

逢甲大學學生報告 ePaper

流感之巔：探索病毒的盛衰潮流

Flu Peaks: Exploring the rise and fall of viruses

作者：洪佩筠、陳柔蓁、賴欣妤、陳筱茹

系級：統計三乙

學號：D1043856、D1089182、D1089178、D1089165

開課老師：劉峰旗 老師

課程名稱：預測分析

開課系所：統計三乙

開課學年：112 學年度第 2 學期

摘要

隨著過去幾年來 COVID-19 疫情成為國際關注的突發公共衛生事件後，這場規模極大的疫情傳染，讓全球認識到傳染病傳播的危險性和對社會所造成的深遠影響。然而除了 COVID-19 疫情之外，流感是一種持續存在且具有潛在爆發風險的傳染疾病，是一個更值得探討的公共衛生議題。因此，本文希望透過此專題研究，深入了解臺灣 2010 年至 2019 年間，每週類流感健保就診人次的趨勢，以及預測在沒有其他疫情影響之下的流感就診人數變化。本文使用衛生福利部疾病管制署在政府資料開放平台所提供的 2010 年 1 月至 2019 年 12 月每週類流感健保門診及住院就診人次統計資料，進行時間序列的建模與預測，並考慮時間序列迴歸模型、ARIMA 模型及指數平滑法三種不同模型方法，利用 Mean Square Error (MSE)、Root Mean Square Error (RMSE)、Mean Absolute Percent Error (MAPE)、Mean Absolute Error (MAE) 和 R 平方等五個指標來比較三種不同方法的預測效果與解釋能力，以選出最佳的預測模型。最後結果顯示，ARIMA 的預測表現最佳，可做為本研究最終預測模型，且從 ARIMA 模型可發現上一週的類流感就診人數與本週呈負相關，而與年的下一週的類流感就診人數與本週呈正相關。此外，還預測 2020 年 1 月至 2020 年 12 月在沒有 COVID-19 發生的情況下每週類流感就診人次分布變化。

關鍵字： 流感、ARIMA 模型、指數平滑法、時間序列迴歸模型

Abstract

As the COVID-19 epidemic has become a public health emergency of international concern in the past few years, this extremely large-scale epidemic has made the world aware of the dangers of the spread of infectious diseases and its profound impact on society. However, in addition to the COVID-19 epidemic, influenza is an infectious disease that persists and has potential outbreak risks, and is a public health issue worthy of discussion. Therefore, this article hopes to use this special study to gain an in-depth understanding of the trend in the number of weekly influenza-like health insurance visits in Taiwan from 2010 to 2019, and to predict the changes in the number of influenza-like visits without the impact of other epidemics. This article uses the statistical data on weekly influenza-like health insurance outpatient and inpatient visits from January 2010 to December 2019 provided by the Department of Disease Control and Prevention of the Ministry of Health and Welfare on the government information open platform, to conduct time series modeling and prediction, and considers three different model methods: sequence regression model, ARIMA model and exponential smoothing method, using Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Percent Error (MAPE), Mean Absolute Error (MAE) and R-squared, etc. Five indicators are used to compare the prediction effects and explanatory capabilities of three different methods to select the best prediction model. The final results show that ARIMA has the best prediction performance and can be used as the final prediction model of this study. From the ARIMA model, it can be found that the number of influenza-like hospitalizations in the previous week is negatively correlated with this week, and is negatively correlated with the number of influenza-like hospitalizations in the next week of the year. The number of people is positively correlated with the week. In addition, changes in the distribution of weekly influenza-like hospital visits from January 2020 to December 2020 in the absence of COVID-19 were also predicted.

Keywords : Influenza, ARIMA model, Exponential smoothing method, Time series regression model

目次

第一章 緒論.....	6
第一節 研究背景與動機.....	6
第二節 研究目的與研究問題.....	6
第三節 資料敘述.....	7
第四節 研究方法.....	7
第二章 文獻探討.....	8
第三章 分析結果.....	9
第一節 敘述統計.....	9
第二節 資料分析結果.....	9
(一) 時間序列迴歸模型(Time series regression model).....	10
1. 模型配適.....	10
2. 參數估計.....	11
3. 診斷分析.....	14
4. 預測結果.....	16
(二) ARIMA 模型.....	18
1. 模型配適.....	19
2. 參數估計.....	20
3. 殘差檢定.....	21
4. 預測結果.....	22
(三) 指數平滑法模型(Exponential smoothing method).....	24
1. 模型配適.....	24
2. 參數估計.....	25
3. 殘差檢定.....	25
4. 預測結果.....	27
第三節 預測結果比較.....	29
第四章 結論與建議.....	30
第一節 結論.....	30
第二節 建議.....	30

表目錄

表 1 敘述統計量.....	9
表 2 Ljung-Box Q test.....	9
表 3 三種時間序列模型的 RMSE.....	10
表 4 Seasonal Dummies+Linear Trend+AR(2)部分參數估計結果	11
表 5 Seasonal Dummies+AR(2) 參數估計結果	13
表 6 預測值及 95%信賴區間	17
表 7 ARIMA(2, 1, 1)參數估計結果(含截距項).....	20
表 8 ARIMA(2, 1, 1)去除截距項的參數估計結果.....	20
表 9 ARIMA(2, 1, 1)預測值及 95%信賴區間.....	23
表 10 Exponential smoothing method 的 Root Mean Square Error	25
表 11 Simple Exponential Smoothing 的參數估計.....	25
表 12 Simple Exponential Smoothing 預測值及 95%信賴區間	27
表 113 預測結果比較.....	29



圖目錄

圖 1 原始資料時間序列圖.....	9
圖 2 Seasonal Dummies+AR(2) 殘差圖	15
圖 3 Seasonal Dummies+AR(2) 殘差 ACF、PACF.....	15
圖 4 Seasonal Dummies+AR(2) 白噪音檢定(左)、單根檢定(右).....	16
圖 5 Seasonal Dummies+AR(2) 預測圖	17
圖 6 原始資料的 ACF、PACF 圖	19
圖 7 原始資料一次差分後時間序列圖.....	19
圖 8 原始資料一次差分後 ACF 圖、PACF 圖	20
圖 9 ARIMA(2, 1, 1)殘差圖	21
圖 10 ARIMA(2, 1, 1)殘差 ACF(左)、PACF(右)圖	22
圖 11 ARIMA(2, 1, 1)白噪音檢定(左)及單根檢定(右).....	22
圖 12 ARIMA(2, 1, 1)預測圖	23
圖 13 指數平滑法殘差圖.....	26
圖 14 指數平滑法的 ACF、PACF 圖	26
圖 15 指數平滑法的白噪音檢定.....	27
圖 16 Simple Exponential Smoothing 預測圖	27

