

國 立 交 通 大 學
運 輸 科 技 與 管 理 學 系 碩 士 班

碩 士 論 文



應用模糊分群法推估路段速率之研究

The Study of Using Fuzzy C-Means Method to Estimate the Link Average Speed

研 究 生： 洪 從 慑
指 導 教 授： 王 晉 元

中 華 民 國 一 百 零 二 年 七 月

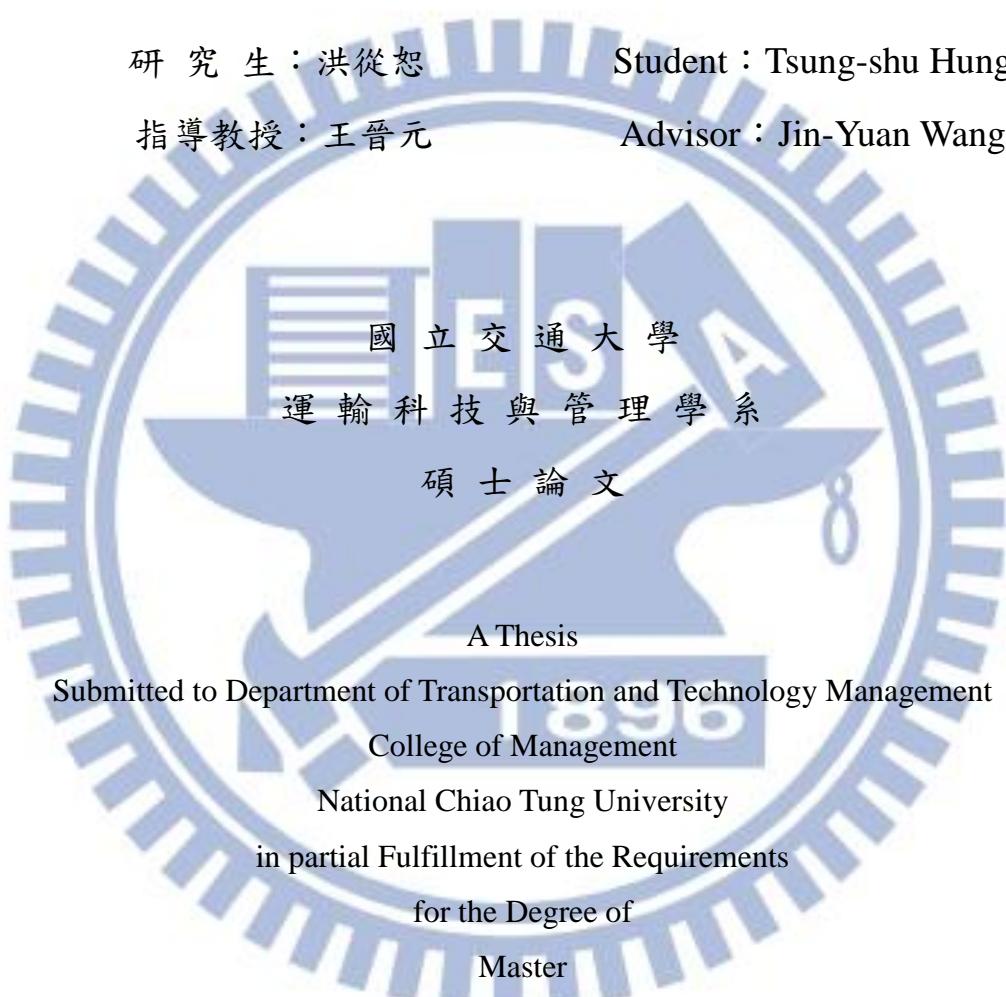
應用模糊分群法推估路段速率之研究
The Study of Using Fuzzy C-Means Method to
Estimate the Link Average Speed

研究 生：洪從恕

Student : Tsung-shu Hung

指 導 教 授：王晉元

Advisor : Jin-Yuan Wang



June 2013

Hsinchu, Taiwan, Republic of China

中 華 民 國 一 百 零 二 年 七 月

應用模糊分群法推估路段速率之研究

研究生：洪從恕

指導教授：王晉元

國立交通大學運輸科技與管理學系碩士班

摘要

在先進旅行者資訊系統(Advanced Traveler Information System, ATIS)中,即時路況所扮演的角色相當重要。特別是在尖峰時間或是道路壅塞時,即時路況的需求更為顯著。透過即時路況資訊的提供,可有效引導用路人選擇合適的路徑與出發間,降低運輸成本並提昇旅行效率。探針車是路況的重要來源之一,但若使用非專用探針車隊,例如公車、計程車等,因其有執行任務的需求,將會回傳與路況不一致的資料；另外,當發生嚴重擁擠而產生走走停停的現象時,所回傳的資料型態與一般狀態有著顯著的不同。本研究透過模糊分群法(Fuzzy C-Means, FCM)處理探針車回傳的資料,並在不同的路況情境下,分別建議合適的分群與計算方式。根據本研究分析結果,當路段旅行平均速率大於30km/hr(或服務水準等級為B級以上)時,可採用「分兩群取高速群」或是「分三群取高速群」的分群方式;而當路段旅行平均速率小於30km/hr(道路等級為C級以下),則建議採用「分三群取中速群」或是「直接平均」。

關鍵詞：即時路況、非專用探針車、FCM、模糊分群法

The Study of Using Fuzzy C-Means Method to Estimate the Link Average Speed

Student : Tsung-shu Hung

Advisor : Jin-Yuan Wang

Department of Transportation Technology and Management
National Chiao Tung University

ABSTRACT

Providing real-time traffic information is a very important and fundamental service in ATIS. Based on this information, travelers can change their departure time or route to decrease travel cost and/or improve travel efficiency. Probe Vehicle is one of the major means to obtain traffic situation. But there are some data filtering issues need to be addressed for the non-dedicated probe vehicles, such as bus and taxi. The data generated while service is ongoing, such as passengers are getting on and off, need to be filtered out since these data do not reflect the real situation. Moreover, when the traffic is severely jammed and having the stop-and-go traffic pattern, special cares are required to rule out the unnecessary data. We proposed a Fuzzy C-Means based method to address the above mentioned issues. Our testing results show that when the average travel speed is higher than 30 km/hr, we suggest dividing the data into two or three groups and take the average of high-speed group. When the average speed is lower than 30km/hr, categorizing data into three groups and takes the middle-speed group is recommended.

Keywords: real-time traffic, probe vehicle, Fuzzy C-Means

誌謝

終於，輪到我寫這一篇誌謝了。兩年的光陰過的飛快，加上四年的大學生涯，在交大的這六年時間，一眨眼居然也接近了尾聲。本篇論文得以完成，首先要感謝的是我的指導老師 王晉元教授，從大四的專題開始，老師對於我的研究，總是循循善誘、不遺餘力的教導其所學，也給了我很多學習上應有的觀念，很感謝老師對於我的想法與作法所表現的尊重，只要思考方式合乎邏輯，執行的結果通常可以獲得老師的肯定，也因此每一次的 meeting 對我而言，都是一件愉悅的事情。此外，除了在學習上的指導，老師也會主動的關心我在生活上的大小事情，如同我在學校的另一個父親，古人說「一日為師，終身為父」以及「亦師亦友」這兩句話，我很慶幸，能夠在我大學以及研究所生涯中能夠真切的體會到，實在難得，對於老師除了感謝之外就是感恩了。

在論文口試期間，感謝 蘇昭銘教授以及 盧宗成教授能夠撥冗閱讀學生的文章，並給予相當寶貴的建議，使得學生的論文研究能夠得以更臻完整，誠摯感謝兩位老師的辛苦付出與指教。

再來就是我實驗室的好夥伴們，感謝 hoho 純予我很多對於人生處事的態度，以及關於電子產品的相關指導，還有不定時的工作機會，讓我對於自己的人生有一些不同的啟發；感謝已經畢業的珮慈學姐還有阿幹，即使畢業了也一樣的關心著我們；感謝老頭，讓我知道學習是對自己負責、不可以渾渾噩噩；感謝育凡、玥心、承勳、邱辰這幾位同甘共苦的好友們，謝謝你們，跟你們聊天總是可以讓我很開心，讓苦悶的研究生活多了幾分開心的因素；感謝小倩、Peter、逸彥、佳伶這幾位可愛的學弟妹們，你們的幽默風趣和處事態度，相當的符合這個 Lab 對你們的期待，也希望能繼續保持下去；感謝 Haru，雖然我知道你看不懂，但你是我生平第一隻可以互動寵物，是你讓我知道，原來貓咪是很可愛的，祝福你在新的環境能夠一樣適應得很好。還要特別感謝地錢，室友當了又當，總是幫我很多又不計較，也會跟我分享一些有趣的八卦，在同學六年的生涯裡面，很高興有你這樣的朋友在身邊，謝謝你。感謝小蔚，在我快畢業的時候才進入了我的生活，跟你在一起總是相當的自在，而屬於我們的心靈默契，更是我不曾有過的經驗，謝謝你，也希望未來的日子裡，還能夠有你繼續陪著我走下去；還有金七國際事業股份有限公司，是一起成長的好夥伴。

最後我要感謝我的家人，謝謝爸爸媽媽、哥哥嫂嫂沛璟和姊姊，不僅在我求學期間提供我不於匱乏的所有資源，並且給予我最大的鼓勵與支持，讓我感受到家是最溫暖的地方，能夠被你們呵護、照顧著是全世界最幸福的事，還有阿公、阿嬤、舅舅、舅媽、阿姨、姨丈以及所有兄弟姊妹等族繁不及備載，你們是我最大的靠山，感謝有你們的疼愛和陪伴才有現在的阿恕。

最後的最後，也就是文末。由於要感謝的人實在太多，謝天真的一點也不為過。感謝天公伯能夠給我生命而存在這個世界、這個地方，進而認識了這麼些人，未來也希望，我自己不能夠照顧自己也能夠照顧周遭的人，讓有我的地方就有歡樂。

