FCU ePaper

逢甲大學學生報告 ePaper

報告題名：

高速鐵路延人公里變動趨勢研究

Trend Analysis of Passenger-Kilometers in High-Speed Rail

作者：吳睿哲、李怡青、謝帛澂、廖家鴻、王亮緯

系級：統計三乙

學號：D1242375、D1223203、D1242362、D1282629、D1242287

開課老師：劉峰旗 副教授

課程名稱：預測分析

開課系所：統計系

開課學年：114 學年度 第 2 學期

商學院

中文摘要

本研究旨在探討台灣高速鐵路自民國 96 年通車以來至民國 114 年間，其整體運周轉量(延人公里)之長期變動趨勢與季節性特徵。研究採用長達 19 年、共 228 筆的月度時間序列資料，運用 ARIMA 模型、時間序列分解法（Winters Multiplicati ve）及指數平滑法（Winters Additive）三種預測方法進行配適與預測。由於實證資料在民國 110 年因 COVID-19 三級警戒遭受劇烈的外在衝擊，本研究特別於 ARIMA 模型中加入介入變數（Intervention Analysis）以吸收極端異常值。結果顯示，含介入變數的 ARIMA$(0,1,1)(0,1,1)_{12}$ 模型在各項誤差評估準則（MSE、RMSE、MAE、MAPE）中表現最佳，能最精準地捕捉高鐵運量的長期上升趨勢與年底的高峰季節性規律。本研究不僅證實了台灣高鐵具備強大的疫後復甦力與長期成長動能，亦透過最佳預測模型為未來高鐵的運能調度與風險控管提供具體的科學建議。

關鍵字：高鐵、延人公里、時間序列分析、ARIMA模型、介入分析、時間序列分解法

Abstract

This study aims to explore the long-term trends and seasonal characteristics of the overall passenger turnover (passenger-kilometers) of the Taiwan High-Speed Rail (THSR) from its commencement of operations in 2007 to 2025. Using 19 years of monthly time series data comprising 228 observations, this research employs three forecasting methods for fitting and prediction: the ARIMA model, the Time Series Decomposition Method (Winters Multiplicative), and the Exponential Smoothing Method (Winters Additive). Because the empirical data suffered a severe external shock in 2021 due to the COVID-19 Level3 alerts, this study specifically incorporates intervention variables (Inter venation Analysis) into the ARIMA model to absorb extreme outliers. The results show that the ARIMA$(0,1,1)(0,1,1)_{12}$ model with intervention variables performs best across various error evaluation criteria (MSE, RMSE, MAE, MAP E), most accurately capturing the long-term upward trend and the year-end p eak seasonal patterns of THSR passenger volume. This study not only confirms THSR's strong post-pandemic resilience and long-term growth momentum but also provides concrete scientific recommendations for future capacity scheduling and risk management through the optimal forecasting model.

Keywords High-Speed Rail, Passenger-kilometers, Time Series Analysis, ARIMA Model, Intervention Analysis, Time Series Decomposition Method

目　次

中文摘要.....1

英語摘要.....2

目次.....3

第一章 前言.....4

　第一節 研究背景與動機.....4

　第二節 研究目的.....4

　第三節 章節介紹.....4

第二章 研究方法.....5

第三章 資料分析.....6

　第一節 資料說明.....6

　第二節 敘述統計.....6

　第三節 資料初探.....7

　第四節 資料分析結果.....10

　第五節 預測結果比較.....21

第四章 結論與建議.....21

參考文獻.....23

附錄.....23

第一章 前言

第一節 研究背景與動機

自民國 96 年 1 月台灣高速鐵路正式通車以來，其憑藉著高速與準點的特性，有效縮短了台灣南北的時空距離，成功建構出「一日生活圈」，並從根本上改變了民眾的通勤、商務與旅遊模式。本研究採用自通車元年至民國 114 年 12 月、橫跨 19 年的月份資料，完整記錄高鐵從新興交通工具成長為台灣西部長途運輸支柱的歷程。在此背景下，我們的研究動機源於高鐵在這 19 年間所經歷的多個重要歷史節點，期盼透過統計數據的變動來觀察其經營韌性。首先是營運初期的成長挑戰，探討高鐵在初期入不敷出且面臨民國 97 年全球金融海嘯的外部環境下，運量卻能穩定增長的原因，其次是擴展新站，研究苗栗、彰化、雲林新增三站的加入，對整體「延人公里」所帶來的結構性影響；最後則是疫情期間的應對，分析在 COVID-19 疫情導致運量斷崖式下跌後，高鐵的運量是否已完全恢復至疫前水準。最終，我們希望藉由這 19 年的縱向數據，深入探討高鐵在經歷長期的成長後，其運量是否已逐漸接近飽和點。

第二節 研究目的

(一)探討高鐵運量之歷史變動特徵

運用敘述統計、時間序列圖、ACF 圖與 PACF 圖，全面剖析台灣高鐵自通車元年至民國 114 年 12 月（共 19 年）的月分運量資料。藉此釐清資料的基礎特徵，並觀察營運初期、金融海嘯、新三站擴建及 COVID-19 疫情等重大歷史節點對運量帶來的結構性變動

(二)評估多種時間序列模型之適用性

針對高鐵運量資料，研究並導入至少三種統計與預測方法，並進行嚴謹的參數估計與殘差診斷分析，以確保模型符合統計假設。

(三)甄選高鐵運量之最佳預測模型

採用各種預測評估準則，對各模型的預測能力進行客觀、量化的優劣比較，從中甄選出最契合高鐵運量趨勢、預測精準度最高的時間序列模型

(四)預測高鐵未來運量與飽和趨勢

透過最佳模型的預測結果與長遠趨勢分析，驗證高鐵在經歷 19 年的成長後，其運量是否已接近飽和點，並以此預測結果為基礎，於結論中為高鐵未來的長途運輸調度、營運韌性及政策規劃提供具體的科學建議。

第三節 章節介紹

首先說明本研究之背景與動機，透過回顧台灣高鐵通車至今所歷經之重要歷史節點，包括金融海嘯、擴建新站以及疫情衝擊等重大事件，進而確立利用時間序列模型評估高鐵運量動態變動之核心研究目的。而在第二章研究方法則針對本研究所採用的三種核心統計模型進行深入的理論介紹，詳細闡述自迴歸整合移動平均模型（ARIMA）、時間序列分解法（Decomposition Method）與指數平滑法（Exponentia l Smoothing Method）

的數學架構、底層邏輯與各自適合之適用情境。接著在第三章資料分析為本篇研究之核心實證部分，內容詳述資料來源與敘述性統計特徵，並透過觀測時間序列圖、自相關與偏自相關圖（ACF/PACF）以及單根檢定進行資料的初步探討；隨後，本章分別展示上述三種預測模型在系統中的參數估計數值與殘差白噪音診斷結果，最終運用 R平方及多項預測誤差評估準則（如 MAPE、RMSE 等）進行嚴謹的綜合效能比較，藉此選出表現最佳之預測模型。最後的第四章結論與建議則系統性地歸納整體實證數據之重要發現，論證高鐵運量在長期增長軌跡下的趨勢特徵與規律季節變動，並據此統計結果，針對未來高鐵在淡旺季的營運運能調度、以及面對外部極端風險時的衝擊緩衝機制，提出具體且具可行性之實務建議。

第二章 研究方法

我們為了在於針對特定時間序列資料進行嚴謹的精準預測與特徵建模。由於這次採用的原始資料屬於月度數據的型態，在時間序列特徵上，往往同時隱含長期的發展趨勢以及每 12 期為一循環週期之強烈月季節性波動。為全面捕捉資料之線性與非線性變動特徵，並建立具備高度客觀性之預測決策模型，我們將利用三種具備學術代表性之時間序列預測方法進行實證分別為：時間序列分解法（Time Series Decomposit ion Method）、指數平滑法（Exponential Smoothing Method），以及 ARIMA 模型（Autoregressive Integrated Moving Average Model）。以下是各方法的介紹:

一、ARIMA 模型（Autoregressive Integrated Moving Average）

自迴歸整合移動平均模型是由 Box 與 Jenkins 所提出的單變量時間序列預測工具其架構主要由自迴歸(p)、差分(d)與移動平均(q)三個部分組成，若原始序列非平穩，需先透過 d 次差分程序將其轉換為平穩序列 W_t，進而建立經典模型公式:

$$W_t = c + \phi_1 W_{t-1} + \cdots + \phi_p W_{t-p} + s_t + \theta_1 s_{t-1} + \cdots + \theta_q s_{t-q}$$

其中 ϕ_i 為自迴歸係數，θ_j 為移動平均係數，s_t 則為當期的隨機衝擊或預測誤差。此模型純粹利用資料自身的歷史軌跡來對未來進行預測，在缺乏外部解釋變數的情況 下，能精準捕捉數據內在的慣性與自相關結構，因此在統計學與金融領域中被廣泛應用於中短期的趨勢預測。

