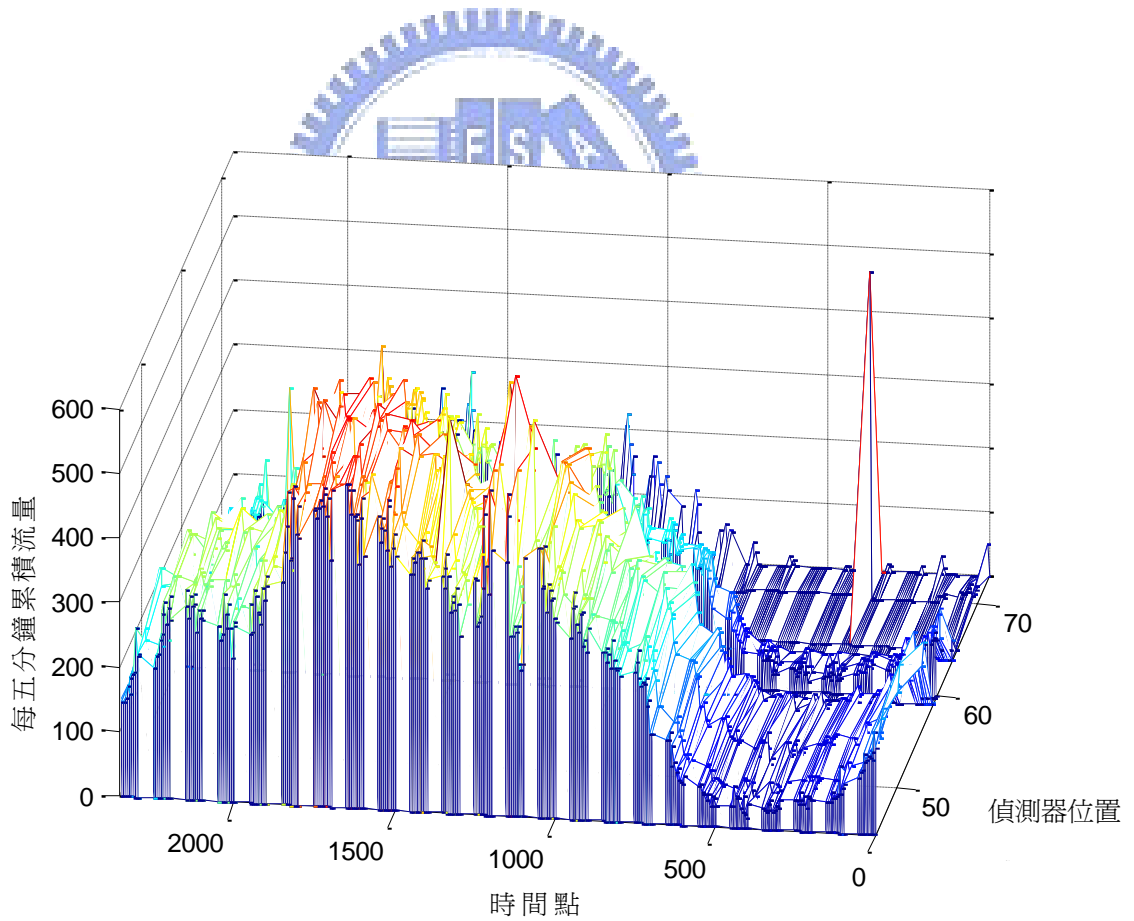
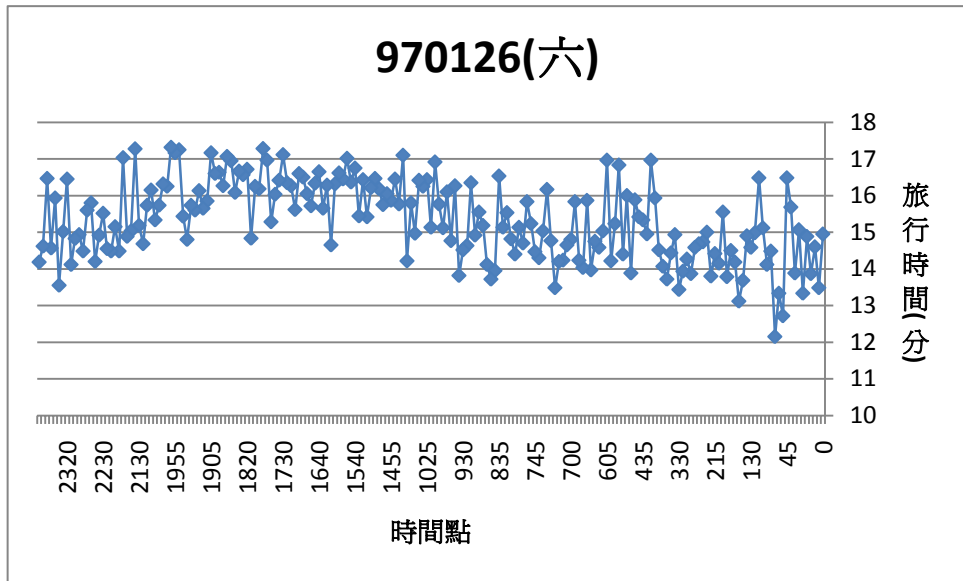


圖 3.6 970126(六)流量與旅行時間對應圖

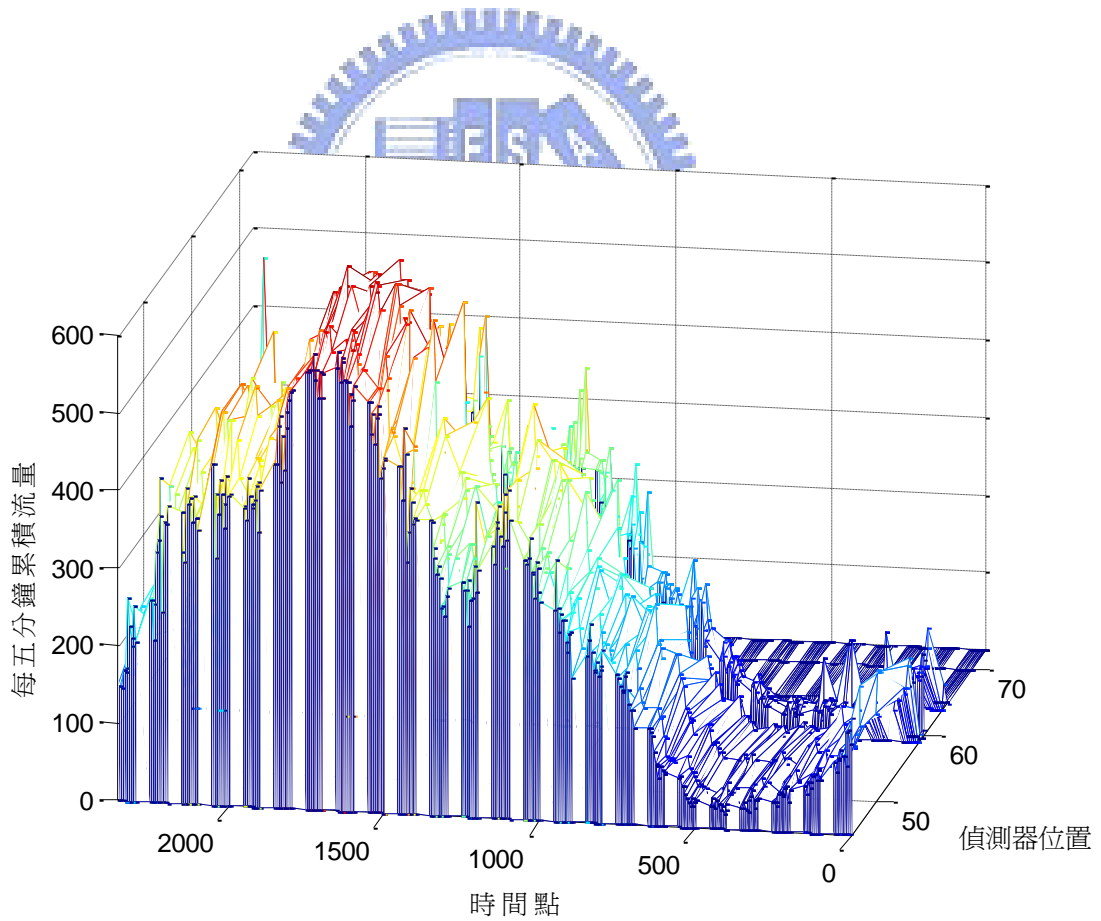
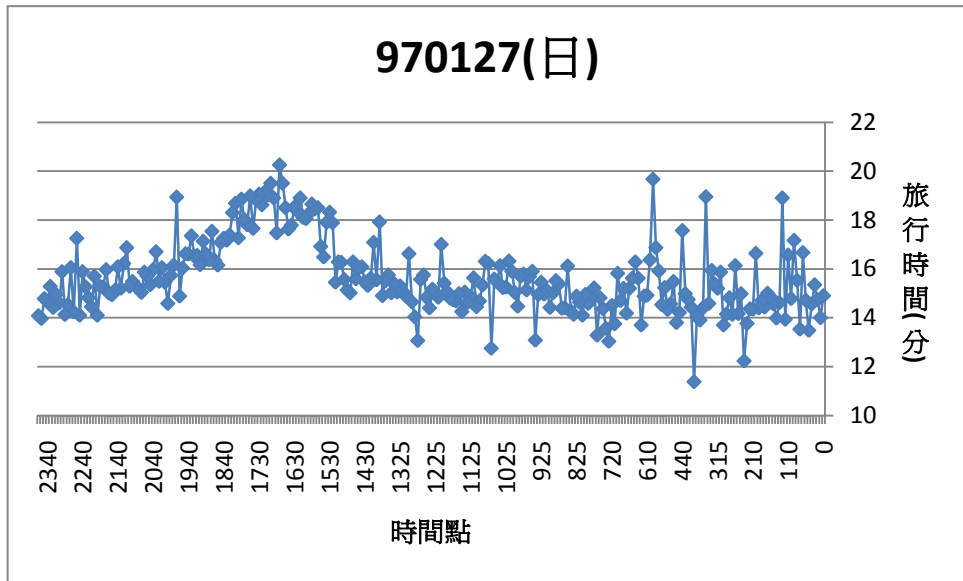


圖 3.7 970127(日)流量與旅行時間對應圖

由以上這幾張圖可以發現以下幾點

A. 當旅行時間較長時，流量資料也有比較大的現象。

流量愈高代表有愈多的車輛使用該路段，因此道路負荷增加而旅行時間增加，因此流量和旅行時間資料之間有關係存在。

B. 當時間點較早時，流量資料較小，旅行時間的變異也比較大。

由於流量較少，因此當有一些車速度較慢，就會造成旅行時間有過慢的現象發生，因此本研究會再找尋方法再進一步將這些資料進行過濾。

C. 偵測器的流量經常出現 0 的情形。

當偵測器出現該情形，有時是出現錯誤資訊(“E”)，有時是提供流量為 0 的資料，然而無法直接判斷是否是沒有車輛經過，抑或是產生了錯誤資訊。因此本研究需要將某些出現過多異常值之偵測器進行過濾。

D. 某些時間點，偵測器流量有過大的情形。

某幾個時間點有資料過大的現象，因此本研究會根據前後時間點的流量判斷是否有異常的現象發生，因此依照各時間點平均流量的 95% 信賴區間進行過濾，如果發現資料在一個或兩個連續時間點有缺漏的現象發生，則利用前後時間點的平均資料進行插補，如果有過長一段時間沒有資料，較視該段時間點沒有資料。

以下兩節將分別針對偵測器資料及收費站資料之處理進行說明及介紹。

3.2 高速公路偵測器資料

目前高速公路偵測器資料主要是在記錄每五分鐘所通過的大型車或小型車的車輛總數、平均通過速率以及佔有率這三項資訊。本研究蒐集從樹林收費站經龍潭收費站到後龍收費站間所有偵測器，以民國 96 年每五分鐘的平均速率及流量資料做為歷史資料。然後利用 97 年的交通資料與歷史資料進行比對，希望可以找出與即時資料相類似的歷史資料。而本研究所要分析的是在半個小時內，交通狀況的變化情形，比較在各個時間點比對前半小時的交通情形下，速率及流量是否有增加或減少的趨勢。最後可以得到各個時間點下的交通資訊趨勢，即為本研究所蒐集進行分析的資料。=

其中樹林到後龍之間，主要有偵測器北上及南下各 24 及 22 組，本研究先以出現錯誤資訊”E”之情形進行統計，發現有相當多缺漏值的出現，本研究將這些出現過多缺漏值之偵測器忽略不計，因此最後本研究將採用 12 組北上及 12 組南下偵測器，以下將分別列出本研究不採用偵測器之缺漏率，及採用偵測器之缺漏率，如表 3.1 及 3.2 所示。其中偵測器名稱代表不同位置之偵測器，如 2s37 即代表國道三號北上路段里程數 37 公里之偵測器。

表 3.1 缺漏嚴重之偵測器代碼及缺漏率

偵測器名稱 (南下)	缺漏率	偵測器名稱 (北上)	缺漏率
2s37	36%	2n59.1	54%
2s41	22%	2n59.8	96%
2s48	13%	2n61	21%
2s58	13%	2n65	58%
2s59.1	12%	2n70	86%
2s71	100%	2n77	46%
2s81	38%	2n80	30%
2s94	61%	2n86	36%

[資料來源：本研究整理]

表 3.2 研究範圍之偵測器代碼及缺漏率

偵測器名稱 (南下)	缺漏率	偵測器名稱(北上)	缺漏率
2s44	16%	2n43	10%
2s49	6%	2n45	10%
2s52	7%	2n47	19%
2s59.6	7%	2n51	11%
2s60	25%	2n57	8%
2s62	21%	2n60	7%
2s67	25%	2n64	25%
2s75	28%	2n73	27%
2s77	30%	2n75	32%
2s80	24%	2n81	26%
2s86	29%	2n93	35%
2s96	24%	2n96	31%

[資料來源：本研究整理]

由於偵測器資料所收集的資料有時和前五分鐘及後五分鐘資料差異過大，有時可能是剛好沒有車輛通過，故流量為零，或是其他的現象。因此本研究將收集 96 年全年的資料視為母體，求得各偵測器每五分鐘速率分佈在 99% 信賴區間。以該信賴區間為資料分佈之上下界將離群值進行過濾。最後再利用該時間點之前後五分鐘資料進行插補，而得到該時間點較合理的交通資訊。本研究所採用之國道三號偵測器信賴區間圖如附錄一所示。

3.3 電子收費系統資料

電子收費系統(ETC)為臺灣 ITS 產業的重要成就，大幅降低了車輛通過收費站的時間成本及收費站營運的人事成本。本研究利用遠通電收所匯整的 ETC 資料，蒐集每一台安裝 ETC 車機之車輛通過各收費站的時間，若車輛通過兩個以上的收費站，進行處理後即可得到這些車輛通過兩收費站間所需要的旅行時間。由於高速公路偵測器的資料是每五分鐘收集一次，因此本研究將整理每五分鐘通過兩兩收費者所需要的平均旅行時間。

然而某一些車輛可能開的過慢，或是中途有離開交流道，因此造成該平均旅行時間過長，因此本研究將收集 96 年通過兩兩收費站間所有車輛資料視為母體，求得各車輛通過兩兩收費站之旅行時間分佈的 99% 信賴區間。若有車輛之旅行時間在該信賴區間外，則視該車為離群值進行過濾，最後可推估較可靠的歷史旅行時間。而國道三號龍潭到樹林北上之旅行時間信賴區間圖如下圖所示，其他日期之比較結果如附錄二所示。

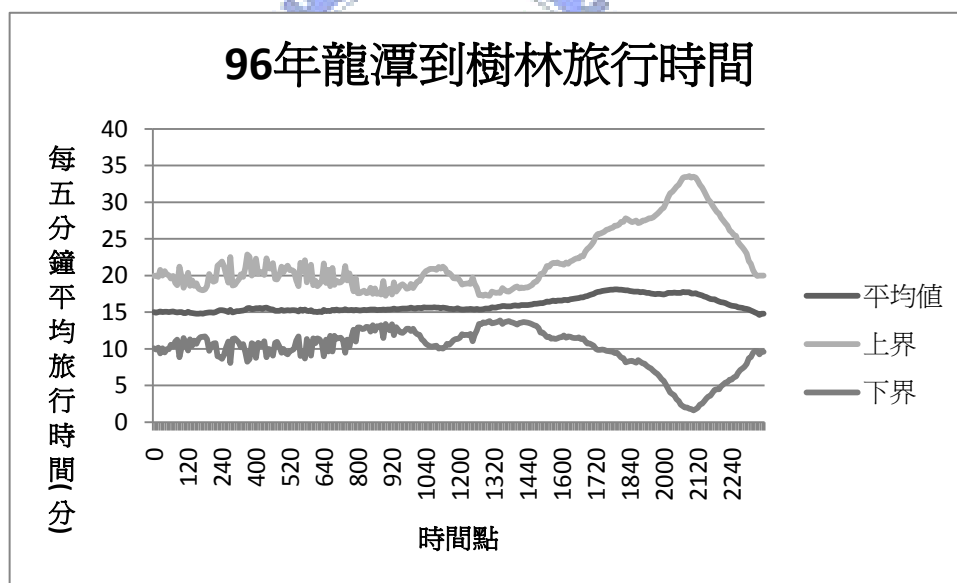


圖 3.8 96 年歷史旅行時間信賴區間圖

[資料來源：本研究整理]

