

利用數位式行車紀錄器資料推估旅行時間之研究

朱松偉¹張家銘²

摘要

本研究採用數位式行車紀錄器之資料，探討高速公路之旅行時間預測，並透過多項軟體工具發展出一套適用於此研究之資料處理流程，以處理這些歷史資料。最後再透過統計 SPSS 10.0 軟體執行複迴歸分析，而得之迴歸模式。本研究採用其資料探討高速公路之旅行時間預測，以達到以下之目的 1.利用數位式行車紀錄器之大量歷史資料，發展一套資料處理流程及 2.預測從圓山交流道到桃園交流道所需的旅行時間。本研究之歷史資料共計超過六千萬筆，數量相當地大，故本研究透過發展一套適用於此研究之資料處理流程，以處理這些歷史資料。研究結果顯示在不分星期二、三、四的情況下，迴歸模式與實際驗證之資料其波動相當一致，僅在時段五跟時段六出現比較大的誤差，最大誤差也不超過一分鐘，故可知本研究之迴歸模式具有相當高的準確性。

關鍵詞：數位式行車紀錄器、旅行時間預測、複迴歸分析。

The Study of Predicting Travel Time by Using Tachograph' s Data

Song-Wei Chu¹ Chia-Ming Chang²

Abstract

There are few people whatever internal or overseas use the data from Tachograph to predict travel time. In this study, the author adopted the data from Tachograph to probe into the travel time prediction of freeway to accomplish the purpose of the study. The objective of this study is to use the extensive recording data from Tachograph to develop a set of data processing. Second, is to predict the spending of travel time from Yuanshan interchange of National Highway No. 1 to Taoyuan interchange. There are 60,387,756 historical data of this study. The study developed a set of data processing which is suitable the research to deal with these historical data. The study used a great deal of historical data construction model which is the Regression model got from Multiple Regression Analysis from Statistics Package Social Sciences 10.0 software. The outcome of the study show that the fluctuation of Regression Model accords with practical data in usual day, nothing but have more deviation in Interval Five and Interval Six. Although have more deviation in Interval Five and Interval Six, the maximum deviation is no more than one minute. So we can say that the Regression Model of the study have good accuracy.

Key words: Tachograph; Travel Time Predict; Multiple Regression Analysis

一、緒論

¹清雲科技大學行銷與流通管理系助理教授（聯絡地址：320 桃園縣中壢市健行路 229 號；電話：03-4581196 轉 7501；E-mail：swchu@cyu.edu.tw）

²清雲科技大學經營管理碩士（聯絡地址：320 桃園縣中壢市健行路 229 號；電話：03-4581196 轉 7501；E-mail：M9434011@cyu.edu.tw）

(一) 研究背景與動機

依據高速公路局統計，各收費站之年平均交通量已超過五億七千輛以上，且逐年成長中。在交通量急遽增加的情況之下，交通問題也相對越來越複雜，用路人不斷的要求提升服務水準，於是，如何提升服務水準便成為現階段政府改善交通的一個重要課題。

一般來說，蒐集即時旅行資料有兩種基本的方法。一為，每隔一段距離，在固定點設置偵測器(Vehicle Detector, VD)。另一，利用在交通流(traffic flow)中行駛之裝有全球衛星定位系統(Global Positioning System, GPS)配備之車輛來蒐集資料，此種裝有全球定位系統配備之車輛也就是探針車輛(Probe Vehicle, PV)。

本研究將探討高速公路之旅行時間預測，以供用路人及商車經營管理之用，並基於本研究對車輛偵測器及探針車輛瞭解，再加上數位式行車記錄器的資料是每零點五秒鐘就產生一筆資料，可詳盡的將行駛中的所有狀況紀錄下來，故決定棄以上兩種蒐集旅行時間資料的方法，而採用數位式行車記錄器的資料來預測旅行時間。