第一章 緒論

第一節 研究背景與動機

隨著 2020 至 2022 年期間 COVID-19 疫情在國際間肆虐，嚴重影響全球經濟脈動，當時成為國際關注的突發性公共衛生事件，這場規模極大的疫情襲擊，讓全球意識到流行性傳染疾病傳播的危險性與對社會的深遠影響。然而，全球每年仍存在一種持續性且具有潛在爆發性的流行傳染疾病——流感，是除了 COVID-19 之外更值得關注的流行病學議題。因此，希望透過此專題研究機會，深入了解臺灣在過去 2010 年至 2019 年間，每週類流感健保就診人次的變化趨勢，以及預測在沒有 COVID-19 疫情影響下的流感就診人數變化，讓我們可以充分掌握每年流感對我國的影響，讓每年的流感疫情對社會的影響降到最低，也讓醫療單位在流感高峰期之前就能提早準備，以有效分配醫療資源與設備。

流感是一種流行性感冒，每年皆在全球各地造成大規模疫情，由於流感病毒經常不斷地演變成新的型態，使得人類對於新型流感病毒缺乏免疫力，導致大量人口在短時間內感染病毒。根據歷史記載，在過去 150 年內，世界共發生了六次流感大流行，其中最嚴重的是 1918 年的西班牙流感，造成全球數千萬人死亡。最近一次流感大流行是 2009 年的 H1N1 流感，造成數十萬人死亡。流感病毒的新型毒株主要從動物傳播至人類，並且可以通過鳥類遷徙或禽類製品的貨運傳播。世界衛生組織(WHO)針對流感大流行採用 6 級預警系統，其中最低級(Phase 1)表示無人類感染動物流感病毒，而最高級(Phase 6)則表示在多個地區人群中出現社區傳播(維基百科，2023)。流感病毒的類型包括禽流感(A/H5N1)、犬流感、馬流感和豬流感(A/H1N1)，流感病毒有 A-D 型，其中只有 A 型及 B 型可以引起季節性流行，近年主要流行的季節性流感病毒以 A(H3N2)、A(H1N1)，以及 B 型流感為主(衛生福利部疾病管制署，2024)。流感可能發生在所有年齡層，但 65 歲以上長者、嬰幼兒、孕婦及慢性病人等為流感高危險族群，萬一感染後容易引起併發症，甚至導致死亡，因此需提高警覺。

由於流感在我國每年的好發期於秋、冬兩季，從每年 11 月開始感染人數逐漸增加，於 12 月至隔年 3 月份達到感染人數的高峰。每年政府皆需花費百億元的預算來購置流感疫苗，提供民眾接種疫苗以預防感染。因此，若能有效掌握流感疫情的影響程度，預估可能受到感染的人數，便可有效分配醫療資源，減少不必要的資源浪費。

第二節 研究目的與研究問題

本專題將透過統計分析軟體 SAS 程式建立時間序列模型，並進行流感就診人

數的分析建模與預測，研究目的可分為以下四個項目：

1. 捕捉 2010 至 2019 年期間，臺灣每週類流感健保就診人次的變化趨勢
2. 掌握流感在一年中的高峰期持續週次
3. 建立預測模型，預測臺灣流感就診人數的變化與週期
4. 比較分析各模型的預測結果，找出最適合的預測模型

根據上述研究目的，我們延伸出下列研究問題：

1. 每週流感人數是否會與前後時段相互影響？
2. 一年中流感高峰期時長大約多久？
3. 在臺灣流感是否有週期性？
4. 最適合用來預測類流感就診人數的模型為何？

透過這些研究問題的探索，期望能夠更好地理解臺灣流感爆發的模式和影響因素，制定更有效的醫療防護策略，進一步保護民眾的健康。

第三節 資料敘述

本次研究資料來源為衛生福利部疾病管制署在政府資料開放平台 (<https://data.gov.tw/dataset/14593>) 所提供的資料，健保門診及住院就診人次統計——類流感，類流感為任何病毒造成與流感一樣嚴重的症狀，但無法證明症狀由流感所引起的疾病。使用的變數為類流感健保就診人次，本研究採用的是週資料，依據資料中同一週內在各縣市中各個性別和各年齡層的人數進行加總，得到該縣市的類流感就診人數，再將各縣市的類流感就診人數加在一起，以計算出全國當週的類流感就診人數，本研究即以此資料進行研究分析，資料蒐集的時間範圍從 2010 年初到 2019 年底，共計 520 週的資料。

第四節 研究方法

在此數據中我們用了三種模型方法進行分析，分別為時間序列迴歸(Time series regression)模型、ARIMA(Autoregressive integrated moving average)模型、指數平滑法(Exponential smoothing method)，此三種方法的詳細說明與應用，將於後續章節中說明。此外，本研究並利用 ACF(Autocorrelation function)、PACF(Partial autocorrelation function)、白噪音檢定(White noise test)及單根檢定(Unit root test)，來觀察序列資料是否存在自相關且是否為平穩序列，以檢驗配適之模型方法是否有效。最後，比較三種模型方法的預測效果，以 MSE、RMSE、MAE、MAPE 及 R 平方等五個指標來進行觀察，並選出表現最佳的模型，做為本研究的最終模型。