二、時間序列分解法（Time Series Decomposition Method）

時間序列分解法主張一個複雜的時間序列變動，是由多個具備明確物理或經濟意義的潛在成分疊加而成，其目的在於將原始序列拆解以看清數據的本質特徵。傳統分解法將序列系統性地拆解為長期趨勢成分（T_t）、景氣循環成分（C_t）、季節成分（S_t）以及隨機不規則成分（I_t），根據數據波動的特性，成分間的組合公式主要分為季節波動不隨時間改變的加法模型:

$$Y_t = T_t + C_t + S_t + I_t$$

以及季節波動隨著長期趨勢放大或縮小的乘法模型:

$$Y_t = T_t \times C_t \times S_t \times I_t$$

此方法能將複雜的數據結構條理化，對於理解數據背後的長期變動規律與擬定營運策略具有高度實用價值。

三、指數平滑法（Exponential Smoothing Method）

指數平滑法是一種奠基於加權移動平均思想的預測技術，其核心觀點認為距離當前越近的歷史觀測值對未來的影響力越大，其最基礎的簡單指數平滑公式為:

$$Y_{t+1} = \alpha Y_t + (1 - \alpha)\hat{Y}_t$$

其中 Y_{t+1} 為下一期的預測值，Y_t 與 $\hat{Y}_t$ 分別為當期的實際觀測值與預測值，而 α 則為介於 0 到 1 之間的平滑常數，決定了新舊數據的加權比重。此方法根據資料特徵的複雜度，進一步延伸出引入趨勢平滑方程的 Holt 線性趨勢模型，以及同時融入趨勢與季節成分的 Holt-Winters 模型。由於該方法公式計算簡便、對新數據反應靈敏且具備高度適應性，因此非常適合應用於需要快速更新預測結果的中短期預測需求。

第三章 資料分析

第一節 資料說明

本次報告的資料來源為行政院主計總處，涵蓋期間自民國96年（2007年）1月起至民國114年（2025年）12月止，共計228筆月資料。資料單位為延人公里（千人公里）係指高鐵客運人數與其行駛公里數之乘積總和，用以衡量整體旅客運輸周轉量。延人公里是運輸量的重要指標，能同時反映載客人數與運輸距離，是評估高鐵營運規模與服務效率的常用衡量單位。

第二節 敘述性計

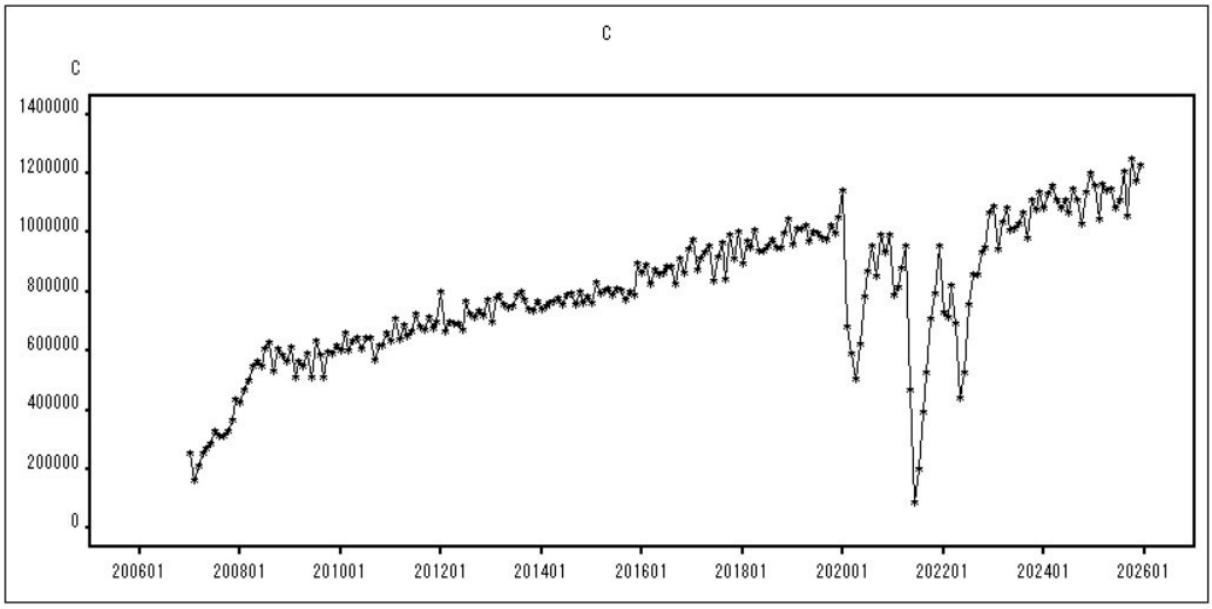
本次資料顯示全期228個月的資料顯示，台灣高鐵延人公里的平均數為793,274千人公里，這代表在過去19年間，台灣高鐵平均每個月約產生79.3萬延人公里的運輸周轉量，若以每日估算，平均每天約有26,100延人公里的運量，顯示高鐵已成為台灣西部走廊重要的長程運輸主力。而本資料的中位數為790,886千人公里，與平均數非常接近，顯示在未考慮時間趨勢前，全期數據的分布大致呈現對稱，群體運量呈現一個相對穩定的核心區間。資料的最小值出現在民國110年（2021年）6月，為89,477千人公里，此時期正值COVID-19疫情三級警戒高峰，運量受到嚴重衝擊；而最大值則出現在民國114年(2025年）10月，為1,248,678千人公里，最大值與最小值之間存在高達 1,159,201千人公里的差距，且最高峰約為最低點的14倍之多，這清楚反映出高鐵運量在疫情衝擊下的劇烈震盪，以及疫情後強勁復甦的成長動能。本研究資料的標準差高達228,960千人公里，如此龐大的標準差說明台灣高鐵的運量並非每個月都穩定維持在相同水準，而是存在極為顯著的月度波動，此特徵強烈支持本研究的初衷：高鐵延人公里在時間軸上的散佈具有高度的不確定性與季節性，因此單憑常規的平均數已不足以掌握其變化規律，必須進一步引入時間序列模型，才能有效捕捉並預測這高達二十餘萬延人公里的月度波動幅度。

第三節 資料初探

本節運用 SAS 軟體對台灣高鐵延人公里時間序列資料進行初步探索與分析，包含時間序列圖、ACF 圖、PACF 圖的視覺化檢視，以及平穩性檢定、模型辨識、參數估計、殘差診斷與短期預測等步驟，目的在於了解資料特性、確認適當的模型形式，並為後續正式模型建構奠定基礎。

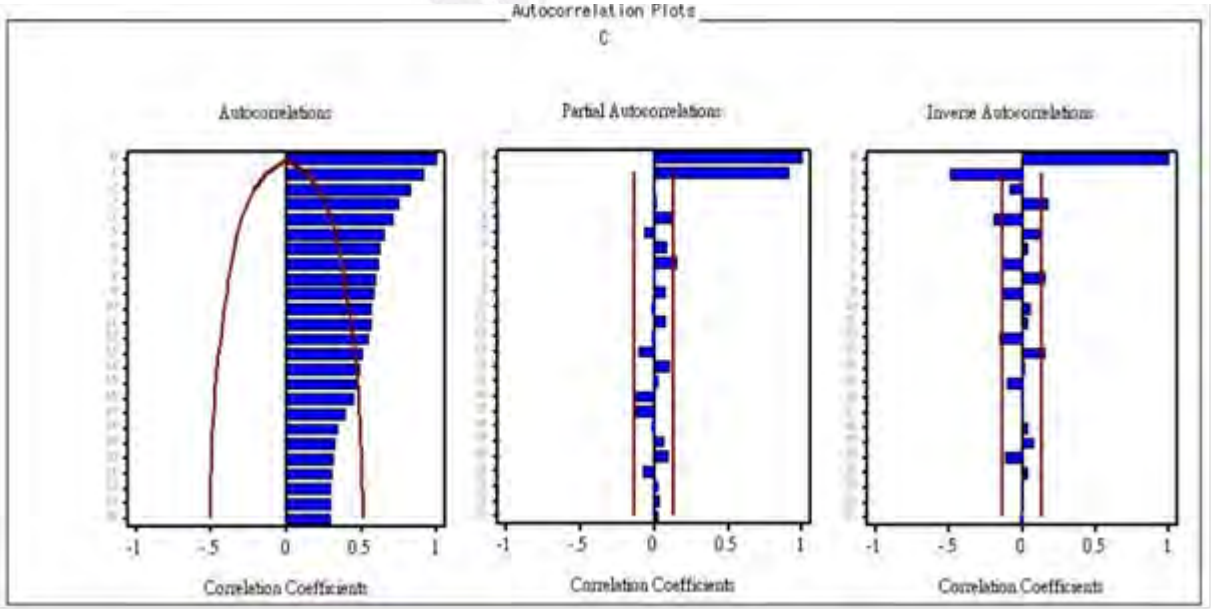
一、 原始資料

首先由原始資料的時間序列圖可見（圖一），台灣高鐵延人公里自民國96年起呈現長期上升趨勢，並伴隨明顯的年度季節性波動，每年春節、暑假及年底連續假期均出現運量高峰。然而，在民國110年（2021年）5月至8月間，因COVID-19三級警戒影響，運量出現斷崖式大幅下跌，此結構性異常需後續進行特別處理。



