

# 第二屆高速公路ETC資料在交通管理 之應用創意競賽-旅行時間預測

## 多維分類與隨機狀態迴歸分析

隊長：錢冠邑

組員：黃士峰、許湘伶、張志浩、李家瑜、黃惠巧  
尤致景、林逸飛、陳嘉緯、呂弘屏

學校：國立高雄大學統計學研究所

# 前言

- 資料蒐集隨著科技發展日新月異，其蒐集種類與質量亦跟著邁入大數據時代。
- 本報告透過高速公路局提供，由高速公路電子收費系統（ETC）所蒐集之資料庫數據，進行國道旅行時間之全路段估計與預測。
- ETC採用無線射頻辨識（RFID）系統，經由前端與後端系統構聯運作，進行高速公路通行費電子收費作業，是目前國內最全面的一項智慧型運輸系統（ITS）應用，除應用於高速公路的電子收費，ETC系統所蒐集的資料更可透過統計等相關分析對車流量、行車速率與旅行時間之估計或預測等進行統計推論。

# 提綱

- ▶ 介紹
  1. 資料來源與概述
  2. 應用軟體
  3. 想法
- ▶ 資料前處理
- ▶ 方法與流程
  1. 階層式分群法
  2. 迴歸模型建立
- ▶ 整合與預測
- ▶ 旅行時間預測結果
- ▶ 總結

# 介紹



1. 資料來源與概述
2. 應用軟體
3. 想法

# 資料來源與概述

- 資料來源:交通部台灣區國道高速公路局「交通資料庫」
- 分析期間採4月16日起至7月15日共三個月之數據資料
- 本分析所採用資料量總計約六千五百萬筆
  
- 所使用的資料變數如下:
  1. 「站間各車種平均旅行時間(T)」 (TDCS\_M04A)
  2. 「站間各車種交通量(V)」 (TDCS\_M04A)
  3. 「站間各車種平均行駛車速」 (TDCS\_M05A)  
轉換為路段速差( $S$ ) = 路段平均行駛速率 - 路段速限 (據該車種速限)
  