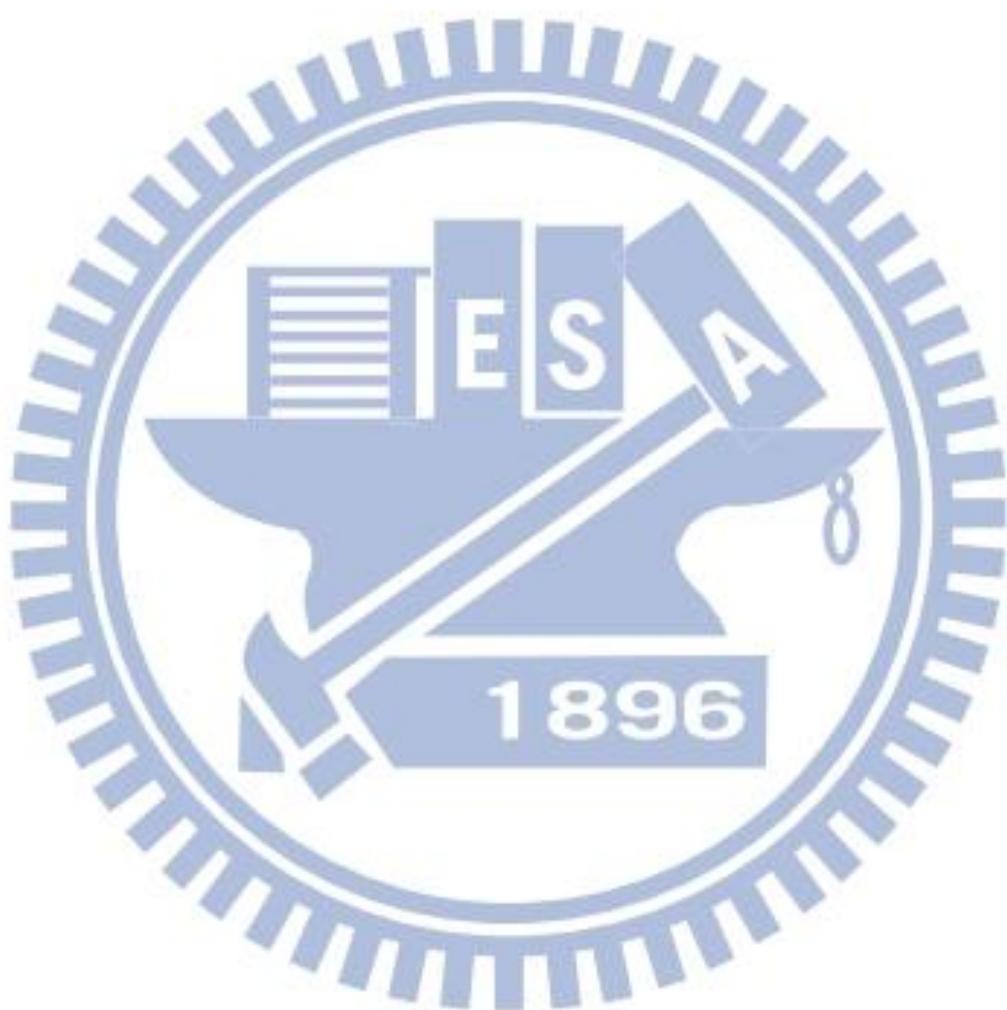
洪從恕 謹致
2013 七月 於 交大 ITS Lab

應用模糊分群法推估路段速率之研究

目錄

摘要.....	i
ABSTRACT.....	ii
誌謝.....	iii
目錄.....	IV
圖目錄.....	VI
表目錄.....	VI
一、 緒論.....	1
1.1. 研究背景與動機.....	1
1.2. 研究目的.....	2
1.3. 研究流程.....	2
二、 文獻回顧.....	5
2.1 探針車與路況推估.....	5
2.2 小結.....	6
三、 研究內容與方法	8
3.1 問題特性分析.....	8
3.2 系統架構.....	9
3.3 群集分析.....	11
3.3.1 模糊分群法 (FUZZY C-MEANS, FCM)	11
3.4 資料分群處理.....	12
3.5 評估指標.....	13
3.5.1 平均絕對值誤差百分比 (MEAN ABSOLUTE PERCENTAGE ERROR, MAPE)	13
3.5.2 均方根誤差 (ROOT MEAN SQUARE ERROR, RMSE)	13
3.5.3 平均絕對誤差(MEAN ABSOLUTE ERROR, MAE)	14
四、 數值分析.....	15
4.1 模擬軟體.....	15
4.2 路網構建.....	15
4.3 模式驗證與分析.....	17
4.3.1 與實際資料比較結果.....	17
4.4 道路情境影響與分群適用時機.....	18
4.5 分車班次之影響.....	21
4.6 回傳頻率之影響.....	24
五、 結論與建議	27

5.1	結論.....	27
5.2	建議.....	27
5.3	建議 FCM 之應用	28
六、	參考文獻.....	29
七、	附錄.....	30



圖目錄

圖 1 研究流程圖	4
圖 2 探針車資料回傳示意圖（離峰）	8
圖 3 探針車資料回傳示意圖（尖峰）	9
圖 4 系統架構圖	10
圖 5 模擬路網圖	16

表目錄

表 1 本研究與文獻差異對照表	7
表 2 MAPE 預測百分比對照表	13
表 3 模擬參數設定表-輸入的車流量	16
表 4 評估結果比較表	17
表 5 速限 50 公里/小時之市區道路服務水準等級劃分標準	19
表 6 分群結果比較表-依道路情境區分	19
表 7 分群方式檢定表-兩群高速與三群高速	21
表 8 分群方式檢定表-三群中速與直接平均	21
表 9 區分車班次比較表-建議分群方式	22
表 10 分車班次統計檢定表-分兩群取高速群	23
表 11 分車班次統計檢定表-分三群取高速群	23
表 12 分車班次統計檢定表-分三群取中速群	23
表 13 區分車班次比較表-其他分群方式	24
表 14 不同頻率下的分析資料表-15 秒	25
表 15 分群方式統計檢定表-分三群取中速群與直接平均	26

一、緒論

1.1. 研究背景與動機

先進旅行者資訊系統（Advanced Traveler Information System, ATIS）為智慧型運輸系統（Intelligent Transportation System, ITS）的子系統之一[1]。透過 GPS、車輛偵測器、影像處理技術等方式所收集到的即時資料，經資訊中心處理後產生出可用資訊，如路況、旅行時間、替代道路、路徑指引、路線（運具選擇）、出發時間等，再經由可變標誌（CMS）、廣播系統、智慧型手機等發佈資訊，提供給使用者，藉以提升旅行時間的效率，減少不必要的等待時間。

在 ATIS 中，即時路況的發佈是相當重要的一環。良好的路況資訊提供，可作為用路人決策的依據，特別是在尖峰時間或是道路壅塞時，即時路況的需求更為重要[1]。透過即時的路況資訊，可引導用路人選擇替代道路或是更改出發時間，來避開壅塞路段與尖峰時間或是依照自身需求選擇合適運具、最短路徑或是以旅行時間最短的方式到達目的地，可有效的降低運輸成本，提升旅行效率。

目前即時路況資料來源除了透過民眾、國道警察等單位之通報外，路況資料收集方式主要有（1）透過探針車（即裝載 GPS 設備之車輛）所回傳的即時行車速度；（2）透過車輛偵測器（Vehicle Detector, VD）提供車速；（3）透過自動車輛辨識（Automatic Vehicle Identification, AVI），所提供之兩點間的旅行時間，再推算該路段的旅行速度。

從設備建置的成本考量而言，透過探針車來收集路況屬於成本較低廉且操作簡單的方式[2]。現行道路若要全面佈設 VD 或是 AVI 設備，需要的設備安裝與維護費用較高昂；反觀探針車，隨著智慧型公共運輸系統的推廣，無論是在國道或是市區道路，均可作為收集路況之管道[3]。因此利用裝載 GPS 設備的探針車來收集路況資料，是一個在短期內可行且既有效率又即時的方式。

國內目前沒有專用的探針車隊，而裝載 GPS 設備的車輛又多為公車、客運、計程車或物流車隊等，為非專用的探針車。由於這些車種具有其主要的運輸目的，因此會產生如載客、巡行、停靠等行為，若將這些車輛所回傳的資料直接予以平均，作為路況發佈的依據，則容易因其特殊的運輸行為，使得整體的平均速度降低，便將無法正確的反應實際路況。

除此之外，當尖峰時間或是發生道路擁擠情況時，由於車輛會產生走走停停的現象，若透過 GPS 定時回傳資料的特性，可能會收集到一些車輛正在加速中的速度資料，而這樣的速度往往與實際的道路狀況並不相符，故需先透過

資料處理，將與實際路況不相符的資料排除，才能使得回傳的資料能夠切確的反應出真實的路況。

Lee 等人[4]曾透過模糊分群法（Fuzzy C-Means, FCM）的分群方式，將計程車回傳的 GPS 資料先建置歷史資料庫，從歷史資料庫裡面依照速度分成高速、中速與低速三群，再利用車輛即時回傳的資料與歷史資料庫的分群結果比對後給予不同的權重值，再將其平均作為路況資訊，該篇研究證明了透過 FCM 的分群方式所處理後的資料，在與實際路況比較的正確程度，優於將回傳資料直接平均。然而，在資料處理方式對於不同的道路情境（如：自由車流情況或擁擠情況等），是否具有通用性並無特別說明，因此，為了解道路情境與分群之間的關係及必要性，本研究將透過實測驗證的方式，針對此點進行更進一步的討論。

對於 FCM 而言，分群數量的決定或是該選擇哪一個群組作為資料使用的依據以及分群數量與不同道路狀況的適用情形，都是值得討論的議題。本研究期望能整理出一套方式，能夠有效的將 FCM 法應用在公車的回傳資料上，使得這些資料值能夠有效的反映出當時候的道路情況，使推估出的路況結果能更符合實際路況。

1.2. 研究目的

本研究希望透過群集分析中的資料分群方式，對於公車輛所回傳的資料進行分析，先了解資料分群的關係，再決定分群之後，各群組之間的資料選擇，再探討不同分群數量與不同的道路情境並找出分群數量改變的適用時機，來處理與實際道路路況不相符的資料，使得探針車所回傳的資料，經處理後的推估結果能更符合實際道路的路況。

1.3. 研究流程

本研究流程如圖 1 研究流程圖，分別說明如下：

確認研究問題

將此研究問題做明確界定，包含研究目的、範圍以及方法。

相關文獻回顧

依照問題性質分類及回顧國內外相關之文獻，了解其研究重點與方向以建立本研究之基礎。

模式方法構建

依照選擇的方法建立可行之路段速度分群模式。

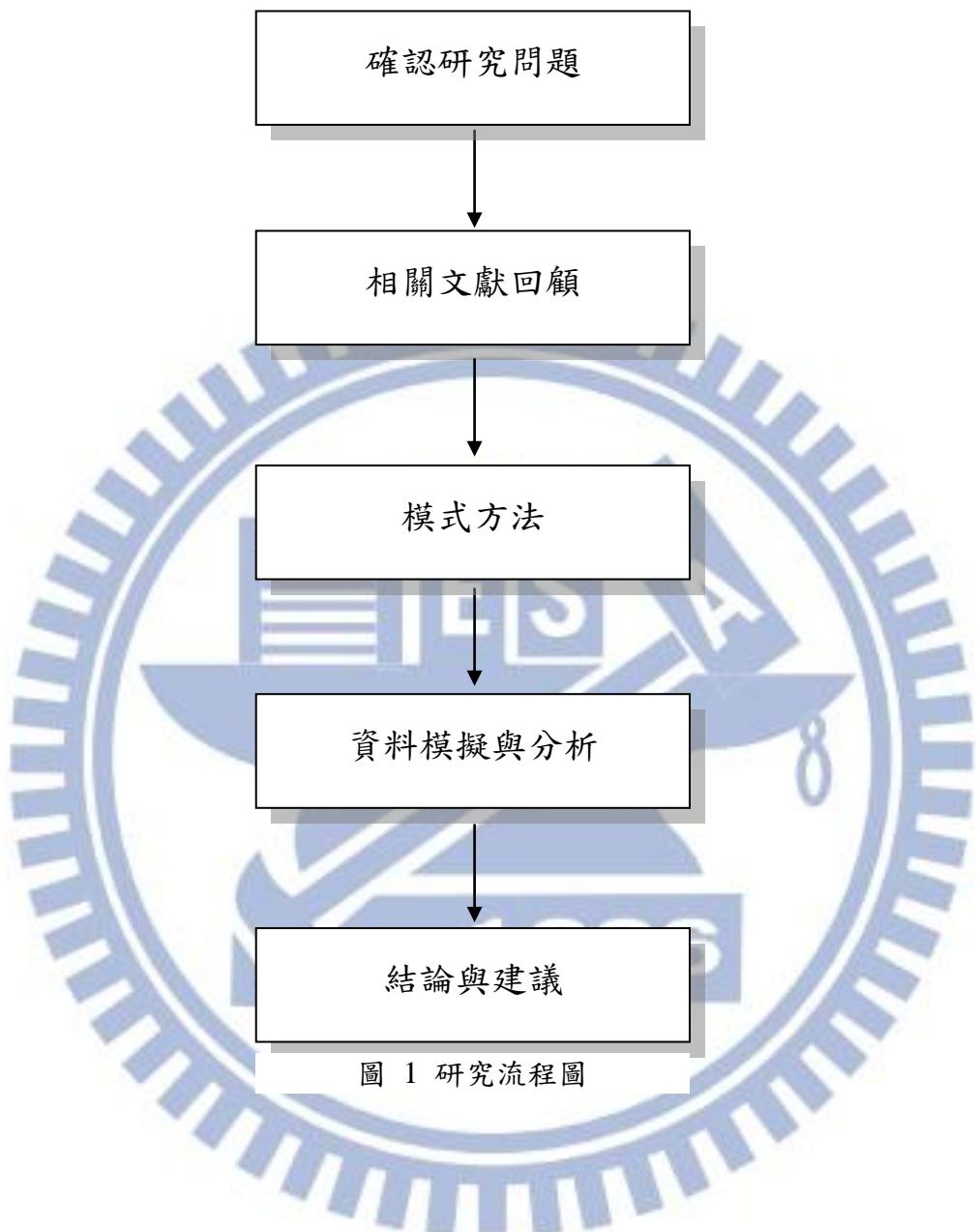
資料模擬與分析

透過模擬軟體模擬不同的道路情境，並根據所收集到之資料進行相關程序分析，以驗證方法之正確性與合理性。

結論與建議

對研究方法、過程與結果做總結，並提出結論與建議。





二、文獻回顧

本研究主要探討非專用探針車輛，針對其回傳的資料，透過群集分析進行處理，用以推估路段速率。在進行研究之前，本章將針對探針車輛相關文獻進行回顧，並根據其結果作為本研究方向之參考。