本研究的研究範圍為國道三號從樹林收費站至後龍收費站之間的旅行時間，在實際調查後發現車輛同時通過樹林、龍潭、後龍，三個收費站的旅行時間比通過兩兩收費站來的少，而且相當多車輛通過樹林及龍潭收費站，因此本研究將收集從樹林到龍潭的歷史旅行時間加上龍潭到後龍的歷史旅行時間，即為樹林到後龍的歷史旅行時間。而收集兩兩收費站的資料以推得三個收費站間的旅行時間，可以得到較多的歷史資訊。舉例說明，當本研究欲蒐集早上 7:00 從樹林出發到達後龍收費站南下的歷史旅行時間為何，資料獲得的方法如下：

1. 蒐集某一天早上 7:00~早上 7:05 通過樹林收費站南下的車輛資訊。
2. 得到這些車輛平均通過龍潭收費站的旅行時間，若需要 33 分鐘。
3. 再蒐集當天早上 7:30~7:35 這五分鐘通過樹林收費站的所有車輛資訊。
4. 得到這些車輛平均通過樹林收費站至後龍收費站所需要的旅行時間，若為 32 分鐘。
5. 將樹林到龍潭以及龍潭到後龍的旅行時間加總，即可判斷某一天早上 7:00 從樹林收費站出發，到達後龍收費站所需要的旅行時間為 65 分鐘。
6. 最後比照以上方法收集各時段及南北向之資料，可以得到當天每五分鐘從樹林(後龍)出發，到達後龍(樹林)所需要的旅行時間。

旅行時間估計流程如下圖所示。

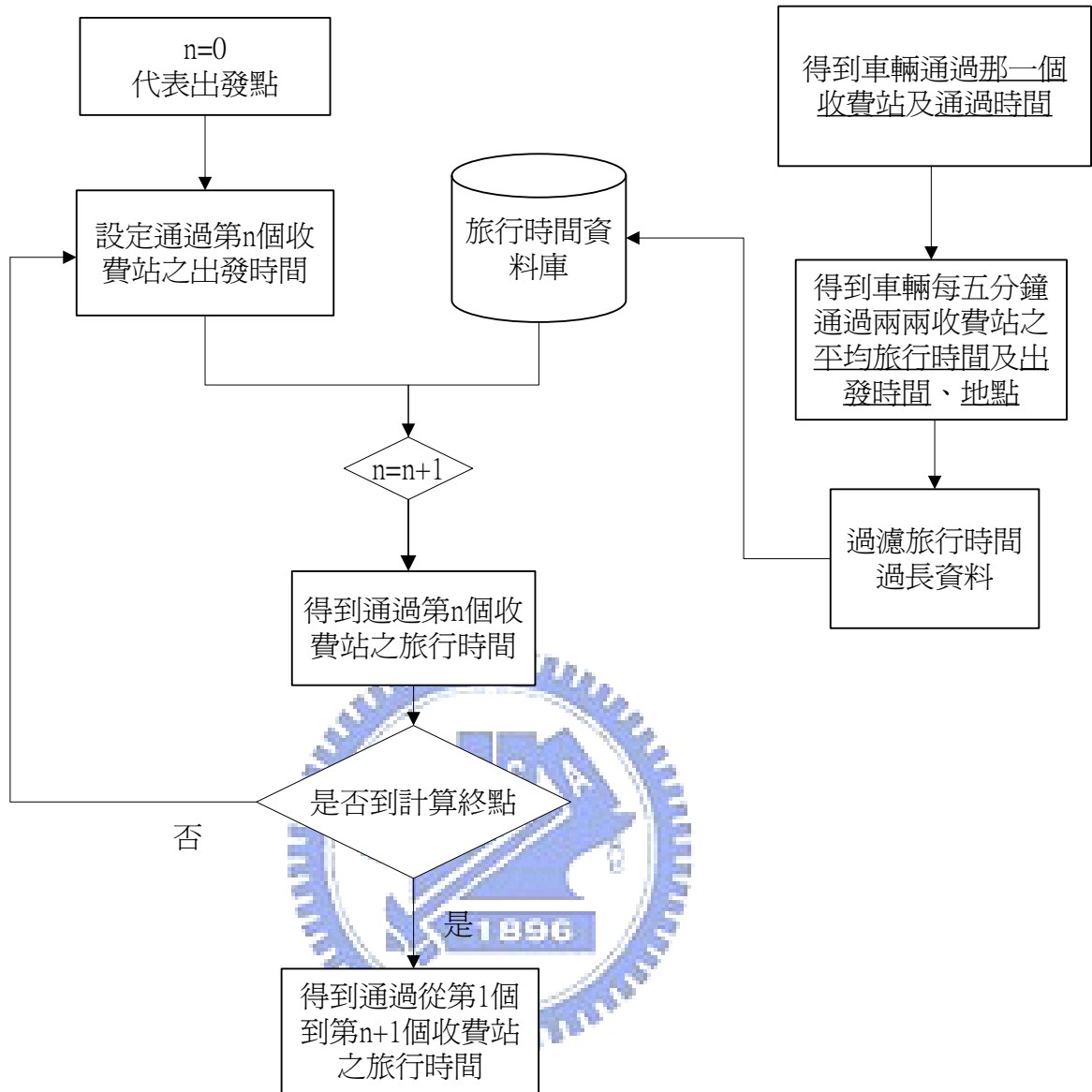


圖 3.9 歷史旅行時間估計流程圖

[資料來源：本研究整理]

3.4 ETC 與 VD 之間關係

本研究的目的是為利用即時偵測器資訊與歷史偵測器資訊進行比對，再利用 ETC 所估計之旅行時間進行預測，以下將討論 ETC 與交通資訊的關係。利用歷史偵測器資訊推估當時之旅行時間以及 ETC 所推估之旅行時間進行比較，觀察兩者是否會反應相同之交通情形。

由於國道高速公路局在民國 98 年開始針對每分鐘一筆交通資訊進行收集與整理，因此本研究由速率資訊與偵測器位置推得各時間點下偵測器間之旅行時間，其推估的流程如下圖所示。

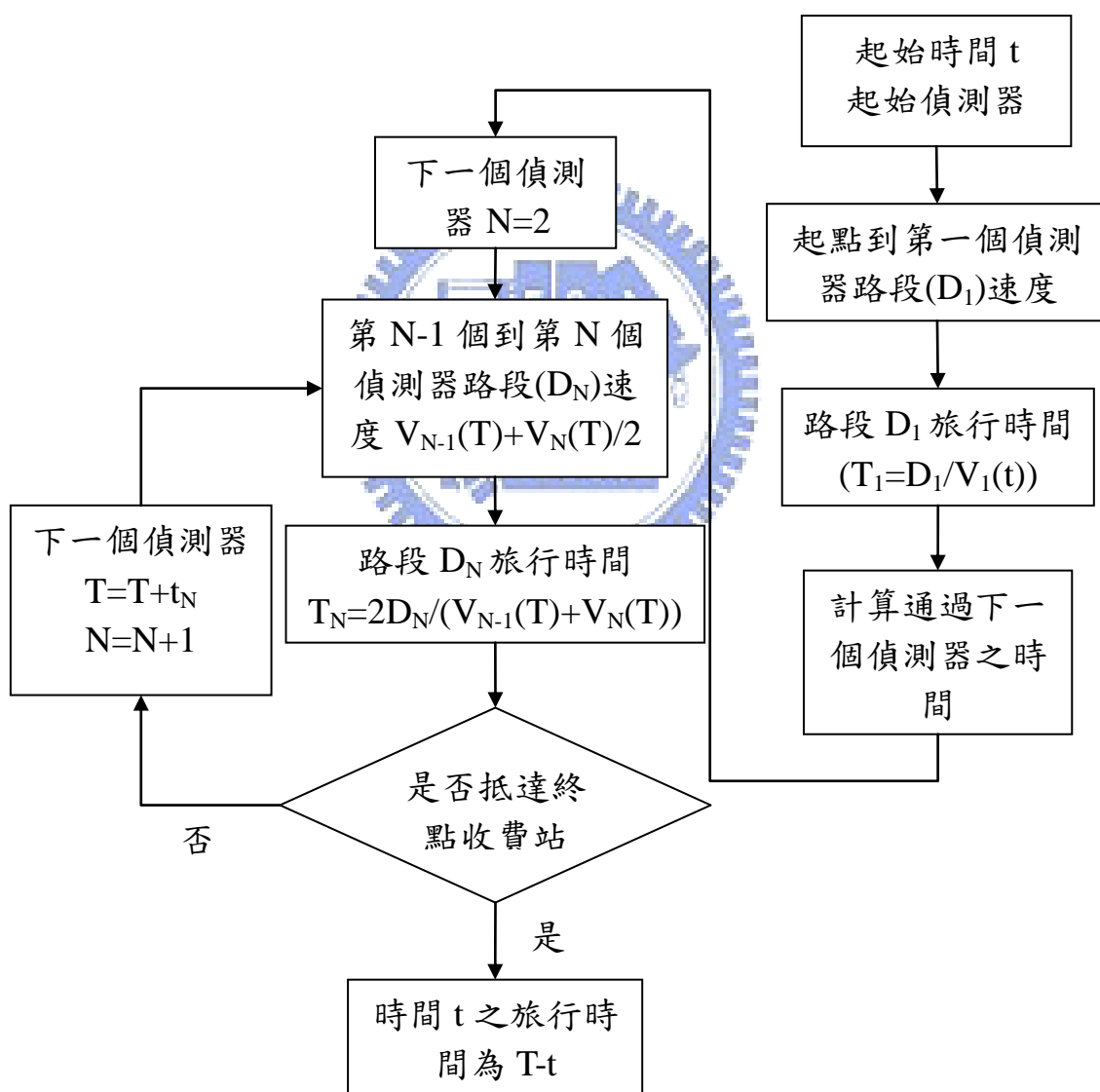


圖 3.10 偵測器推估歷史旅行時間流程圖
[資料來源：本研究整理]

本研究收集 5 月 12 日至 5 月 25 日之偵測器資訊及 ETC 資訊進行比較，以 5 月 14 日國道三號由樹林到龍潭為例，如下圖所示。由於電子收費不一定每分鐘都有車輛經過，故旅行時間仍以每五分鐘推估一次。

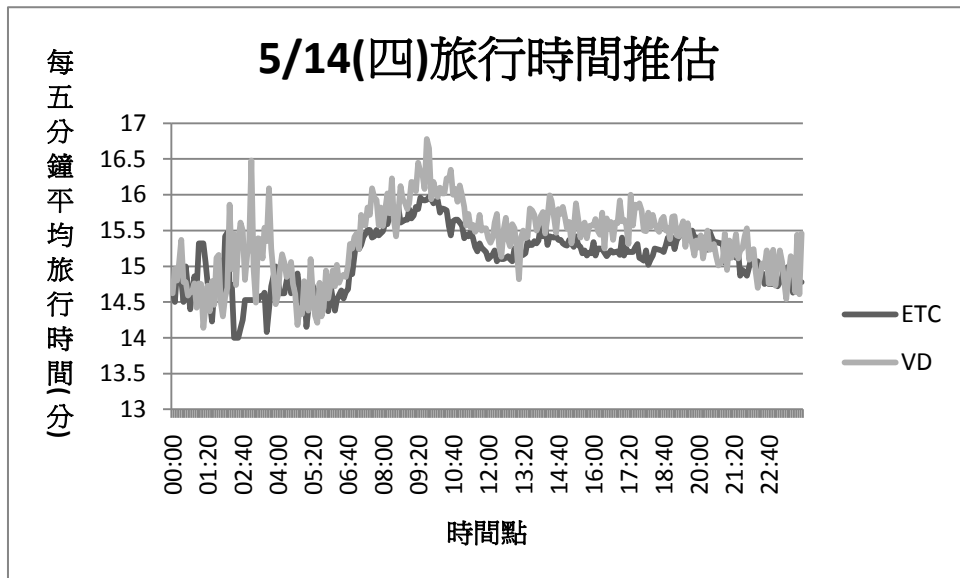


圖 3.11 偵測器與 VD 歷史旅行時間推估圖
[資料來源：本研究整理]

由上圖可以發現偵測器(VD)的變異較 ETC 來的大，由於偵測器速率的變異較大，所以推估出來的旅行時間變異也隨之較大，但是兩者有大約有相同的趨勢，因此兩者有類似的關係。

第四章 旅行時間預測模式介紹

本研究的目的是在於如何利用即時的交通資訊預測旅行時間，由於旅行時間會隨著交通情形的不同而有不同的現象，若道路較為擁塞，旅行時間也會隨之上升，因此本研究將利用 k -NN 找出交通資訊與旅行時間的關係。以下將針對 k -NN 模式進行介紹，並且說明模式所需要的參數。

4.1 k -NN 介紹

k 最近鄰法(k -nearest neighbor method)是一套利用無母數迴歸發展而成的方法，主要的貢獻在於由已知的歷史樣本中，找出與輸入值相似的樣本，本研究希望可以利用該方法以找尋與即時樣本相類似之歷史樣本。假設在相同的交通趨勢變化下會有相同的旅行時間，因此利用即時的交通資訊比對歷史交通資訊，可以利用本方法辨識出那些歷史時間下擁有類似的交通資訊，並由該歷史時間所對應的旅行時間預測目前的旅行時間。模式流程如下：

- 由目前的交通資訊(速度、流量)，找出類似的歷史資訊
- 觀察某路段各偵測器所得半小時的交通資訊變化情形
- 得到這些類似的歷史旅行時間，預測目前的旅行時間

4.2 k -NN 參數設定

由 2.2 節文獻回顧中，學者 Robinson 和 Polak(2005)將 k -NN 方法進行實作提出如果要建立一個準確的 k -NN 模式，首先要先決定以下這四個要素：1、決定模式中的特徵向量。2、決定判斷分群的樣本數量。3、利用加權法減少 k -NN 模式的誤差。4、設定距離量度(distance metric)。因此以下將針對這四點進行一步說明。

1. 決定模式中的特徵向量。

k -NN 的目標是在找尋最接近即時資訊的歷史資訊，本計畫利用偵測器即時可收集到的資訊(流量、速率)做為特徵向量，由這些特徵進行比對，找出與即時資訊相類似的歷史資訊。

2. 利用加權法減少 k -NN 模式的誤差。

由於本計畫將同時比對流量與速率資訊的差異，由於這兩個變數的單位不一致，所以需要再設定權重以減少因單位差異而造成的誤差。本研究希望在尖鋒時刻下仍具有相當的準確率，因此以尖鋒時刻通過各偵測器之平均流量與平均速率做為權重設定的判斷方式。

由於速率和流量單位的差異，而且平均流量較平均速率較大，因此流量的變化也隨之較大，因此若未設定權重，可能速率變化沒有流量的變化來的大，而忽略了交通異常的發生，並且也可能在類似交通樣本的判定上發生誤差，因此需要設定權重值。而根據 Clark(2003)所提出權重的設定方法，將利用尖鋒時刻的平均速率及流量進行標準化。假設在尖鋒時段下平均五分鐘流量約為每五分鐘 500 輛汽車，而平均速率約為每小時 70 公里。權重的設定為平均流量乘上 $1/500$ 而平均速率為乘上 $1/70$ ，將標準化後的歷史資訊與即時資訊進行比對，以找出較類似於即時資訊之歷史資訊。

舉例來說：若即時速度及流量為 80、200，而有兩筆歷史資訊，分別為 90、200 及 80、250。則差異值分別為 $(90-80)/70=1/7$ ， $(250-200)/500=1/10$ ，則第二筆歷史資訊較接近。

3. 設定距離量度(distance metric)。

將即時資訊與各分群資訊的差異加總，即可得到樣本與各分群的特徵差異距離，而距離量度為判斷即時資料與那一組特徵資料相類似的一種判斷指標，當距離量度小於某一個門檻值，則判斷該即時資訊與某一組歷史資料類似。然而距離量度的判斷目前並沒有一個很好的準則來對各種資料進行距離量度的判斷，在面對不一樣的環境條件，分析人員必須自行建立一套準則來做處理，因此隨著不同的特徵分群，將會採用不同的距離量度以判斷即時資料與那一個歷史資料類似。

本研究將挑選最佳的類似樣本數做為旅行時間預測的門檻，因此本研究將以該樣本數做為距離量度的門檻。

4. 決定判斷類似樣本的數量。

本計畫將比對每半個小時下每五分鐘通過各偵測器的流量變化和平均速度變化，再與歷史資料進行比對，比對出該半小時的交通狀況較近似於那些歷史資料，再從這些歷史資料中挑選最接近的 k 筆資料進行旅行時間預測。