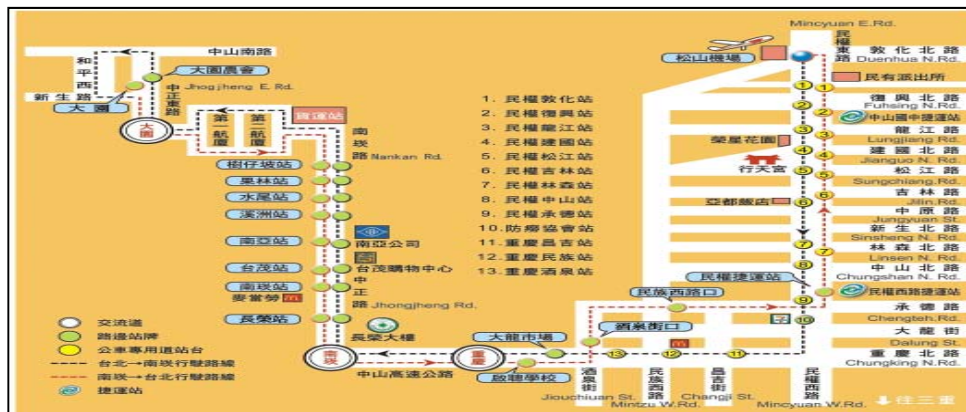
(二) 研究目的

從上述研究背景與動機的討論，可知道運用數位式行車記錄器的資料來預測旅行時間，具備有維護成本較低、記錄之資料較詳盡及應用層面較廣泛等優勢。故本研究欲達之目的如下：

1. 利用數位式行車記錄器所記錄下來之大量歷史資料，發展一套資料處理流程。
2. 使用數位式行車紀錄器之大量歷史資料，預測從中山高速公路圓山交流道到台北交流道所需花的旅行時間是多少，甚至預測從圓山交流道到高雄交流道所需花的旅行時間。

(三) 研究對象與範圍

本研究之數位式行車紀錄器之大量歷史資料，乃由與清雲科技大學行車運籌管理研究中心有長期合作的亞通客運提供，從98年1月到4月，資料筆數已經接近五億筆，而本研究選擇一月跟四月的每各禮拜非假日的二三四，排除四月五日的清明節，共24天，並且經過篩選過後，資料共有五百萬筆。亞通客運是屬於大型車，行駛固定的路線（除少數部分接受包車），其行駛路線包含桃園地區、中山高速公路和台北地區，且時間越尖峰，班次就越多，目前總共有24輛車輛裝有數位式行車紀錄器，且自98年1月起，每輛車，每隔0.5秒蒐集1筆資料並紀錄於行車紀錄器中，累積大量歷史連續資料，時空解像力較高，可應用於預測旅行時間之上。其數位式行車紀錄器所蒐集到的資訊，包含有車號、時間、經度、緯度及速度……等等。亞通客運其行駛路線如圖一所示：



圖一 亞通客運台北到桃園路線圖

由於亞通客運行駛範圍包含桃園地區、中山高速公路和台北地區，範圍太過於大，故本研究為專精於旅行時間預測方面，將研究範圍縮小，設定中山高速公路圓山交流道（23.2KM）到桃園交流道（49.1KM）約 26 公里的範圍間，完全排除桃園跟台北地區之市區資料，其行駛路線會由圓山上交流道，其間經過台北交流道、三重交流道、五股交流道、泰山收費站及林口交流道，最後下桃園交流道。

二、文獻探討

(一) 旅行時間預測模式

預估旅行時間最常見的方式有兩種：第一種是利用模擬程式，模擬出駕駛人行為之假設性資料，再透過假設性資料預估出旅行時間；第二種是利用即時或事後所偵測之資料，來進一步預估旅行時間。上述兩種方法中以第二種方式較為常見，故本研究亦採用第二種方式進行高速公路旅行時間預估。