圖一 原始資料的時間序列圖

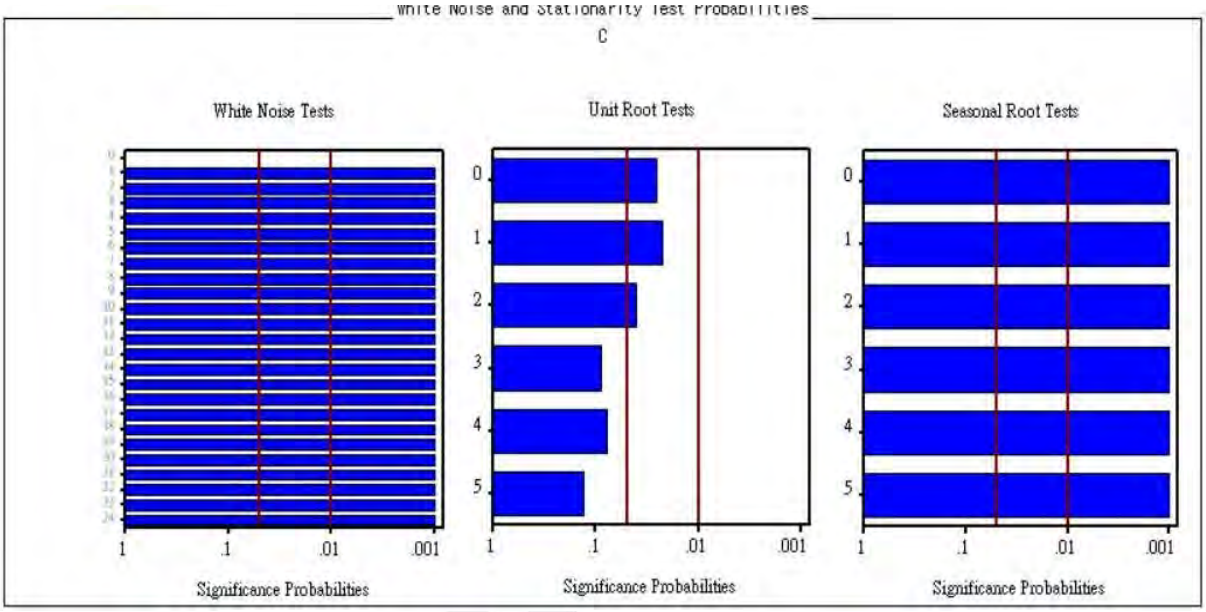
觀察自相關圖（圖二）可知，原始序列的 ACF 圖呈現明顯的 Tailing-off（緩慢衰減）型態，多個滯後期均超出顯著界線，顯示序列具有高度的慣性與非平穩特性；PACF 圖則在 Lag 1 附近出現較明顯的截斷現象，顯示當期運量深受前一期影響。綜合時間序列圖與 ACF/PACF 圖的結果，判斷原始序列為非平穩序列，需進行差分處理。



圖二 原始資料的自相關圖

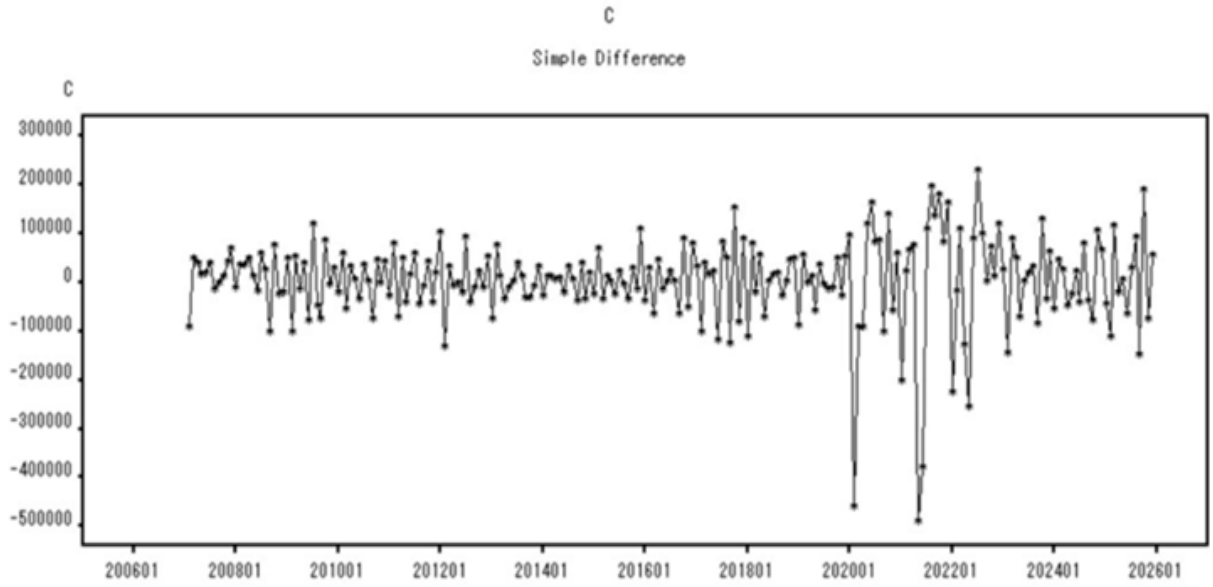
進一步觀察自相關圖（圖二），原始序列的 ACF 圖呈現明顯的 Tailing-off（緩慢衰減）型態，多個滯後期均超出顯著界線，顯示序列具有高度的慣性與非平穩特性；PACF 圖則在 Lag 1 附近出現較明顯的截斷現象，顯示當期運量深受前一期影響。綜合時間序列圖與 ACF/PACF 圖的結果，判斷原始序列為非平穩序列，需進行差分處理。

圖三 原始資料的白噪音檢定與單根檢定結果



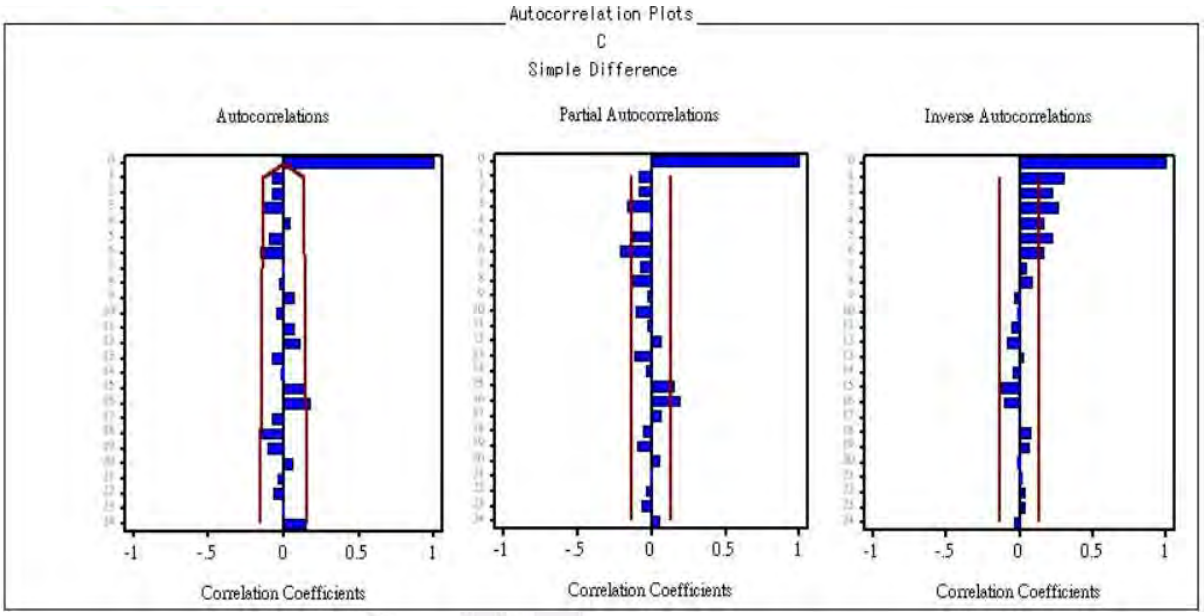
白噪音檢定與單根檢定結果（圖三）顯示，白噪音檢定的p-value 均小於0.001，強烈拒絕原始序列為白噪音的假設；單根檢定（Unit Root Test）的 p-value 均大於 0.05，無法拒絕非平穩假設，確認原始序列為非平穩（I(1)）。因此本研究對序列進行一階差分，以達到平穩狀態。經上述資料初探，本研究初步辨識適合以 AR 為主、結合季節性成分的 SARIMA模型進行後續分析，為正式模型配適與預測奠定基礎。

二、 一次差分



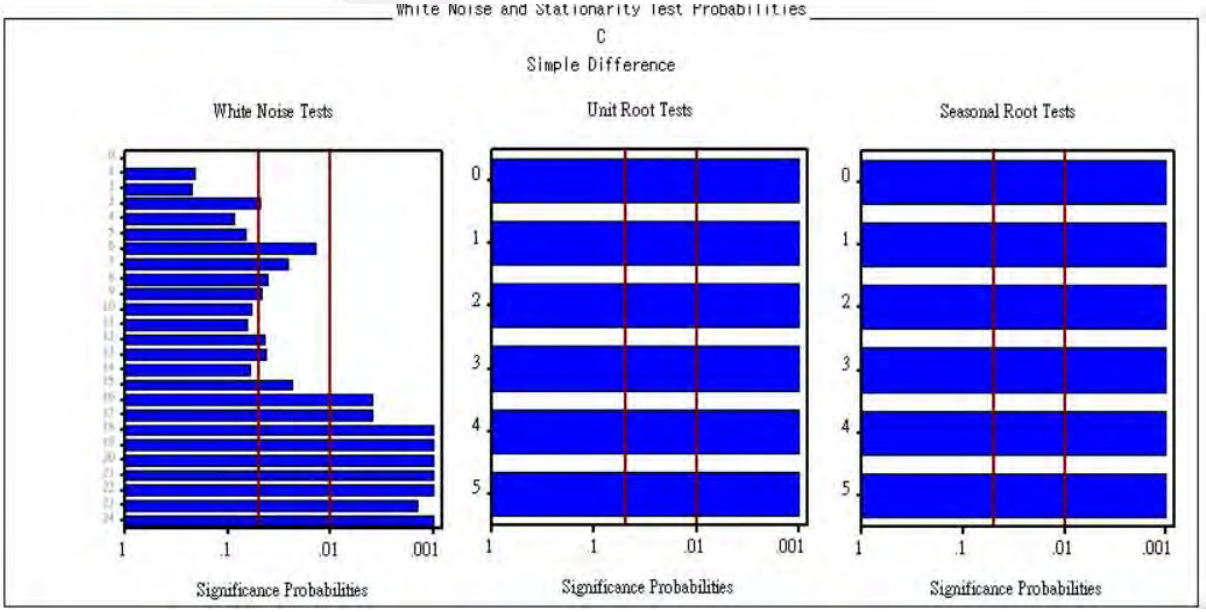
圖四 一次差分的時間序列圖

為確認原始序列的非平穩特性並使其達到平穩，本研究對原始資料進行一階差分處理。以下針對差分後的序列進行進一步分析，一階差分後的時間序列圖（圖四）顯示，序列已在0附近波動，整體均值接近0，長期趨勢已基本消除，達到平穩序列的特性。雖然2020年至2022年間仍可觀察到較大幅度的震盪，但整體已無明顯向上或向下的趨勢。