根據 Fukunag[34]於 1973 年時提出 k -NN 於資料分群下，每一群最佳的數量，若該資料滿足常態分配，則各群最佳數量如下式所示。

$$k^* = \left[\frac{(n+2)^2 (n-2)^{2+n/2}}{\Gamma^{4/n} \left(\frac{n+2}{2}\right) n^{1+n/2} (n^2 - 6n + 16)} \right]^{n/n+4} \times N^{4/n+4}$$

n : 代表樣本維度
 N : 代表樣本總數

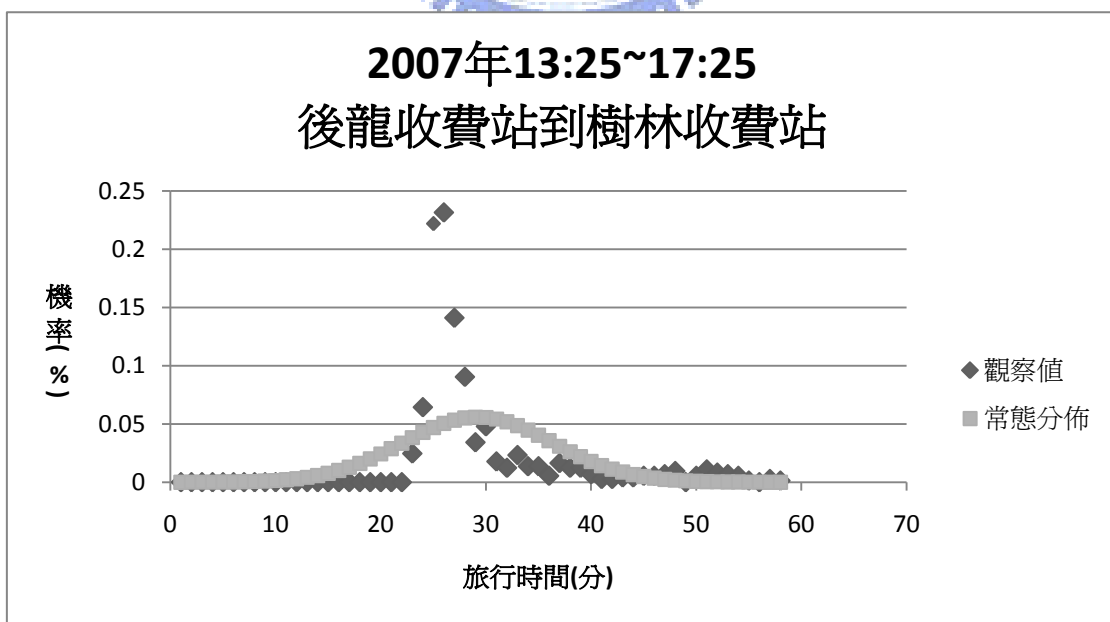
本研究將針對旅行時間分配進行驗證，判斷是否滿足常態分配。驗證各旅行時間的機率是否滿足常態分配，各旅行時間機率如圖 4.1 所示。本研究驗證方法為利用卡方適合度檢定進行驗證。驗證公式如下：

$$\sum_{i=1}^n \frac{(e_i - o_i)^2}{e_i}$$

,

其中 e 代表在常態分佈下，各旅行時間下有幾天有相同的旅行時間
 o 代表實際各旅行時間下，有幾天有相同的旅行時間
 n 代表總共有幾天，考慮 1 年的樣本，因此 $n=365$ 天

圖 4.1 旅行時間對應常態分配分佈圖



由上圖發現，在後龍到樹林間旅行時間大部分有 21 種的出現的可能 (21 分~42 分)

檢定的假設在 95% 信賴水準下結果如下所示：

H_0 : 滿足常態分佈

H_1 : 不滿足常態分佈

$$\sum_{i=1}^n \frac{(e_i - o_i)^2}{e_i} = 2107.87 > X_{(21-1)}^2 = 31.41$$

因此旅行時間的分佈拒絕常態分佈(H_0)的假設。

由於 Robinson 和 Polak[31](2005)比較不同的類似樣本數進行比較，挑選在平均絕對誤差率(MAPE)及均方根誤差(RMSE)均最小的最佳的樣本數，發現當歷史資料庫有 15,000 筆下，挑選 400 筆類似樣本進行預測可以得到最小的誤差，因此本研究參考該方法，利用 MAPE 針對本研究對後龍到樹林間旅行時間預測結果進行驗證，嘗試在各種不同的 k 值下，旅行時間預測的誤差率為何。

下圖為比較為在統計 97 年 3 月的預測結果後，列出不同的 k 值所預測的旅行時間與實際旅行時間的誤差百分比，結果當 $k=400$ 時為誤差最小的結果，驗證文獻回顧中 Robinson 所得到的結果。

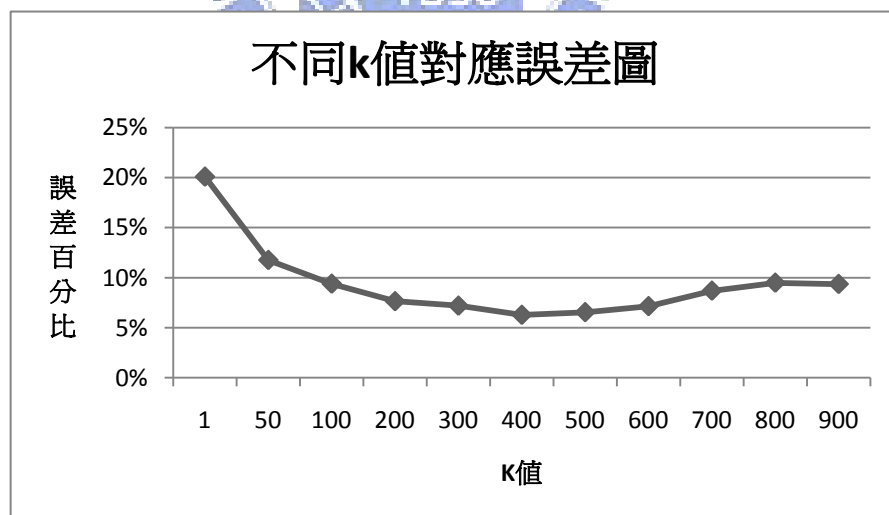


圖 4.2 不同 k 值對應誤差圖

[資料來源：本研究整理]

4.3 k -NN 模式建立

在分類及分群之後，本計畫目標將即時資訊和歷史資訊相比對，找尋即時資訊和那一天的歷史資訊相類似。根據文獻回顧中，Clark 於 2003 年提出利用歷史的流量及佔有率以預測即時的流量及佔有率。本篇文章利用速率及流量作為交通資訊，找尋歷史與即時資訊的平方差加總後，最相近的 400 組樣本進行分析。由於這些樣本的歷史資訊與即時資訊類似，因此可利用這些樣本的歷史資訊來預測交通狀況，而判斷的模式建立如下：

$$tss = \sum_{i=1}^L \sum_{j=1}^T \left[w_q (q_{ij}^r - q_{ij}^h)^2 + w_v (v_{ij}^r - v_{ij}^h)^2 \right] \quad (1)$$

L ：比對範圍的總偵測器數量。

T ：比對的時間長度，本計畫將比對一個小時的交通變化。

q_{ij}^r ：即時各時間點(i)、各偵測器下(j)之流量。

q_{ij}^h ：歷史各時間點(i)、各偵測器下(j)之流量。

v_{ij}^r ：即時各時間點(i)、各偵測器下(j)之速率。

v_{ij}^h ：歷史各時間點(i)、各偵測器下(j)之速率。

w_q ：流量的權重。 w_v ：速率的權重。

tss ：總差異和。

在找到類似的分群特徵後，本研究將收集歷史資料中滿足該特徵下實際日期，並代入歷史時間資料庫後，得到這些日期的歷史旅行時間平均值，即可預測各時間點之旅行時間。本研究同時考慮速度與流量，希望可以避免將尖鋒時段因擁塞而流量較小以及離峰時段流量較少的現象視為同一現象，並且本模式希望在擁塞時段仍有較好的預測能力，因此以擁塞時段之平均速率及流量做為權重的設定值。

4.4 本研究所設定之 k -NN 模型

本篇研究與之前研究有以下兩點差異，1.設定不同的偵測器參數。2.增加時間的維度。以下將分別介紹這兩點差異。

1.設定不同的偵測器參數

除了不同的交通資訊需要不同的參數以進行標準化，不同的路段的資料變異也有相當的差異，所以也應該需要設定不同的參數進行標準化以提升準確率，因此將 3.3.3 節之公式(1)進行修正。

由式(1)可以得知判斷資料差異的方法為下式所示：

$$tss = \sum_{i=1}^L \sum_{j=1}^T \left[w_q (q_{ij}^r - q_{ij}^h)^2 + w_v (v_{ij}^r - v_{ij}^h)^2 \right]$$

若以兩組偵測器為例，則可將式(1)改寫如下：

$$tss = \sum_{i=1}^L \sum_{j=1}^T \left[w_q (q_{ij}^r - q_{ij}^h)^2 + w_v (v_{ij}^r - v_{ij}^h)^2 \right]$$

$$= \sum_{j=1}^T \begin{bmatrix} q_{1j}^r - q_{1j}^h \\ q_{2j}^r - q_{2j}^h \\ v_{1j}^r - v_{1j}^h \\ v_{2j}^r - v_{2j}^h \end{bmatrix}^T \cdot \begin{bmatrix} w_q & 0 & 0 & 0 \\ 0 & w_q & 0 & 0 \\ 0 & 0 & w_v & 0 \\ 0 & 0 & 0 & w_v \end{bmatrix} \cdot \begin{bmatrix} q_{1j}^r - q_{1j}^h \\ q_{2j}^r - q_{2j}^h \\ v_{1j}^r - v_{1j}^h \\ v_{2j}^r - v_{2j}^h \end{bmatrix} .$$

然後本研究再加入同時考慮不同偵測器所帶來的差異而重新調整權重值，因此將利用不同偵測器在不同的交通資訊的變異數(σ^2)，考慮將總差異和(tss)的歐氏距離(Euclidean distance)進行標準化，最後再針對即時資訊與歷史資訊的差異進行標準化進行求解。

因此最後 k -NN 模型將如下式所示：

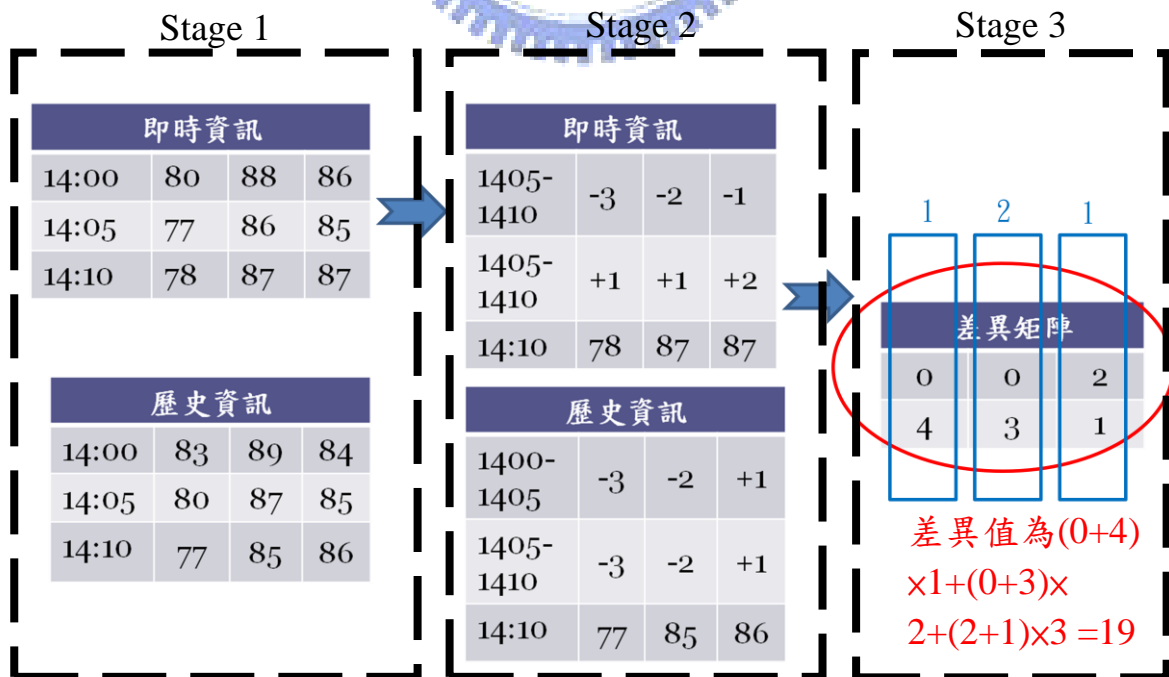
$$tss = \sum_{i=1}^L \sum_{j=1}^T \begin{bmatrix} q_{1j}^r - q_{2j}^h \\ q_{2j}^r - q_{2j}^h \\ \vdots \\ q_{ij}^r - q_{ij}^h \\ v_{1j}^r - v_{1j}^h \\ v_{2j}^r - v_{2j}^h \\ \vdots \\ v_{ij}^r - v_{ij}^h \end{bmatrix}^T \cdot \begin{bmatrix} \sigma_{q1}^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_{q2}^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{v1}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{v2}^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \ddots \end{bmatrix} \cdot \begin{bmatrix} q_{rij} - q_{mij} \\ q_{r2i} - q_{m2i} \\ \vdots \\ v_{r1i} - v_{m1i} \\ v_{r2i} - v_{m2i} \\ \vdots \end{bmatrix} \dots\dots(2)$$

2. 增加時間的維度

之前的文獻大多都只有一個時間點對一個時間點進行比較，進而比較兩個時間點之間的差異，可稱之為單時間點 k -NN 法。但是本研究希望可以比較在一段時間的推移下，交通資訊的改變。因此希望可以分析在半個小時之間交通狀況的變化下，是否有某一歷史時間點有相同的交通狀況變化。故本研究比較是否交通狀況逐漸變順暢或是逐漸變擁擠，利用特徵比對 (pattern Recognition) 的方法進行搜索，可稱之為多時間點 k -NN 法。

舉例說明，如果針對一組三個時間點、三個偵測器的資料進行比對，在下午兩點鐘進行比較，則差異值的求法可分為三個階段(stage)如下圖說明。

圖 4.3 差異值說明圖



[資料來源：本研究整理]

首先在第一個階段得到即時或歷史各時間點之交通資訊，接著在第二階段建立交通資訊的變化趨勢矩陣，第三階段為利用第二階段兩個時間點交通趨勢矩陣，由絕對差異得到兩時間點比對之差異矩陣。最後利用該差異矩陣求出歷史與即時交通資訊的差異值。在最後一個階段需要加入一個考慮因素，由於不希望比對出的交通資訊趨勢反應著不同的交通狀況。舉例說明，像是交通從流量很少到流量正常，與流量正常到擁擠，是反應著兩種交通狀況，但是可能反應著相同的交通變化趨勢，因此需要再加入一個起初點。而起始點為即時與歷史時間點當時的交通資訊，並找出起始點之絕對差異，再由起始點之差異值與交通趨勢之差異進行內積，並考慮不同的偵測器所比對之差異值，再將這些差異值加總，即為比對的差異值。因此由上例，歷史的 14:00 與即時的 14:00 這兩個時間點的差異值為 $(0+4) \times 1 + (0+3) \times 2 + (2+1) \times 3 = 19$ 。最後比較所有歷史時間點與即時資訊的差異值，比對出差異最小的結果進行預測。

本研究時間比對範圍設定為 30 分鐘，以每五分鐘視為一個時間點，故比對範圍為六個時間點。而比對之時間門檻(time window)為比對與當下時間點的前後兩小時，故以 24 個時間點做為時間門檻，希望可以找出與目前類似的歷史時間點。因此公式改寫如下。

$$\begin{aligned}
 \text{tss}_y = \sum_{hi=y-6}^y \sum_{ri=x-6}^x \sum_{j=1}^L (|q_x^j - q_y^j|) \times & \left(\begin{bmatrix} q_{ri}^j - q_{hi}^j \\ q_{ri}^j - q_{hi}^j \\ q_{ri}^j - q_{hi}^j \\ q_{ri}^j - q_{hi}^j \\ q_{ri}^j - q_{hi}^j \end{bmatrix} \cdot \text{diag}(\sigma_q^j) \cdot \begin{bmatrix} q_{ri}^j - q_{hi}^j \\ q_{ri}^j - q_{hi}^j \\ q_{ri}^j - q_{hi}^j \\ q_{ri}^j - q_{hi}^j \\ q_{ri}^j - q_{hi}^j \end{bmatrix} \right) + \\
 (|v_x^j - v_y^j|) \times & \left(\begin{bmatrix} v_{ri}^j - v_{hi}^j \\ v_{ri}^j - v_{hi}^j \\ v_{ri}^j - v_{hi}^j \\ v_{ri}^j - v_{hi}^j \\ v_{ri}^j - v_{hi}^j \end{bmatrix} \cdot \text{diag}(\sigma_v^j) \cdot \begin{bmatrix} v_{ri}^j - v_{hi}^j \\ v_{ri}^j - v_{hi}^j \\ v_{ri}^j - v_{hi}^j \\ v_{ri}^j - v_{hi}^j \\ v_{ri}^j - v_{hi}^j \end{bmatrix} \right) \dots (3)
 \end{aligned}$$

其中 hi 代表歷史時間點，而 ri 代表即時時間點。 x 代表當下的時間點，每五分鐘代表一個時間點。假設當日晚上 00:00 為第一個時間點，則 $x=0$ ，而 00:05 為第二個時間點，故 $x=1$ ，依此類推。而 y 代表比對的時間點(即即時時間點的前後兩小時)，故 y 的上下界為歷史每一天的 $x-24 < y < x+24$ 時間點，最後將 L 個偵測器所比對之資料進行加總。

本研究將比對在每半個小時歷史資訊與即時向前推半個小時資訊的差異，以找出在半個小時歷史資訊和即時資訊的變化，如果兩者變化相近，則代表這兩個時間點上有相似的交通資訊，並且以兩個小時做為時間門檻。舉例來說，若欲預測今天早上九點的旅行時間，則以今天早上八點半到早上九點交通資訊，搜索歷史早上六點半到七點、七點到七點半...到早上十點半到早上十一點的交通資訊，每天各九組資料進行比對，找出最相似的交通資訊。