第二章 文獻探討

在流感的相關研究文獻中，透過流感發病病例數據進行研究的文章有，莊淑如(2012)探討了在流感在大流行時以及流感在後大流行時民眾對於戴口罩這個行為有何差別，結果顯示在流感大流行期間，民眾有類流感症狀時，他們願意戴口罩比例明顯高於後大流行時，因此流感病例數的變化也可能會因為其他因素造成波動，在處理數據上應參考當時環境，決定看是否需要加入某些影響點，讓數據顯現出更真實的一面。周書任(2017)將流感分成健康期、隱藏期、發病期、康復期四個階段，透過流感確診者在這四個階段可能的發文情形來分析流感的趨勢，利用線性迴歸製作模型進行預測，以英國為主要研究地區，使用 Twitter 與 Google 關鍵字搜尋熱度為主要分析內容。而吳姿瑩等人(2018)則透過分群分析將流感症狀嚴重程度進行分級，並繪製管制圖來監控疫情何時會爆發以及嚴重程度，並結合地理資訊系統分析台灣各地區的疫情真實情況。我國疾管署現階段也使用時間序列分析圖來推估流感疫情未來的趨勢，但缺乏仍預警的功能。

在採用時間序列分析方法進行分析的相關文獻方面，黃彥嘉等人(2019)透過時間序列分析、ARIMA 模型來判斷低溫和流感的關係，並且以大台北地區作為代表，發現低溫與流感間為負相關，表示溫度越低，流感就診比例越高。此外利用濕度、降水量以及放假天數三項分別與流感進行分析，流感就診比例與放假天數呈正相關；流感就診比例與濕度呈正相關；流感就診比例與降雨量呈負相關。鄒佳蓁等人(2021)使用 Python 程式進行時間序列分析，以流感的急診人數、門診人數、併發重症本土病例數、死亡人數進行研究，分析調查全國流感的感染情況。研究中使用 ARIMA、Crosto、Holt-Winter' seasonal 三種預測模型，其中在門診就診人數中 ARIMA 模型具有良好的表現。查文婷等人(2021)採用 Spearman 秩相關係數對氣象因子和流感發病數進行相關性分析，並以湖南省流感併發數為基礎進行時間序列分析、ARIMA 模型以及加入氣象因子後的 ARIMAX 模型來判斷氣象因素與流感發病的關係，研究結果發現流感發病數與月平均氣溫、平均日照時數、平均日降水量呈負相關，與月平均氣壓、平均相對溼度、平均最大風速呈正相關，好發期為冬春季。加入氣象因子的 ARIMAX 模型較 ARIMA 模型預測效果更好，但兩個模型對於預測效果皆不佳，需加入更多因素完善模型。錢晨嗣等人(2023)使用 SARIMA 模型對上海市流感就診百分比的發病趨勢進行預測，利用 2015 年至 2019 年，共 212 週的數據進行建模，過程中發現翊年中流感就診百分比呈雙峰分布，主要為夏季(7 至 9 月)與冬季(12 至 2 月)高峰，藉此文獻，我們在後續建模時可觀察資料是否存在週期性，並嘗試季節差分。

第三章 分析結果

第一節 敘述統計

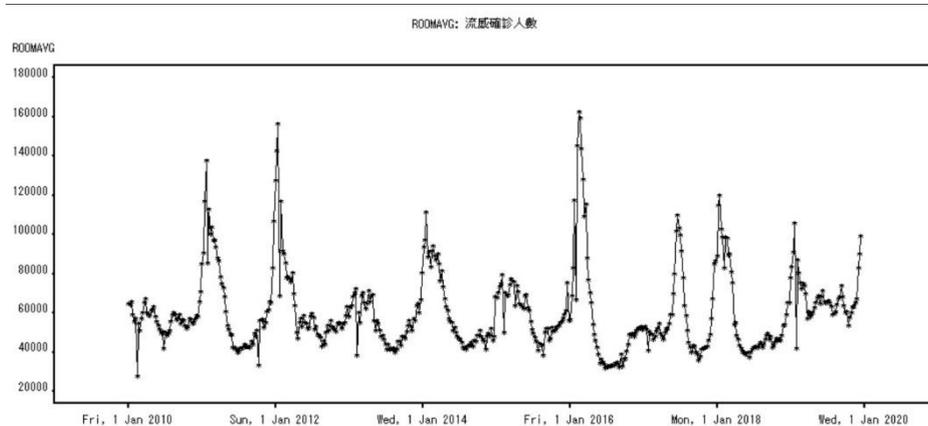


圖 1 原始資料時間序列圖

從圖 1 的時間序列圖來看，可發現資料的平均數並不平穩，但變異數看起來較為穩定。

表 1 敘述統計量

平均數	中位數	變異數	標準差	最大值	最小值
60332.17	54942	422492147	20554.61	162267	27854

我們透過 Ljung-Box Q test 檢定數據是否存在自相關性，Ljung-Box Q test 的統計為

$$Q = n(n+2) \sum_{h=1}^k \frac{\rho_h^2}{n-h}$$

進行假設檢定：

- $$\begin{cases} H_0 : \text{時間序列數據沒有自相關性，即所有滯後期的自相關係數均為 } 0。 \\ H_1 : \text{時間序列數據存在自相關性，即至少有一個滯後期的自相關係數不為 } 0。 \end{cases}$$

表 2 Ljung-Box Q test

Box-Ljung test					
X-squared	2805.6	df	52	p-value	< 0.001

從表 2 的檢定結果來看，p-value 小於 0.05 代表我們有足夠證據拒絕虛無假設 H_0 ，意味著，資料存在自相關性，需進一步對資料進行模型的配適，以捕捉資料的自我相關性。

第二節 資料分析結果

在資料分析中，使用了時間序列迴歸模型(Time Series Regression)、ARIMA 模型和指數平滑法模型(exponential smoothing method)三種模型進行預測 2020 年

流感確診人數，並比較三種模型何者最優。進行模型配適時，須檢測殘差是否隨機且平穩，來確定此模型是否適合用來進行預測。我們透過 ACF 及 PACF 來檢測殘差是否存在自相關，檢測時，ACF 及 PACF 需小於二倍標準差，才能說明殘差不存在自我相關。透過白噪音檢定判定殘差是否隨機，白噪音檢定的假設為：

$$\begin{cases} H_0 : \text{序列為白噪音，殘差為隨機。} \\ H_1 : \text{序列不為白噪音，殘差不隨機。} \end{cases}$$

若 p-value 大於 0.05 便無足夠證據拒絕虛無假設(H_0)，代表殘差為隨機的。我們透過單根檢定檢測殘差是否平穩，單根檢定的假設為：

$$\begin{cases} H_0 : \text{具有單根，殘差不平穩。} \\ H_1 : \text{不具有單根，殘差平穩。} \end{cases}$$