(二) 使用自動辨識系統之資料預測高速公路之旅行時間

Laurence R. Rilett 等之研究針對高速公路路徑旅行時間預測之信賴區間估計提出一種方法論，並使用泰勒級數近似法來估計路徑旅行時間的期望平均數及變異數，此外亦引入路段預測平均數及變異數。

張慶麟之研究是以 AVI 方式，針對高速公路「路段中無施工」、「無事故且無道路縮減」的三車道基本路段來進行旅行時間的預估。

邱裕鈞等之研究利用多元迴歸技術建立在不同交通情境下兩者間的關係式。

(三) 使用探測車之資料預測高速公路之旅行時間

吳佳峰之研究主要透過車輛歷史旅行資料再加上 GPS 之衛星定位系統資訊來預估車輛旅行時間。

吳金杰此研究以微觀角度之車流模擬程式產生車輛偵測器與探測車資料，自行建立一套能夠融合車輛偵測器與探測車資料，並構建浮動加總旅行時間

預測模式。

(四) 使用偵測器之資料預測高速公路之旅行時間

張修榕此研究經由實際調查方式蒐集高速公路重要參數，以構建高速公路基本路段車流模擬模式，接著透過兩階段進行高速公路旅行時間之預測。

黃裕文之研究主要為了構建高速公路施工路段車流模擬模式、建立施工路段旅行時間預測模式及分析施工路段封閉車道長度的不同對於旅行時間之影響。

綜合上述所有文獻可知，利用即時或事後所偵測之資料，來進一步預估旅行時間的資料蒐集方式有四種，有自動車輛辨識、探測車、偵測器及行車紀錄器，而前三種資料蒐集方式運用在預估高速公路之旅行時間上，已經相當普遍，惟獨不同的是，大家使用研究方法的不同。目前數位式行車紀錄器之資料運用，都在研究管理司機的駕駛行為及車輛器材損壞等方面，鮮少有使用數位式行車紀錄器的資訊來預測旅行時間。在旅行時間預估研究方面，國內則因相關技術未能發展至一定水準，促使相關研究常侷限於有限的交通資料。本研究有鑑於此，故採用亞通客運所裝設之數位式行車紀錄器所得之大量資料，並配合上 GIS 軟體 ArcView，預測高速公路之旅行時間。

三、研究方法

由於數位式行車記錄器原本之歷史資料共有六千萬筆，數量相當地大，故本研究針對於此，發展出一套適用於此方面研究之資料處理流程。

- (一) 轉換資料：由於檔案格式為 SOR 的檔案，必須透過 HAS-101 數位式行車紀錄器所設計之解碼程式，將原本 SOR 的檔案格式轉換成 TXT 純文字檔的格式。
- (二) 合併檔案：由於數位式行車紀錄器之資料是以每台數位式行車紀錄器作為每天之記錄，故需要透過 AeroMerge 1.0 將各檔案合併成一個檔案。
- (三) 撰寫 Visual Basic 程式語言：因為數位行車記錄器是每零點五秒產生一筆資料並傳至資料庫，如此會造成資料處理上一種極大的負擔，所以本研究利用透過 Visual Basic 的程式語言去找出每六秒為一筆的資料，來當本研究之研究樣本。
- (四) 建立完整路段資訊庫：將每六秒為一筆之資料，利用匯入的方式轉進去 ACCESS，以構建完整資訊庫。
- (五) 資料庫匯出，轉成 DBF 檔格式：為利於地理資訊軟體讀取，故須先將 ACCESS 資料庫裡之資料，匯出成 DBF 檔案格式。
- (六) 環域分析 (Buffer Analysis)：本研究使用之地理資訊系統軟體為 Arcview，為了找出本研究之研究範圍資料，故使用 Arcview 之環域分析工具進行分析。