4.5 模式預測流程

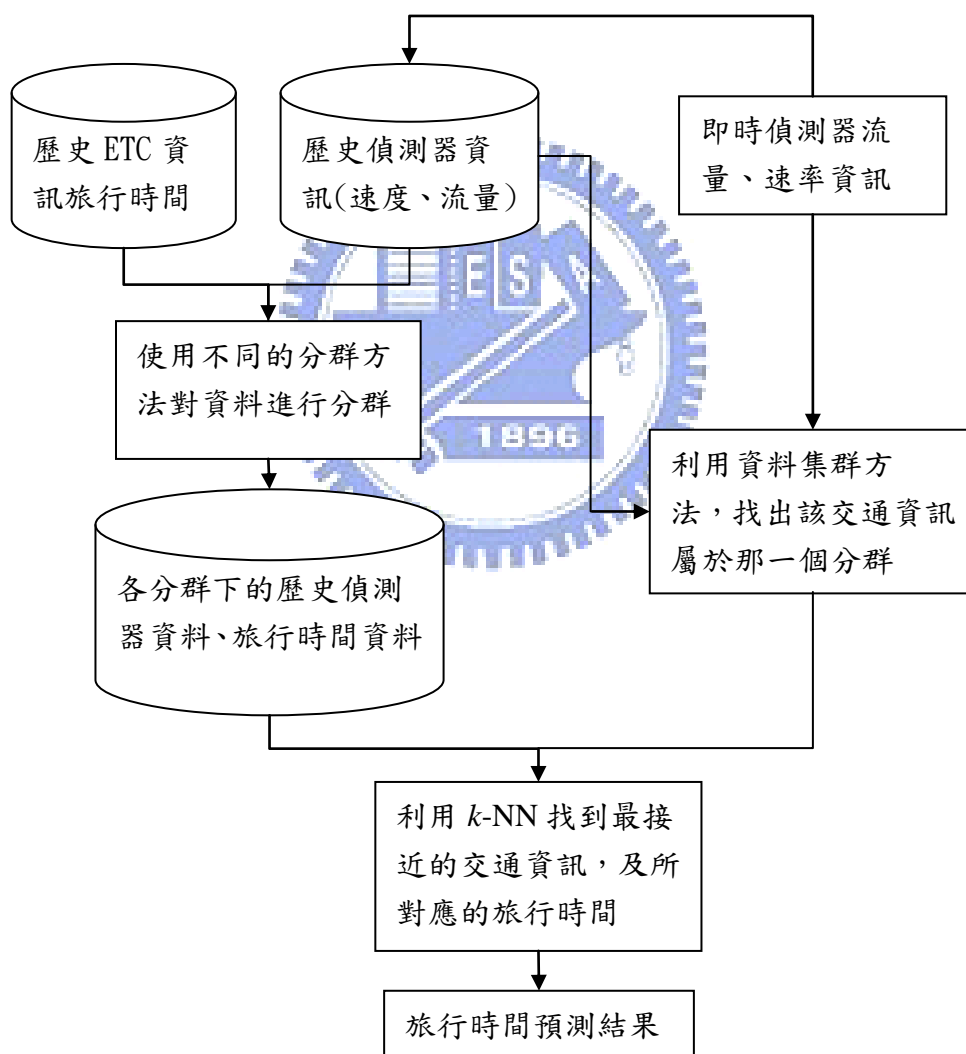


圖 4.4 預測流程架構圖

[資料來源：本研究整理]

第五章 實例測試

5.1 研究範圍

近年來我國為了促進產業升級、調整產業結構及改善國內投資環境，因此持續積極發展具高科技與高附加價值產業，仿倣先進國家發展工作之作法，以設置科學園區或科技工業區等以協助國內企業技術水準、產品品質與市場競爭力。因此行政院國家科學委員會於民國 69 年設立了「新竹科學工業園區」，接著為了建立我國發展資訊軟體及知識密集工業，因此經濟部工業局與臺北市政府產業發展局於民國 88 年開發了「南港科學工業園區」及於民國 91 年將「內湖科學工業園區」由輕工業區轉為科技園區，這三大工業園區建立了北台灣的科技發展重鎮。

而目前新竹科學園區至台北內湖科技園區及南港軟體工業園區之間溝通仍以陸運交通為主，其中又以私有小客車旅次居多。然而由於台北到新竹之間的交通相當繁忙，由表 5.1 所示，對於國道三號而言，小型車通過南港、新竹之間的收費站(樹林收費站、龍潭收費站)，於民國 97 年八月經統計通過次數達 6,494 千輛次，佔該月所有小客車通過收費站的 43.09%，若於該範圍再加上後龍收費站，則比例佔該月的 52.11%。故通過該範圍的車輛數超過所有小客車通過國道三號之收費站車次的二分之一，因此本研究以預測從樹林到後龍的旅行時間做為實例進行分析，該範圍如圖 5.1 所示。



圖 5.1 研究範圍高速公路路線圖

[資料來源：交通部臺灣區國道高速公路局]

表 5.1 臺灣地區國道三號高速公路各收費站通行車輛次數

民國 97 年 8 月	合計	小型車		大客貨車		聯結車	
		數量	百分比	數量	百分比	數量	百分比
七 堵	676	597	4.13%	15	1.61%	63	7.34%
樹 林	4,524	4,135	28.61%	221	23.69%	168	19.58%
龍 潭	2,359	2,093	14.48%	146	15.65%	120	13.99%
後 龍	1,576	1,304	9.02%	99	10.61%	174	20.28%
大 甲	1,631	1,342	9.29%	107	11.47%	182	21.21%
名 間	1,332	1,248	8.63%	64	6.86%	20	2.33%
古 坑	870	795	5.50%	54	5.79%	21	2.45%
白 河	871	790	5.47%	59	6.32%	23	2.68%
善 化	912	821	5.68%	58	6.22%	33	3.85%
田 寮	1,039	909	6.29%	85	9.11%	45	5.24%
竹 田	452	419	2.90%	26	2.79%	8	0.93%

[資料來源：交通部統計處[35]]

(單位：千輛次)

5.2 旅行時間預測結果

本研究將先以較短之距離進行測試，並測試在不同日期別，或是不同星期別進行分群下，預測結果是否有所差異。最後再將該結果應用是較長路段進行測試，並進行討論。

5.2.1 龍潭收費站到樹林收費站測試結果

本研究在較短距離的測試主要以 97 年 3 月 4 日進行討論，將分別測試在單時間點 k -NN 法與多時間點 k -NN 是否在預測效率有差異，以及測試在不同的分群下(依星期別、依日期別)結果的好壞。

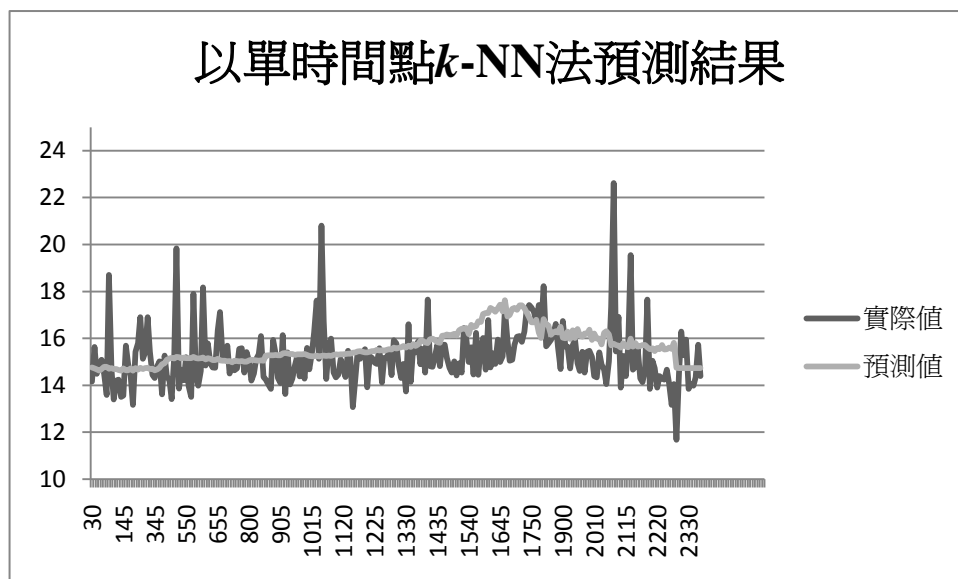


圖 5.2 龍潭到樹林依單時間點 k -NN 法旅行時間預測圖
[資料來源：本研究整理]

上圖為利用單時間點 k -NN 法所得到的結果，其中最大誤差為 23.29%，而平均誤差為 7.57%，以下為多時間點 k -NN 法所得到的誤差結果圖，分別以星期別或日期別進行討論，如圖 5.3 及圖 5.4 所示。

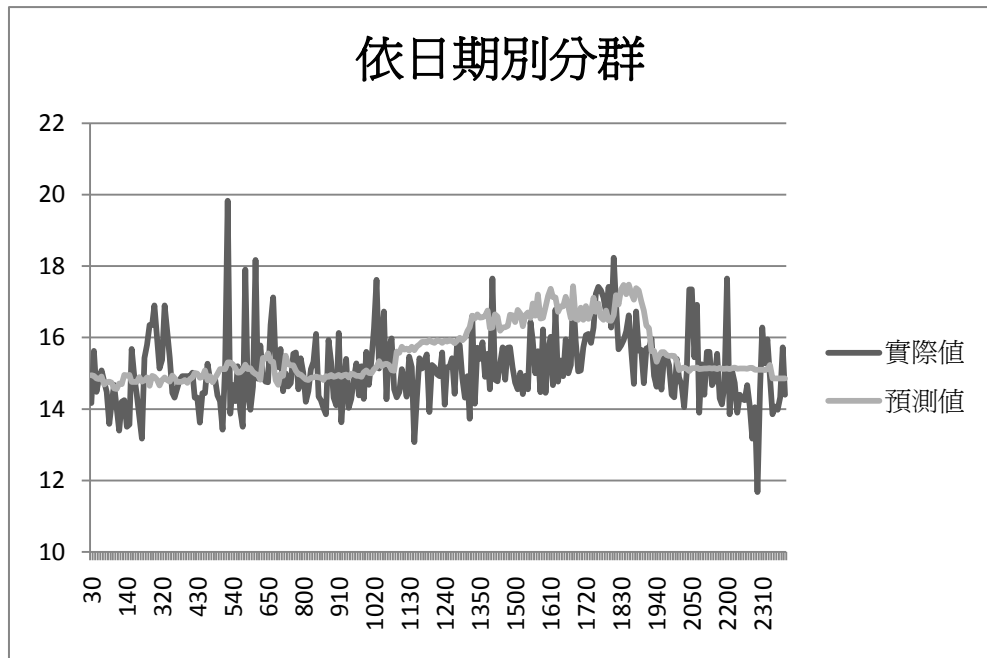


圖 5.3 龍潭到樹林依日期別分群旅行時間預測圖
[資料來源：本研究整理]

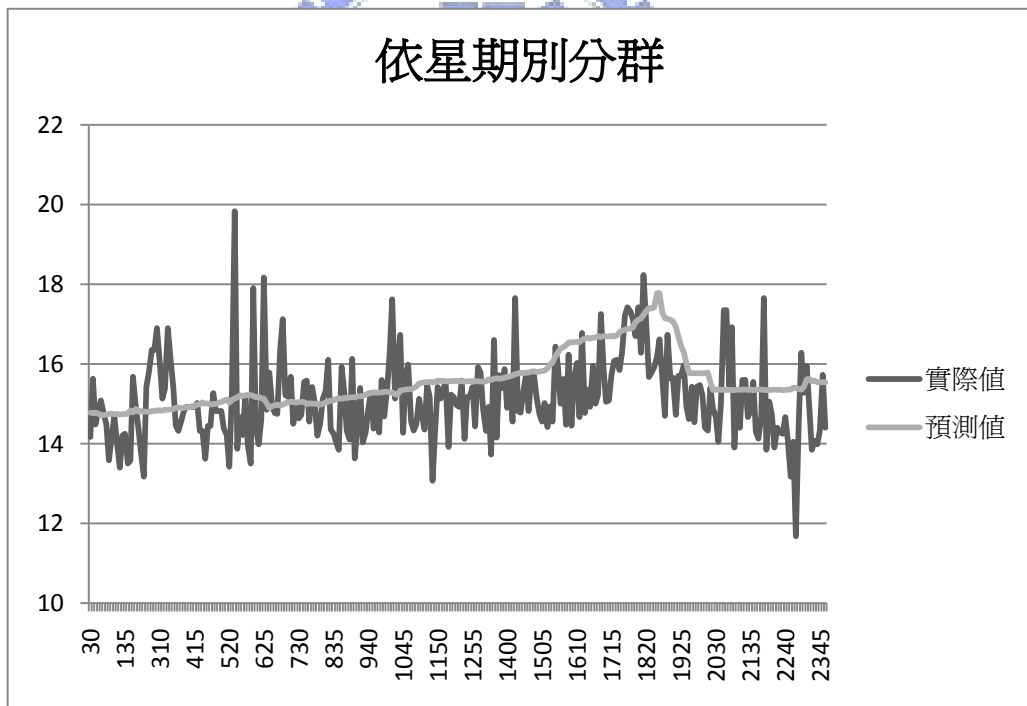


圖 5.4 龍潭到樹林依星期別分群旅行時間預測圖
[資料來源：本研究整理]

兩收費站間預測結果誤差結果整理如下：

依星期別分群：最大誤差：18.89%，平均：6.52%。

依各日分群：最大誤差：19.69%，平均：6.88%

由以上結果發現多時間點 k -NN 不論是以星期別分群或是以日期別分群，結果均比單時間點 k -NN 為佳。而若針對多時間點 k -NN 法而言，以星期別進行分群，不管是最大誤差或是平均誤差上，均較以日期別分群而言有改善的情形，因此本研究將利用星期別分群進行較長距離的驗證(樹林到後龍)，測試本模式的預測結果。

5.2.2 後龍收費站到樹林收費站測試結果

本研究隨機從民國 97 年 3 月中抽取六天的資料進行旅行時間預測。目前本研究將考慮不同的星期別下進行分群。由 ETC 的資料可得該天的旅行時間實際值，與各模式所預測的旅行時間進行比較可得以下的結果。以下列出利用 k -NN 模式從後龍收費站到樹林收費站的旅行時間，其結果如圖 5.4~5.9 所示。

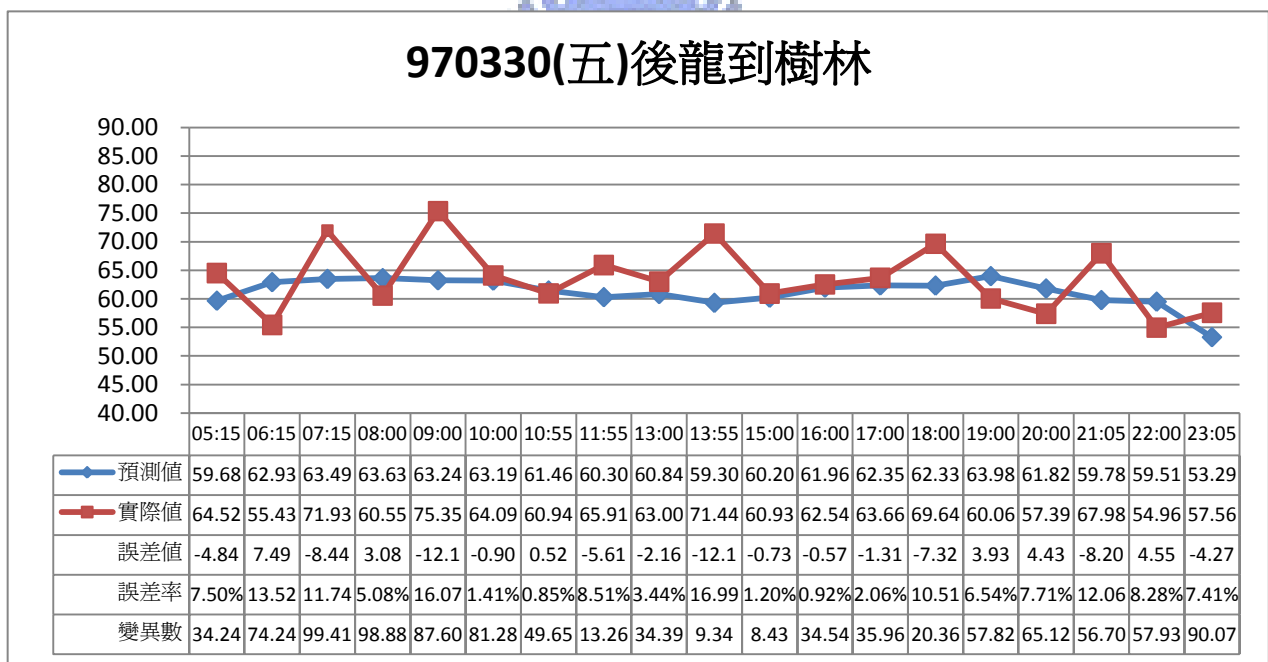


圖 5.5 k -NN 模式旅行時間預測圖(970330)