若 p-value 小於 0.05，我們便有足夠證據拒絕虛無假設(H_0)，代表殘差為平穩的。

(一)時間序列迴歸模型(Time series regression model)

透過觀察圖 1 的數據配置時間序列迴歸模型，進行 RMSE(Root mean square error)比較後選擇 RMSE 最小的模型，並且判斷該模型參數估計的結果是否顯著。選定模型後，透過進行診斷分析以確保模型的適用性，透過殘差圖來判斷資料是否適合使用時間序列迴歸模型。觀察 ACF 和 PACF 圖，判斷殘差是否存在自我相關。如果殘差存在顯著的自我相關，則模型可能需要進一步改進。白噪音檢定來判定殘差是否存在自我相關。若殘差通過白噪音檢定，則殘差不存在自我相關且模型擬合較好。單根檢定以判定殘差是否平穩。若殘差平穩，則模型的假設條件較好，適合進行時間序列分析。

1. 模型配適

由於從圖 1，我們觀察到資料具有季節性，因此配適季節虛擬變數(Seasonal dummies)模型。此外，從表 2 也可發現，資料存在自我相關，因此，本研究接續考慮配適誤差項具有自我相關(AR)的模型，並嘗試了 Seasonal Dummies+IAR(1,1)、Seasonal Dummies+AR(2)和 Seasonal Dummies+Linear Trend+AR(2)三種不同時間序列迴歸模型，比較個模型 RMSE，最後選出 RMSE 最小的模型做為要進行分析比較的模型。

從表 3 中可發現 RMSE 最低者為 Seasonal Dummies+Linear Trend+AR(2)的模型，但因為從表 4 中可發現，此模型的 Linear Trend 部份，其參數估計結果並不顯著，因此，進一步考慮 Seasonal Dummies+AR(2)的模型，並從表 5 可發現 Seasonal Dummies+AR(2)模型的參數估計較為合理且主要模型參數皆顯著，後續即以該模型進行分析。

表 3 三種時間序列模型的 RMSE

Time series regression	RMSE
------------------------	------

Seasonal Dummies+IAR(1,1)	8856.6
Seasonal Dummies+AR(2)	8674.9
Seasonal Dummies+Linear Trend+AR(2)	8665.2

表 4 Seasonal Dummies+Linear Trend+AR(2)部分參數估計結果

ROOMAVG：流感確診人數				
Seasonal Dummies + Linear Trend + AR(2)				
Model Parameter	Estimate	Std. Error	T	Prob> ITI
Seasonal Dummy 35	-23522	6731	-3.4946	0.0005
Seasonal Dummy 36	-22208	6670	-3.3293	0.0009
Seasonal Dummy 37	-20523	6601	-3.1093	0.002
Seasonal Dummy 38	-22341	6520	-8.4263	0.0007
Seasonal Dummy 39	-22105	6429	-3.4386	0.0006
Seasonal Dummy 40	-21719	6324	-3.4346	0.0006
Seasonal Dummy 41	-23724	6204	-3.8241	0.0001
Seasonal Dummy 42	-20909	6067	-8.4465	0.0006
Seasonal Dummy 43	-19659	5910	-3.3266	0.0009
Seasonal Dummy 44	-18970	5729	-3.311	0.001
Seasonal Dummy 45	-18921	5521	-3.4269	0.0007
Seasonal Dummy 46	-19572	5279	-8.7073	0.0002
Seasonal Dummy 47	-16879	4998	-3.377	0.0008
Seasonal Dummy 48	-15548	4657	-3.3384	0.0009
Seasonal Dummy 49	-11420	4272	-2.6729	0.0078
Seasonal Dummy 50	-8593	3705	-2.3192	0.0208
Seasonal Dummy 51	-2217	3291	-0.6734	0.501
Linear Trend	12.44166	20.2196	0.6153	0.5386
Model Variance (sigma squared)	83776774			

2. 參數估計

透過表 5 的參數估計結果，我們可將 Seasonal Dummies+AR(2)模型的估計的時間序列迴歸模型表示如下：

$$\hat{y}_t = 69958 + 8313W_1 + 16205W_2 + 23567W_3 + 20218W_4 + 6694W_5 + 15231W_6 + 11223W_7 + 15365W_8 + 10693W_9 + 11535W_{10} + 8889W_{11} + 8126W_{12} + 1106W_{13} - 4150W_{14} - 4828W_{15} - 8394W_{16} - 11043W_{17} - 11707W_{18} - 13724W_{19} - 15584W_{20} - 14873W_{21} - 15373W_{22} - 16159W_{23} - 14812W_{24} - 15150W_{25} - 17162W_{26} - 17028W_{27} - 18273W_{28} - 20488W_{29} - 22797W_{30} - 22883W_{31} - 23444W_{32} - 24140W_{33} - 24925W_{34} - 23456W_{35} - 22135W_{36} - 20445W_{37} - 22257W_{38} - 22017W_{39} - 21627W_{40} - 23629W_{41} -$$

$$20812W_{42} - 19561W_{43} - 18871W_{44} - 18823W_{45} - 19476W_{46} - 16787W_{47} - 15462W_{48} - 11341W_{49} - 8523W_{50} - 2160W_{51} + \varepsilon_t,$$