2.1 探針車與路況推估

Taylor [5]等人認為衡量交通系統的表現好壞有相當多的方式，但從用路人的觀點而言，交通壅塞是最主要關心的對象。因此提出 7 種量化的方式，來衡量交通壅塞的情形。而該篇研究也指出，相較於安裝在地面的車輛偵測器，透過 GPS 探針車來收集資料，是一種更便宜又方便的替代方案。該研究利用裝載 GPS 設備的小客車，針對其回傳的經緯度與速度資料，進行三週的資料收集，並根據尖離峰的路況資料進行分析，研究發現透過 GPS 的所回傳的資料，不僅是一個正確且符合成本效益的方式，且對於路況壅塞時也可提供可靠的資訊。

Lee 等人[6]認為路段速率彼此獨立、量測方式較簡易且客觀，可應用在轉換成旅行時間或是衡量服務水準。因此提出了一個衡量方式，使得在不同交通條件的情況下，能夠了解探針車的比例與推估路段速率的誤差關係。該篇研究利用新加坡超過 10,000 輛裝載 DGPS 的計程車實際回傳的資料值作為資料來源，並透過 INTERGRATION 模擬軟體，模擬不同情境並產生資料，採用空間平均速度(Space-Mean-Speed)方式計算路段平均速度。在相同的起訖點間透過 6 種不同的交通量以及 6 種不同的探針車輛佔有比例，進行研究測試。研究結果發現，在路網中若探針車的比例佔整個路網的 4-5% 或至少有 10 輛探針車經過該路段，在 95% 的信心水準中，速率推估的絕對誤差將會小於 5km/hr。

Uno 等人[2]利用公車作為探針車輛，並透過其回傳的資料值，用來分析旅行時間的變異性以及量測道路的服務水準。該篇研究針對公車在執行任務時特有的行為表現（如停靠站牌），說明必須透過一套處理方式，將公車進入（或離開）停靠站時所產生的旅行時間或是速度資料予以過濾，才能校正旅行時間的推估結果，透過該篇研究的過濾方式，可以判斷出進入（離開）停靠站的回傳資料，且準確度可高達 92%，因此對於分析旅行時間的變異性與道路的服務水準的結果，皆產生了良好的表現。

Lee 等人[4]在推估路段速度的資料分類上指出 GPS 的資料特性與路面上的車輛偵測器所傳回的資料型態不同，因此需針對 GPS 的資料特性發展一套新的使用方法，該篇研究利用模糊分群法（Fuzzy C-Mean, FCM）演算法進行資料測試，將資料分為高速、中速、低速三群，資料來源從裝載 GPS 設備的計程車輛取得，以每 5 分鐘的方式收集資料筆數，該篇研究認為如果單純以 GPS 收集

到的資料進行時間速度平均（Time-Mean-Speed）作為發佈的資訊，在發佈上會有所影響，故針對資料速度屬於高速的部份為優先考量，並進行資料測試比對。研究結果發現，經 FCM 方法所處理後的資料，與實際的路段速度資料進行比較後，其結果優於將 GPS 速度資料直接平均。

2.2 小結

經由文獻回顧可知，利用探針車輛作為推估路段速率的資料來源，是一種便宜且方便的替代方案，也常用於路況推估、旅行時間或是衡量道路的服務水準。然而，如果資料來源從非專用的探針車中取得（如公車、客運、計程車等），由於探針車本身具有其應執行的任務，則會產生與一些不適用資料，需先予以處理方能使用。

Uno 等人利用自訂的過濾規則，當公車速度小於 3km/hr 且距離公車站牌小於 23.7m 時，所產生的資料予以濾除，作為調整後續推估旅行時間變異的方法，但如果遇到公車因前方停靠站有違停車輛或有其他公車已先行停靠，則會有提前停靠或是延後停靠的可能性產生，而此時的過濾方法效果將可能有限。而 Lee 等人對於計程車回傳資料做過 FCM 的分群分析，但並沒有明確地說明此方法所適用的路況條件，無法有效的說明此種分群方式的適用性，只透過數據證明應用 FCM 方法的結果會優於將資料直接平均。

因此，本研究認為以非專用探針車所回傳的 GPS 資料中，採用合適的分群方法，應可以有效的區分探針車因執行任務（如停靠站牌），所產生與一般車輛行為不相符的低速資料，以提升資料回傳處理後的準確度。

由於實際的路況速度資料不易取得，且為了可分析各種不同情境下的路況，故本研究將以模擬軟體產生路況及探針車的回傳資料，並參考 Lee 等人所應用的 FCM 分群方式進行資料分析，本研究與 Lee 的文獻差異如表 1 所示，本研究選用的資料來源類型為公車，並透過 FCM 的方式處理即時回傳的資料，並討論在不同車流情境下的模擬內容，了解分群方式的適用性，最後以預測評估指標以及統計的方式加以驗證其結果的優劣，使分析的結果能更接近實際的路況資料。

表 1 本研究與文獻差異對照表

	Lee et al (2006)	本研究
資料來源種類	計程車	公車
研究方法與內容	FCM 處理歷史資料 + 即時資料	FCM 直接處理即時資料
分群數	三群	兩群、三群
適用情況討論		有
參考路網	實際	實際
結果評估方式		評估指標



三、研究內容與方法

3.1 問題特性分析

本研究利用具有 GPS 設備的公車所回傳的資料進行研究，然而公車並非專用的探針車輛，利用其回傳的資料提供路況時，將會產生一些問題，例如公車本身具有載客的任務，所以當公車在執行任務時，常會有不同於正常車輛的行為表現（如停車載客等），若是直接將回傳的速率資料加以平均作為路段的平均速率，可能會造成推估結果與實際路段速率不相符的誤差產生，進而影響推估速率的結果表現。

圖 2 為離峰時段探針車所回傳的資料表現，其中縱軸為速率，橫軸為時間，該圖顯示公車每秒回傳一次的速率資料結果，從圖中可以觀察出回傳的速率資料不僅有明顯的高低起伏且探針車的最高速率在 40km/hr 以上，符合一般公車在離峰時的速度表現，而圖中低速或是速度 0 的點則是因為探針車因為載客或是遇號誌所產生的停等紀錄。

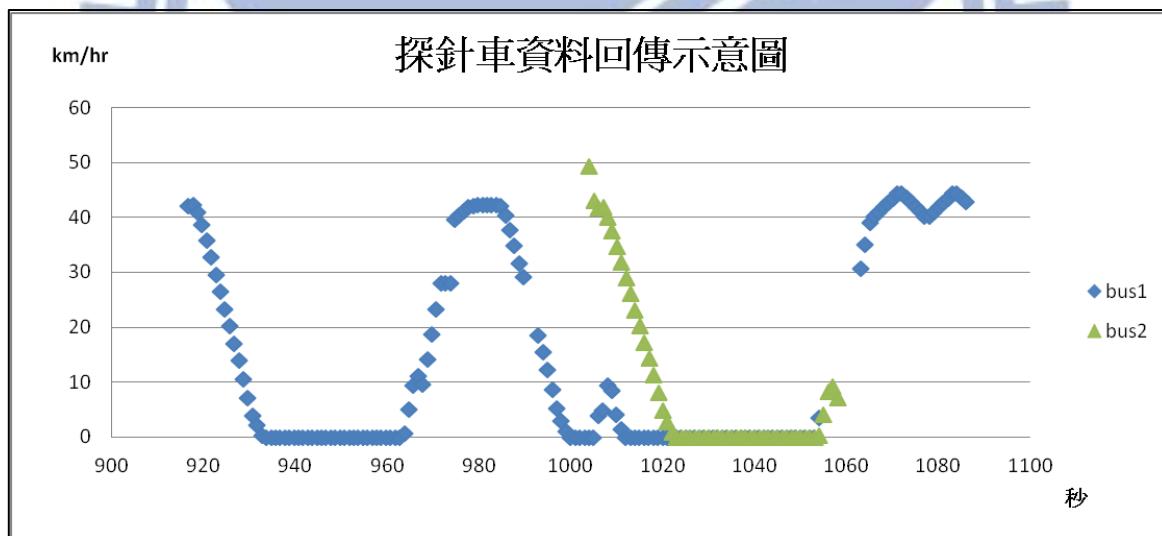


圖 2 探針車資料回傳示意圖（離峰）

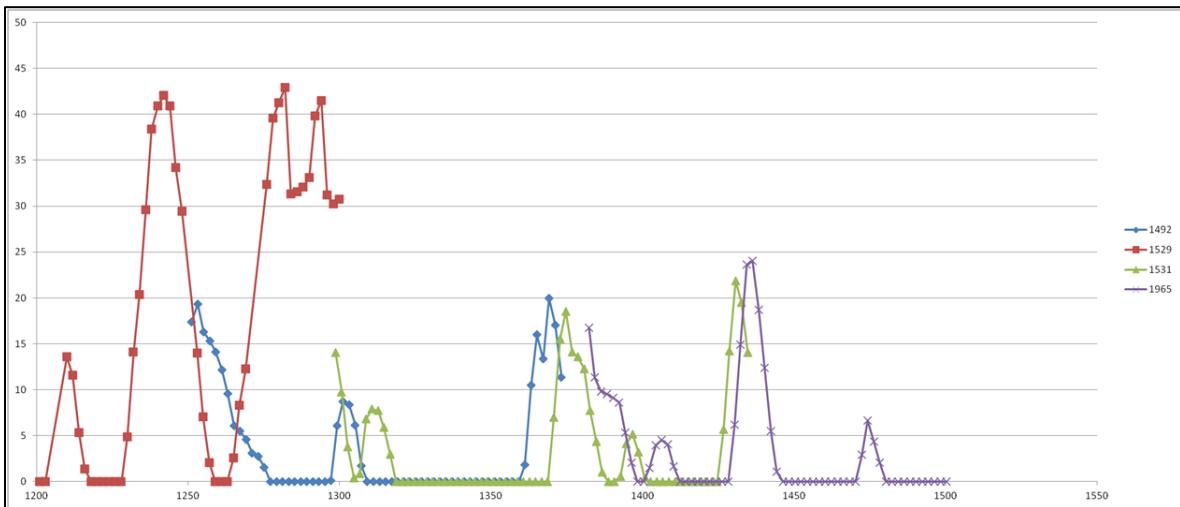


圖 3 探針車資料回傳示意圖（尖峰）

此外，尖峰時間資料的回傳表現亦有所不同，如圖 3 所示，同樣地，橫軸代表時間，縱軸代表速率，該圖顯示同樣為公車每秒回傳一次的速率資料結果，右邊圖例則代表探針車的編號，共有 4 台探針車。從圖 3 可以發現，除了編號 1529 的探針車所回傳資料因低速或速率 0 的資料比例較少，且最高速率皆在 40km/hr 以上，較貼近離峰時間的表現之外，其他三輛探針車回傳的資料值，速度為 0 的資料佔整體回傳量的比例相當高且最高速率皆小於 25km/hr，即為道路情況壅塞時所產生的回傳現象。