- (七) 資料轉換並進行統計檢定：由於從 Arcview 之環域分析得到研究範圍之資料為 DBF 格式，故需將資料轉成文字檔格式或 Excel 可讀取之格式，以利對其進行統計軟體 SPSS 之 T 檢定及建構高速公路之資料庫。
- (八) 剔除異常值和極端值：由於 GPS 是具有誤差的，故在利用 Arcview 選取研究範圍所需之資料時，有一些偏離之值會被涵蓋進來，此時須將其剔除。由於資料具有連續性，故當一筆或多筆資料並不具有連續性時，或雖具有連續性但未跑完全程，則須直接剔除，此為本研究之異常值判斷方法。
- (九) 判別南北向：經過本研究篩選過濾之後，只留下上交流道及下交流道兩筆資料，並配合上交流道經緯度之判斷，如此便可判別南北向出北上和南下資料。
- (十) 區分時段及樞紐分析：本研究依照交通部運研所高速公路路段速率推估之研究，將每天分為七個不同時段，以利透過樞紐工具進行分析。
- (十一) 建立高速公路資訊庫：進行完上述之步驟後，可將處理完之資料匯入 ACCESS 資料庫，建構符合本研究之空間及時間之資訊庫。
- (十二) 旅行時間：經過前述之步驟，接下即可利用 Excel 算出歷史資料之旅行時間，即可算出各交流道間之旅行時間及整段研究範圍之旅行時間。

四、模式建構

由於本研究主要是要預測旅行時間，故本章將利用統計複迴歸分析，建構預測旅行時間之模式。

(一) 變數選取：

本研究之資料來源，將從已建立之一月及四月高速公路資訊庫選取，選取 2009 年 1 月 6 日至 2009 年 4 月 30 日之一月及四月星期二、三、四資料，而資訊庫之資料具有北上與南下的資料，故本研究將選取北上路段作為分析。資訊庫裡的資料包含有車號、時段、星期、回報時間、旅行時間、速度、經度、緯度等等，本研究將選取星期、時段、旅行時間來作為複迴歸分析之輸入變數。

(二) 模式定義：

由於進行複迴歸分析時，須先行定義自變數與依變數，故本研究將旅行時間定義為依變數 (Y)，將星期、時段定義為自變數 (X)。複迴歸模式如下：

$$Y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_{10} x_{10} + \varepsilon$$

Y：旅行時間； α ：常數項； ε ：誤差項； x_1 ：星期二； x_2 ：星期三； x_3 ：星期四； x_4 ：時段一（0 點~7 點）； x_5 ：時段二（7 點~9

點)； x_6 ：時段三(9點~11點)； x_7 ：時段四(11點~15點)； x_8 ：
 時段五(15點~17點)； x_9 ：時段六(17點~19點)； x_{10} ：時段七(19
 點~24點)

(三) 模式建構：

將運用統計軟體 SPSS 10.0 進行複迴歸分析，以建立一套旅行時間的預測模型，並期能釐清研究變數間關係，以及如何對於依變數的變異提出一套具有最合理解釋。以下則是利用一月及四月之高速公路資訊庫資料進行複迴歸分析所跑出來之結果，如表一。

依照表一之結果，本研究之迴歸式如下：

$$Y = 1060338 + 12.125x_1 + 23.198x_2 + 39.855x_3 + (-90.871x_4) + 105.754x_5 + 135.788x_6 + 6.352x_7 + 46.876x_8 + 320.755x_9 + 59.421x_{10}$$

表一 係數分析表

模式	未標準化係數		標準化係數	顯著性
	B 之估計值	標準誤	Beta 分配	
常數	1060.338	14.216		0.000*
星期二	12.125	11.674	0.031	0.000*
星期三	23.198	13.158	0.038	0.000*
星期四	39.855	13.149	0.065	0.002*
時段一	-90.871	22.354	-0.089	0.000*
時段二	105.754	17.512	0.146	0.000*
時段三	135.788	20.686	0.148	0.000*
時段四	6.352	11.324	0.026	0.000*
時段五	46.876	20.465	0.052	0.021*
時段六	320.755	20.154	0.361	0.000*
時段七	59.421	17.400	0.083	0.001*