圖五 差分後的自相關圖

進一步觀察差分後的自相關圖（圖五），ACF 圖在 Lag 1 呈現非常強的正相關，之後數期雖逐漸衰減，但仍有若干滯後期超出顯著界線，顯示存在短期自相關結構；PACF 圖則在低階數有明顯尖峰，且在較高階（如 Lag 6 與 Lag 16 附近）亦出現若干顯著相關，初步顯示適合加入自迴歸（AR）成分。



圖六 原始資料的白噪音檢定與單根檢定結果

而在白噪音檢定與單根檢定結果(圖六)顯示，單根檢定的 p-value 均小於 0.05，已拒絕非平穩假設，確認一階差分後序列達到平穩狀態；然而白噪音檢定的結果顯示

多數 p-value 仍小於 0.05，代表差分後序列仍存在可被模型解釋的自相關結構，尚未達到白噪音，需進一步建立 ARIMA 模型來捕捉剩餘的規律。

第四節 資料分析結果

一、ARIMA 模型

本研究以 SAS 軟體建構ARIMA 模型型，針對台灣高鐵延人公里資料進行分析。由於原始時間序列在民國110年（2021年）5月至8月間受到COVID-19三級警戒影響，出現明顯的結構性斷裂與斷崖式下跌，若未妥善處理將導致模型參數估計產生嚴重偏誤。因此，本研究進行介入分析（Intervention Analysis），建立兩項虛擬變數以吸收疫情衝擊：一是Pulse變數（脈衝變數），用以捕捉110年6月出現的歷史性最低點單一極端衝擊；二是Step變數（階段變數），用以反映110年5月至8月期間運量整體向下位移的結構性影響。此介入處理有效降低了異常值對模型的干擾，大幅提升了模型的穩健性與預測能力。經模型辨識與多次測試後，最終選定之最適模型為含介入變數的 ARIMA$(0,1,1)(0,1,1)_{12}$ 模型（含 Pulse 及 Step 介入變數處理疫情衝擊）。模型估計結果如下：

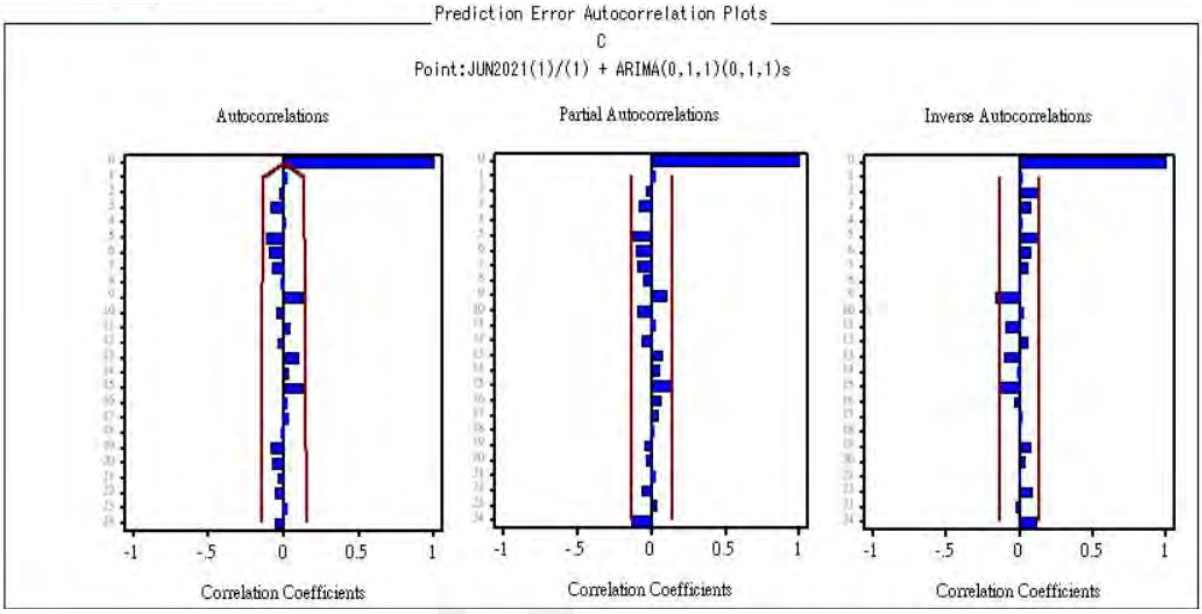
表一 、二 Point:JUN2021(1)/(1) + ARIMA$(0,1,1)(0,1,1)_{12}$ 干預模型估計結果

Model Parameter	Estimate	Std. Error	T	Prob > |T|
Intercept	-414.81784	769.0486	-0.5394	0.5902
Moving Average, Lag 1	0.34658	0.0670	5.1763	<.0001
Seasonal Moving Average, Lag 12	0.84407	0.0504	16.7364	<.0001
Point:JUN2021(1)/(1)	-552494	71849	-7.6897	<.0001
Point:JUN2021(1)/(1) Num1	198301	87284	2.2719	0.0241
Point:JUN2021(1)/(1) Den1	0.58756	0.1128	5.2111	<.0001
Model Variance (sigma squared)	5764091365	.	.	.

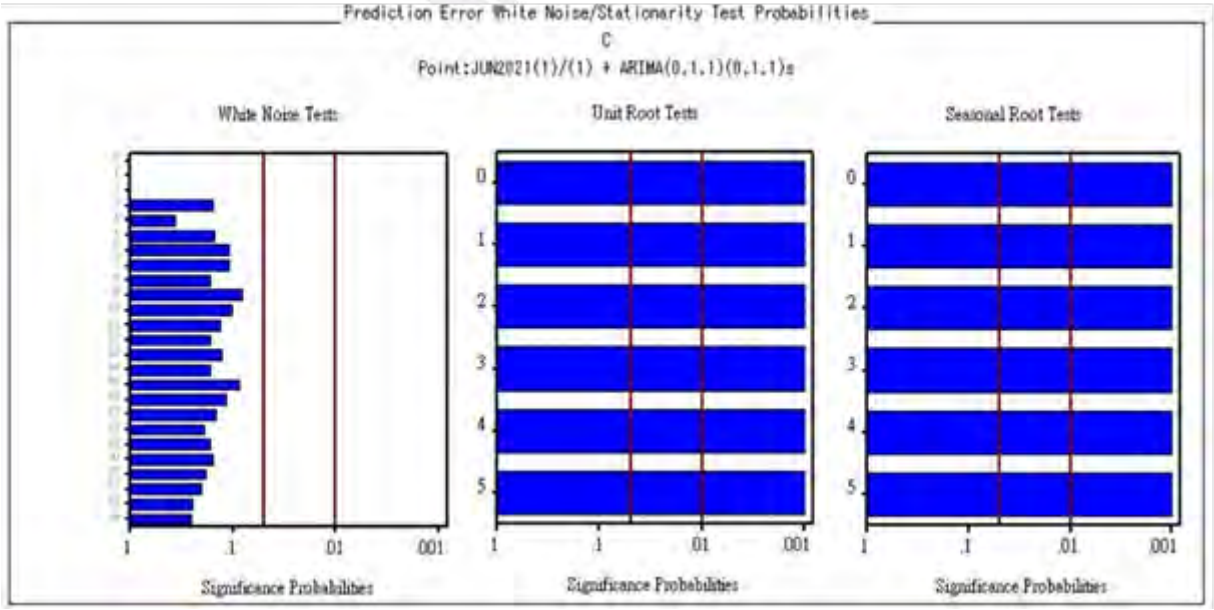
Statistic of Fit,Value	Value
Mean Square Error	5762878227
Root Mean Square Error	75913.6
Mean Absolute Percent Error	6.56755
Mean Absolute Error	48795.6
R-Square	0.853

參數估計表可知季節性移動平均項 SMA(1,12) 的係數為 0.84407（p-value < 0.0001），呈現高度統計顯著性，顯示高鐵延人公里具有非常穩定的年度季節性規律。移動平均項 MA(1) 係數為 0.34658（p-value < 0.0001），亦達顯著水準。此外，疫情相關的介入變數（Point:JUN2021 的 Pulse 與 Step 變數）均呈現統計顯著，成功捕捉了民國110年疫情期間的異常衝擊，提升了模型的穩健性。模型的 R-Squ are 達 0.853，顯示整體擬合效果良好。

在殘差診斷方面，從圖七和圖八，我們發現預測誤差的自相關圖（ACF 與 PACF）顯示殘差已無明顯的系統性相關結構，白噪音檢定結果亦顯示多數滯後期的 p-value 大於 0.05，確認殘差接近白噪音序列，代表模型已充分提取資料中的趨勢、季節性及疫情衝擊等重要資訊，模型適配性良好。



圖七 ARIMA 模型的自相關圖



圖八 ARIMA 模型的的白噪音檢定與單根檢定結果

利用建立之含介入變數的 SARIMA(0,1,1)(0,1,1)$_{12}$ 模型，對民國115年高鐵延人公里進行短期預測，並與實際觀測值進行比較。模型最終方程式如下：