因此，藉由上述兩項觀察結果可以推論，在離峰時，若直接平均所有的資料回傳值，將會明顯低估整個路段的平均速率，造成嚴重的誤差，將無法確實反應實際的道路路況；而在尖峰時間，車輛容易因為塞車而產生走走停停的現象，由於無法明確判斷回傳的資料值，是否為正在加速中的速度，便無法切確表示其路況。所以，若能夠透過一套方法處理因載客或是遇路口號誌所產生停等的低速資料，亦或是能有效的保留符合實際路況的資料，對於路況推估的結果都將會有所幫助。

3.2 系統架構

針對路況推估計算之方法，本研究的系統架構如圖 4 所示。說明如下。

(1) 公車之 GPS 定位資料

針對每一公車在路段速率更新時間內，所回傳的所有紀錄時間、經緯度、速度等定位資料。

(2) 擷取路段資料

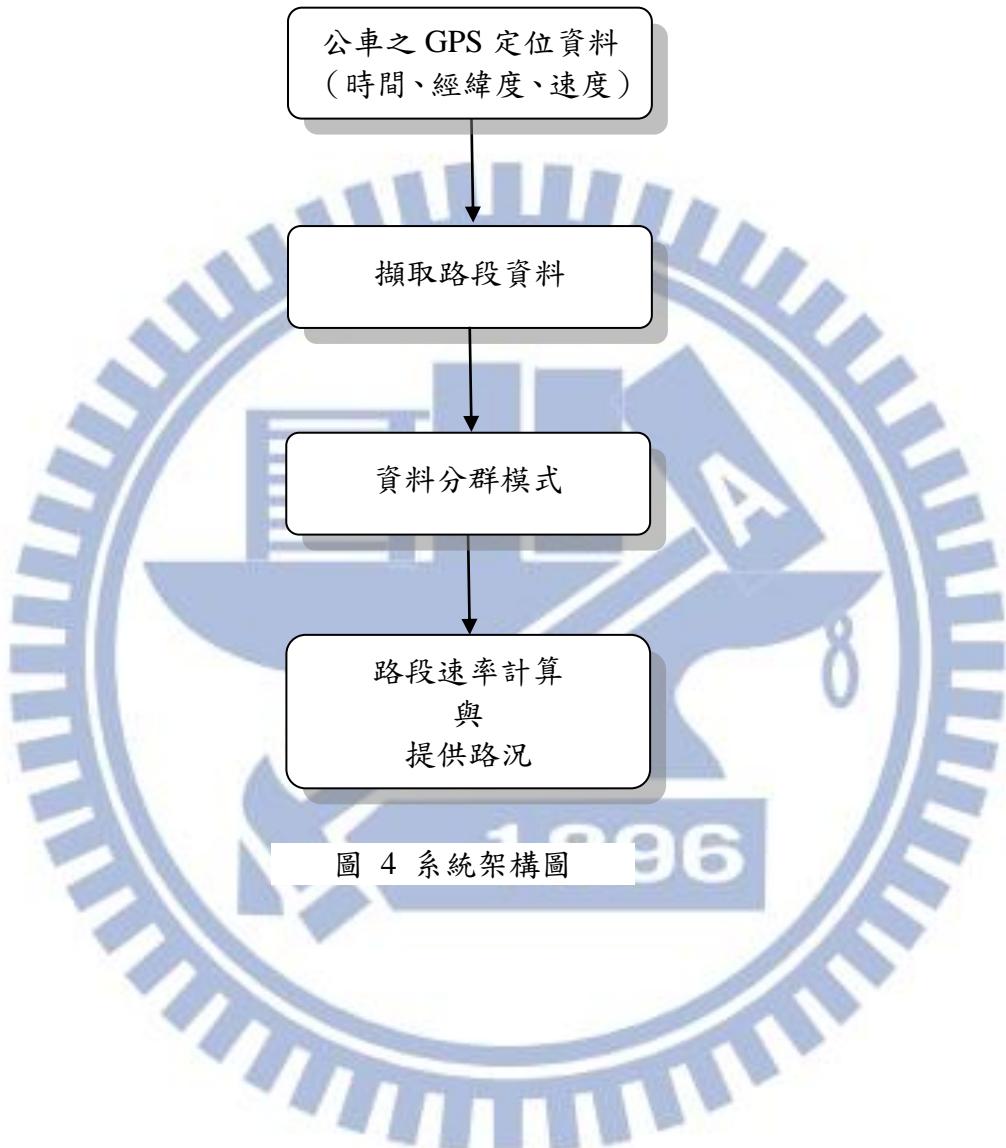
針對特定路段擷取所有經過的公車資料。由於本研究主要在提供路段速率，所以必須先選取好特定的路段，就本研究而言，路段的可視為一固定長度的市區幹道，其中可包含多個路口，並在路段速率更新時間內，將所有經過該路段的公車資料擷取出來。

(3) 資料分群模式

收集完所有路段中的公車資料之後，進行資料群集分析，透過選擇不同群組內的資料，找出最符合該路況的速度群。

(4) 路段速率計算與提供路況

最後將所選擇的群組資料進行平均，作為提供的路況速率。



3.3 群集分析

群集分析 (Cluster Analysis) 是資料探勘 (Data Mining) 中常用的方式之一。其目的是透過資料彼此間的”距離(Distance)”或是”相似程度 (Similarity)”進行分群，期望區分出同一群集內的資料同質性高，而不同群集間差異性高的分群結果，由於分群過程中是依據資料屬性與彼此間的距離程度或是相似程度來區分，單純討論資料之間的相互關係，屬於較客觀的計量方式[7]。

由於探針車資料回傳的特性，會因為其執行任務的不同，而產生與一般車輛不同的行為表現，因此本研究希望透過分群方式，能有效的區分速度差異，留下較可能反映實際道路情況的速度群來代表路況。回顧文獻後發現 FCM 法可適用於此一情況，故本研究將以 FCM 為主要的分群方式，並應用此方法在以公車作為探針車所回傳的資料上，將分群結果與實際路況做比較，並透過不同的道路情境找出分群數量的適用時機。

3.3.1 模糊分群法 (Fuzzy C-Means, FCM)

模糊分群法 (Fuzzy C-Means) 簡稱 FCM，是根據 C-Means (或 K-Means) 衍生而來的分群法，由 Dunn 在 1973 年提出[8]，並由 Bezdek[9]在 1981 改良成目前常見之方法架構。Fuzzy C-Means，其目標函數定義如同 C-Means，但其原本的權重矩陣 W 不再是二位元矩陣，而是應用了模糊理論的概念，使得每一輸入向量不再僅歸屬於某一特定的群集，並透過權重的迭代計算其歸屬程度來表現屬於各群集的程度。

Fuzzy C-Means 所使用到的參數與表示意義如下：

w_{ji} ：點對各群中心點的權重，其中 i 為群組， j 表示第 j 筆速度資料。

C_i ：第 i 群的中心點

C_s ：各群的中心點

X_j ：表示第 j 點的速度資料

J ：目標值

ε ：誤差值

m ：為大於 1 的權重指數，通常設定為 2

Fuzzy C-means 的應用步驟如下[4]：

- (1) 決定群組數：事先決定速度資料分群的群組數量，即要分成多少群。
- (2) 決定所有點對各群組的權重：每筆回傳的速度資料，會根據群組數量，而有不同個數的權重值，其中起始權重值為任意給定，而每筆速度資料與各群組間的權重值和為 1，如式 1。

$$\sum_{i=1}^k w_{ji} = 1, \forall j = 1, \dots, n \quad (1)$$

(3) 計算中心點：依照每筆速度資料與權重值的乘積以及所有權重值間的關係，找出群組的中心點，如式 2。

$$C_i = \frac{\sum_{j=1}^n w_{ji}^m X_j}{\sum_{j=1}^n w_{ji}^m} \quad (2)$$

(4) 計算目標值 J ：根據每筆速度資料與各群組中心點間的距離以及權重值，計算出目標值 J ，如式 3。

$$J = \sum_{i=1}^k J_i = \sum_{i=1}^k \left(\sum_{j=1}^n w_{ji}^m \|X_j - C_i\|^2 \right) \quad (3)$$

(5) 停止條件為 $|J_n - J_{n+1}| \leq \varepsilon$ ：當前次的目標值 J 與當次的目標值 J 兩者之間的絕對誤差小於設定的 ε ，則停止運算。

(6) 若未達停止條件則更新此點對中心點之權重，如式 4，並回到步驟 4：

$$w_{ji} = \frac{1}{\sum_{s=1}^k \left(\frac{\|X_j - C_i\|}{\|X_j - C_s\|} \right)^{\frac{2}{m-1}}} \quad (4)$$

3.4 資料分群處理

本研究採用的分群方式有兩種，分別是分兩群與分三群，並以人工的方式設定之。分兩群進行分群分析的原因為兩群為群集分析作業中，最小單位的分群數量；而分三群則是考慮若道路情況擁擠時，在車輛產生走走停停的情況下，若單只分成高低速兩群時，高速群可能會紀錄到車輛正在加速中的資料，而低速群則會紀錄速率低速或是 0 的資料，因兩群間的差異大，不易代表真實的路況，故考慮在兩群中間多增加一個群組來進行測試。

分兩群之後，本研究依其平均速度的不同，將平均速度較高的稱「高速群」，平均速度較低的稱「低速群」；同樣地，在分三群之後，依其平均速度的不同以「低速群」、「中速群」、「高速群」稱之。

由於分三群的方式所產生的「低速群」、「中速群」、「高速群」已足以提供本研究進行測試與討論，由於每一群的資料量可能不足，將會出現無法正確分群的可能，且分群之後尚有群組資料選擇的問題，若分群數量越多將會造成後續群組間資料選擇的困難降低計算效率，因此本研究選擇以分兩群與分三群為主，不再往後續分。

3.5 評估指標

在資料分析後的推估結果上，本研究採用文獻上常見的三種預測評估指標作為評估方式，分別是：平均絕對值誤差率（Mean Absolute Percentage Error, MAPE）、均方根誤差（Root Mean Square Error, RMSE）與平均絕對誤差（Mean Absolute Error, MAE）作為依據，由於前兩者屬於相對誤差值，不受單位與數值大小之影響，判斷依據較為客觀；而後者為絕對指標值，可用以判定估計值與實際值之間的差距程度，屬於較主觀之評估指標。以下將個別介紹這些方法：

3.5.1 平均絕對值誤差百分比（Mean Absolute Percentage Error, MAPE）

又稱為平均絕對誤差率，用以作為預測模式優劣之評估指標。主要因為 MAPE 為相對數值，不受測量值與預估值單位及大小之影響，能夠客觀得獲得估計值與實際值間之差異程度，可應用於本研究中了解推估結果的準確程度，其計算方式如式 5 所示：

$$MAPE = \frac{1}{M} \sum_{i=1}^M \left| \frac{x(i) - y(i)}{x(i)} \right| \times 100\% \quad (5)$$

其中， M 為樣本數

$x(i)$ 為第 i 項的實際值

$y(i)$ 為第 i 項的估計值

MAPE 值若越接近 0 表示估計效果越佳。此外，Lewis[10]則認為 MAPE 為最有效之評估指標，並訂定相關評估標準，如表 2 所示：

表 2 MAPE 預測百分比對照表

MAPE(%)	說明
<10	高準確的預測
10-20	優良的預測
20-50	合理的預測
>50	不準確的預測

3.5.2 均方根誤差（Root Mean Square Error, RMSE）

均方根誤差為各次估計誤差平方和的平均值方根，可以了解模式估計的可靠性，若均方根誤差越小，表示估計可靠程度越高；相反地，若均方根誤差越大，則表示模式估計可靠程度越低，可藉由此方法，了解本研究推估結果的可靠程度，其計算方式如式 6 所示：

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^M [x(i) - y(i)]^2}{M}} \quad (6)$$

其中， M 為樣本數

$x(i)$ 為第 i 項的實際值

$y(i)$ 為第 i 項的估計值

3.5.3 平均絕對誤差(Mean Absolute Error, MAE)