*：代表 95%顯著之水準

(四) 配適度檢測及模式校估：

經過複迴歸分析之後，得到預測旅行時間之模式，但模式在使用之前須檢測其配適度及進行模式校估，以使模式的解釋力及精確度提高。以下將對模式進行檢測及校估：

1.F 檢定值：

F 檢定值係檢定係數是否具有穩定性，係數的穩定性亦代表迴歸方程式之配適程度，若配適度高則表示其模式可行，若配適度低則表示其模式需修正。

H_0 ：迴歸係數等於零

H_1 ：迴歸係數不全為零

若 $F^* > F_{0.025}$ ，則拒絕 H_0 ，表示係數不全為零，可解釋依變數與自變數之間具有線性迴歸之關係。

若 $F^* < F_{0.025}$ ，則接受 H_0 ，表示係數等於零，可解釋依變數與自變數之間具不具有線性迴歸之關係。

在 $\alpha=0.025$ 的顯著水準下，模式中之 F 值檢定結果 $F=38.76547 > F_{0.025}$ ，因此模式拒絕 H_0 之假設，故表示該模式係數不全為零，亦配適度高，則表此模式可行。如表二：

表二 模式之變異數分析

模式	平方和	自由度	F 檢定	顯著性
迴歸	27325710	10	38.76547	0.000*
殘差	167573315	2373		
總和	194899025	2383		

*：代表達 95% 之顯著水準

2. 共線性：

若變數間具有共線性問題，其變異數膨脹因素(Variance Inflation Factors, VIF)的極大值將大於 10，允差(Tolerance)的極小值將小於 0.1，故本研究將以此為標準，檢定迴歸模式是否須作校正。如表三所示，允差皆大於 0.1 並且變異數膨脹因素也都小於 10，故發展出的模式不具有共線性，不需校正即可使用。

表三 共線性統計量之判斷

共線性統計量		
變數	允差	變異數膨脹因素(VIF)
星期二	0.732	1.288
星期三	0.790	1.266
星期四	0.790	1.265
時段一	0.758	1.319
時段二	0.620	1.613
時段三	0.720	1.390
時段四	0.685	1.399
時段五	0.713	1.402
時段六	0.705	1.418
時段七	0.616	1.624

五、模式驗證

根據第四章所發展出的迴歸模式，本章將對其迴歸模式進行驗證，驗證之資料選擇亞通客運數位式行車紀錄器之 2009 年 1 月 6 日-2009 年 1 月 15 日之星期二、三、四北上資料，共計六天。在進行驗證前，須先把 2009 年 1 月 6 日-2006 年 1 月 15 日之實際歷史資料進行如第三章之資料處理，繼而進行驗證程序。

(一) T 檢定

在資料處理過程中需經過統計檢定，主要原因是必須先釐清資料之間是否沒有差異，而具有一致性，若具有一致性才可進行。以下將利用統計軟體

SPSS 10.0 針對 2009 年 1 月 6 日-2009 年 1 月 15 日進行獨立樣本 T 檢定：

$$H_0 : \mu_0 = \mu_1$$

$$H_1 : \mu_0 \neq \mu_1$$

在 $\alpha=0.05$ 的顯著水準下，結果顯示 $F=1.557 < F(1-0.05)$ ，且顯著性(雙尾)為 $0.212 > \alpha=0.05$ ，故不顯著，因此模式接受 $H_0 : \mu_0 = \mu_1$ 之假設，表示本研究之模式推估出路段行駛速率之資料與北上資料庫之一星期資料是沒有差異的，具有一致性。