$$\varepsilon_t = 0.58128\varepsilon_{t-1} + 0.29304\varepsilon_{t-2} + a_t,$$

$$\hat{\sigma}^2 = 83664855,$$

其中 W_i 表示一年當中第 i 週的虛擬變數，各週虛擬變數的定義如下：

$$\begin{aligned} W_1 &= \begin{cases} 1, & \text{if week 1} \\ 0, & \text{otherwise} \end{cases} & W_2 &= \begin{cases} 1, & \text{if week 2} \\ 0, & \text{otherwise} \end{cases} & W_3 &= \begin{cases} 1, & \text{if week 3} \\ 0, & \text{otherwise} \end{cases} \\ W_4 &= \begin{cases} 1, & \text{if week 4} \\ 0, & \text{otherwise} \end{cases} & W_5 &= \begin{cases} 1, & \text{if week 5} \\ 0, & \text{otherwise} \end{cases} & W_6 &= \begin{cases} 1, & \text{if week 6} \\ 0, & \text{otherwise} \end{cases} \\ W_7 &= \begin{cases} 1, & \text{if week 7} \\ 0, & \text{otherwise} \end{cases} & W_8 &= \begin{cases} 1, & \text{if week 8} \\ 0, & \text{otherwise} \end{cases} & W_9 &= \begin{cases} 1, & \text{if week 9} \\ 0, & \text{otherwise} \end{cases} \\ W_{10} &= \begin{cases} 1, & \text{if week 10} \\ 0, & \text{otherwise} \end{cases} & W_{11} &= \begin{cases} 1, & \text{if week 11} \\ 0, & \text{otherwise} \end{cases} & W_{12} &= \begin{cases} 1, & \text{if week 12} \\ 0, & \text{otherwise} \end{cases} \\ W_{13} &= \begin{cases} 1, & \text{if week 13} \\ 0, & \text{otherwise} \end{cases} & W_{14} &= \begin{cases} 1, & \text{if week 14} \\ 0, & \text{otherwise} \end{cases} & W_{15} &= \begin{cases} 1, & \text{if week 15} \\ 0, & \text{otherwise} \end{cases} \\ W_{16} &= \begin{cases} 1, & \text{if week 16} \\ 0, & \text{otherwise} \end{cases} & W_{17} &= \begin{cases} 1, & \text{if week 17} \\ 0, & \text{otherwise} \end{cases} & W_{18} &= \begin{cases} 1, & \text{if week 18} \\ 0, & \text{otherwise} \end{cases} \\ W_{19} &= \begin{cases} 1, & \text{if week 19} \\ 0, & \text{otherwise} \end{cases} & W_{20} &= \begin{cases} 1, & \text{if week 20} \\ 0, & \text{otherwise} \end{cases} & W_{21} &= \begin{cases} 1, & \text{if week 21} \\ 0, & \text{otherwise} \end{cases} \\ W_{22} &= \begin{cases} 1, & \text{if week 22} \\ 0, & \text{otherwise} \end{cases} & W_{23} &= \begin{cases} 1, & \text{if week 23} \\ 0, & \text{otherwise} \end{cases} & W_{24} &= \begin{cases} 1, & \text{if week 24} \\ 0, & \text{otherwise} \end{cases} \\ W_{25} &= \begin{cases} 1, & \text{if week 25} \\ 0, & \text{otherwise} \end{cases} & W_{26} &= \begin{cases} 1, & \text{if week 26} \\ 0, & \text{otherwise} \end{cases} & W_{27} &= \begin{cases} 1, & \text{if week 27} \\ 0, & \text{otherwise} \end{cases} \\ W_{28} &= \begin{cases} 1, & \text{if week 28} \\ 0, & \text{otherwise} \end{cases} & W_{29} &= \begin{cases} 1, & \text{if week 29} \\ 0, & \text{otherwise} \end{cases} & W_{30} &= \begin{cases} 1, & \text{if week 30} \\ 0, & \text{otherwise} \end{cases} \\ W_{31} &= \begin{cases} 1, & \text{if week 31} \\ 0, & \text{otherwise} \end{cases} & W_{32} &= \begin{cases} 1, & \text{if week 32} \\ 0, & \text{otherwise} \end{cases} & W_{33} &= \begin{cases} 1, & \text{if week 33} \\ 0, & \text{otherwise} \end{cases} \\ W_{34} &= \begin{cases} 1, & \text{if week 34} \\ 0, & \text{otherwise} \end{cases} & W_{35} &= \begin{cases} 1, & \text{if week 35} \\ 0, & \text{otherwise} \end{cases} & W_{36} &= \begin{cases} 1, & \text{if week 36} \\ 0, & \text{otherwise} \end{cases} \end{aligned}$$

$$\begin{aligned}
 W_{37} &= \begin{cases} 1, & \text{if week 37} \\ 0, & \text{otherwise} \end{cases} & W_{38} &= \begin{cases} 1, & \text{if week 38} \\ 0, & \text{otherwise} \end{cases} & W_{39} &= \begin{cases} 1, & \text{if week 39} \\ 0, & \text{otherwise} \end{cases} \\
 W_{40} &= \begin{cases} 1, & \text{if week 40} \\ 0, & \text{otherwise} \end{cases} & W_{41} &= \begin{cases} 1, & \text{if week 41} \\ 0, & \text{otherwise} \end{cases} & W_{42} &= \begin{cases} 1, & \text{if week 42} \\ 0, & \text{otherwise} \end{cases} \\
 W_{43} &= \begin{cases} 1, & \text{if week 43} \\ 0, & \text{otherwise} \end{cases} & W_{44} &= \begin{cases} 1, & \text{if week 44} \\ 0, & \text{otherwise} \end{cases} & W_{45} &= \begin{cases} 1, & \text{if week 45} \\ 0, & \text{otherwise} \end{cases} \\
 W_{46} &= \begin{cases} 1, & \text{if week 46} \\ 0, & \text{otherwise} \end{cases} & W_{47} &= \begin{cases} 1, & \text{if week 47} \\ 0, & \text{otherwise} \end{cases} & W_{48} &= \begin{cases} 1, & \text{if week 48} \\ 0, & \text{otherwise} \end{cases} \\
 W_{49} &= \begin{cases} 1, & \text{if week 49} \\ 0, & \text{otherwise} \end{cases} & W_{50} &= \begin{cases} 1, & \text{if week 50} \\ 0, & \text{otherwise} \end{cases} & W_{51} &= \begin{cases} 1, & \text{if week 51} \\ 0, & \text{otherwise} \end{cases}
 \end{aligned}$$