即平均每筆估計值與實際值的絕對殘差值，可以用來檢視本研究推估結果的誤差範圍，其計算方式如式 7 所示：

$$\text{MAE} = \frac{1}{M} \sum_{i=1}^M |x(i) - y(i)| \quad (7)$$

其中， M 為樣本數

$x(i)$ 為第 i 項的實際值

$y(i)$ 為第 i 項的估計值

由於本研究利用分群分析進行路況資料推估，並將結果與實際資料進行比較，為能多方面了解推估的結果，因此利用上述三種評估方式不同的特性，包含推估的準確程度(MAPE)、可靠程度(RMSE)以其誤差範圍(MAE)，作為本研究的評估指標。

四、數值分析

本研究主要利用非專用探針車所回傳的 GPS 資料進行分析，並透過分群分析方法處理與一般車輛不同行為（如載客停等）所回傳的資料，並依照不同道路情境找出做分群數量改變的適用時機。本研究擬以車流模擬軟體對於一般市區道路之車流行為進行模擬，並設定公車為裝載 GPS 設備的非專用探針車輛，作為 GPS 資料回傳的來源，藉此模擬方式討論在不同的道路情境下，分群數量與分群後群組之間的資料選擇，對於 GPS 所推估的路段速率有何影響，期望能有效利用探針車回傳的資料推估出較準確的路況。

透過車流模擬軟體來取得探針車的資料，為文獻上常用的工作法[6, 11, 12]。由於實際測試無法針對某些特殊狀況進行實驗，亦或是調查的時間及成本太高，而採用模擬的方式不僅可以節省成本同時也可產生出各種不同情境之資料以做測試。因此本研究將透過車輛模擬軟體來產生分群分析所需之資料。

4.1 模擬軟體

本研究使用 VISSIM 模擬軟體作為資料來源產生之工具。由於 VISSIM 屬於微觀模式的車流模擬軟體，可以針對單一車輛的個別行為與現象以及與其他車輛之間的關係作為分析基礎，著重於車輛與車輛之間的個別運作處理，可以提供個別車輛的期望速率、間距、相對速度、行進方向等參數，同時也可紀錄車輛在路口、路段的軌跡與駕駛特性[13]。由於本研究以公車作為非專用探針車，並透過其回傳的資料進行分析處理，考慮到本研究內容的特性，是以個別公車回傳的資料作為分析對象，所使用之模擬軟體必須要能針對不同種類之車種進行不同的設定，同時能夠對於不同道路狀況的參數也可進行調整，以便模擬出本研究所需之資料，由於 VISSIM 所具備的條件與本研究內容需求相符，故採用 VISSIM 模擬軟體做為主要的資料來源產生工具。

4.2 路網構建

本研究透過車輛模擬軟體 VISSIM 5.10 來構建道路路網，參數的資料來源為實際路口調查後所建置之路網資料。調查地點為新北市土城區中央路三段，由大暖路口起往東至大安路口止為模擬路段。

模擬路網圖如圖 5，模擬的道路長度約 800 公尺，在調查路段路口前後各約 100 公尺皆設有公車停靠站共兩站，編號 1 到 6 則為軟體中車流量的輸入點。

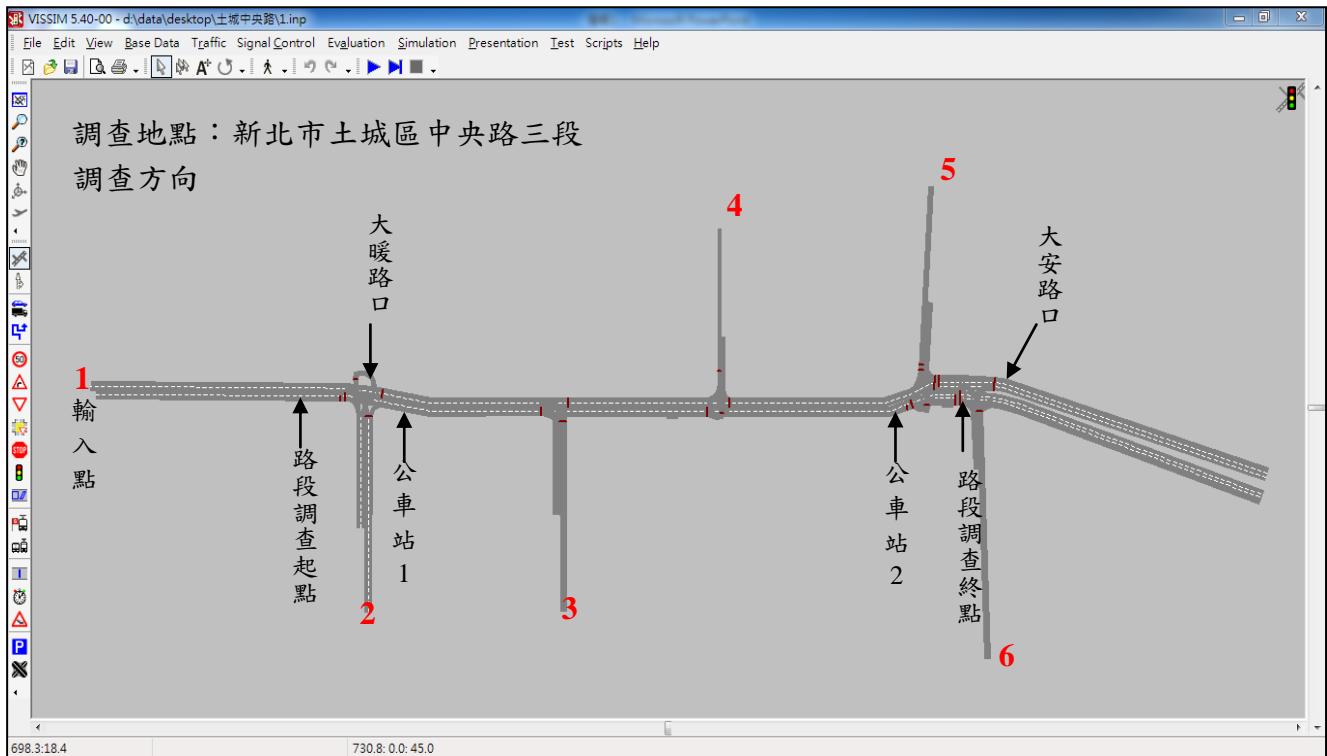


圖 5 模擬路網圖

該路網的實際車流比例，經調查計算後為小客車 48.2%、機車 47.9%、客運公車 3.9%；在模擬軟體中，可藉由透過改變路段的車流量來模擬不同情境下的路況。相關模擬的車流量資料，可參照表 3，編號為構建完成的路網資料中，進行車流量調整的位置，共 6 個輸入點；原始車流量為實際調查的車流量參數資料；改變車流量則為本研究自行輸入至模擬軟體中的車流量，改變的方式以原始車流量為基準，並以倍數增加、減半或是改變部份進入點的車流量等方式任意進行調整，共產生 6 種不同的模擬結果。

表 3 模擬參數設定表-輸入的車流量

編號	1	2	3	4	5	6	模擬結果
原始車流量	101	1227	164	107	430	180	1 & 2
改變車流量	303	2454	492	321	1290	360	3
	101	2454	164	107	430	180	4
	50	613	80	50	225	90	5
	100	1227	164	107	430	180	6

註：模擬次數 1 與 2，為不同的起始亂種數子。

(單位：車輛/小時)

本研究以每次模擬時間為 60 分鐘，每 5 分鐘為單位輸出資料，共 12 筆資料，並以相同方式模擬 6 種不同車流量的資料，故產生 72 個案例資料。每個案例資料均包含公車所回傳的速度資料以及該路段上所有小客車的平均旅行速率，並依小客車的平均旅行速率作為實際資料以及道路情境，以便後續作業討論在不同道路情境下的推估結果。

4.3 模式驗證與分析

4.3.1 與實際資料比較結果

為測試 FCM 的方法如在不同的道路情境下使用，是否具有通用性，因此本研究將 72 筆案例中的公車回傳資料，透過 FCM 法分成「分兩群取低速群」、「分兩群取高速群」、「分三群取低速群」、「分三群取中速群」、「分三群取高速群」及不使用 FCM 法將資料直接予以平均的「直接平均」等六種方式，個別進行資料處理與分群，並將不同案例產生的推估結果與小客車平均旅行速率進行比較。由於本研究是以公車作為探針車，並分析其回傳的資料，將推估的結果發佈給小客車的使用者，故以小客車車速作為研究比較對象。

分群後及直接平均的結果則採文獻上常用的三種評估方式進行分析評估，分別為 MAPE、RMSE 以及 MAE 等三種方式，分析的詳細數據請參照附錄，本節僅針對全部案例資料的評估結果進行說明，評估結果如表 4 所示。

表 4 評估結果比較表

分群方式		MAPE	RMSE	MAE(km/hr)
兩群	低速群	94%	38.4	33.9
	高速群	60%	13.9	10.8
三群	低速群	99%	38.7	34.6
	中速群	27%	13.7	9.1
	高速群	74%	15.6	12.3
直接平均		35%	16.8	13.2

透過平均絕對誤差百分比(MAPE)可以了解方法預測的準確程度，根據表 4 MAPE 的結果可以發現，除了「分三群取中速群」以及「直接平均」的 MAPE 值分別為 27% 與 35%，對照 Lewis 預測表後，屬於合理的預測範圍之外，其他的分群方式的 MAPE 值皆大於 50%，屬於不合理的預測。因此，就 MAPE 的結果而言，對於所有案例而言，透過 FCM 以「分三群取中速群」推估的結果較為準確。

透過均方根誤差(RMSE)的結果可以了解模式預測的可靠程度。除了「分

兩群取低速群」及「分三群取低速群」為最高 38.4 及 38.7 之外，其他的結果為「分兩群取高速群」13.9、「分三群取中速群」13.7、「分三群取高速群」15.6 以及「直接平均」16.8。由於，RMSE 的值越小，表示其可靠程度越佳，因此比較所有方式的 RMSE 結果可以發現，對於所有案例而言，「分三群取中速群」的結果是整體表現中最好的結果。

透過平均絕對誤差 (MAE) 的結果，可以了解推估結果的誤差範圍，除了「分兩群取低速群」及「分三群取低速群」誤差都大於 30km/hr 之外，其他方式的誤差範圍結果分別為「分兩群取高速群」10.8km/hr、「分三群取中速群」9.1km/hr、「分三群取高速群」12.3km/hr 以及「直接平均」13.2km/hr。比較所有分群方式的 MAE 可以發現，對於所有案例而言，以「分三群取中速群」的誤差範圍為最小。

雖然綜合以上三項評估方式的結果可以發現，以「分三群取中速群」的表現似乎為三項評估指標中最佳的結果，然而，在比較與「分兩群取高速群」的方式之後發現，「分兩群取高速群」的 MAPE 為 60%，雖然遠高於「分三群取中速群」的 27%，但其 RMSE 却只有 13.9 與「分三群取中速群」的 13.7 相當接近，MAE 的誤差範圍兩者間也不到 2km/hr。這樣的結果若對照評估指標的分析內容會發現，「分兩群取高速群」的預測範圍雖然不合理，但其可靠程度卻與「分三群取中速群」的結果差不多，兩者有相互矛盾的現象，因此本研究推測透過 FCM 的分群方式來推估路況，如果在不同的道路情境下使用並不具有通用性，但由於目前仍無法有明確的結論，故本研究將依照不同的道路情境進行討論，探討道路情境與分群方式的關係。

4.4 道路情境影響與分群適用時機

從評估指標的結果顯示，若不考慮道路情境（即道路的壅塞程度），在推估整體路段速率的表現上，將會產生較大的誤差。因此本研究希望透過不同道路情境與分群方式的關係，找出在不同道路情境中分群方式的適用時機。