表四 獨立樣本 t 檢定

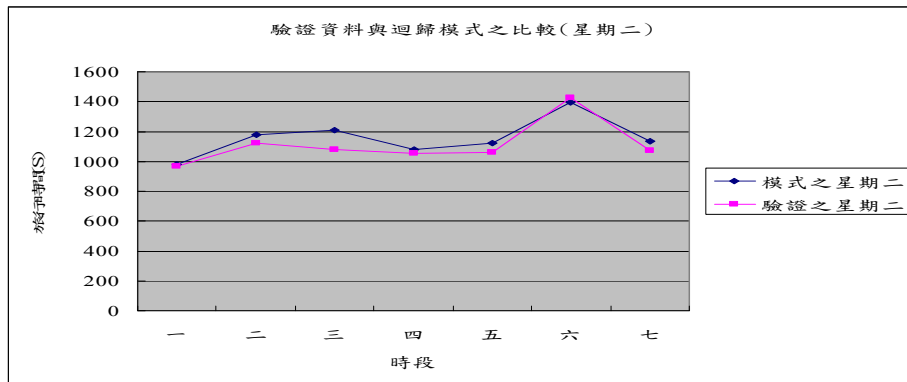
	變異數相等的 Levene 檢定		平均數相等的 t 檢定		
	F 檢定	顯著性	t	自由度	顯著性(雙尾)
假設變異數相等	1.557	0.212	-0.521	2388	0.602
不假設變異數相等			-0.526	1350.838	0.599

(二) 驗證資料

以下將處理完的 2009 年 1 月 6 日-2009 年 1 月 15 日之北上路段資料當驗證資料及本研究所發展出的迴歸模式之平均旅行時間進行比較：

1. 星期二之驗證資料與迴歸模式：

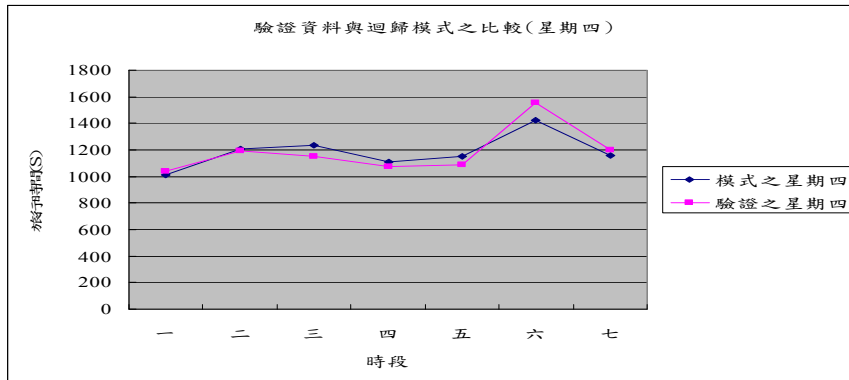
將星期二之驗證資料與迴歸模式相比較，可看出迴歸模式與驗證資料之間的差異並不大，只有些微的差距。如圖二



圖二 驗證資料與迴歸模式之比較 (一)

2. 星期四之驗證資料與迴歸模式：

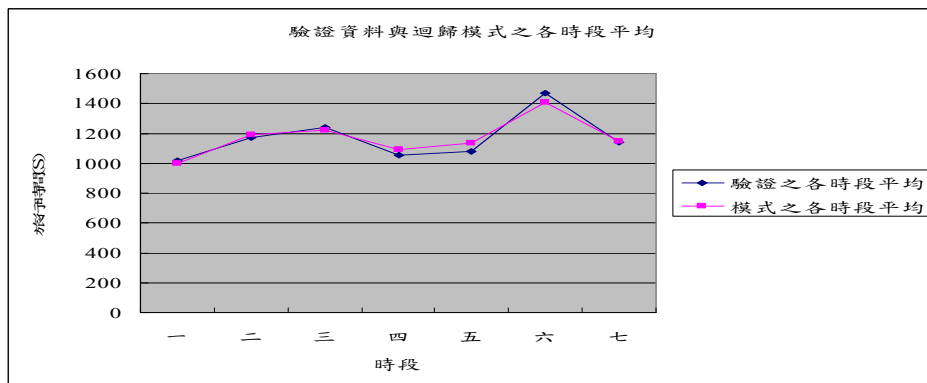
將星期四之驗證資料與迴歸模式相比較，可看出迴歸模式其可準確預測出旅行時間之波動，並且與驗證資料相去不遠，只有些微差距。如圖三