- 採用交通資料庫以每五分鐘為時間單位統整之數據資料，則單日共計288筆

# 應用軟體

- ▶ R語言: 主要應用軟體
  - 資料整理
  - 資料分析
  - 資料繪圖
  - 輸出與建檔
  
- ▶ Excel: 輔助應用軟體



Excel

# 想法

## ▶ 各小路段

- 例如：01F0248S(圓山-台北) 至01F0264S(台北-三重)  
為一小路段

## 旅行時間、交通量會有特定趨勢：

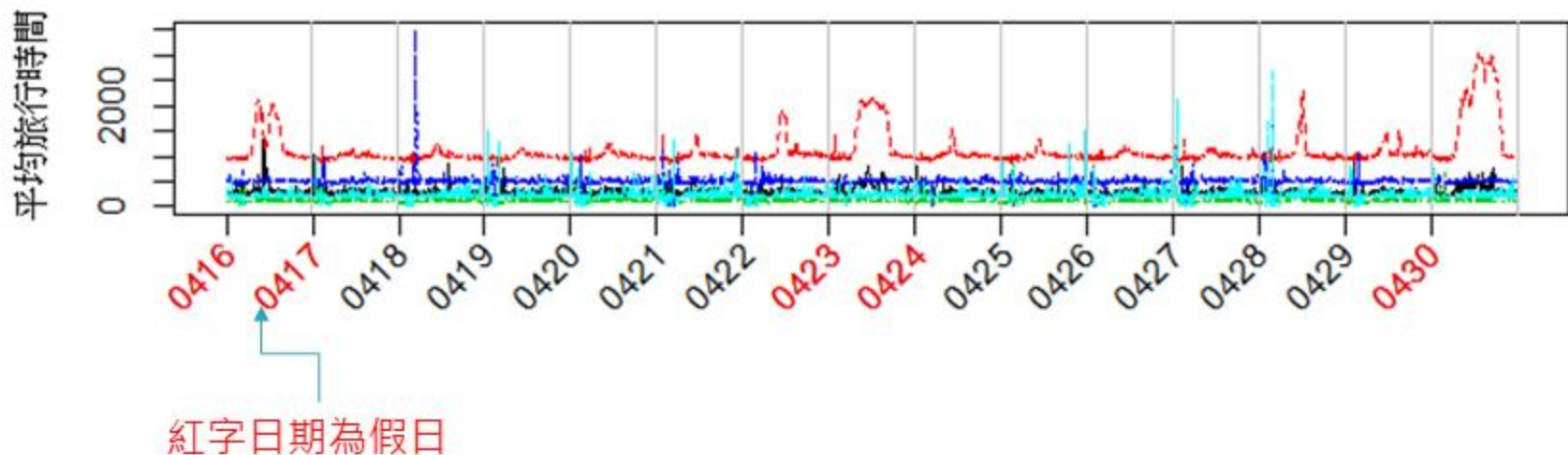
- 白天車流量較大，使得旅行時間增加
- 夜間車流量較白天來的少

## 平日、假日的趨勢亦有特定趨勢：

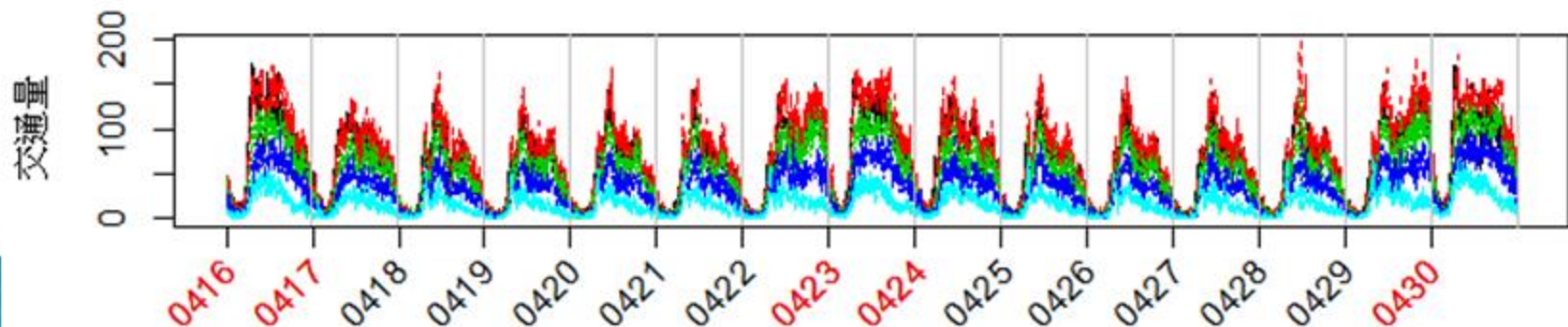
- 假日接近中午時，旅行時間變得特別長

# 小路段的趨勢(國五南下-南港至蘇澳)

## 各天平均旅行時間表現(南下)



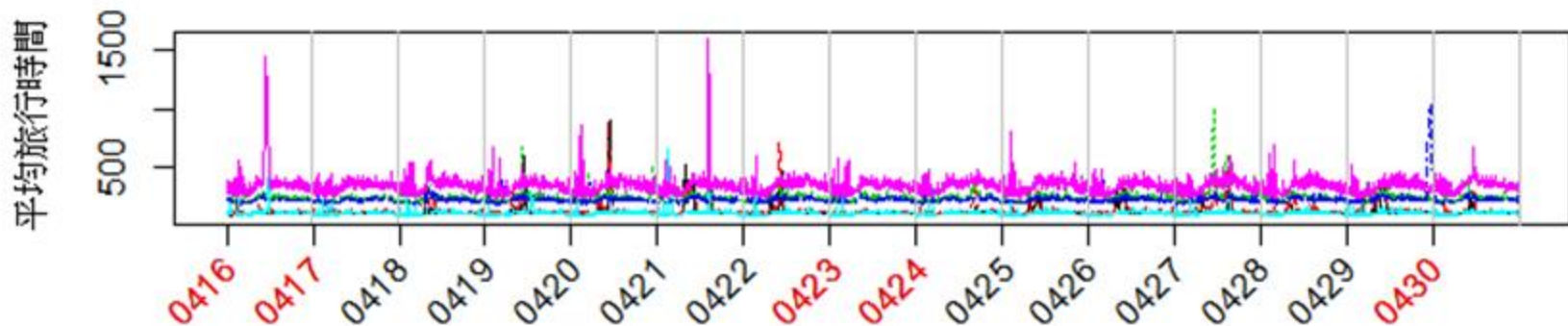
## 各天交通量表現(南下)



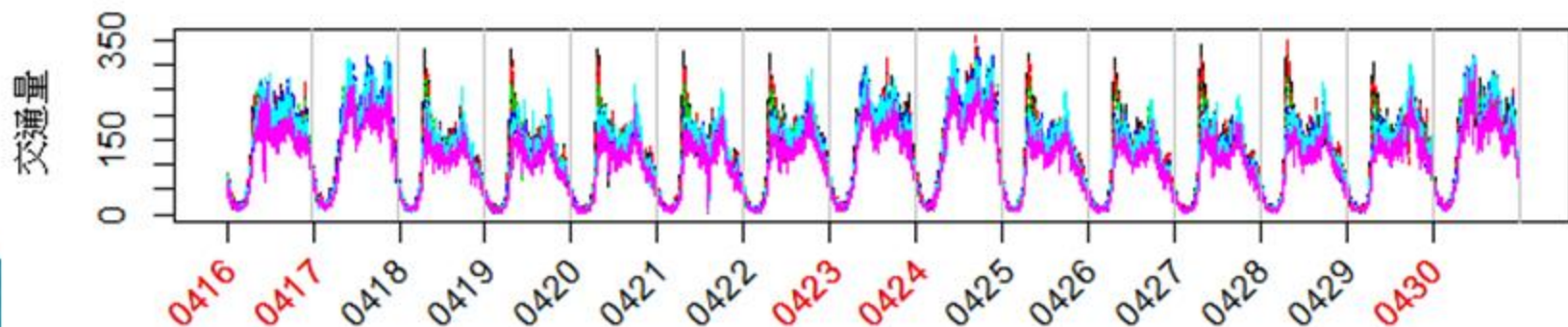