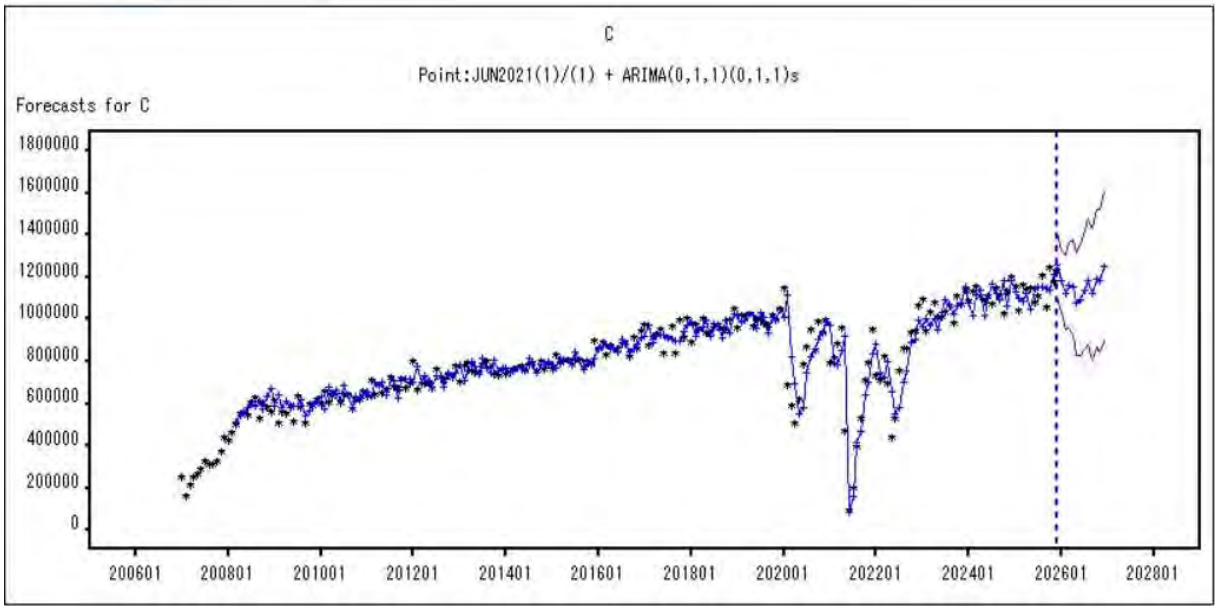
$$(1-B)(1-B^{12})Y_t = (1+0.34658B)(1+0.84407B^{12})\epsilon_t - 552.494 \cdot P_t - 198.301 \cdot S_t$$

其中 P_t為 Pulse 介入變數（捕捉110年6月極端低點），S_t 為 Step 介入變數（反映110年5–8月階段性衝擊）。

表三　ARIMA 模型 實際值與模型預測值比較表

統計期	實際值	SAS模型預測值	預測誤差	絕對百分比誤差
115年2月	1,178,337	1,125,461	52,876	**4.49%**
115年3月	1,225,645	1,161,933	63,712	**5.20%**
115年4月	1,210,082	1,151,121	58,961	**4.87%**
平均	**1,204,688**	**1,146,172**	**58,516**	**4.85%**

實際值與模型預測值之比較如表三所示。結果顯示，115年2月至4月之平均絕對百分比誤差（MAPE）為4.85%，整體預測表現尚屬合理。其中2月至4月模型均呈現低估現象，但仍能有效捕捉整體上升趨勢與季節性波動。為進一步評估模型之短期預測能力，本研究針對115年2月至12月進行事前預測，其預測結果如圖九與表四所示，模型預測值呈現穩定上升趨勢，與歷史季節性波動型態一致；115年2月至12月預測區間（95%信心水準）寬度適　中，顯示預測具有一定穩定性。圖十一則更清楚顯示，模型在垂直虛線（2026年6月）後的預測路徑能延續歷史上升趨勢，並未出現明顯偏離，顯示本研究所建構之SARIMA模型具備良好的短期預測能力。



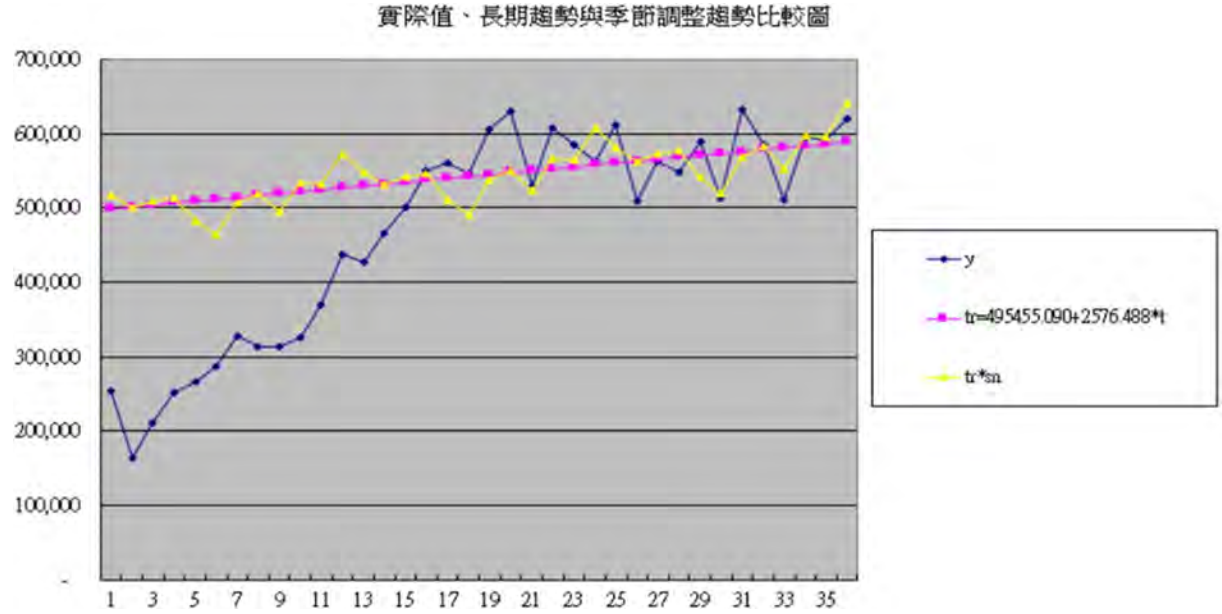
圖九 ARIMA 模型對民國 115 年高鐵延人公里之預測走勢圖

表四 ARIMA模型對民國115年2月至12月高鐵延人公里預測值

DATE	PREDICT	U95	L95
202602	1125461	1303215	947707
202603	1161933	1364542	959324
202604	1151121	1375852	926389
202605	1076778	1321642	831915
202606	1095429	1358891	831967
202607	1141478	1422309	860647
202608	1181979	1479166	884792
202609	1119747	1432435	807058
202610	1193864	1521321	866407
202611	1183909	1525496	842321
202612	1251519	1606675	896362

二、時間序列分解法（Time Series Decomposition Method）

本研究採用時間序列分解法對高鐵延人公里資料進行分析。首先透過傳統分解方法（Classical Decomposition）估計長期趨勢與季節指數，再進一步應用 Winters Multiplicative 乘法模型進行預測。首先我們採用傳統時間序列分解法對高鐵延人公里資料進行分析。首先利用12期移動平均（12-period Moving Average）計算中心化移動平均（CMA），以估計長期趨勢成分，並進一步計算各月份季節指數。



圖十實際值、長期趨勢與季節調整趨勢比較圖

從圖十可知估計之長期線性趨勢（Trend = 495,455.09 + 2,576.49×t），黃色折線則為趨勢乘以季節指數後之值（tr × sn）。由圖可明顯看出，高鐵延人公里資料同時存在穩定的長期上升趨勢（每月平均增加約2,576人公里）與明顯的季節性波動。黃色線能有效捕捉季節高峰與低谷，顯示分解模型具有良好的適配性。



圖十一　高鐵延人公里季節指數圖

而在圖十一為各月份的季節指數。由圖可知，12月與1月季節指數最高（約1.08與1.03），顯示為季節高峰；6月季節指數最低（約0.90），為季節低谷。此季節型態與高鐵運量受寒暑假及年節影響的特性一致。基於傳統分解結果，本研究進一步採用Winters Multiplicative 乘法指數平滑模型進行短期預測。該模型可同時考量水平（Level）、趨勢（Trend）與季節性（Sea sonality）成分。