表 5 Seasonal Dummies+AR(2) 參數估計結果

Model Parameter	Estimate	Std. Error	T	Prob> TI
Intercept	69958	5235	13.3648	<.0001
Autoregressive, Lag 1	0.58128	0.0443	18.1202	<.0001
Autoregressive, Lag 2	0.29304	0.0444	6.5936	<.0001
Seasonal Dummy 1	8313	3337	2.4912	0.0131
Seasonal Dummy 2	16205	3746	4.3258	<.0001
Seasonal Dummy 3	23567	4304	5.4755	<.0001
Seasonal Dummy 4	20218	4683	4.3169	<.0001
Seasonal Dummy 5	6694	5020	1.3385	0.1830
Seasonal Dummy 6	15231	5298	2.8751	0.0042
Seasonal Dummy 7	11223	5537	2.0269	0.0432
Seasonal Dummy 8	15365	5743	2.6754	0.0077
Seasonal Dummy 9	10693	5922	1.8058	0.0716
Seasonal Dummy 10	11535	6077	1.8982	0.0583
Seasonal Dummy 11	8889	6213	1.4308	0.1532
Seasonal Dummy 12	8126	6332	1.2833	0.2
Seasonal Dummy 13	1106	6436	0.1718	0.8636
Seasonal Dummy 14	-4150	6527	-0.6358	0.5252
Seasonal Dummy 15	-4828	6606	-0.7308	0.4652
Seasonal Dummy 16	-8394	6676	-1.2575	0.2092
Seasonal Dummy 17	-11043	6736	-1.6395	0.1018
Seasonal Dummy 18	-11797	6787	-1.738	0.0829
Seasonal Dummy 19	-13724	6831	-2.0089	0.0451
Seasonal Dummy 20	-15584	6869	-2.2689	0.0237
Seasonal Dummy 21	-14873	6899	-2.1558	0.0316
Seasonal Dummy 22	-15373	6924	-2.2204	0.0269

Seasonal Dummy 23	-16159	6942	-2.3277	0.0204
Seasonal Dummy 24	-14812	6955	-2.1295	0.0337
Seasonal Dummy 25	-15150	6963	-2.1758	0.0301
Seasonal Dummy 26	-17162	6965	-2.4639	0.0141
Seasonal Dummy 27	-17028	6963	-2.4457	0.0148
Seasonal Dummy 28	-18273	6954	-2.6276	0.0089
Seasonal Dummy 29	-20488	6941	-2.9519	0.0033
Seasonal Dummy 30	-22797	6922	-3.2937	0.0011
Seasonal Dummy 31	-22883	6897	-3.318	0.001
Seasonal Dummy 32	-23444	6865	-3.4148	0.0007
Seasonal Dummy 33	-24140	6828	-3.5355	0.0004
Seasonal Dummy 34	-24925	6783	-3.6747	0.0003
Seasonal Dummy 35	-23456	6730	-3.485	0.0005
Seasonal Dummy 36	-22135	6669	-3.3189	0.001
Seasonal Dummy 37	-20445	6599	-3.098	0.0021
Seasonal Dummy 38	22257	6519	-3.4143	0.0007
Seasonal Dummy 39	22017	6427	-3.426	0.0007
Seasonal Dummy 40	21627	6321	-3.4214	0.0007
Seasonal Dummy 41	23629	6201	-3.8105	0.0002
Seasonal Dummy 42	20812	6064	-3.4323	0.0007
Seasonal Dummy 43	19561	5906	-3.3119	0.001
Seasonal Dummy 44	18871	5726	-3.2959	0.0011
Seasonal Dummy 45	18823	5517	-3.4116	0.0007
Seasonal Dummy 46	19476	5275	-3.6921	0.0002
Seasonal Dummy 47	16787	4994	-3.3615	0.0008
Seasonal Dummy 48	15462	4653	-3.3229	0.001
Seasonal Dummy 49	11341	4268	-2.6571	0.0082
Seasonal Dummy 50	8523	3701	-2.3028	0.0217
Seasonal Dummy 51	2160	3287	-0.6569	0.5115
Model Variance (sigma squared)	83664855			

3. 診斷分析

為了解 Seasonal Dummies+AR(2)模型的殘差是否符合對誤差的假設，即希望殘差為不存在自我相關且平穩的，透過 ACF、PACF 檢測及白噪音檢定殘差是否存在自相關，以及透過單根檢定檢測殘差是否平穩序列。從圖 2 中，可以觀察到殘差大部分為隨機跳動的，顯示無自我相關，但在 2016 年至 2018 年間則出現

明顯的自我相關。從圖 3 的 ACF 及 PACF 圖來觀察，可發現在 lag51 和 lag52 的地方仍存在顯著的自我相關，代表殘差為達到隨機。且從圖 4 的白噪音檢定(左圖)結果中，也可發現 p-value 小於 0.05，代表殘差非白噪音，存在自我相關。而單根檢定(右圖)的結果則顯示，p-value 皆小於 0.05，拒絕殘差存在單根，代表序列已達到平穩的狀態。

由於此模型的 RMSE 已是所嘗試配適的時間序列迴歸模型中最小的，但殘差診斷的結果仍存在自相關，因此，本研究認為受到此模型設定的限制，這份資料並不適合使用時間序列迴歸模型。