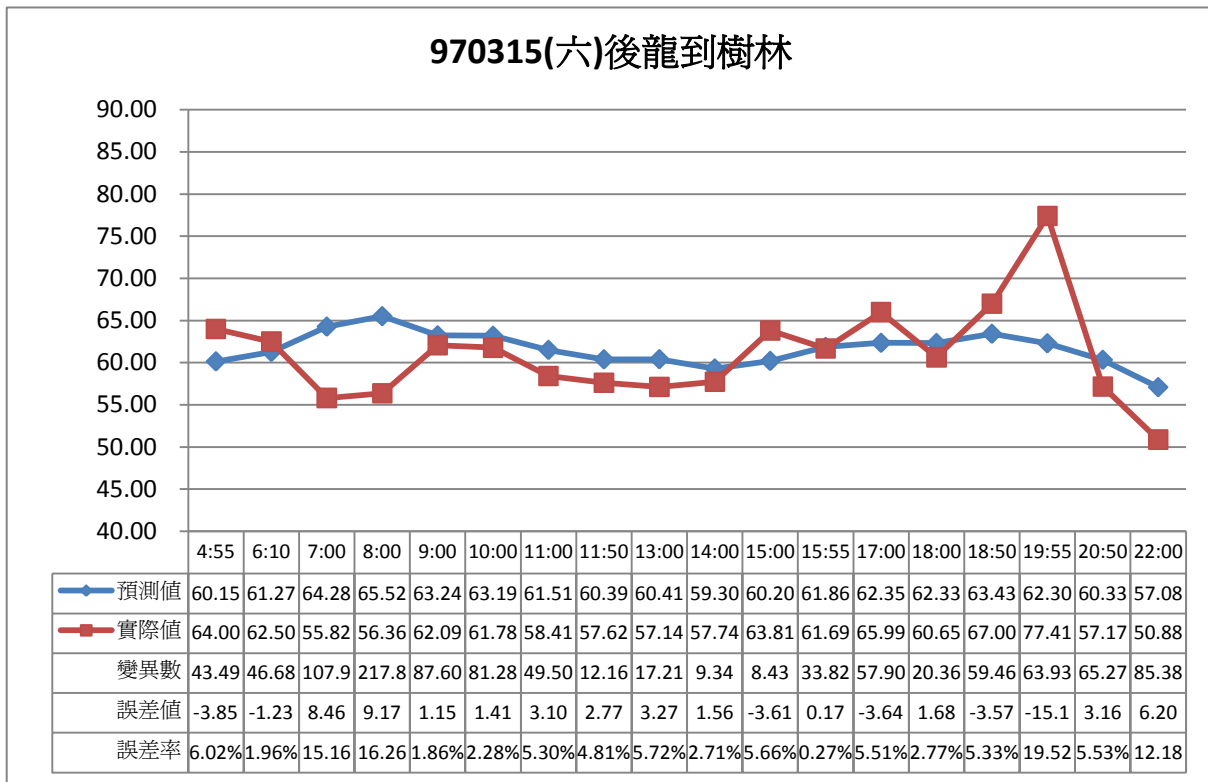


圖 5.6 k -NN 模式旅行時間預測圖(970315)

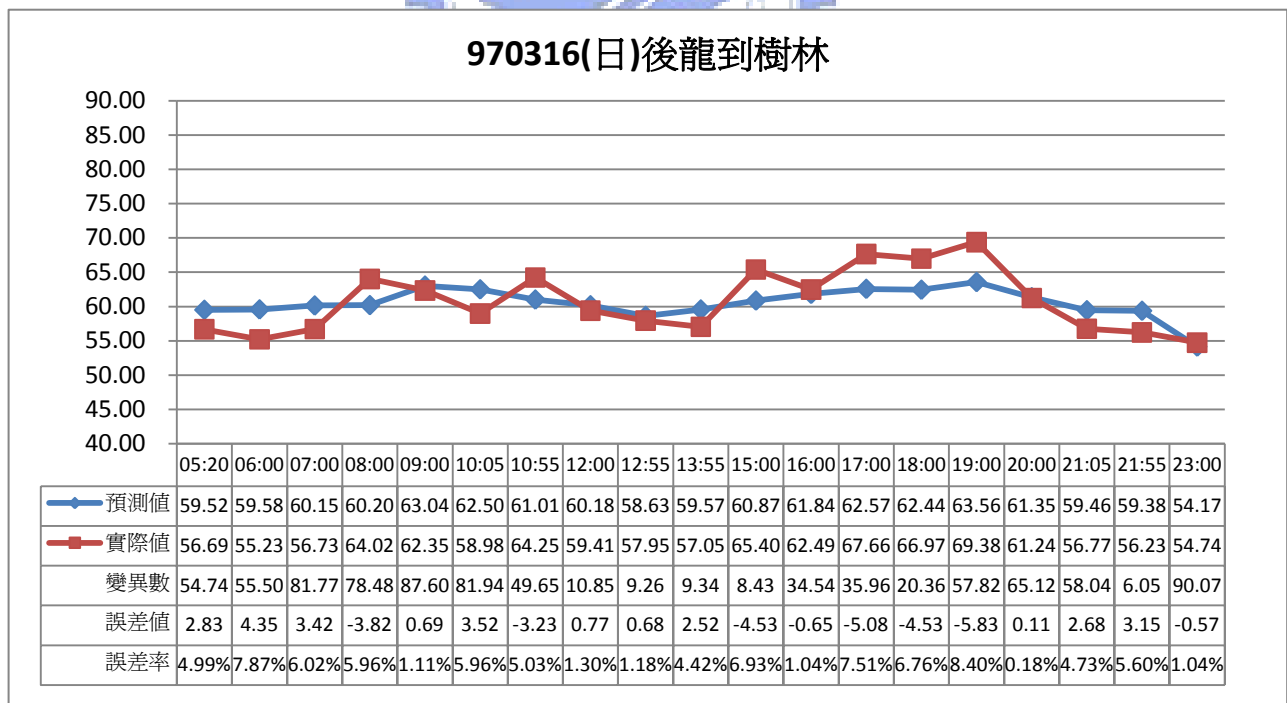


圖 5.7 k -NN 模式旅行時間預測圖(970316)

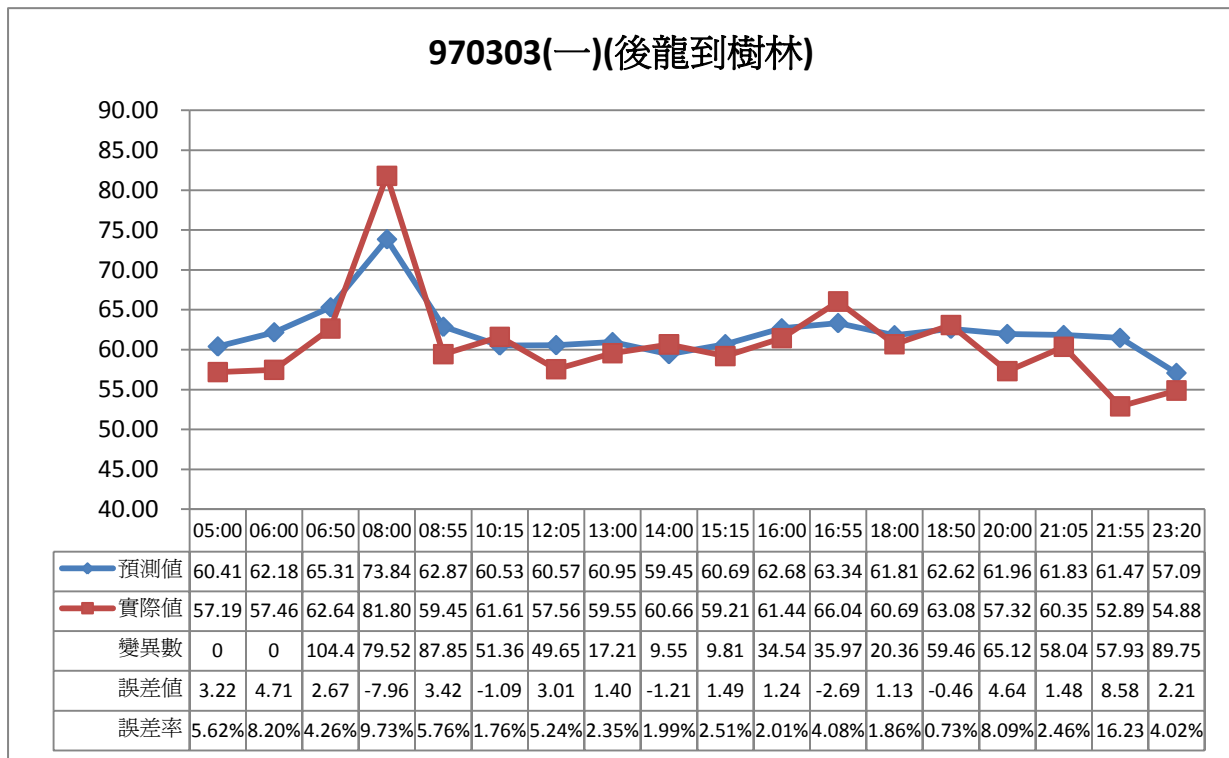


圖 5.8 k -NN 模式旅行時間預測圖(970303)

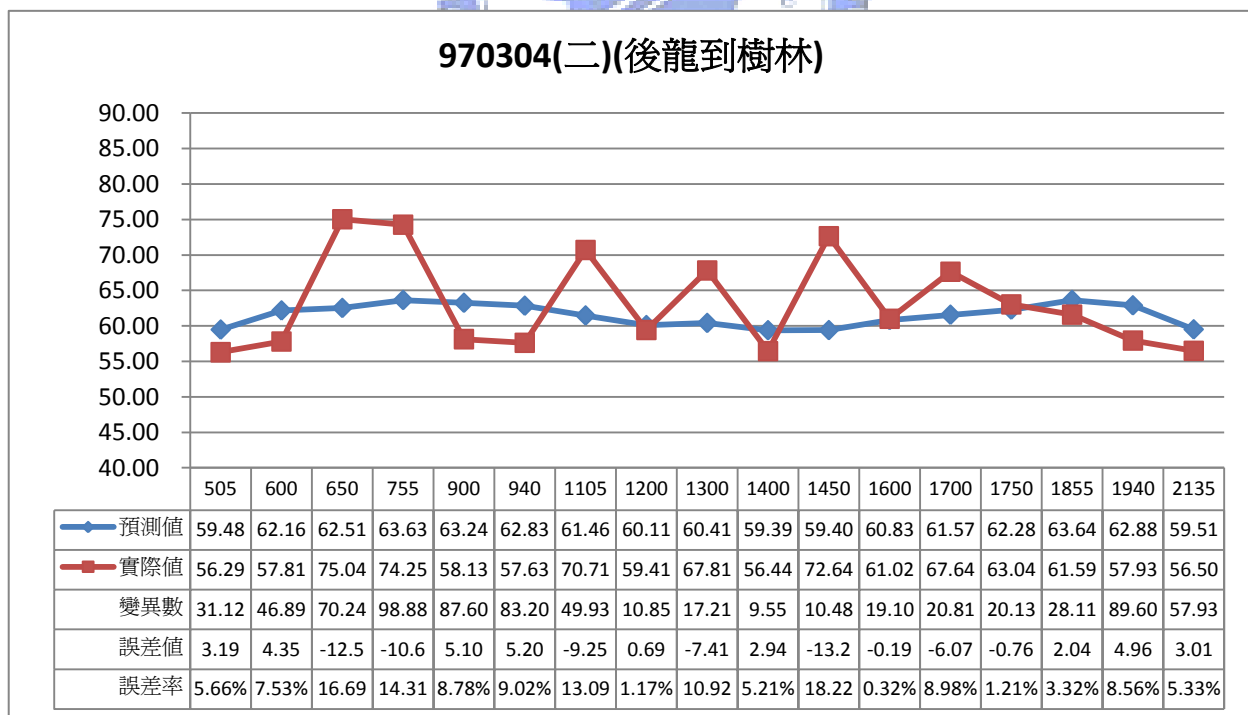


圖 5.9 k -NN 模式旅行時間預測圖(970304)

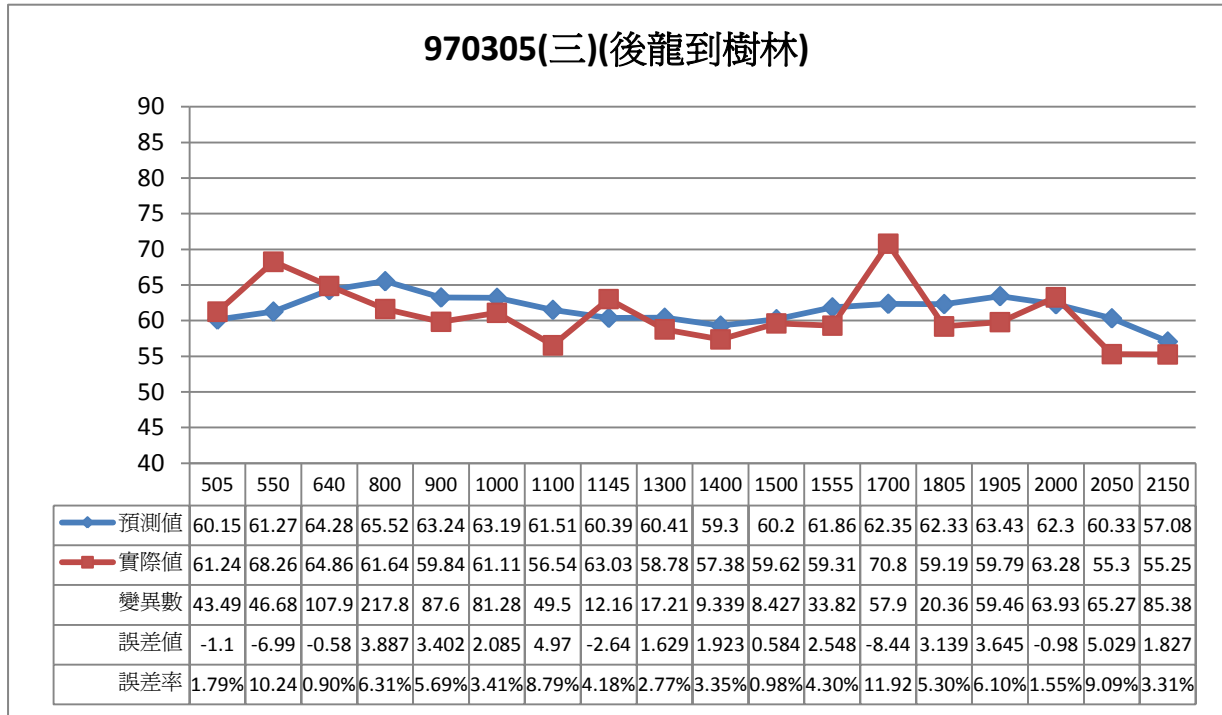


圖 5.10 k -NN 模式旅行時間預測圖(970305)

[資料來源：本研究整理]

而利用 k -NN 模式從後龍收費到樹林收費站的旅行時間預測結果與實際值的誤差情形及各類似樣本的變異數如表 5.2 所示，其他天的預測結果請參閱附錄三。

表 5.2 k-NN 模式旅行時間預測整理(970330)

時間	預測值 (分)	實際值 (分)	誤差值 (分)	誤差率 (%)	變異數
05:15	59.68	64.52	-4.84	7.50%	34.24
06:15	62.93	55.43	7.49	13.52%	74.24
07:15	63.49	71.93	-8.44	11.74%	99.41
08:00	63.63	60.55	3.08	5.08%	98.88
09:00	63.24	75.35	-12.11	16.07%	87.60
10:00	63.19	64.09	-0.90	1.41%	81.28
10:55	61.46	60.94	0.52	0.85%	49.65
11:55	60.30	65.91	-5.61	8.51%	13.26
13:00	60.84	63.00	-2.16	3.44%	34.39
13:55	59.30	71.44	-12.14	16.99%	9.34
15:00	60.20	60.93	-0.73	1.20%	8.43
16:00	61.96	62.54	-0.57	0.92%	34.54
17:00	62.35	63.66	-1.31	2.06%	35.96
18:00	62.33	69.64	-7.32	10.51%	20.36
19:00	63.98	60.06	3.93	6.54%	57.82
20:00	61.82	57.39	4.43	7.71%	65.12
21:05	59.78	67.98	-8.20	12.06%	56.70
22:00	59.51	54.96	4.55	8.28%	57.93
23:05	53.29	57.56	-4.27	7.41%	90.07

[資料來源：本研究整理]

由上例發現平均誤差率為 7.16%，最大誤差率為 16.99%，但是最大誤差仍有過大的問題，因此還可以有改善的空間存在。而 k -NN 模式中，取最相似之 400 筆交通資訊以預測旅行時間，而這些交通資訊所對應之旅行時間之變異數少則為 8.43，而大則為 99.41，這些變異數的變化可能是由於某些車輛短暫離開高速公路或是行駛速率過慢，而造成旅行時間過長，而且由於其異常現象並沒有非常顯著，所以較難被完全過濾掉。

由於高速公路的偵測器不一定每一個偵測器每五分鐘都能回傳資訊，所以造成旅行時間預測上常常會有缺值的情形發生。並且由於不一定每五分鐘都有車輛從龍潭收費站經過後龍收費站，因此本研究每小時進行一次預測。



第六章 結論與建議

(1) 由以上結果發現，短距離旅行時間預測之誤差約為 6.5% 而長距離旅行時間預測之誤差約為 7.1%，因此本方法在兩個及三個收費站間旅行時間之預測績效沒有太大的差異。