對於不同道路情境的區分，本研究參考交通部 2011 年最新版的台灣公路容量手冊中[14]，市區道路服務水準等級劃分表，依其提供之平均旅行速率，作為道路情境的區分標準。

本研究所模擬的區域路段，根據實際調查結果發現其限速為 50km/hr，參考公路容量手冊後，得到的道路服務等級如表 5 所示：

表 5 速限 50 公里/小時之市區道路服務水準等級劃分標準

平均旅行速率 V(km/hr)	服務水準等級
$V \geq 35$	A
$30 \leq V < 35$	B
$25 \leq V < 30$	C
$20 \leq V < 25$	D
$15 \leq V < 20$	E
$V < 15$	F

資料來源：2011 台灣公路容量手冊

Underwood 在其研究中發現，可以透過觀察路段平均旅行速率是否小於 30km/hr，來衡量道路是否產生擁擠現象[15]。因此本研究根據模擬出來的資料案例結果，參考 Underwood 的擁擠衡量建議，以 30km/hr 作為分界，合併服務水準 A、B 的資料個數，而合併後的樣本數為 42 個，其餘 C、D、E、F 服務水準內的資料個數合併，合併後的樣本數為 30 個，兩者皆符合統計上大樣本數的需求。

經過合併後的案例資料，進行模式重新驗證與分析，得到的結果如表 6 所示。

表 6 分群結果比較表-依道路情境區分

平均旅行速率	分群方法		MAPE	RMSE	MAE(km/hr)
> 30 km/hr	兩群	低速群	100%	46.9	46.1
		高速群	14%	9.9	7.2
	三群	低速群	100%	47.0	46.2
		中速群	23%	16.7	11.7
		高速群	14%	9.0	6.9
	直接平均		38%	20.7	18.4
$\leq 30 \text{ km/hr}$	兩群	低速群	84%	13.7	12.3
		高速群	142%	19.1	17.2
	三群	低速群	97%	15.3	14.0
		中速群	34%	5.5	4.6
		高速群	179%	23.0	21.8
	直接平均		29%	4.8	4

將所有案例資料依照本研究的分類之後，以平均旅行速率 30km/hr 作為分界速率，其中平均旅行速率為模擬資料中的小客車平均旅行速率，而分群方法則是利用不同分群方式分析公車回傳資料所產生的推估結果。

從 MAPE 的表現上，可以很明顯的看出，當平均旅行速率大於 30km/hr 時(即道路服務等級為 B 級以上)，採用以「分兩群取高速群」及「分三群取高速群」的方式效果表現最好，其 MAPE 的平均結果皆為 14%，屬於優良的預測範圍；在 RMSE 的表現則是「分三群取高速群」為所有比較方式中最小的 9.0，而且其 MAE 的平均誤差範圍約在 6.9km/hr，為所有比較方式中最小；而當平均旅行速率小於或等於 30km/hr 時，則是「直接平均」的結果較佳，其 MAPE 的平均結果為 29%，屬於合理的預測範圍，RMSE 也是所有比較方式中最小的 4.8，MAE 的平均誤差範圍約在 4km/hr 左右。

然而，在平均旅行速率大於 30km/hr 時，「分兩群取高速群」及「分三群取高速群」的表現結果相當接近，而且在平均旅行速率小於 30km/hr 時「分三群取中速群」與「直接平均」的位差範圍也很接近，為求嚴謹，本研究利用統計檢定的方式來檢驗，比較「分兩群取高速群」與「分三群取高速群」的結果，對於路段平均速率大於 30km/hr 時的結果是否存在顯著的差異以及「分三群取中速群」與「直接平均」，對於路段平均速率小於等於 30km/hr 時的結果是否也存在顯著的差異。由於本研究的樣本數大於 30 個，故採用 Z 檢定中的信賴區間差異比較，來比較兩兩數列之間與路段平均速率的差異性。

原始的信賴區間在 95% 的情況下，Z 值的計算方式為 $Z = 1.96 \times \frac{\sigma}{\sqrt{n}}$ ，其中 σ 為變異數，n 為觀察個數。但由於本研究數列間彼此較為相似，故將兩數列的相關性一併考慮進去，使方程式變為 $Z = \sqrt{2 \times (1 - \rho)} \times 1.96 \times \frac{\sigma}{\sqrt{n}}$ ，其中 ρ 則為兩數列間的相關性。

本研究先假設兩比較數列存在差異性，則當相關性越高 Z 值會越小，將越不容易拒絕假設，亦即此兩數列之間具有顯著的差異，反之當相關性越低則 Z 值會越大，假設將被拒絕，表示此兩數列之間不具有顯著差異。因此可藉由兩數列的相關性來決定 Z 值大小，並透過兩數列的平均誤差來比較兩數列是否有顯著的差異，當計算出的 Z 值小於兩數列間的平均誤差，則表示兩數列之中存在差異性且與實際值平均絕對誤差較小的數列表現較佳；若 Z 值大於兩數列差的平均值，則表示兩數列間並無明顯的差異。因此透過此方式，即可作為兩數列的優劣程度判斷的依據。

表 7 為分群方式檢定表，其中 MAE 為平均絕對誤差 (Mean Average Error)，MAE1 為「分兩群取高速群」與路段平均速率之間的平均絕對誤差，MAE2 則是「分三群取高速群」與路段平均速率之間的平均絕對誤差；而 diff(1&2) 為 MAE1 與 MAE2 的差值。

表 7 分群方式檢定表-兩群高速與三群高速

平均旅行速率 > 30km/hr	結果：無差別		
統計檢定	MAE1	MAE2	diff(1&2)
mean	7.17	6.88	0.29
std	6.95	5.85	4.09
Z 值 (95%)	2.84	2.41	1.23

從表 7 的結果可以發現，當平均旅行速率大於 30km/hr 的情況下，其 95% 下的 Z 值為 1.23 大於兩數列差的平均值 0.29，表示其結果並無顯著的差異，亦即「分兩群取高速群」與「分三群取高速群」在路段平均速率大於 30km/hr 的情況下，其推估結果是沒有差別的。同樣地，表 8 為「分三群取中速群」與「直接平均」的檢定結果，當平均旅行速率小於等於 30km/hr 時，其 95% 下的 Z 值為 1.70 大於兩數列差的平均值 0.47，表示其結果並無顯著差異，亦即「分三群取中速群」與「直接平均」在路段平均速率小於 30km/hr 的情況下，其推估結果同樣也沒有差別。

表 8 分群方式檢定表-三群中速與直接平均

平均旅行速率 $\leq 30\text{km/hr}$	結果：無差別		
統計檢定	三群中速	直接平均	diff(1&2)
mean	4.47	4.00	0.47
std	3.30	2.69	4.53
Z 值 (95%)	1.80	2.36	1.70

綜合以上的結果可發現，分群後群組間的資料選擇的推估結果與路段平均速率有關，故需先考慮道路情境，就本研究的結果，採用 Underwood 的建議在 30km/hr (或是道路服務水準等級為 B 時) 作為一個區分點，當平均旅行速率高於此區分點，分群方式可採用「分兩群取高速群」或是「分三群取高速群」作為推估路況的方式，當低於此區分點則可採用「分三群取中速群」或是「直接平均」作為推估路況的方式，方能使推估結果更為準確。

4.5 分車班次之影響

在資料處理的過程中，通常是將所有通過路段的車輛資料，以不分車輛班次的情況收集後再進行處理，然而相同路段上，每輛公車的行為可能有所不同，(如：有些公車會靠站載客，有些則會過站不停)若將所有車輛資料合併在一起處理，對於車輛的行為表現可能會失真，造成資料處理上產生不對等的結果，因此本研究又將回傳資料區分為「分車班次」與「不分車班次」，並以「分車」與「不分車」稱之，以避免不同車輛的不同行為彼此干擾並探討此方法的處理結果是否有

所差異。

「分車」表示將所有公車資料，依照車班次分開，再做分群；而「不分車」則表示將所有公車資料，以不分車班次全部合併後，再做分群。無論分車班次與否，資料的總筆數並不會改變，而是改變處理的資料量。

本研究在區分道路情境之後，確認合適的分群方式，再進行區分車班次的處理，整理後的結果如表 9 所示，從表 9 可以看出，根據三項評估指標的結果可以發現，分車班次與否對於推估的結果並無明顯的差異，如在路段速度大於 30km/hr 的情況下，以「分兩群取高速群」的方式，「分車」的 MAPE 為 15%，「不分車」的為 14%，RMSE 的平均結果分別為「分車」9.5 與「不分車」9.9，MAE 平均則為「分車」7.2 與「不分車」7.3，以「分三群取高速群」的方式，「分車」與「不分車」的 MAPE 皆為 14%，RMSE 的平均結果分別為「分車」9.3 與「不分車」9.0，MAE 平均則為「分車」6.7 與「不分車」6.9；小於 30km/hr 的情況，以「分三群取中速群」的方式，「分車」與「不分車」的差異亦差別不大，本研究同樣利用統計檢定的方式進行檢驗，證明其是否存在顯著的差異性，以探討分車班次與路段推估的影響程度。

表 9 區分車班次比較表-建議分群方式

平均旅行速率	分群方法		MAPE	RMSE	MAE(km/hr)
>30 km/hr	兩群	分車	15%	9.5	7.2
		不分車	14%	9.9	7.3
	三群	分車	14%	9.3	6.7
		不分車	14%	9.0	6.9
	直接平均	-	38%	20.7	18.4
	≤30km/hr	三群	分車	36%	6.0
		中速	不分車	34%	5.5
		直接平均	-	29%	4.8
					4

表 10、表 11 與表 12 為不同分群方式下的分車班次統計檢定比較表，其中 MAE 為平均絕對誤差 (Mean Average Error)，MEA1 為「分車班次」與路段平均速率之間的平均絕對誤差，MEA2 則是「不分車班次」與路段平均速率之間的平均絕對誤差；而 diff (1&2) 為 MEA1 與 MEA2 的差值。

表 10 分車班次統計檢定表-分兩群取高速群

平均旅行速率 > 30km/hr		結果：無差別		
分二群 (高速群)	統計檢定	MAE1	MAE2	diff(1&2)
	mean	7.20	7.30	0.10
	std	6.28	6.77	1.70
	Z 值 (95%)	2.57	2.79	0.22

表 11 分車班次統計檢定表-分三群取高速群

平均旅行速率 > 30km/hr		結果：無差別		
分三群 (高速群)	統計檢定	MAE1	MAE2	diff(1&2)
	mean	6.71	6.88	0.17
	std	6.52	5.85	5.06
	Z 值 (95%)	3.54	2.41	1.67

表 12 分車班次統計檢定表-分三群取中速群

平均旅行速率 \leq 30km/hr		結果：無差別		
分三群 (中速群)	統計檢定	MAE1	MAE2	diff(1&2)
	mean	4.69	4.64	0.05
	std	3.85	3.25	4.16
	Z 值 (95%)	2.09	2.85	1.81

在平均旅行速率大於 30km/hr 的情況下，無論在哪一種分群方式下，分車班次與否兩者間都沒有顯著的差異。從表 10 與表 11 的結果可以發現，以「分兩群取高速群」的方式，在 95%下的 Z 值為 0.22 大於兩數列差的平均值 0.10，表示其結果並無顯著的差異；同樣地，以「分三群取高速群」的方式，在 95%下的 Z 值為 1.67 大於兩數列差的平均值 0.17，表示其結果並無顯著的差異。

而在平均旅行速率小於等於 30km/hr 的情況下，無論分車班次與否兩者間亦無顯著差異。根據表 12 的結果，以「分三群取中速群」的方式，在 95%下的 Z 值為 1.81 同樣大於兩數列差的平均值 0.05，表示其結果並無顯著的差異。

為求嚴謹，本研究也將不建議使用的分群方式，依道路情境進行區分車班次的比較，了解結果是否有所改善，其區分車班次的比較如表 13 所示，從結果可以看出，對於不建議使用的分群方式，在區分車班次之後的推估結果，其改善程度有限甚至沒有改善，而有改善的結果，也不及本研究先前所建議的分群方式。