# 小路段的趨勢(國一北上-鼎金系統至台南)

## 各天平均旅行時間表現(北上)



## 各天交通量表現(北上)



# 想法

- ▶ 將平日、假日(含連續假期)分開建模與分析
- ▶ 將各小路段
  - Step 1. 依照每五分鐘為單位時間進行分群
  - Step 2. 再按各群資料進行建模與分析
- ▶ 透過時間推移演算法得出指定路段旅行時間

# 資料前處理



1. 新設交流道整合
2. 車種加權
3. 平滑化插補

# 資料前處理-新設交流道整合

## 大灣交流道南下匝道完工 下午通車



中廣新聞網 - 2016年7月6日 上午 11:13

### 相關內容



大灣交流道南下匝道完工 下午通車

歷經20多年規劃與建，終於完工的國道一號「大灣交流道」，繼去年底開放北上匝道通車之後，南下匝道也完成建置，今天下午4點開放通車，施工單位舉辦隆重的通車典禮；交通部長賀陳旦親自出席通車典禮，希望大灣交流道能夠與永康、仁德發揮更好的功能，工程完工後也會落實「售後服務」。

大灣交流道規劃「南出」、「北入」2條匝道，高公局簡報指出，「大灣交流道」距離台南火車站、成大醫院等地區，比永康、仁德兩個交流道近許多，初步評估可以節省用路人10分鐘至15分鐘繞行市區時間。