(2) 由於本研究目前沒有演算法直接根據交通資訊判斷是否該時段發生事故，因此在資料過濾時，即把一些速度過慢、旅行時間過長之情形判斷為離群值而進行過濾，因此未來如果加入事件資料庫，可以增加判斷事件發生的機制，而建立在事件發生時之旅行時間資料庫，將有助於在事故發生時進行旅行時間預測。

(3) 本研究目前對於參數的設定如權重的設定、門檻值的設定及距離量度的給定，由於目前沒有一篇論文是完全利用 k -NN 方法進行旅行時間預測，並且目前參數值均以比較方法進行搜索，像是時間門檻的設定，本研究根據平均旅行時間大部分落於(42~64 分間)，故本研究以半小時做為時間的門檻，找尋在每半小時下，最接近的歷史資料。希望未來可以有進一步的研究，探討這些參數要如何決定。

(4) 目前本研究僅以歷史資料進行過濾及插補，然而若應用於實際案例中，如果偵測器有離群值的產生，或是資料缺漏，就無法得到準確的資料。

(5) 未來可以再多取一些樣本，驗證是否本模式有足夠的解釋意義。

參 考 文 獻

- [1] Wachs, M., “Relationships between drivers’ attitudes toward alternate routes and driver and route characteristics” ,Highway Research Record , vol.197, pp.70-87, 1967.
- [2] van Berkum, E., and P. van der Mede, “The Impact of Traffic Information: Dynamics in Route and Departure Time Choice, Delft University Press, Delft, the Netherlands” , PhD. thesis of the Transport and Planning Department, Faculty of Civil Engineering and Geosciences, Delft University of Technology, 1993.
- [3] 運輸安全組, “北臺灣科技走廊智慧型系統規劃與開發” , 交通部運輸研究所 , 2008年6月。
- [4] Kwon, J., B. Coifman, and P. Bickel, “Day-to-Day Travel-Time Trends and Travel-Time Prediction from Loop-Detector Data” , Transportation Research Record: Journal of the Transportation Research Board, Vol.1717, pp.120-129,2000.
- [5] Sen, A., P. Thakuriah, and X.Q. Zhu, “Frequency of Probe Reports and Variance of Travel Time Estimates” , Journal of Transportation Engineering, Vol. 123, No. 4, pp. 290-297, 1997.
- [6] Rice, J. and E. van Zwet, “A Simple and Effective Method for Predicting Travel Times on Freeways” , IEEE Transactions on Intelligent Transportation Systems, Vol. 5(3), pp. 200-207, 2004.
- [7] Yang, J.S. and D. Billings , “Application of The ARIMA Models to Urban Roadway Travel Time Prediction” ,SMC IEEE International Conference, Vol.3, pp.2529-2534, 2006.
- [8] Vapnik, V., “Predicting Time Series with Support Vector Machine” , Springer LNCS 1327, pp.999-1004.
- [9] Wu, C.H., J.M. Ho, and D.T. Lee, “Travel time prediction with support vector regression” , IEEE Transactions on Intelligent Transportation Systems, Vol.5, NO. 4, pp. 276-281, 2004.

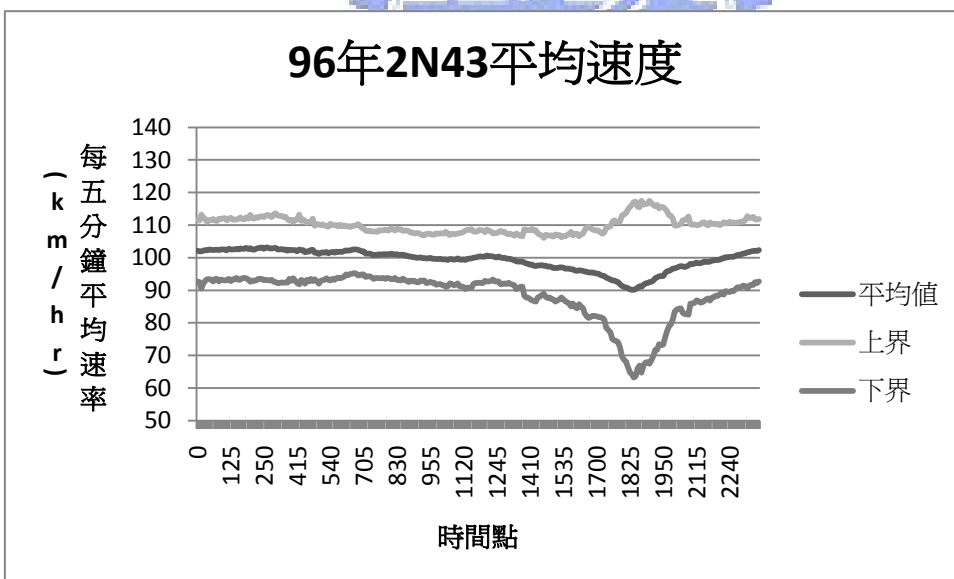
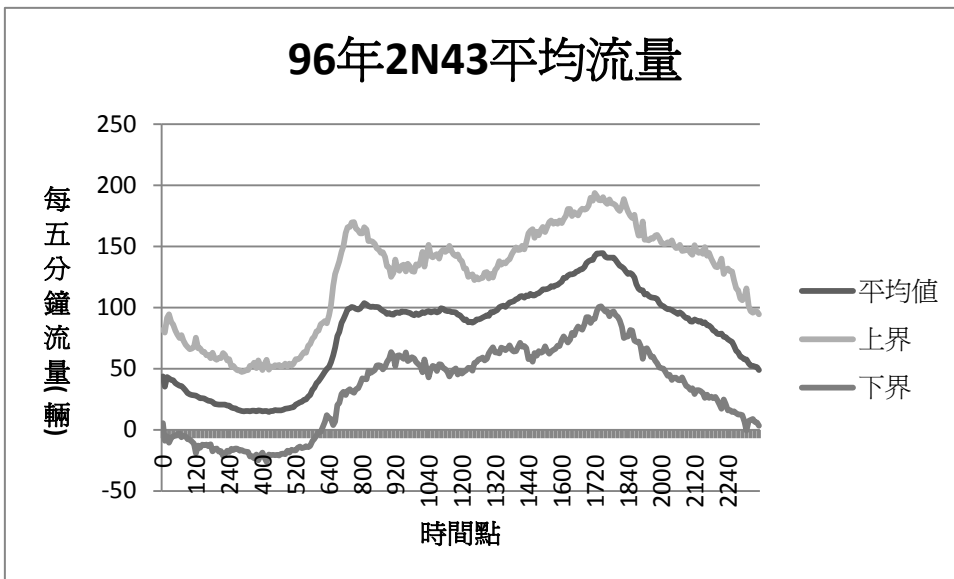
- [10] Yoshikazu, O., K. Toshihiro, and S. Shigehito, "Online-Learning Type of Traveling Time Prediction Model in Expressway", Intelligent Transportation System, ITSC IEEE Conference, pp. 350-355, 1998.
- [11] 李季森, 「應用探測車法預測高速公路旅行時間」, 中央大學土木工程研究所, 碩士論文, 2001。
- [12] 張修榕, 「高速公路旅行時間之研究」, 中央大學土木工程研究所, 碩士論文, 2001。
- [13] 黃裕文, 「高速公路施工路段旅行時間預測之研究」, 中央大學土木工程研究所, 碩士論文, 2003。
- [14] 溫志元, 「高速公路進口匝道匯流路段旅行時間研究」, 中央大學土木工程研究所, 碩士論文, 2002。
- [15] 林士傑, 「高速公路旅行時間預測模式之研究-類神經網路之應用」, 成功大學交通管理科學研究所, 碩士論文, 2001。
- [16] 吳佳峰, 「有 GPS 資訊提供下之車輛旅行時間預估模式之研究」, 交通大學運輸工程與管理系, 碩士論文, 2001。
- [17] 李穎, 「國道客運班車旅行時間預測模式之研究」, 成功大學交通管理科學研究所, 碩士論文, 2002。
- [18] 張慶麟, 「應用自動車輛辨識預測高速公路路段旅行時間」, 中央大學土木工程研究所, 碩士論文, 2002。
- [19] 魏健宏, 林士傑, 李穎「高速公路客運車輛旅行時間預測之實證評析」, 運輸計劃季刊, 第 32 卷, 第 4 期, 651~679 頁, 2003。
- [20] Chang, G.L., T. Junchaya, and A.J. Santiago, "A real-Time Network Traffic Simulation Model For ATMS Applications: Part I - Simulation Methodologies", Journal of Intelligent Transportation Systems, pp. 227-241, 1994.
- [21] Johnston, C.M., and A.T. Chronopoulos, "The parallelization of a Highway Traffic Flow Simulation", Frontiers of Massively Parallel Computation, pp.192-199, 1999.

- [22] Kachani, S., and G. Perakis, “Second-Order Fluid Dynamics Models for Travel Travel Times in Dynamic Transportation Networks” , Proceedings on IEEE Intelligent Transportation Systems Conference, pp. 251-256, 2001
- [23] Kiesling, T., and J. Luthi, “Towards Time-Parallel Road Traffic Simulation,” Principles of Advanced and Distributed Simulation, pp. 7-15, 2005.
- [24] Benedetti, J., “On the Nonparametric Estimation of Regression Functions” , Journal of the Royal Statistical Society, Series B, Vol. 39, pp. 248-253, 1977.
- [25] Stone, C., “Consistent Nonparametric Regression” , The Annals of Statistics, Vol. 5, pp. 595-645, 1977.
- [26] Tukey, J., “Exploratory Data Analysis” , Reading, MA: Addison Wesley, 1977.
- [27] Altman, N. S., “An Introduction to Kernel and Nearest Neighbor Nonparametric Regression” , The American Statistician, Vol. 46, NO. 3, pp. 175-185, 1992.
- [28] Davis, G. A., and N. L. Nihan, “Nonparametric Regression and Short-Term Freeway Traffic Forecasting” , Journal of Transportation Engineering, Vol. 117, NO. 2, pp. 178-188, 1991.
- [29] Smith, B., and M. Demetsky, “Traffic Flow Forecasting: Comparison of Modeling Approaches” , Journal of Transportation Engineering, Vol.123, NO.4, pp. 261 - 266, 1997.
- [30] Smith, B., B. Williams, and K. Oswald, “Comparison of Parametric and Non-parametric Models for Traffic Flow Forecasting” , Transportation Research Part C, Vol. 10(4), pp. 303 - 321, 2002.
- [31] Clark, S., “Traffic Prediction using Multivariate Nonparametric Regression” , Journal of Transportation Engineering, Vol. 129(2), pp. 161-168, 2003.

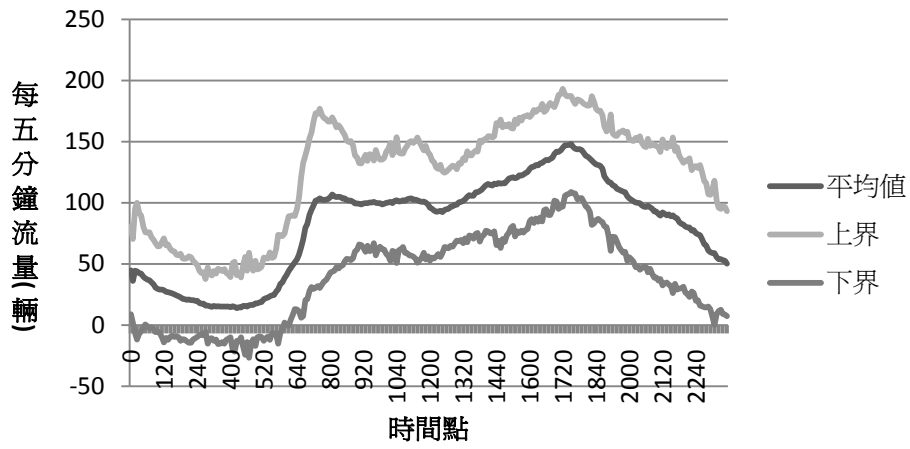
- [32] Robinson, S., and J. Polak, “Modeling Urban Link Travel Time with Inductive Loop Detector Data by Using the k -NN Method” , Journal of the Transportation Research Board, Vol. 1935, pp. 47-56, 2005.
- [33] Chang, G.L., N. Zou, and J. Wang, “Development and Field Evaluation of a Real-Time Travel Time Prediction System” , Maryland Department of Transportation State Highway Administration, Final Report, 2006.
- [34] Fukunaga, K., and L. Hostetler , “Optimization of k -Nearest-Neighbor Density Estimates ,” IEEE Transactions on Information Theory, Vol.19, pp.320-326, 1973.
- [35] 交通部統計處， “交通統計月報” ，交通部，2008。