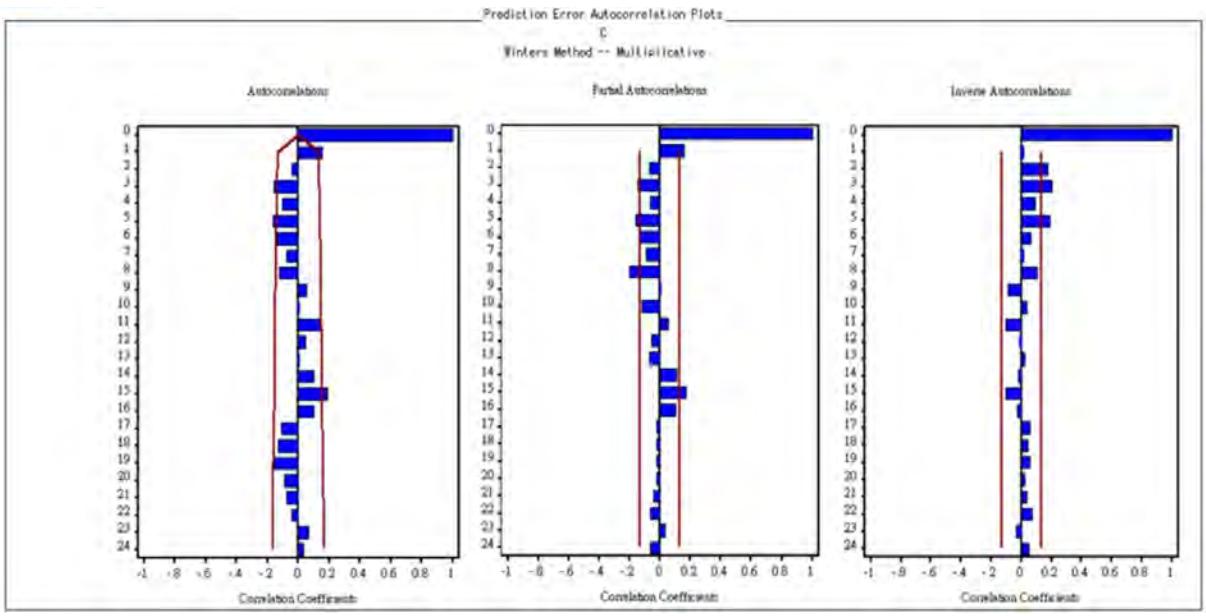
表五、六 Winters Multiplicative 模型的參數估計

Model Parameter	Estimate	Std. Error	T	Prob > |T|
LEVEL Smoothing Weight	0.79966	0.0406	19.6911	<.0001
TREND Smoothing Weight	0.00100	0.0066	0.1515	0.8797
SEASONAL Smoothing Weight	0.20647	0.0489	4.2245	<.0001
Residual Variance (sigma squared)	7537619955	.	.	.
Smoothed Level	1188992	.	.	.
Smoothed Trend	3167	.	.	.
Smoothed Seasonal Factor 1	0.97991	.	.	.
Smoothed Seasonal Factor 2	0.94139	.	.	.
Smoothed Seasonal Factor 3	0.97478	.	.	.
Smoothed Seasonal Factor 4	1.00368	.	.	.
Smoothed Seasonal Factor 5	0.97726	.	.	.
Smoothed Seasonal Factor 6	0.90837	.	.	.
Smoothed Seasonal Factor 7	1.04727	.	.	.
Smoothed Seasonal Factor 8	1.07222	.	.	.
Smoothed Seasonal Factor 9	0.97431	.	.	.
Smoothed Seasonal Factor 10	1.02192	.	.	.
Smoothed Seasonal Factor 11	0.98690	.	.	.
Smoothed Seasonal Factor 12	1.03425	.	.	.

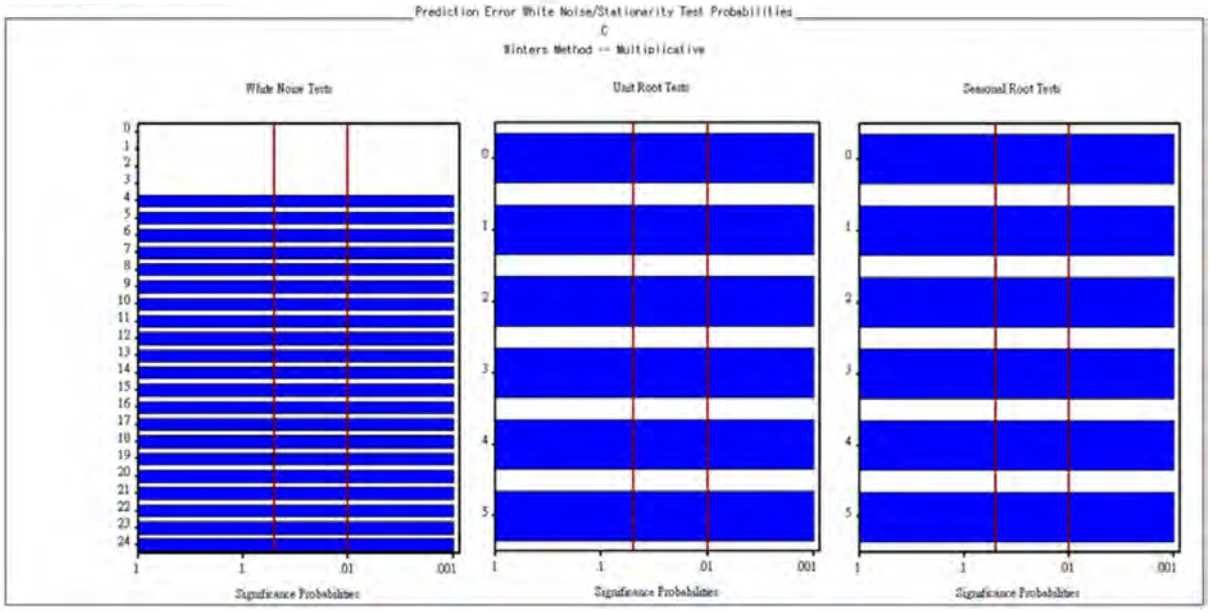
Statistic of Fit, Value	Value
Mean Square Error	7438440745
Root Mean Square Error	86246.4
Mean Absolute Percent Error	9.70001
Mean Absolute Error	54458
R-Square	0.857
Statistic of Fit, Value	48795.6

參數估計顯示，Level 平滑權重為 0.79966，T 值 = 19.6911，p-value < 0.0001，達到高度統計顯著水準，顯示台灣高鐵延人公里具有明顯的整體水平結構；Seasonal 平滑權重為 0.20647，T 值 = 4.2245，p-value < 0.0001，同樣呈現高度統計顯著，反映出高鐵運量存在穩定的季節性規律。而 Trend 平滑權重為 0.00100p- value = 0.8797，未達統計顯著水準，顯示長期趨勢變化相對平緩。整體而言，模型的 R^2 達 0.857，顯示 Winters 乘法分解模型對原始資料具有良好的擬合效果。

為確認模型適用性，本研究進一步進行殘差診斷分析。由下方圖十二自相關圖可知，各滯後期(Lag)之相關係數大多落於顯著界線範圍內，未出現明顯的系統性自相關現象。此外，在圖十三的殘差診斷方面，白噪音檢定結果顯示各滯後階數之 p-value 均顯著小於 0.0 5，拒絕殘差為白噪音之虛無假設。這顯示殘差序列中仍存在未被模型完全提取的系統性自相關結構，模型仍有配適不足(Underfitting)之現象。然而，單根檢定強烈拒絕單位根存在（p < 0.05），顯示序列具備定態性。後續研究建議調整模型階數或轉換其他計量模型以優化預測效果。



圖十二 Winters Multiplicative模型的自相關圖



圖十三 Winters Multiplicative 模型的白噪音檢定與單根檢定結果

為進一步評估 Winters Multiplicative 模型之短期預測能力，本研究針對民國115年2月至12月進行事前預測，且將115年前三個月的實際資料進行比較，其結果如表七、八及圖十四所示。

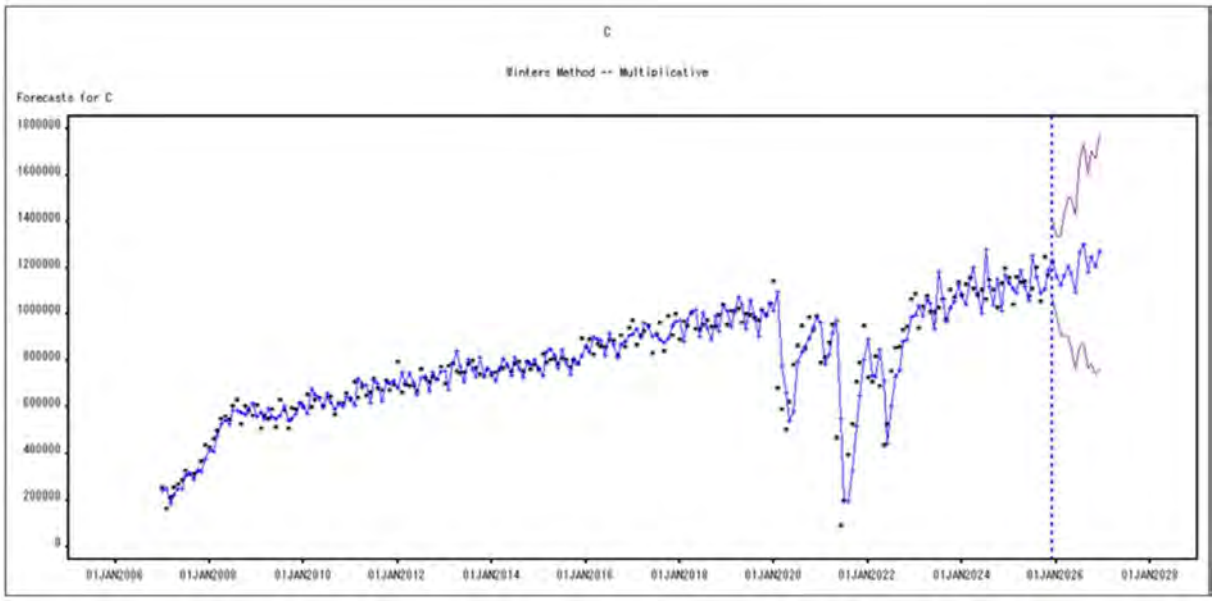
表七 Time Series Decomposition Method 實際值與模型預測值比較表

統計期	實際值	SAS模型預測值	預測誤差	絕對百分比誤差
115年2月	1,178,337	1,125,266	53,071	**4.50%**
115年3月	1,225,645	1,168,274	57,371	**4.68%**
115年4月	1,210,082	1,206,089	3,993	**0.33%**

平均	1,204,688	1,166,543	38,145	3.17%

表八 Time Series Decomposition Method民國115年2月至12月高鐵延人公里預測值

DATE	PREDICT	U95	L95	_LEVEL_	_TREND_	_SFACTOR_
01FEB2026	1125266	1339924	910607	1195327	3167	0.9414
01MAR2026	1168274	1427588	908960	1198495	3167	0.9748
01APR2026	1206089	1504917	907262	1201662	3167	1.0037
01MAY2026	1177432	1499714	855151	1204830	3167	0.9773
01JUN2026	1097303	1428715	765891	1207997	3167	0.9084
01JUL2026	1268418	1670084	866751	1211165	3167	1.0473
01AUG2026	1302033	1734980	869087	1214332	3167	1.0722
01SEP2026	1186223	1605030	767416	1217500	3167	0.9743
01OCT2026	1247421	1706541	788301	1220667	3167	1.0219
01NOV2026	1207806	1672762	742849	1223835	3167	0.9869
01DEC2026	1269025	1774361	763689	1227002	3167	1.0342