交通部長賀陳旦表示，工程完工後更要有售後服務，未來要保證交通要順暢，用路人能即時得到更便利資訊。賴清德說，原有的永康和仁德兩個交流道車流量多，每到年節交通流量爆量，市府如臨大敵，希望以後交通會更順暢。

高公局拓建工程處指出，大灣交流道南下出口通車後，相關行車動線有部分限制，請用路人遵行。(中廣記者 劉怡伶)

上游測站	下游測站
------	------

01F3252N (台南-大灣)	01F3227N (大灣-永康)
---------------------	---------------------

01F3286N (仁德系統-台南)	01F3252N (台南-大灣)
-----------------------	---------------------

整合為

1. 旅行時間、交通量

⇒ 相加

2. 車速(速差)

⇒ 平均

上游測站	下游測站
------	------

01F3286N (仁德系統-台南)	01F3227N (大灣-永康)
-----------------------	---------------------

# 資料前處理

## ▶ 車種加權

- 由於各車種在不同時間點，不同日期使用高速公路的頻率不同，故我們引入一權重來平衡此一影響。
- 各車種權重：

$$\frac{\text{第}i\text{時刻該車種交通量之和}}{\text{第}i\text{時刻所有車種交通量之和}}, i = 1, \dots, 288$$

## ▶ 平滑化插補

- 然而由於並未所有時間與路段皆會有車輛通過，以致平均旅行時間、交通量等之變數資料紀錄為0
- 以最近鄰居(KNN)演算法進行平滑化補正，採用欲插補數據位置最鄰近其他資訊進行插補

# 方法與流程



1. 階層式分群法
2. 迴歸模型建立

# 分群

- ▶ 採用階層式分群法(Hierarchical Clustering)  
將小路段依時間分群以降低組內變異
- ▶ 以集結式群集方法，來進行階層式分群
  - 先將每一個組資料視為單一個群集，然後將各群集依序合併，而其合併順序視合併後之組內總變異數之大小而定

# 模型建立

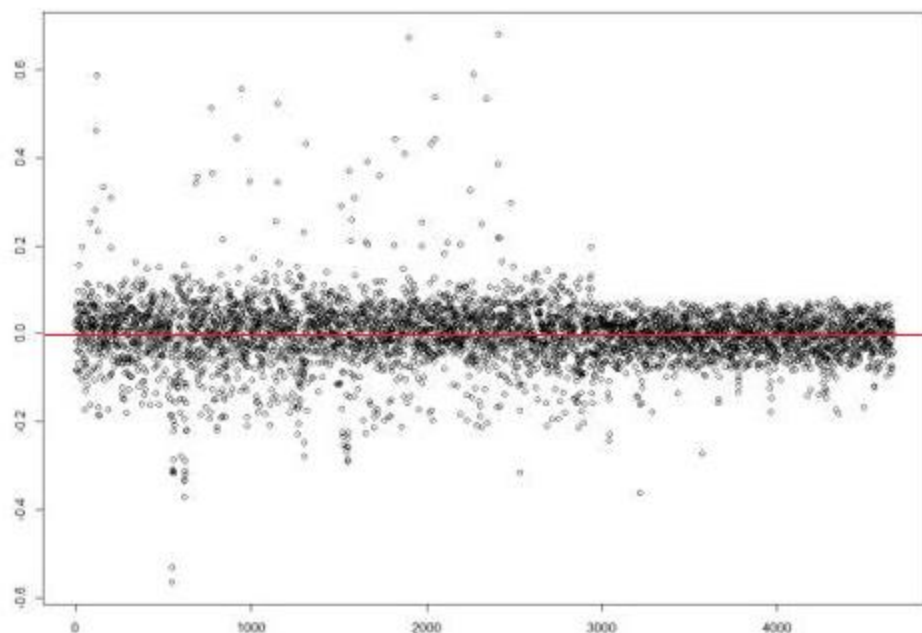
- ▶ 對前述分群後之資料，分別對各路段建立交通量、路段速差與星期幾的指標函數對旅行時間之迴歸模型
- ▶ 將模型建立後所得旅行時間估計值扣掉真實值，可得初步的殘差估計