圖三 驗證資料與迴歸模式之比較 (二)

3. 驗證資料與迴歸模式之各時段平均：

將驗證資料與迴歸模式之星期二三四的每各時段各自加總起來並平均，可以看出迴歸模式之預測旅行時間在各時段是相當的不錯，皆只有些微的差異，故本研究所發展的迴歸模式有不錯的預測效果。



圖四 驗證資料與迴歸模式之各時段比較

六、結論與建議

(一) 結論

綜合上述，本研究之結論有以下幾點：

1. 一般的研究都未處理如此大量的歷史資料，故本研究發展出一套嚴謹的資料處理方法，適用於處理大量歷史資料的研究，可作為後續研究者參考。
2. 本研究所建構之模式，乃利用數位式行車紀錄器之大量歷史資料，透過統計 SPSS 10.0 軟體執行複迴歸分析，而得之迴歸模式。
3. 本研究將發展出來的迴歸模式與實際驗證之資料做比較，得知在不分星期二、三、四的情況下，也就是我們定義的平常日，迴歸模式與實際驗證之資料其波動相當一致，僅在時段五六出現比較大的誤差，最大誤差也不超過一分鐘，故可知本研究之迴歸模式具有相當高的準確性。

(二) 建議

本研究雖然已有不錯的成果，惟受限於時間與研究限制，仍有幾點建議可供後續研究者研究，分別說明如下：

1. 增長時間範圍：本研究因受限於研究時間與人力，將研究範圍設定為中山高速公路北上 23.2KM-49.2KM，但其中的變數仍相當的多，宜將時間範圍增長，將各時段之樣本數提高，如此應可增加迴歸模式的精準度。
2. 加入影響因素：在道路行駛時會遇到許多的狀況，如尖峰、離峰、時段等，而這些狀況會影響到旅行時間，但本研究均未將其影響因素納入考慮，故建議後續研究者，可以將其納入考量。
3. 納入小客車資料：本研究僅對亞通客運的大客車資料做分析，但此迴歸模式將無法預測小客車之旅行時間，故建議可以將小客車之資料加入分析，以讓模式發展出來可適用於各種車種。

參考文獻

1. Laurence R. Rilett, Dongjoo Park, Byron Gajewski, 1999, *Estimating Confidence Interval for Freeway Corridor Travel Time Forecasts*, 6th World Congress on Intelligent Transport Systems, Toronto, Canada.
2. Turner, S. M., and Holdener, D. J., 1995, *Probe Vehicle Sample Size for Real-Time Information: The Houston Experience Proc.*, Vehicle Navigation and Information Sys. (VNIS) Conf., IEEE, New York, pp. 3-9.
3. 交通部統計處，2006 年，交通統計月報。
4. 交通部運輸研究所，2005 年，高速公路路段速率推估之研究。
5. 吳佳峰，2003 年，有 GPS 資訊提供下之車輛旅行時間預估模式之研究，交通大學交通運輸工程研究所碩士論文。
6. 吳金杰，2005 年，融合偵測器與探測車資料預測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。
7. 邱裕鈞、艾嘉銘、廖梓淋，2006 年，利用車輛偵測器偵測速率資料建立空間旅行速率之預測模式，九十五年道路交通安全與執法研討會論文集，桃園：中央警察大學交通學系，頁 495-506。
8. 亞通客運，<http://www.yatungbus.com.tw/>
9. 張慶麟，2003 年，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。
10. 張修榕，2001 年，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。
11. 黃裕文，2003 年，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。