圖十四 Winters Multiplicative 模型對民國115年高鐵延人公里之預測走勢圖

由表七可知，115年2月至4月實際值與 SAS 模型預測值之平均絕對百分比誤差（MAPE）為3.17%，較先前有所改善，其中4月預測值已非常接近實際值（誤差僅0.33%），圖十及表八顯示模型預測能力穩定。Winters Multiplicative 模型預測值呈現穩定上升趨勢，並維持與歷史一致的季節性波動型態。預測區間寬度適中，115年下半年延續成長動能， 12月預測值達最高峰，符合高鐵運量受年節影響之特性。最終我們以 Winters Multiplicative 模型進行時間序列分解與預測，推論出的預測方程式為：

$$\hat{Y}_{t+h} = (L_t + h \cdot T_t) \times S_{t+h-12}$$

其中 L_t、T_t、S_t 分別為經指數平滑更新之水平、趨勢與季節成分。模型參數估計結果為 α=0.78、δ=0.01、γ=0.26，R^2 高達 0.955，顯示良好的適配效果。

三、指數平滑法（Exponential Smoothing Method）

指數平滑法採用 SAS 軟體之 Winters 方法（Winters Method - Additive）進行指數平滑分析，以分解台灣高鐵延人公里資料的水平（Level）、趨勢（Trend）與季節性（Seasonal）成分。相較於乘法模型，加法模型假設季節性波動幅度相對穩定，適合用來補充分析高鐵運量之結構特性。

表九、十 Exponential Smoothing Method 模型的參數估計

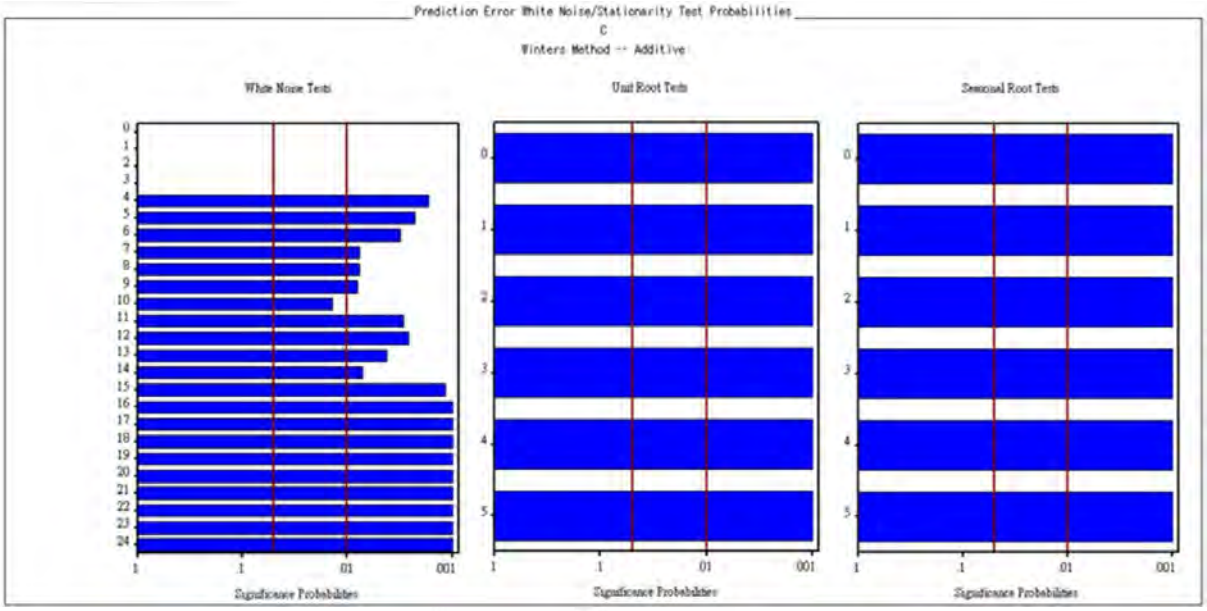
Model Parameter	Estimate	Std. Error	T	Prob>|T|
LEVEL Smoothing Weight	0.97352	0.0476	20.4404	<.0001
TREND Smoothing Weight	0.00100	0.0070	0.1419	0.8873
SEASONAL Smoothing Weight	0.00100	0.4078	0.002452	0.9980
Residual Variance (sigma squared)	6278603100	.	.	.
Smoothed Level	1152022	.	.	.
Smoothed Trend	3184	.	.	.
Smoothed Seasonal Factor 1	20675	.	.	.
Smoothed Seasonal Factor 2	-13032	.	.	.
Smoothed Seasonal Factor 3	446.88503	.	.	.
Smoothed Seasonal Factor 4	3439	.	.	.
Smoothed Seasonal Factor 5	-41417	.	.	.
Smoothed Seasonal Factor 6	-64870	.	.	.
Smoothed Seasonal Factor 7	-12251	.	.	.
Smoothed Seasonal Factor 8	13503	.	.	.
Smoothed Seasonal Factor 9	-32791	.	.	.
Smoothed Seasonal Factor 10	30366	.	.	.
Smoothed Seasonal Factor 11	19093	.	.	.
Smoothed Seasonal Factor 12	76845	.	.	.

Statistic of Fit	Value
Mean Square Error	6195989901
Root Mean Square Error	78714.6
Mean Absolute Percent Error	9.28040
Mean Absolute Error	52851.4

R-Square	0.881

參數估計結果如表九、十顯示，Level 平滑權重高達 0.97352（p-value < 0.0001），表示模型主要依賴近期觀測值更新水平成分；Trend 平滑權重為 0.00100（p-value = 0.8 873），未達統計顯著，顯示長期趨勢變化相對平緩；Seasonal 平滑權重為0.09100。整體而言，模型 R^2 達 0.881，MAPE 為 9.28%，顯示 Winters 加法模型對高鐵延人公里資料具有良好的擬合效果。相較於乘法模型，加法模型更適合用來觀察季節性波動幅度相對穩定的特性。

在殘差診斷方面，從圖十五顯示Winters Additive 模型的預測誤差自相關圖顯示殘差已無明顯的系統性相關結構；而從圖十六的殘差診斷方面，白噪音檢定結果顯示各滯後階數之 p-value 均顯著小於 0.05，拒絕殘差為白噪音之虛無假設。這顯示殘差序列中仍存在未被模型完全提取的系統性自相關結構，模型仍有配適不足(Underfitting)之現象。然而，單根檢定強烈拒絕單位根存在（p < 0.05），顯示序列具備定態性。後續研究建議調整模型階數或轉換其他計量模型以優化預測效果。

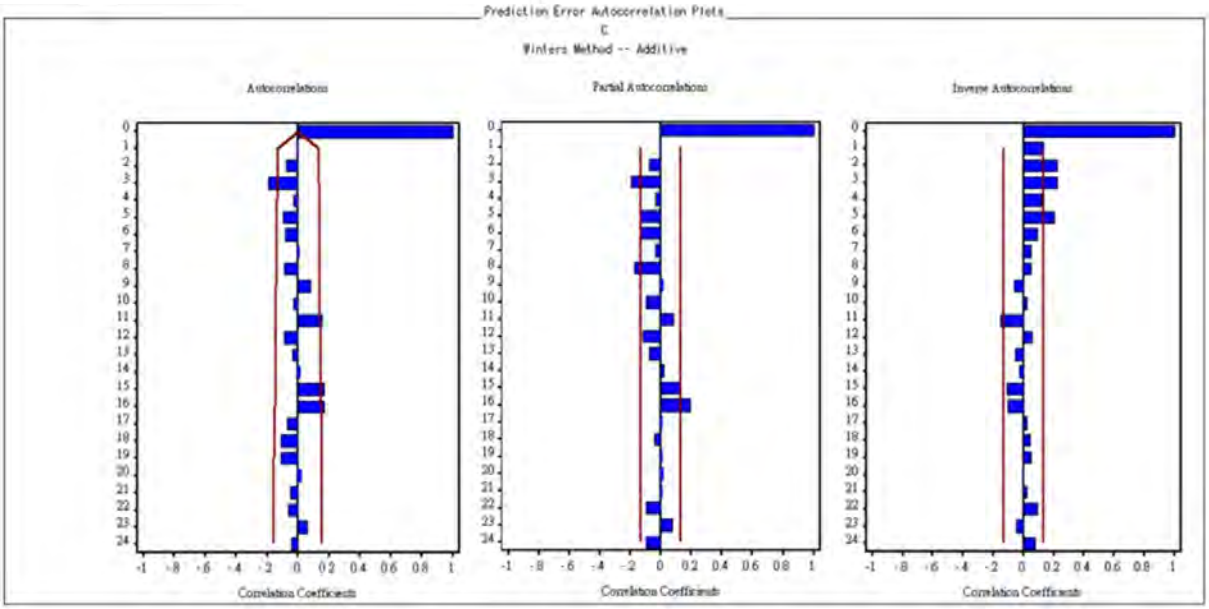


圖十五 Exponential Smoothing Method 模型的自相關圖

圖十六 Exponential Smoothing Method 模型的的白噪音檢定與單根檢定結果

而透過表十一、十二和圖十七可知 ，Winters Additive 模型在115年2月至4月的平均絕對百分比誤差（MAPE）為3.82%，整體預測表現良好，其中2月誤差最低（2.80%）。由預測走勢圖與預測資料表可見，模型能有效捕捉長期上升趨勢（每月成長約3,184人公里）與季節性波動，115年下半年預測值持續成長，12月達到高峰，符合高鐵運量受年節影響之特性。