$$\hat{\varepsilon} = \text{旅行時間} - \widehat{\text{旅行時間}}$$



# 模型建立 (conti.)

- ▶ 再以集結式群集方法對上一步所得到的初步殘差估計進行分群，進行模型重配適，下圖為模型重配適後之殘差



重配適模型之殘差

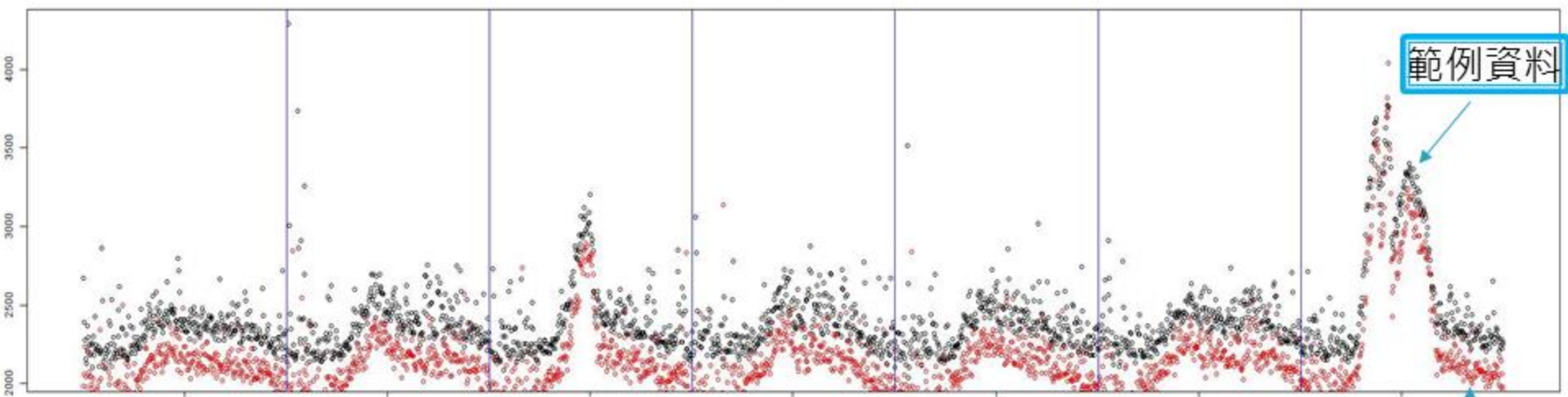
# 整合與預測



1. 全路段旅行時間推估
2. 旅行時間預測結果

# 全路段旅行時間推估

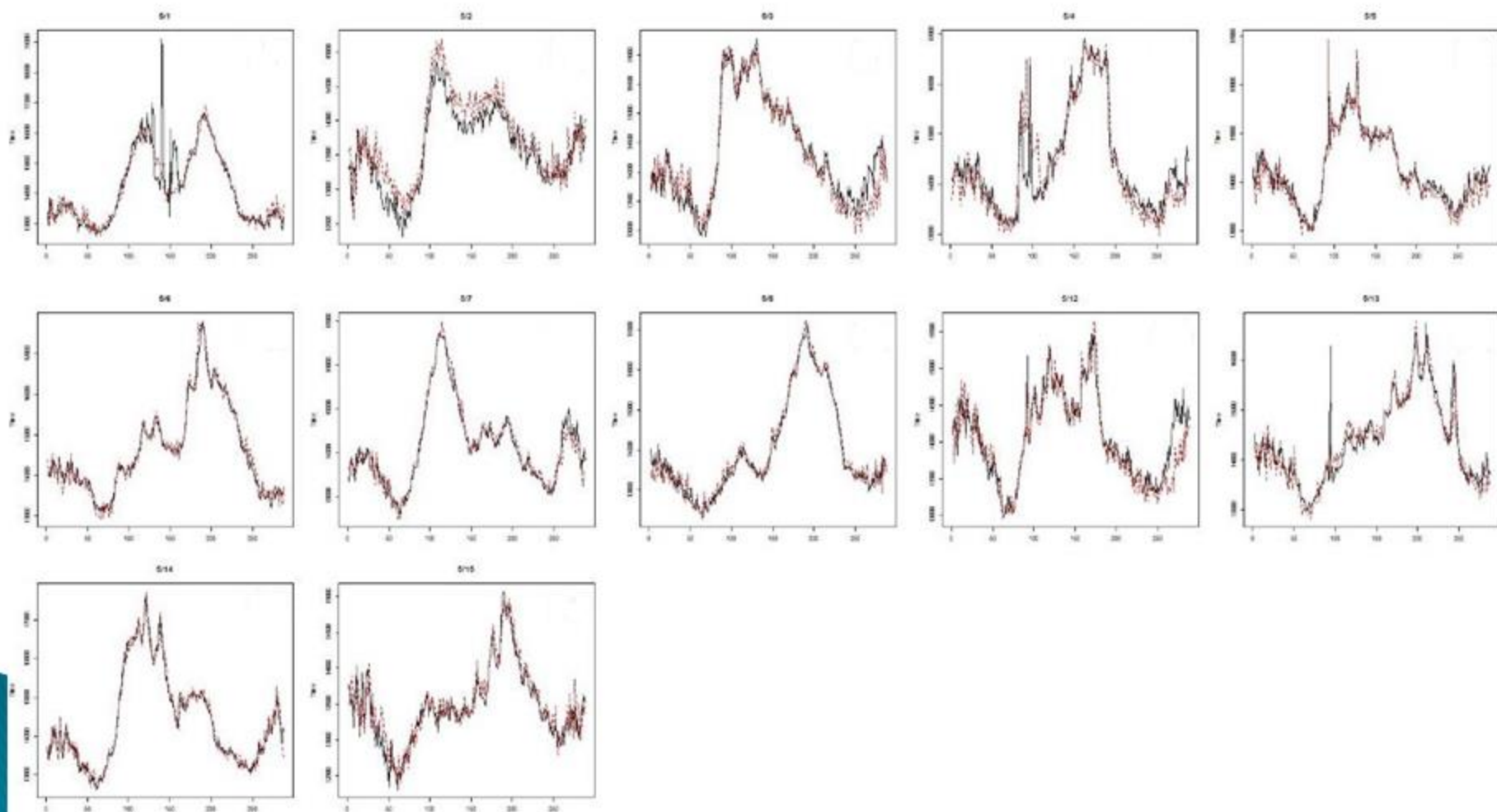
- 透過時間推移演算法得出指定路段旅行時間 (以國五南下為例)



- 因部分測站並非架設在交流道出入口上，故須進行修正，本方法採用**常數修正**，**透過與範例資料之差**，計算所需修正值並根據時間區段分為**凌晨**與**非凌晨**時段之修正值。

# 旅行時間預測結果

## ▶ 與主辦單位之範例資料比較



# 總結

- ▶ 本報告結合階層式分群法與迴歸模型分析進行各路段旅行時間之預測
  1. 依據歷史高速公路使用情況，將資料分為平日、假日兩群分別進行預測
  2. 針對不同車種分別進行加權及插補後整理成包含各路段之「旅行時間」、「車流量」與「路段速差」資料結構
  3. 採用階層式分群法進行分群，並將分群所得之資料分別進行迴歸模型分析
  4. 依照各路段預測結果，將之整合成指定路段之旅行時間預測值

謝謝您的聆聽!