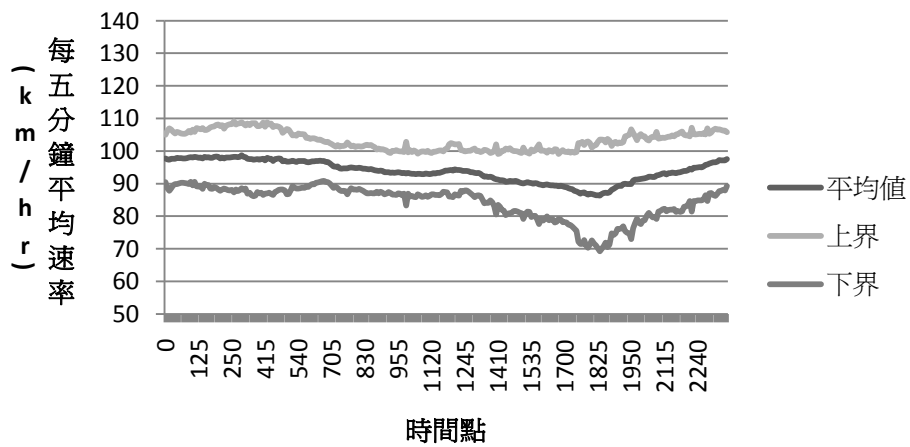
附錄一 各偵測器交通資訊之信賴區間



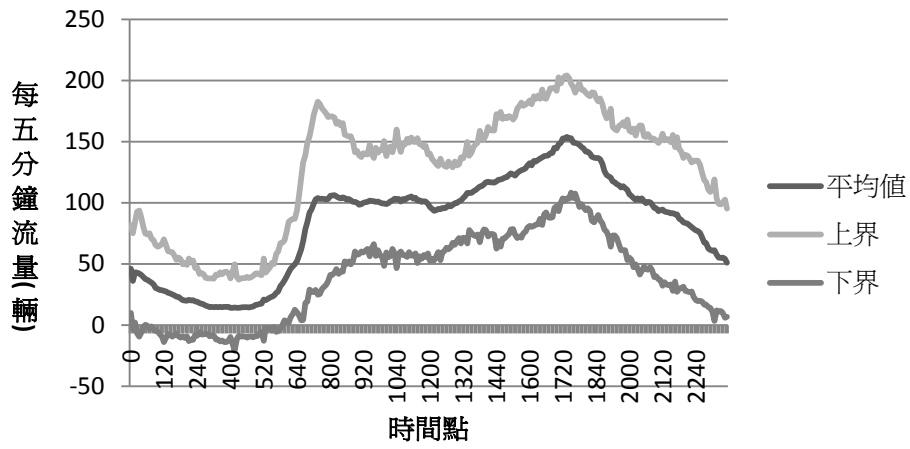
96年2N45平均流量



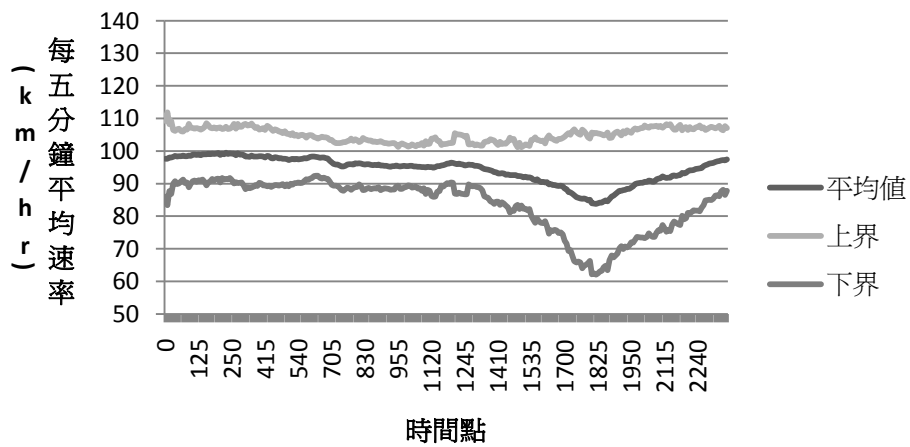
96年2N45平均速度



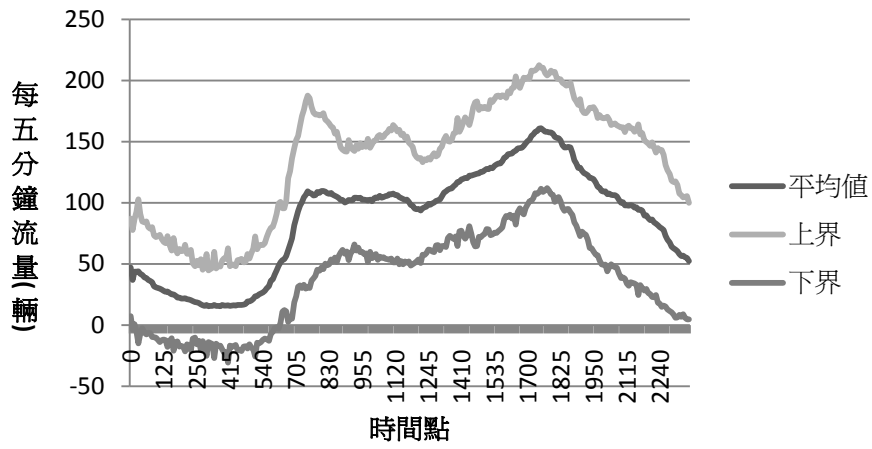
96年2N47平均流量



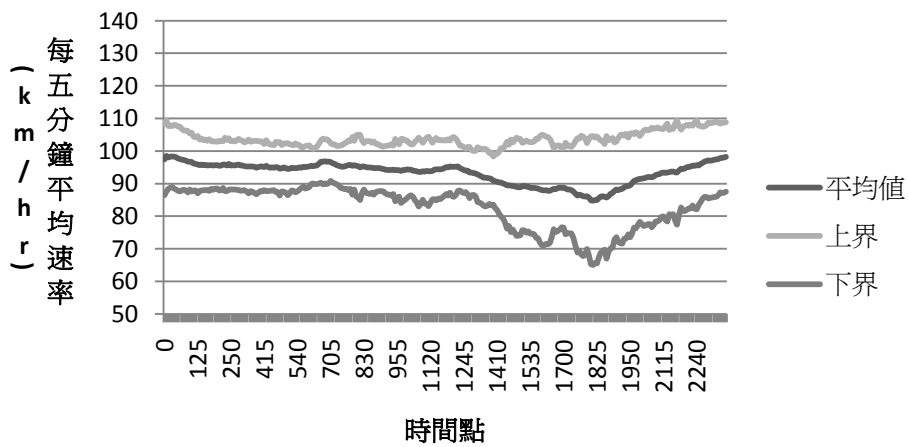
96年2N47平均速度



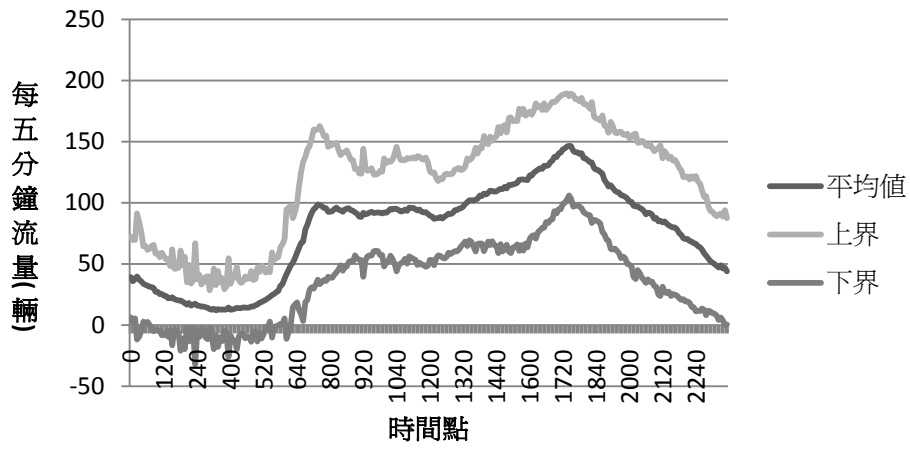
96年2N51平均流量



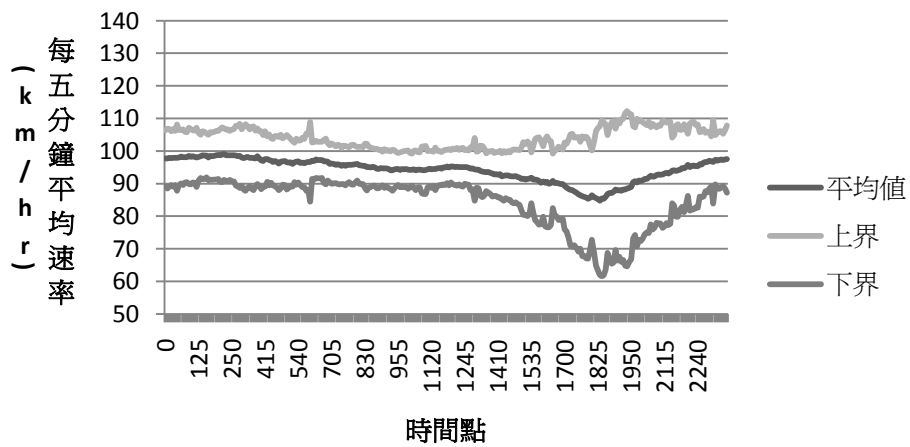
96年2N51平均速度



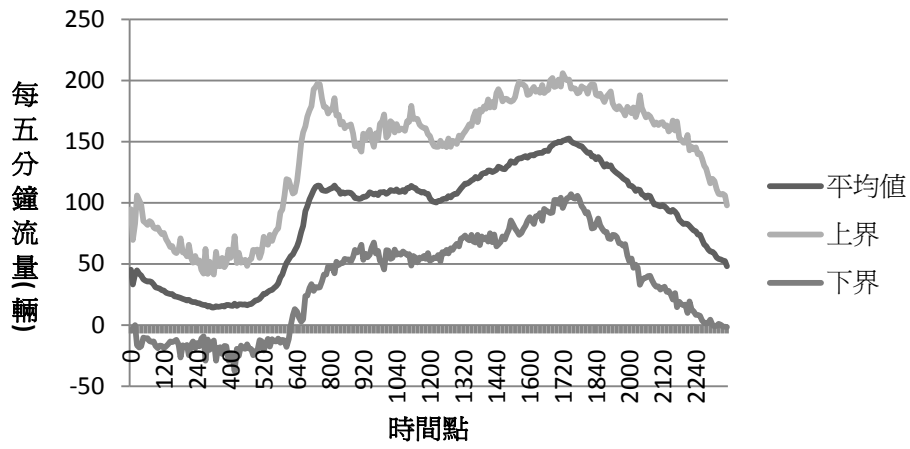
96年2N57平均流量



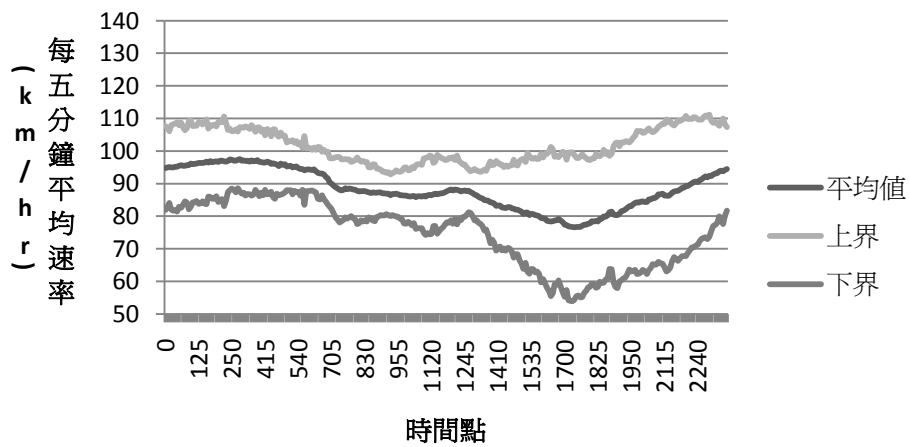
96年2N57平均速度



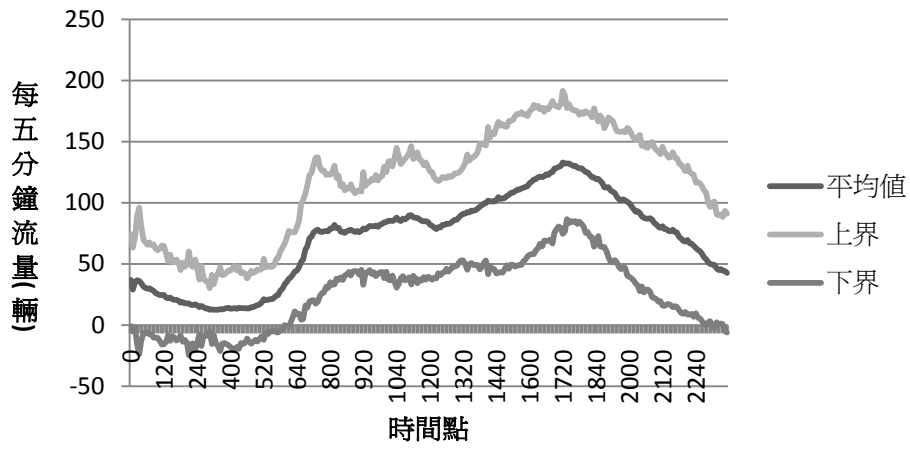
96年2N60平均流量



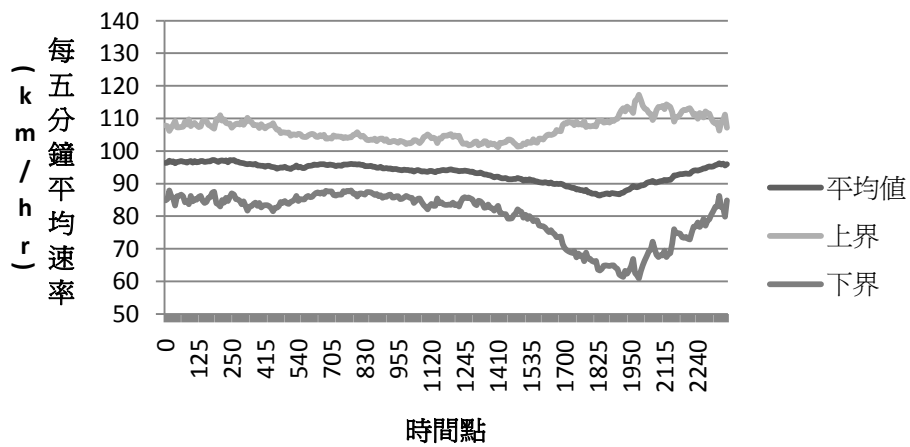
96年2N60平均速度



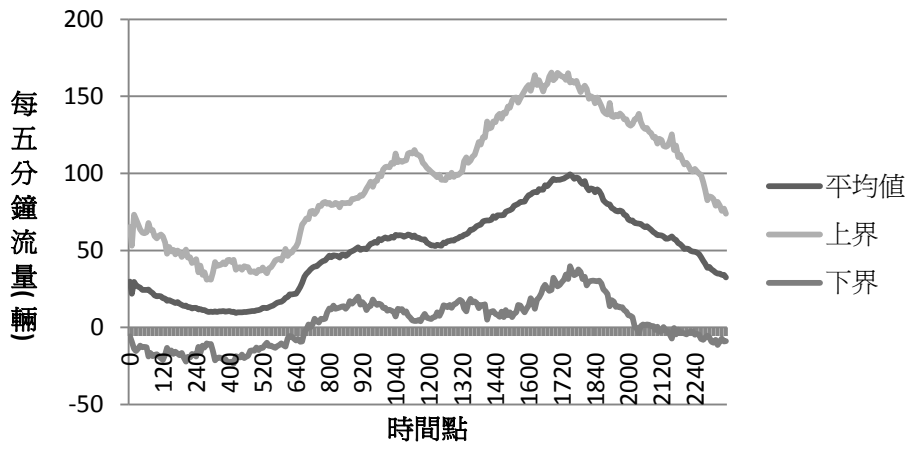
96年2N64平均流量



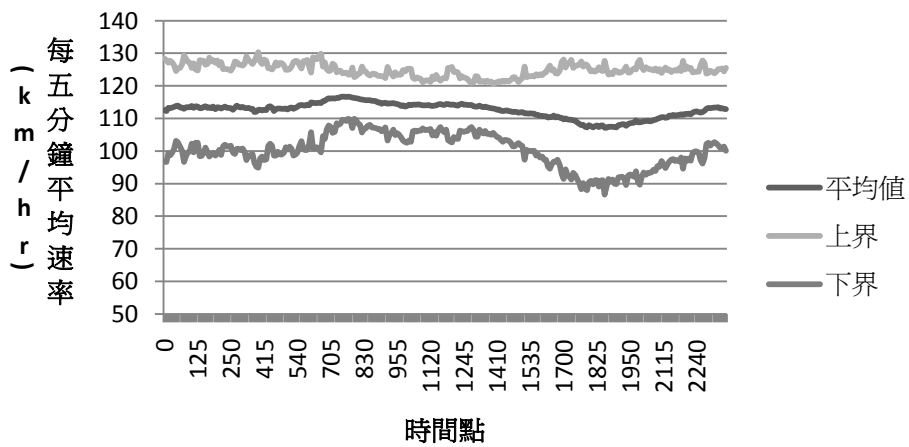
96年2N64平均速度

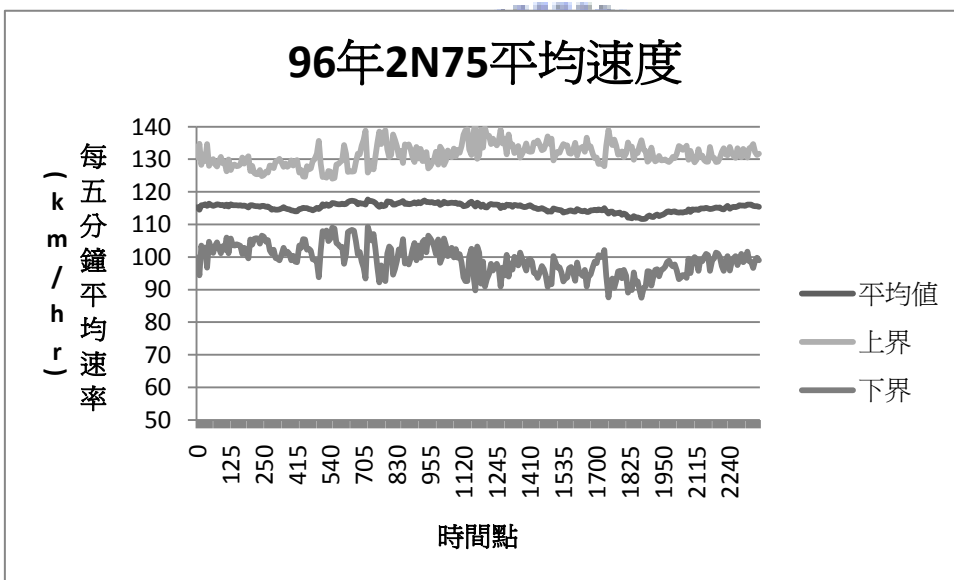
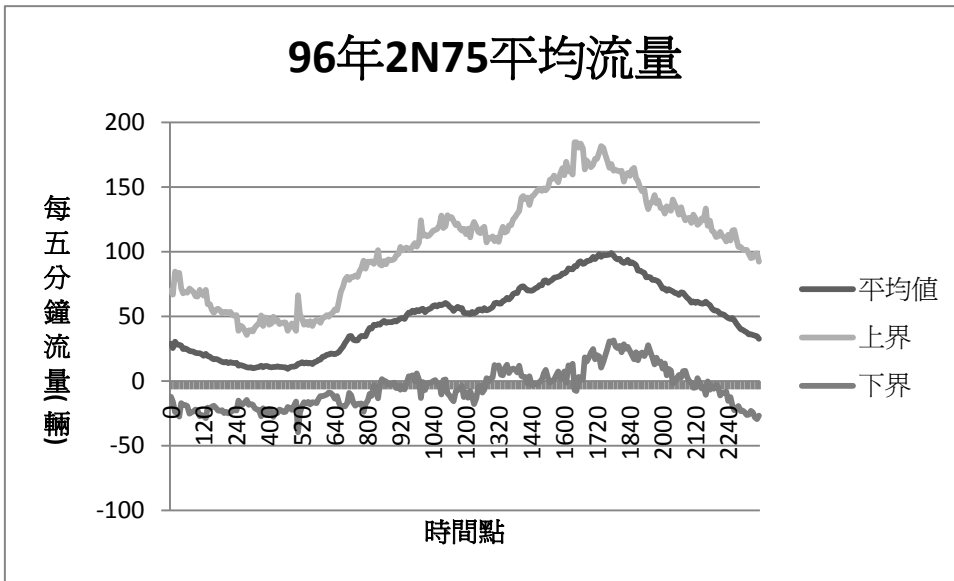