表 13 區分車班次比較表-其他分群方式

平均旅行速率	分群方法		MAPE	RMSE	MAE(km/hr)
> 30 km/hr	兩群 低速	分車	100%	46.9	46.1
		不分車	100%	46.9	46.1
	三群 低速	分車	99%	46.7	46.0
		不分車	100%	47.0	46.2
	三群 中速	分車	23%	14.6	11.5
		不分車	23%	16.7	11.7
$\leq 30\text{km/hr}$	兩群 低速	分車	92%	14.4	13.2
		不分車	84%	13.7	12.3
	兩群 高速	分車	92%	12.9	11.5
		不分車	142%	19.1	17.2
	三群 低速	分車	100%	15.7	14.4
		不分車	97%	15.3	14.0
	三群 高速	分車	113%	16.8	14.7
		不分車	179%	23.0	21.8

因此，根據上述的統計檢定結果可以判斷，當路段平均旅行速率大於 30km/hr 的情況下，「分兩群取高速群」及「分三群取高速群」的方式，在進行資料分群分析之前，分車班次與否與推估結果並無顯著差異；同樣地，當平均旅行速率小於 30km/hr 的情況下，以「分三群取中速群」的方式，在進行分群分析之前，分車班次與否與推估結果同樣無顯著差異，亦即是區分車班次對於推估結果並無影響，且亦無法改善非本研究建議的分群方法之推估結果。

4.6 回傳頻率之影響

由於本研究的對象是以裝載 GPS 設備的公車所回傳的資料進行分群分析，因此本研究認為訊號資料的回傳頻率或是發車頻率都可能會影響到分析的結果，然而改變回傳頻率相較於改變發車頻率，在實際處理上較為可行且簡易，故本研究選擇透過不同的回傳頻率來檢視回傳的結果是否存在其差異性。

原本的公車回傳頻率根據實際調查參數結果為 30 秒回傳一次，為了解回傳頻率的影響結果，由於資料回傳頻率多寡與後續的通訊成本有關，根據我國交通部運輸研究所的研究資料顯示，回傳頻率 5 秒與 15 秒的差異性並不大[16]，考量未來將之應用於實際路況的可能性，因此本研究採用 15 秒回傳一次作為測試，並在其他參數條件不變的情況下依照不同的道路情境，以相同的分群方式進行分析並與原本回傳頻率為 30 秒一次的資料進行檢定，得到的結果如表 14 所示：

表 14 不同頻率下的分析資料表-15秒

平均旅行速率	分群方法	分群結果-30sec			分群結果-15sec			統計檢定結果
		MAPE	RMSE	MAE	MAPE	RMSE	MAE	
>30 km/hr	兩群 (高速群)	14%	9.9	7.2	14%	9.3	7.2	無差別
	三群 (中速群)	23%	16.7	11.7	19%	12.8	9.6	無差別
	三群 (高速群)	14%	9.0	6.9	15%	9.9	7.4	無差別
	直接平均	38%	20.7	18.4	55%	27.7	26.1	有顯著差異
$\leq 30\text{km/hr}$	兩群 (高速群)	142%	19.1	17.2	85%	11.3	9.7	有顯著差異
	三群 (中速群)	34%	5.5	4.6	21%	4.7	2.7	有顯著差異
	三群 (高速群)	179%	23.0	21.8	92%	11.9	10.9	有顯著差異
	直接平均	29%	4.8	4	35%	6.9	5.5	有顯著差異

從表 14 可以看出，在增加回傳頻率之後，與原先的回傳頻率相互比較之下，在平均旅行速率大於 30km/hr 的情況下，雖有評估指標各有增減但統計結果除「直接平均」之外，均無顯著差異；而在平均旅行速率小於等於 30km/hr 的情況下，MAPE、RMSE 以及 MAE 值皆有明顯的下降，且統計上的結果也都有顯著的差異。

在平均旅行速率小於等於 30km/hr 的情況下，以回傳頻率為 15 秒一次時，「分三群取中速群」的方式所得到的結果，其三項的評估指標皆比直接平均來的佳，MAE 的平均誤差的範圍為 2.7km/hr 同樣也相對較小許多，本研究將其與同樣回傳條件下的「直接平均」進行檢定，發現兩者存在顯著的差異，結果如表 15 所示。

表 15 分群方式統計檢定表-分三群取中速群與直接平均

平均旅行速率 $\leq 30\text{km/hr}$	結果：有顯著差異		
統計檢定	直接平均	三群中速	diff(1&2)
mean	5.50	2.72	2.78
std	4.20	3.90	4.36
Z 值 (95%)	3.68	2.12	0.97

由上表的檢定結果可以得知，在平均速率小於 30km/hr 的情況下，以回傳頻率為 15 秒一次時，「分三群取中速群」的推估結果表現更為優異。因此，本研究可以推論出當路段的平均速率降低時，需要較高的回傳頻率或是在不改變回傳頻率的情況下增加探針車的發車班次，才能較正確的反映出實際的路況，這也符合 Chen & Chein[11]在研究中的結果，即在道路條件在極輕微 (very light) 或是極壅塞 (very heavy) 時，由於兩者的路段平均速率變數相較於穩定車流時來得大，因此需要的探針車比例也相對較高。



五、結論與建議

本研究針對透過裝載 GPS 設備的公車作為非專用探針車輛，並根據其回傳的速度資料進行研究，配合實際調查的資料以及車流模擬軟體來建置路網，將收集到的資料透過模糊分群法（FCM）進行群集分析以及利用三種評估指標與統計檢定作為評定推估結果的準確程度，主要探討的內容有（1）道路情境對於資料分析的影響及分群的適用時機（2）分車班次對於資料處理的影響（3）資料回傳頻率的影響，共三個部份，經本研究實驗分析之後，所得到的結論與建議如下：

5.1 結論

- (1) 根據本研究的結果，Fuzzy C-Means 亦可使用在即時資料的分析與應用上，可減少因收集歷史資料所產生的時間與成本。
- (2) 對於不同的道路情境，資料處理的方式亦應有所不同。根據本研究的結果，若不考慮道路情境，將所有資料透過 FCM 分群方式作為處理，從評估指標的結果可以發現，會產生結果相互矛盾的現象。
- (3) 從道路情境的角度而言，本研究採用 Underwood 的建議，以 30km/hr 作道路情境區分速率，當路段旅行平均速率大於 30km/hr(或服務水準等級為 B 級以上)時，可採用「分兩群取高速群」或是「分三群取高速群」的資料分群方式會有最佳的表現，MAPE 的結果屬於優良的預測，平均誤差約在 7km/hr 左右；而當路段旅行平均速率小於 30km/hr (道路等級為 C 級以下)，則採用「分三群取中速群」或是「直接平均」，MAPE 的結果也屬於合理的預測，平均誤差約在 4km/hr 左右。
- (4) 對於是否區分車輛行為再進行分群分析，經本研究透過統計檢定後發現，在道路情境區分出分群方式之後，無論分車班次與不分車班次，在推估的結果上，並沒有顯著的差異，表示是否先分車班次，再進行群集分析，對於路段速率的推估結果並無影響。
- (5) 資料回傳頻率對於路段推估速度亦有影響。特別是在路段旅行平均速率小於 30km/hr(或服務水準等級為 C 級以下)時候，根據本研究測試結果，如果能夠提昇回傳頻率（如本研究測試，將回傳頻率提升為 15 秒回傳一次），則以「分三群取中速群」的方式，會有最好的推估結果，MAPE 同樣屬於合理的預測，但平均誤差可以下降到 2.7m/hr 左右，表示提高回傳頻率可以提高推估路段速率的準確程度。
- (6) 無論是尖峰或是離峰的情境，皆可以使用 FCM 方式進行資料處理，經本研究的測試結果而言都會較直接平均來的準確。

5.2 建議

- (1) 本研究透過實際資料測試，證明模糊分群法（FCM）亦可用在以公車作為非專用探針車的回傳資料上，由於非專用的探針車量種類多樣，如果

- 能透過此一方法找出適合每一種非專用探針車的資料處理方式，對於未來採用 GPS 資料分析的工作將會有所助益。
- (2) 由於實際資料收集不易，若未來能夠完整收集公車回傳的歷史資料，並與 Lee 進行分析比較，以探討兩種不同資料來源的處理結果。
 - (3) 本研究將以裝載 GPS 的公車回傳資料進行路段速率推估，發現若依照道路情境進行不同分群會有不錯的效果，未來也可考慮依照推估的結果進行旅行時間的推估，期望能使旅行時間推估結果更為準確。
 - (4) 在路段速率推估方面，本研究只測試 800 公尺左右的路段長度，期望未來可以延長推估的長度，或是將此方法應用在高速公路的路況或是旅行時間的推估上，因高速公路的車輛不會受到號誌影響，故在推估上的限制較小，推估的結果可能會有更好的表現。

5.3 建議 FCM 之應用

對於 FCM 方式之應用，根據本研究的測試結果提出建議步驟，說明如下：

- (1) 在尚未了解道路平均旅行速率的情況下，可先採用「分兩群取高速群」或是「分三群取高速群」來推估路段速率，並依此作為平均旅行速率。建議採用的理由是，在未知路段旅行速率的情況下可先視為自由車流下的速率，若實際路況情境與假設的結果相反為擁擠車流，則公車回傳的最高速度也定會被限制，因此可以藉此改變下一個時段的推估方式，使結果趨向準確。
- (2) 在了解道路情境之後，便可依照本研究的方式，依照不同的道路情境，使用不同的分群方式，便可使推估出的路段速率結果更為準確

六、參考文獻

1. 交通部運輸研究所, *台灣地區智慧型運輸系統綱要計畫* (2004 年版摘要本). 2004, 台北: 交通部.
2. Uno, et al., *Using Bus Probe Data for Analysis of Travel Time Variability*. Journal of Intelligent Transportation Systems, 2009. **13**(1): p. 2-15.
3. de Fabritiis, C., R. Ragona, and G. Valenti. *Traffic Estimation And Prediction Based On Real Time Floating Car Data*. in *Intelligent Transportation Systems, 2008. ITSC 2008. 11th International IEEE Conference on*. 2008.
4. Lee, S.-H., B.-W. Lee, and Y.-K. Yang, *Estimation of Link Speed Using Pattern Classification of GPS Probe Car Data*, in *Computational Science and Its Applications - ICCSA 2006*, M. Gavrilova, et al., Editors. 2006, Springer Berlin Heidelberg. p. 495-504.
5. D'Este, G.M., R. Zito, and M.A.P. Taylor, *Using GPS to Measure Traffic System Performance*. Computer-Aided Civil and Infrastructure Engineering, 1999. **14**(4): p. 255-265.
6. Long Cheu, R., C. Xie, and D.-H. Lee, *Probe Vehicle Population and Sample Size for Arterial Speed Estimation*. Computer-Aided Civil and Infrastructure Engineering, 2002. **17**(1): p. 53-60.
7. Han, J., M. Kamber, and J. Pei, *Data mining: concepts and techniques*. 2006: Morgan kaufmann.
8. Dunn, J.C., *A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters*. Journal of Cybernetics, 1973. **3**(3): p. 32-57.
9. Bezdek, J.C., *Pattern Recognition with Fuzzy Objective Function Algorithms*. 1981: Kluwer Academic Publishers. 256.
10. Lewis, E.B., *Control of body segment differentiation in Drosophila by the bithorax gene complex*. Prog Clin Biol Res, 1982. **85 Pt A**: p. 269-88.
11. Chen, M. and S. Chien, *Determining the Number of Probe Vehicles for Freeway Travel Time Estimation by Microscopic Simulation*. Transportation Research Record: Journal of the Transportation Research Board, 2000. **1719**(1): p. 61-68.
12. Ygnace, J.-L., et al., *Travel Time Estimation on the San Francisco Bay Area Network Using Cellular Phones as Probes*, 2000.
13. viSSIM. Available from: <http://www.vissim.com/>.
14. 林豐博, et al., *2011 年臺灣公路容量手冊*. 2011: 交通部運輸研究所.
15. Underwood, R.T., *Acceleration Noise and Traffic Congestion*. 1968: Traffic engineering & control.
16. 周家慶 and 吳玉珍, *我國都市交通控制發展與推動策略*. 2008: 交通部運輸研究所.