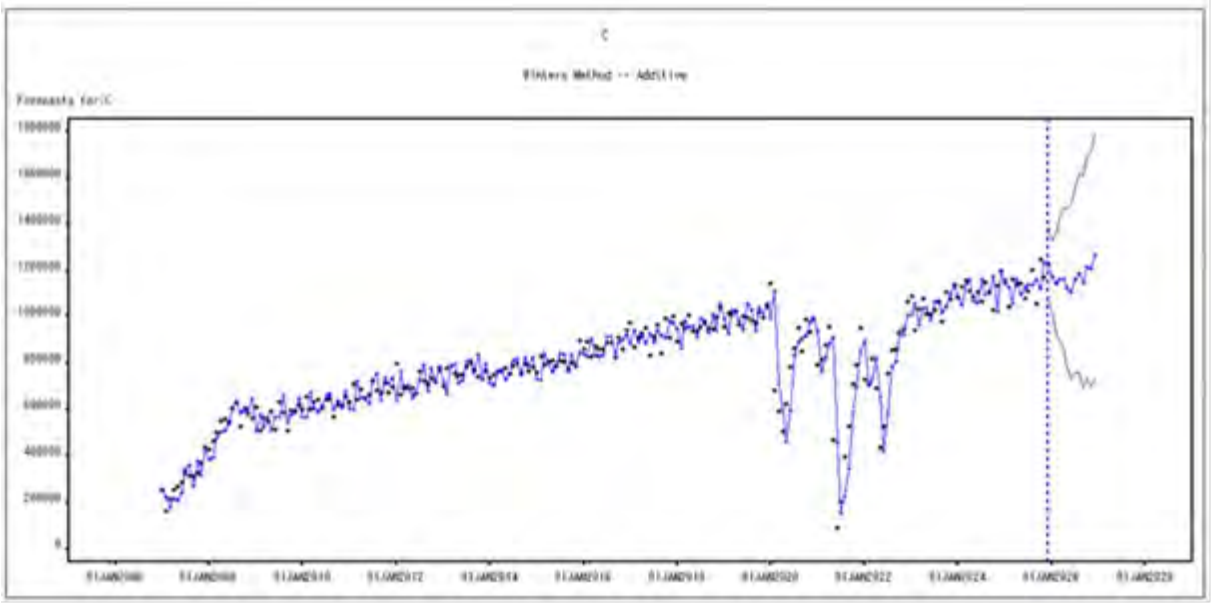
表十一 Exponential Smoothing Method 模型的實際值與模型預測值比較表



統計期	實際值	SAS模型預測值	預測誤差	絕對百分比誤差
115年2月	1,178,337	1,145,359	32,978	2.80%
115年3月	1,225,645	1,162,021	63,624	5.19%
115年4月	1,210,082	1,168,197	41,885	3.46%
平均	1,204,688	1,158,526	46,162	3.82%

表十二 Exponential Smoothing Method 模型的民國115年2月至12月高鐵延人公里預測值

DATE	PREDICT	U95	L95	_LEVEL_	_TREND_	_SFACTOR_
01FEB2026	1145359	1362207	928510	1158390	3184	-13032
01MAR2026	1162021	1426546	897496	1161574	3184	446.8850
01APR2026	1168197	1473106	863289	1164758	3184	3439
01MAY2026	1126525	1467128	785921	1167942	3184	-41417
01JUN2026	1106256	1479215	733296	1171126	3184	-64870
01JUL2026	1162059	1564839	759278	1174310	3184	-12251
01AUG2026	1190997	1621592	760402	1177494	3184	13503
01SEP2026	1147887	1604655	691118	1180678	3184	-32791
01OCT2026	1214228	1695798	732658	1183862	3184	30366
01NOV2026	1206138	1711339	700938	1187046	3184	19093
01DEC2026	1267075	1794893	739258	1190230	3184	76845



圖十七 Exponential Smoothing Method 模型預測走勢圖

採用 Winters Additive 加法指數平滑模型進行預測，其最終預測方程式為：

$$\hat{Y}_{t+h} = L_t + h \cdot T_t + S_{t+h-12}$$

其中 L_t、T_t、S_t 分別代表經指數平滑更新之水平成分、趨勢成分與季節成分。模型參數估計結果顯示 Level 平滑權重高達 0.97352，顯示模型高度依賴近期觀測值；Trend 平滑權重為 0.00100（趨勢變化較平緩）。整體而言，該模型 R^2 達 0.881，MAPE 為 3.82%，具備良好的短期預測能力。

第五節 預測結果的比較

為評估本研究所採用之 ARIMA 模型、時間序列分解法（Winters 乘法模型）與指數平滑法(Winters 加法模型)的預測表現，我們採用了判定係數(R^2)、均方誤差(MSE)、均方根誤差（RMSE）、平均絕對誤差（MAD/MAE）及平均絕對百分比誤差（MAPE）等多項準則進行綜合比較。

表十三 各預測方法的比較表

評估準則	ARIMA	時間序列分解法	指數平滑法
R平方	0.853	0.857	0.881
MSE（均方誤差）	5,762,878,227	7,438,440,745	6,195,989,901
RMSE	75,913.6	86,246.4	78,714.6
MAE/MAD	48,795.6	54,458.0	52,851.4
MAPE	6.567%	9.700%	9.280%
115年2-4月短期預測MAPE	4.85%	3.17%	3.82%

分析結果顯示，雖然指數平滑法在整體模型的解釋力上表現最佳（R^2高達0.881），且時間序列分解法在針對民國115年2至4月的短期預測中擁有最好的精準度（MAPE僅 3.17%），但若從整體歷史資料的配適誤差來看，ARIMA$(0,1,1)(0,1,1)_{12}$ 模型因為成功加入了脈衝（Pulse）與階段（Step）介入變數來吸收疫情期間的極端衝擊，使其在 MSE（約57.6億）、RMSE（75,913.6）、MAD（48,795.6）以及整體的 MAPE（6.567%）等絕對與相對誤差指標上皆為三種方法中最低的。綜合以上評估，考量到高鐵運量極易受到如疫情等重大外部事件的劇烈干擾，具備介入變數處理能力的 ARIMA 模型能有效降低異常值干擾、提供整體誤差最小且最穩健的配適效果，因此本研究將其評選為預測高鐵延人公里變動趨勢之最佳時間序列模型

第五章 結論與建議

一、研究結論

本研究透過分析民國 96 年至 114 年間的高鐵延人公里歷史資料，得出以下關於高鐵運量發展的關鍵結論：

1. 長期穩定成長與尚未飽和

高鐵延人公里呈現穩定的線性上升趨勢，平均每個月能增加約 2,576 至 3,184 千人公里的運量且從 115 年的預測走勢圖來看，運量曲線仍保持向上延伸，並未出現增長停滯的水平狀，顯示高鐵經過 19 年的營運，其長途運輸需求仍具備成長動能，運量尚未完全觸及飽和點

2. 強烈的年度季節性波動

高鐵運量具備極度規律的季節特性。每年的 12 月與 1 月為運量最高峰（季節指數高達 1.03-1.08），而 6 月則為全年度的運量最低谷（季節指數約 0.90）這與國人年底返鄉、跨年及農曆春節的長途移動需求高度吻合。

3. 面對極端外部衝擊展現復甦韌性

民國 110 年 5 月至 8 月的 COVID-19 三級警戒，導致運量出現歷史性的斷崖式暴跌（從常態的 79 萬跌至 8.9 萬）然而，在疫情趨緩後，數據顯示運量呈現 V 型反轉，並迅速回歸到原有的長期上升軌道，證明高鐵在台灣西部運輸網路中具有不可替代的剛性需求

二、政策與實務建議

基於上述資料分析與模預測結果，本研究對台灣高鐵未來的營運提出以下建議：

1. 依據季節指數進行精準運能調度

鑑於每年 12 月至隔年 1 月為絕對的運量高峰，且模型預測未來的年底需求將持續突破歷史新高，建議高鐵應提早於第四季啟動車輛保修與人力盤點，並在年底至春節期間實施最大化增班計畫。相對地，可將年度大規模維修安排在運量最低谷的 6 月進行，以達到資源最佳配置

2. 針對淡季實施差異化行銷策略

針對 6 月等傳統淡季，建議高鐵可擴大推出離峰優惠、大學生專案或結合地方觀光推動套裝行程，以刺激額外需求，填補淡季的運能閒置

3. 建立極端風險的營運緩衝機制

此次研究顯示，如疫情般的極端事件會對運量造成劇烈破壞。即使高鐵具備強大復甦力，仍建議管理層需參考此次 ARIMA 介入變數所估算出的衝擊規模，制定未來應對大型公共衛生或天然災害的營運降載 SOP 與財務儲備計畫，強化企業營運韌性。

參考文獻

行政院主計總處。（2026）。總體統計資料庫：高鐵客運量統計值。中華民國統計資訊網。https://nstatdb.dgbas.gov.tw/dgbasall/webMain.aspx?sys=100&funid=qryout&funid2=A070103010&outmode=8&ym=11501&ymt=11504&cycle=41&outkind=4&compmode=0&ratenm=%u7D71%u8A08%u503C&fldlst=010&codlst0=0100000&compmode=0&rr=q9242x&&rdm=R9999779

附錄一　組內工作分配表

姓名	占比	工作內容
吳睿哲	25%	設計主題大綱、書面報告撰寫、敘述統計、簡報製作、報告、EXCEL製作
李怡青	15%	簡報製作、報告
王亮緯	15%	簡報製作、報告
謝帛澂	30%	SAS模型操作 書面報告修改
廖家鴻	15%	簡報製作、報告