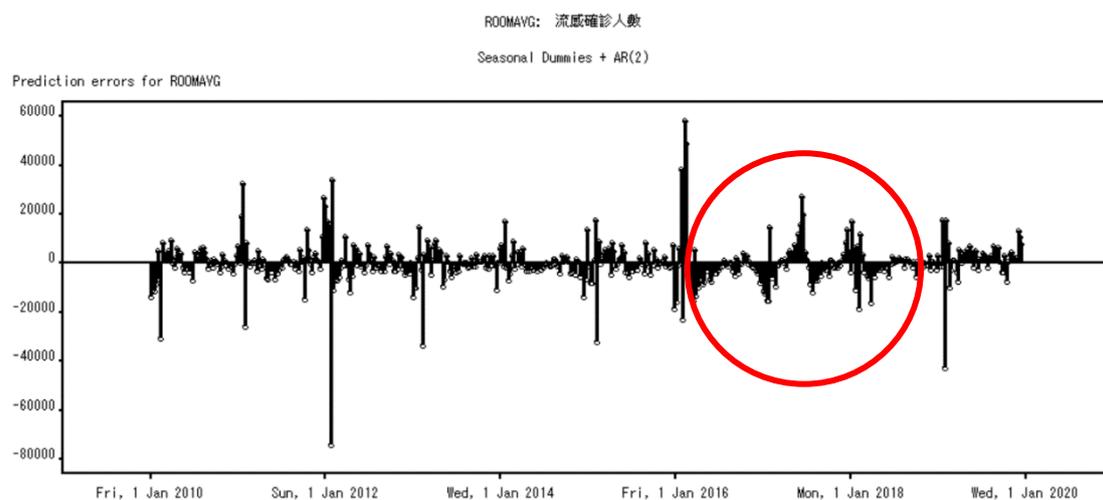


圖 2 Seasonal Dummies+AR(2) 殘差圖

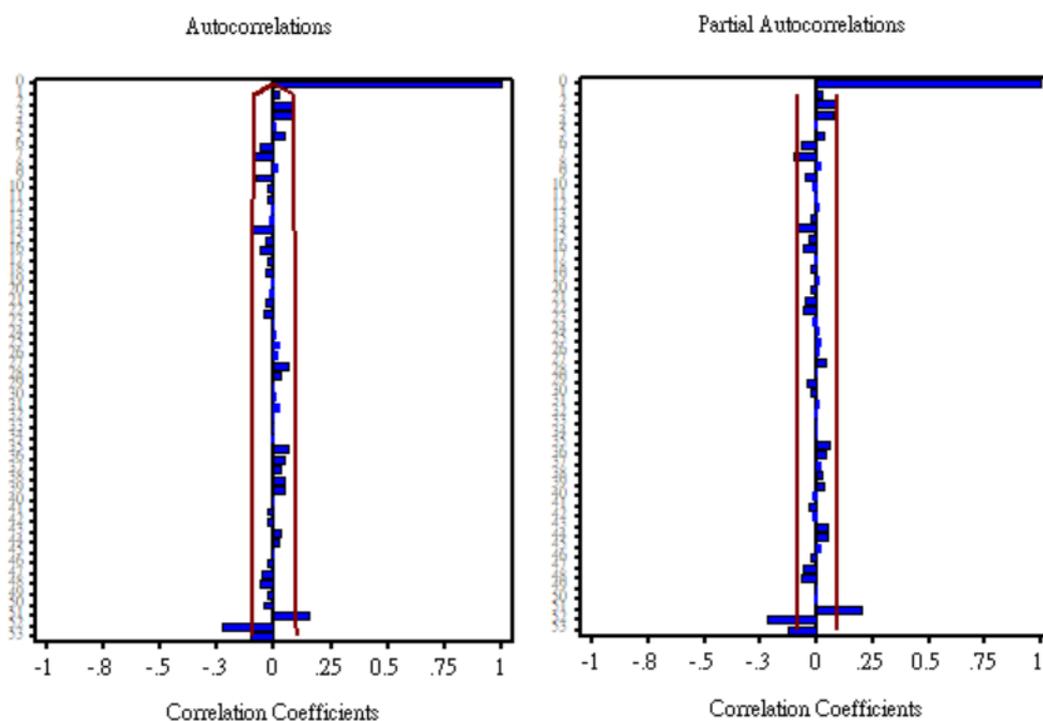


圖 3 Seasonal Dummies+AR(2) 殘差 ACF、PACF

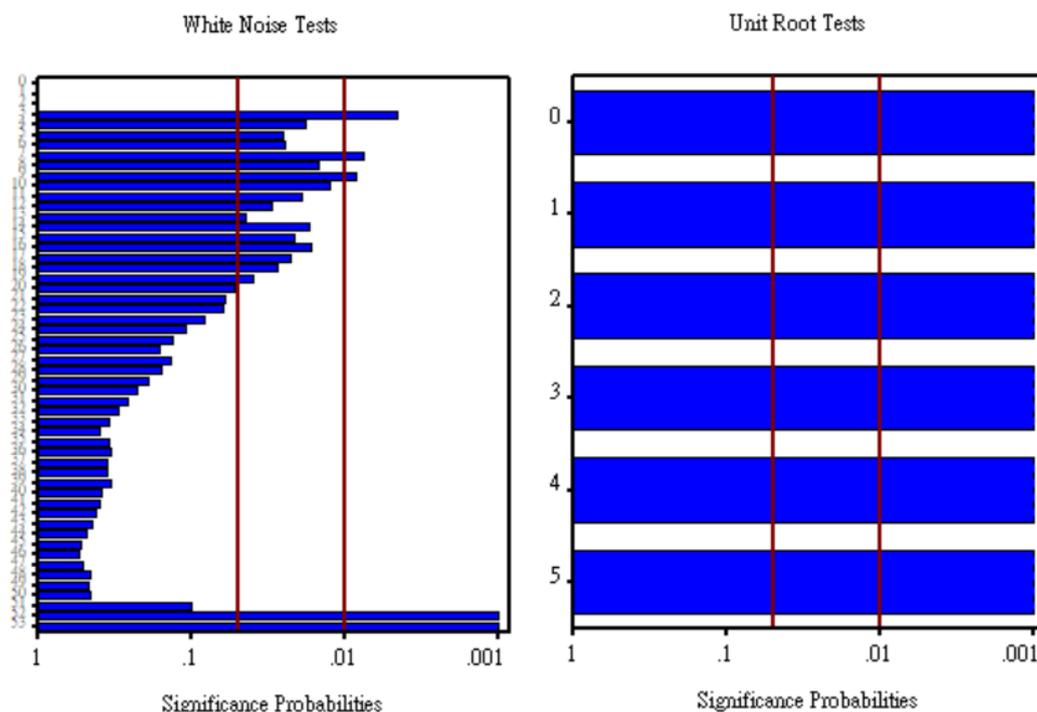


圖 4 Seasonal Dummies+AR(2) 白噪音檢定(左)、單根檢定(右)

4. 預測結果

若我們仍進一步利用 Seasonal Dummies+AR(2)模型來進行下一年度(2020年)的流感就診人數預測，並在圖 5 中呈現出預測值(紫色線)及 95%預測區間(上界為黃色線，下界為灰色線)，可發現 2020 年初，即 1 月至 3 月，流感就診人數仍處於高峰期，4 月後才逐漸趨緩下降。最後，在表 6 中為此模型預測的預測值及 95%的信賴區間。