96年2N73平均流量

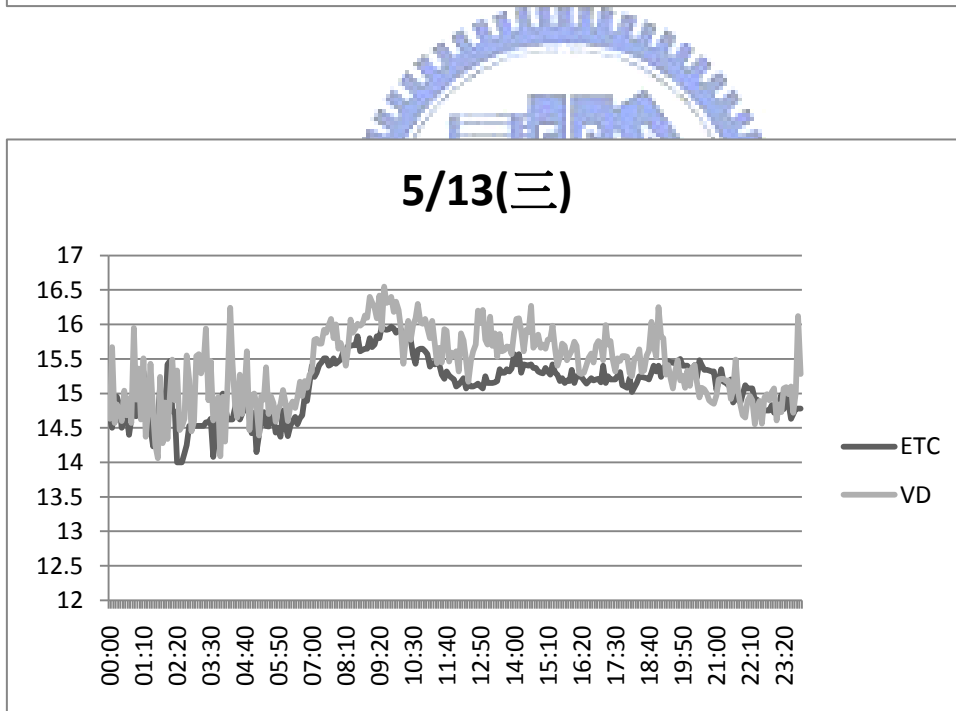
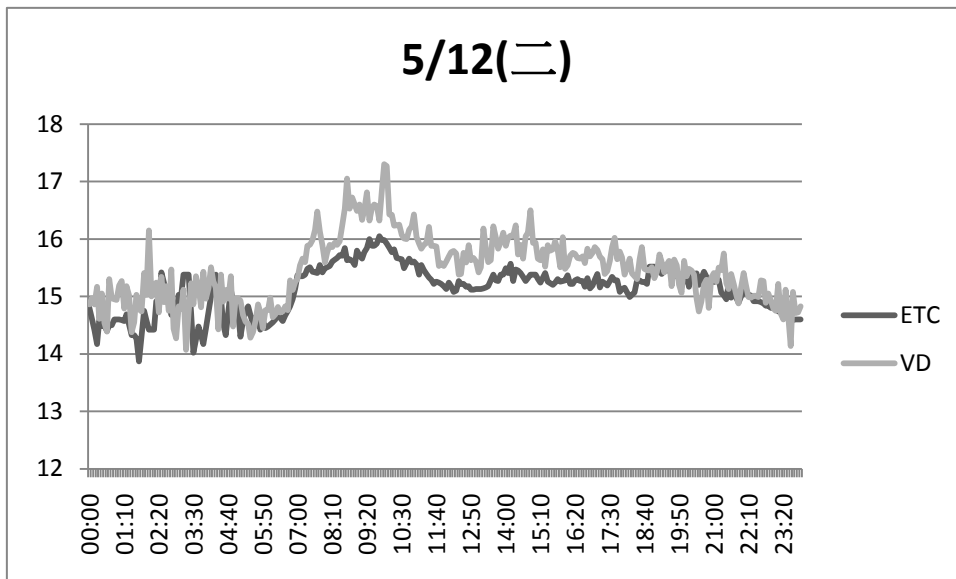


96年2N73平均速度





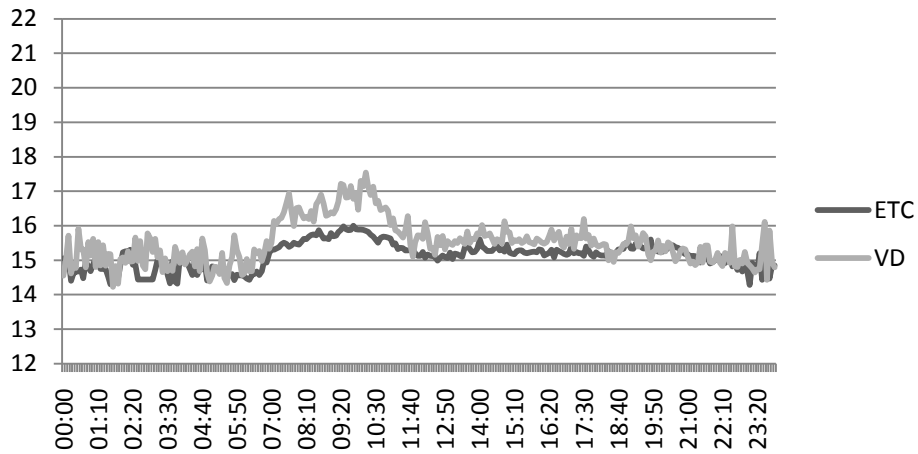
附錄二 VD 與 ETC 推估旅行時間比較圖



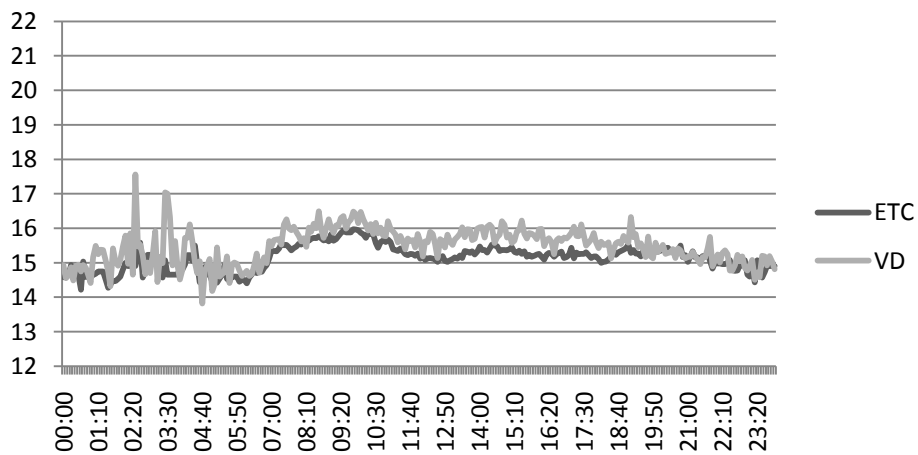
5/18(一)



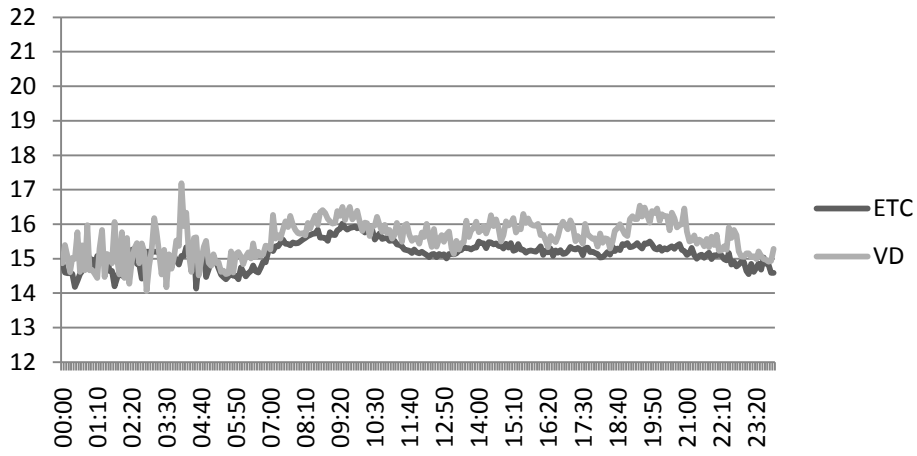
5/19(二)



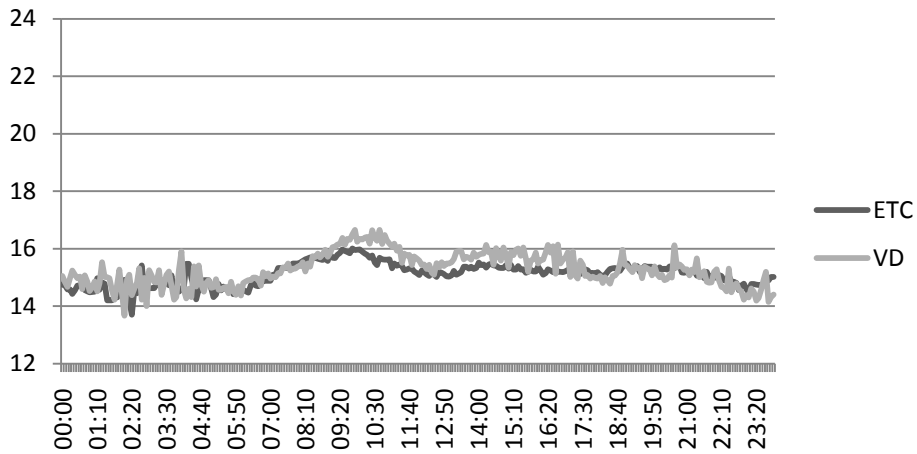
5/20(三)



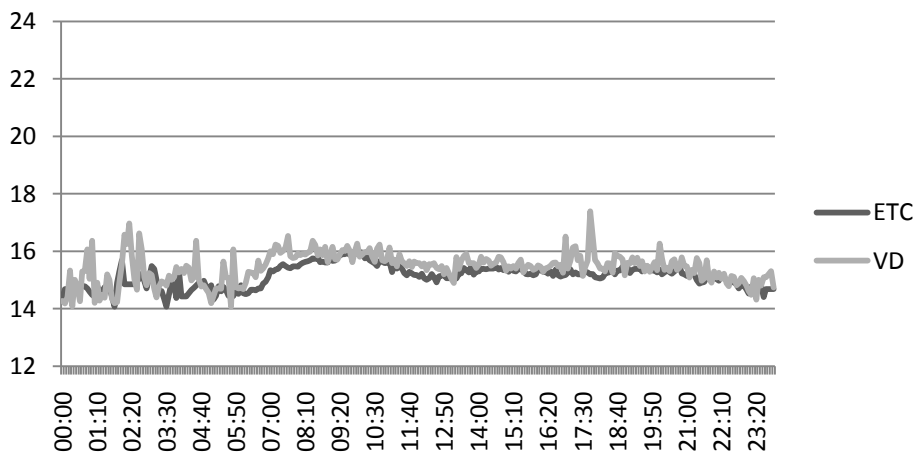
5/22(五)



5/24(日)



5/25(一)



附錄三 旅行時間預測結果整理

k-NN 模式旅行時間預測整理(970303(一))

時間	預測值 (分)	實際值 (分)	誤差值 (分)	誤差率 (%)	變異數
05:00	59.68293	57.19167	2.491261	4.36%	40.64368
06:00	63.25098	57.46481	5.786163	10.07%	76.91641
07:00	65.30463	62.635	2.669632	4.26%	104.4071
08:00	74.24098	81.79902	-7.55804	9.24%	79.51625
09:00	63.80902	59.44722	4.3618	7.34%	87.85285
10:00	61.92293	61.61204	0.310889	0.50%	51.35536
12:00	61.46049	57.55833	3.902154	6.78%	49.65348
13:00	60.40829	59.55	0.85829	1.44%	17.21197
14:00	59.38634	60.65655	-1.27021	2.09%	9.552464
15:00	60.34634	59.20833	1.138009	1.92%	9.814049
16:00	61.96195	61.44095	0.521006	0.85%	34.5413
17:00	62.22366	66.0358	-3.81214	5.77%	35.97134
18:00	62.32585	60.68653	1.639323	2.70%	20.35702
19:00	63.43122	63.07972	0.351497	0.56%	59.46309
20:00	61.81756	57.32143	4.496135	7.84%	65.11891
21:00	59.77585	60.35	-0.57414	0.95%	58.04229
22:00	59.50829	52.88667	6.621624	12.52%	57.93485
23:00	53.24927	54.87917	-1.6299	2.97%	89.74966

k-NN 模式旅行時間預測整理(970304(二))

時間	預測值 (分)	實際值 (分)	誤差值 (分)	誤差率 (%)	變異數
05:00	59.47756	56.29166	3.185897	5.66%	31.12386
06:00	62.16195	57.80833	4.353617	7.53%	46.89345
07:00	62.51463	75.03999	-12.5254	16.69%	70.23805
08:00	63.62829	74.25445	-10.6262	14.31%	98.8784
09:00	63.23927	58.13461	5.104655	8.78%	87.59885
10:00	62.83171	57.63333	5.198376	9.02%	83.20334
11:00	61.46024	70.71439	-9.25415	13.09%	49.92535
12:00	60.10707	59.41459	0.692488	1.17%	10.85384
13:00	60.40829	67.81333	-7.40504	10.92%	17.21197
14:00	59.38634	56.44295	2.94339	5.21%	9.552464
15:00	59.4022	72.63712	-13.2349	18.22%	10.47545
16:00	60.82512	61.01743	-0.1923	0.32%	19.10081
17:00	61.56561	67.64052	-6.07491	8.98%	20.81106
18:00	62.27854	63.04219	-0.76365	1.21%	20.13344
19:00	63.63585	61.59391	2.041944	3.32%	28.1084
20:00	62.88195	57.92578	4.956174	8.56%	89.60308
21:00	59.50829	56.49889	3.009407	5.33%	57.93485

k-NN 模式旅行時間預測整理(970305(三))

時間	預測值 (分)	實際值 (分)	誤差值 (分)	誤差率 (%)	變異數
05:00	60.14854	61.24444	-1.09591	1.79%	43.48507
06:00	61.2722	68.26111	-6.98891	10.24%	46.68024
07:00	64.2778	64.85834	-0.58053	0.90%	107.9457
08:00	65.52463	61.6375	3.887137	6.31%	217.803
09:00	63.23927	59.8375	3.401766	5.69%	87.59885
10:00	63.19073	61.10568	2.08505	3.41%	81.27873
11:00	61.50927	56.53944	4.969824	8.79%	49.50127
12:00	60.39366	63.03	-2.63634	4.18%	12.1592
13:00	60.40829	58.77917	1.629126	2.77%	17.21197
14:00	59.30146	57.37889	1.922576	3.35%	9.339123
15:00	60.2	59.61593	0.584071	0.98%	8.427365
16:00	61.85902	59.31088	2.548145	4.30%	33.82436
17:00	62.35341	70.79547	-8.44206	11.92%	57.8961
18:00	62.32585	59.18652	3.139331	5.30%	20.35702
19:00	63.43122	59.78662	3.644599	6.10%	59.46309
20:00	62.29561	63.2753	-0.97969	1.55%	63.93373
21:00	60.33049	55.30129	5.029203	9.09%	65.26557

k-NN 模式旅行時間預測整理(970308(六))

時間	預測值 (分)	實際值 (分)	誤差值 (分)	誤差率 (%)	變異數
05:00	59.67732	64.51667	-4.83935	7.50%	34.23878
06:00	62.92585	55.43333	7.49252	13.52%	74.24427
07:00	63.48805	71.93	-8.44195	11.74%	99.41214
08:00	63.62829	60.55227	3.07602	5.08%	98.8784
09:00	63.23927	75.35152	-12.1122	16.07%	87.59885
10:00	63.19073	64.09461	-0.90388	1.41%	81.27873
11:00	61.46049	60.94392	0.516572	0.85%	49.65348
12:00	60.3022	65.91349	-5.61129	8.51%	13.25599
13:00	60.8378	63.00206	-2.16425	3.44%	34.39317
14:00	59.30146	71.44	-12.1385	16.99%	9.339123
15:00	60.2	60.92981	-0.72981	1.20%	8.427365
16:00	61.96195	62.53651	-0.57456	0.92%	34.5413
17:00	62.34976	63.66306	-1.3133	2.06%	35.96177
18:00	62.32585	69.64307	-7.31721	10.51%	20.35702
19:00	63.98463	60.05606	3.928573	6.54%	57.8174
20:00	61.81756	57.39223	4.425334	7.71%	65.11891
21:00	59.78268	67.9803	-8.19762	12.06%	56.70156
22:00	59.50829	54.95548	4.552814	8.28%	57.93485
23:00	53.29341	57.56012	-4.26671	7.41%	90.06584

k-NN 模式旅行時間預測整理(970316(日))

時間	預測值 (分)	實際值 (分)	誤差值 (分)	誤差率 (%)	變異數
05:00	59.51843	56.68889	2.829542	4.99%	54.73985
06:00	59.58098	55.23333	4.347646	7.87%	55.50018
07:00	60.14941	56.73333	3.416078	6.02%	81.77313
08:00	60.20216	64.02055	-3.8184	5.96%	78.47968
09:00	63.03804	62.34528	0.692759	1.11%	87.59885
10:00	62.4998	58.9837	3.5161	5.96%	81.94361
11:00	61.01176	64.24514	-3.23338	5.03%	49.65348
12:00	60.18314	59.40857	0.774568	1.30%	10.85384
13:00	58.63373	57.95139	0.682336	1.18%	9.255779
14:00	59.56922	57.04548	2.523737	4.42%	9.339123
15:00	60.86608	65.39675	-4.53067	6.93%	8.427365
16:00	61.83902	62.48769	-0.64867	1.04%	34.5413
17:00	62.57333	67.6551	-5.08176	7.51%	35.96177
18:00	62.44118	66.9667	-4.52552	6.76%	20.35702
19:00	63.55843	69.38439	-5.82596	8.40%	57.8174
20:00	61.35157	61.23906	0.112513	0.18%	65.11891
21:00	59.45588	56.77148	2.684405	4.73%	58.04229
22:00	59.38078	56.23139	3.149392	5.60%	6.054254
23:00	54.17157	54.74088	-0.56931	1.04%	90.06584

簡 歷



姓名：蔡繼光

籍貫：臺灣省基隆市

生日：民國 72 年 2 月 16 日

學歷：民國九十八年七月國立交通大學運輸科技與管理學系碩士班畢業

民國九十四年七月國立交通大學運輸科技與管理學系學士班畢業

電郵信箱：bright7722@gmail.com