七、附錄

FCM 分群與實際資料結果-MAPE

case	MAPE					
	分兩群		分三群			直接平均
	低速群	高速群	低速群	中速群	高速群	
1	100%	34%	100%	99%	34%	54%
2	100%	39%	100%	47%	35%	64%
3	100%	37%	100%	41%	35%	56%
4	100%	33%	100%	37%	22%	66%
5	100%	28%	100%	32%	15%	50%
6	100%	33%	100%	45%	27%	57%
7	100%	24%	100%	31%	4%	46%
8	100%	26%	100%	39%	20%	44%
9	100%	23%	100%	35%	16%	44%
10	100%	14%	100%	33%	7%	57%
11	100%	23%	100%	23%	23%	40%
12	100%	21%	100%	40%	12%	43%
13	100%	10%	100%	10%	10%	45%
14	100%	13%	100%	20%	6%	40%
15	100%	25%	100%	34%	21%	37%
16	100%	17%	100%	33%	10%	58%
17	100%	6%	100%	6%	6%	39%
18	100%	17%	100%	30%	11%	38%
19	100%	8%	100%	18%	18%	37%
20	100%	6%	100%	11%	6%	59%
21	100%	11%	100%	23%	5%	26%
22	100%	10%	100%	20%	5%	36%
23	100%	1%	100%	5%	16%	32%
24	100%	0%	100%	24%	8%	34%
25	100%	2%	100%	9%	17%	45%
26	100%	1%	100%	6%	18%	34%
27	100%	12%	100%	25%	6%	33%
28	100%	5%	100%	12%	14%	62%
29	100%	8%	100%	11%	7%	32%
30	100%	7%	100%	10%	2%	31%

31	95%	3%	100%	83%	3%	32%
32	100%	5%	100%	10%	2%	29%
33	100%	3%	100%	5%	1%	28%
34	100%	6%	100%	20%	0%	26%
35	100%	4%	100%	4%	27%	41%
36	100%	4%	100%	5%	1%	27%
37	100%	2%	100%	5%	7%	25%
28	100%	0%	100%	0%	0%	26%
29	100%	3%	100%	12%	9%	15%
40	100%	14%	100%	14%	14%	43%
41	100%	16%	100%	10%	33%	21%
42	100%	4%	100%	2%	19%	15%
43	100%	22%	100%	22%	22%	16%
44	100%	13%	100%	0%	19%	14%
45	100%	19%	100%	8%	24%	35%
46	95%	42%	100%	69%	42%	9%
47	99%	38%	98%	10%	56%	40%
48	85%	53%	98%	33%	53%	24%
49	95%	26%	95%	7%	58%	16%
50	76%	56%	100%	31%	56%	1%
51	86%	50%	97%	21%	73%	33%
52	79%	102%	93%	12%	121%	33%
53	98%	90%	98%	52%	167%	42%
54	92%	102%	92%	65%	158%	17%
55	80%	81%	92%	18%	206%	42%
56	70%	180%	99%	6%	180%	30%
57	81%	103%	93%	20%	138%	4%
58	99%	173%	99%	107%	189%	27%
59	95%	88%	98%	22%	154%	51%
60	73%	185%	96%	38%	185%	13%
61	84%	77%	88%	18%	95%	37%
62	64%	288%	99%	27%	288%	35%
63	91%	122%	100%	34%	142%	42%
64	90%	116%	95%	73%	303%	39%
65	70%	279%	100%	15%	279%	50%
66	81%	324%	96%	44%	324%	7%
67	81%	295%	100%	8%	295%	65%

68	66%	278%	100%	23%	278%	14%
69	97%	63%	97%	23%	137%	23%
70	61%	292%	100%	46%	352%	9%
71	88%	148%	96%	81%	274%	1%
72	100%	89%	99%	52%	89%	45%
Average	94%	60%	99%	27%	74%	35%

FCM 分群與實際資料結果-RMSE

case	RMSE					直接平均	
	分兩群		分三群				
	低速群	高速群	低速群	中速群	高速群		
1	4330	515	4356	4238	515	1267	
2	4316	660	4316	942	533	1747	
3	4176	582	4176	709	525	1291	
4	4109	454	4109	552	199	1798	
5	3624	286	3624	380	85	912	
6	3388	377	3388	687	250	1116	
7	3043	176	3043	288	5	649	
8	2920	200	2920	447	115	554	
9	2767	142	2767	346	72	524	
10	2559	50	2559	275	14	840	
11	2559	130	2559	130	130	411	
12	2510	112	2510	404	37	471	
13	2448	24	2448	24	24	496	
14	2388	41	2388	92	9	391	
15	2359	150	2359	274	107	330	
16	2288	68	2288	251	21	775	
17	2261	9	2261	9	9	348	
18	2220	66	2220	199	26	314	
19	2156	15	2156	68	73	290	
20	2131	9	2131	24	7	732	
21	2046	23	2046	105	4	138	
22	2023	20	2023	81	6	262	
23	2000	0	2000	6	49	208	
24	1978	0	1978	110	12	233	
25	1956	1	1956	15	56	397	
26	1934	0	1934	7	66	227	

27	1934	28	1934	121	7	216
28	1928	6	1928	27	40	740
29	1924	12	1924	24	8	192
30	1871	9	1871	18	1	181
31	1611	2	1792	1227	2	183
32	1736	4	1736	17	1	143
33	1727	2	1727	4	0	131
34	1691	7	1691	66	0	117
35	1591	3	1591	2	117	262
36	1559	2	1559	4	0	114
37	1528	0	1528	4	7	98
28	1499	0	1499	0	0	104
29	1491	1	1491	21	13	34
40	1470	29	1470	29	29	273
41	1421	35	1421	14	156	64
42	1401	2	1401	1	49	31
43	1227	57	1227	57	57	33
44	1154	19	1154	0	44	24
45	1066	38	1066	7	63	129
46	912	174	1011	475	174	9
47	812	119	806	9	259	134
48	597	227	791	91	230	47
49	648	48	648	3	242	19
50	395	214	680	65	214	0
51	260	88	335	15	190	39
52	163	273	224	4	386	29
53	229	193	229	65	665	41
54	183	227	183	92	542	6
55	129	135	173	6	866	36
56	94	625	191	1	630	18
57	122	200	163	7	355	0
58	177	547	177	210	650	13
59	155	135	165	8	410	46
60	90	581	156	24	576	3
61	115	97	126	5	148	22
62	58	1162	137	10	1162	17
63	101	181	122	14	245	22

64	96	161	106	64	1094	18
65	52	824	106	3	824	27
66	68	1080	95	20	1080	1
67	66	871	100	1	871	42
68	40	696	90	5	696	2
69	73	32	73	4	147	4
70	25	579	68	14	839	1
71	43	122	51	36	417	0
72	54	42	52	15	42	11
Average	38.4	13.9	38.7	13.7	15.6	16.8

FCM 與實際資料結果-MAE

case	MAE					直接平均
	分兩群		分三群			
	低速群	高速群	低速群	中速群	高速群	
1	65.8	22.7	66.0	65.1	22.7	35.6
2	65.7	25.7	65.7	30.7	23.1	41.8
3	64.6	24.1	64.6	26.6	22.9	35.9
4	64.1	21.3	64.1	23.5	14.1	42.4
5	60.2	16.9	60.2	19.5	9.2	30.2
6	58.2	19.4	58.2	26.2	15.8	33.4
7	55.2	13.3	55.2	17.0	2.2	25.5
8	54.0	14.1	54.0	21.1	10.7	23.5
9	52.6	11.9	52.6	18.6	8.5	22.9
10	50.6	7.1	50.6	16.6	3.8	29.0
11	50.6	11.4	50.6	11.4	11.4	20.3
12	50.1	10.6	50.1	20.1	6.1	21.7
13	49.5	4.9	49.5	4.9	4.9	22.3
14	48.9	6.4	48.9	9.6	2.9	19.8
15	48.6	12.3	48.6	16.6	10.4	18.2
16	47.8	8.2	47.8	15.8	4.6	27.8
17	47.5	2.9	47.5	2.9	2.9	18.6
18	47.1	8.1	47.1	14.1	5.1	17.7
19	46.4	3.9	46.4	8.2	8.6	17.0
20	46.2	3.0	46.2	4.9	2.6	27.1
21	45.2	4.8	45.2	10.2	2.0	11.7
22	45.0	4.5	45.0	9.0	2.4	16.2

23	44.7	0.3	44.7	2.4	7.0	14.4
24	44.5	0.1	44.5	10.5	3.5	15.3
25	44.2	1.0	44.2	3.9	7.5	19.9
26	44.0	0.2	44.0	2.7	8.1	15.1
27	44.0	5.3	44.0	11.0	2.6	14.7
28	43.9	2.4	43.9	5.2	6.3	27.2
29	43.9	3.5	43.9	4.9	2.9	13.9
30	43.3	3.1	43.3	4.3	0.8	13.5
31	40.1	1.3	42.3	35.0	1.3	13.5
32	41.7	2.1	41.7	4.2	1.0	12.0
33	41.6	1.3	41.6	2.0	0.4	11.5
34	41.1	2.6	41.1	8.1	0.1	10.8
35	39.9	1.6	39.9	1.6	10.8	16.2
36	39.5	1.5	39.5	2.1	0.5	10.7
37	39.1	0.7	39.1	2.1	2.7	9.9
28	38.7	0.2	38.7	0.2	0.2	10.2
29	38.6	1.0	38.6	4.6	3.6	5.8
40	38.3	5.4	38.3	5.4	5.4	16.5
41	37.7	5.9	37.7	3.7	12.5	8.0
42	37.4	1.6	37.4	0.9	7.0	5.5
43	35.0	7.6	35.0	7.6	7.6	5.7
44	34.0	4.3	34.0	0.0	6.6	4.9
45	32.6	6.2	32.6	2.6	8.0	11.3
46	30.2	13.2	31.8	21.8	13.2	3.0
47	28.5	10.9	28.4	3.0	16.1	11.6
48	24.4	15.1	28.1	9.5	15.2	6.8
49	25.5	6.9	25.5	1.8	15.5	4.4
50	19.9	14.6	26.1	8.1	14.6	0.2
51	16.1	9.4	18.3	3.9	13.8	6.2
52	12.8	16.5	15.0	2.0	19.6	5.4
53	15.1	13.9	15.1	8.1	25.8	6.4
54	13.5	15.1	13.5	9.6	23.3	2.5
55	11.4	11.6	13.2	2.5	29.4	6.0
56	9.7	25.0	13.8	0.8	25.1	4.2
57	11.1	14.1	12.8	2.7	18.8	0.5
58	13.3	23.4	13.3	14.5	25.5	3.6
59	12.5	11.6	12.9	2.8	20.2	6.8

60	9.5	24.1	12.5	4.9	24.0	1.7
61	10.7	9.9	11.2	2.3	12.2	4.7
62	7.6	34.1	11.7	3.2	34.1	4.1
63	10.0	13.5	11.0	3.8	15.7	4.6
64	9.8	12.7	10.3	8.0	33.1	4.2
65	7.2	28.7	10.3	1.6	28.7	5.2
66	8.2	32.9	9.7	4.5	32.9	0.7
67	8.1	29.5	10.0	0.8	29.5	6.5
68	6.3	26.4	9.5	2.2	26.4	1.3
69	8.6	5.6	8.6	2.0	12.1	2.1
70	5.0	24.1	8.2	3.8	29.0	0.8
71	6.6	11.0	7.2	6.0	20.4	0.1
72	7.3	6.5	7.2	3.8	6.5	3.3
Average	33.9	10.8	34.6	9.1	12.3	13.2

單位 : km/hr