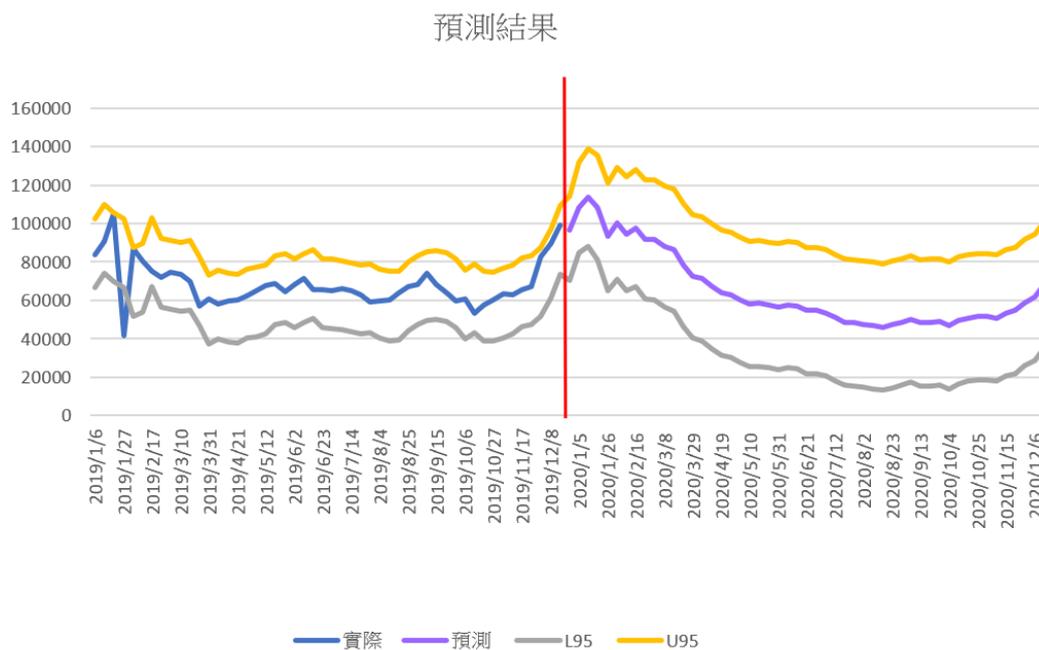


圖 5 Seasonal Dummies+AR(2) 預測圖

表 6 預測值及 95%信賴區間

時間	預測值	L95	U95
2020/1/5	108264	84644	131885
2020/1/12	113589	88082	139097
2020/1/19	108315	81296	135334
2020/1/26	93076	64897	121255
2020/2/2	100052	70952	129151
2020/2/9	94633	64802	124465
2020/2/16	97497	67079	127915
2020/2/23	91670	60780	122560
2020/3/1	91466	60195	122738
2020/3/8	87873	56293	119454
2020/3/15	86253	54421	118084
2020/3/22	78457	46422	110493
2020/3/29	72500	40298	104701
2020/4/5	71186	38849	103523
2020/4/12	67045	34597	99492
2020/4/19	63876	31338	96414
2020/4/26	62651	30040	95263
2020/5/3	60298	27626	92970
2020/5/10	58052	25331	90773

2020/5/17	58414	25652	91175
2020/5/24	57597	24803	90392
2020/5/31	56525	23704	89347
2020/6/7	57614	24771	90458
2020/6/14	57041	24180	89903
2020/6/21	54818	21941	87694
2020/6/28	54759	21871	87648
2020/7/5	53341	20443	86240
2020/7/12	50969	18062	83875
2020/7/19	48517	15604	81430
2020/7/26	48303	15384	81222
2020/8/2	47625	14702	80548
2020/8/9	46824	13897	79751
2020/8/16	45943	13013	78873
2020/8/23	47326	14394	80259
2020/8/30	48568	15634	81503
2020/9/6	50189	17253	83125
2020/9/13	48312	15375	81249
2020/9/20	48494	15556	81432
2020/9/27	48831	15892	81771
2020/10/4	46782	13842	79722
2020/10/11	49556	16616	82497
2020/10/18	50768	17827	83710
2020/10/25	51423	18481	84365
2020/11/1	51439	18497	84381
2020/11/8	50757	17815	83699
2020/11/15	53420	20477	86362
2020/11/22	54722	21779	87664
2020/11/29	58821	25878	91764
2020/12/6	61619	28677	94562
2020/12/13	67966	35023	100909

(二)ARIMA 模型

ARIMA 模型是一種時間序列預測模型，能夠捕捉資料的趨勢、季節性、週期性等特徵。在配適 ARIMA 模型前，需要檢查資料是否存在平均數及變異數不穩的情況，從圖 1 已判斷資料並不存在變異數不穩，但平均數可能存在不穩，因此，進一步從 ACF 圖判斷是否存在不穩，若存在不穩需進行差分。

之後透過觀察 ACF 圖及 PACF 圖來判斷 ARIMA 模型的配適，並透過殘差的 ACF 圖、PACF 圖及白噪音檢定來判斷殘差是否存在自我相關，最後，透過單根檢定判定殘差是否平穩。

1. 模型配適

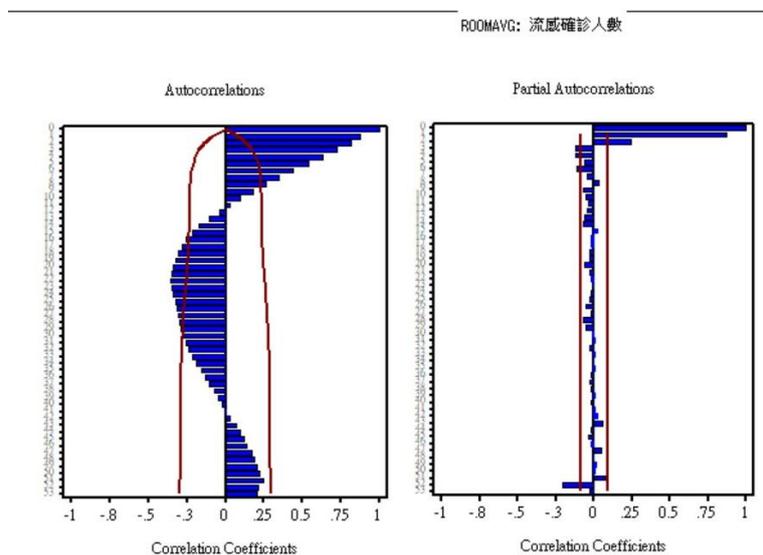


圖 6 原始資料的 ACF、PACF 圖

從圖 6 的 ACF 圖可觀察到 ACF 呈現 dies down slowly 的現象，從而可以判斷此資料並未達到平穩，需要進行一次差分。

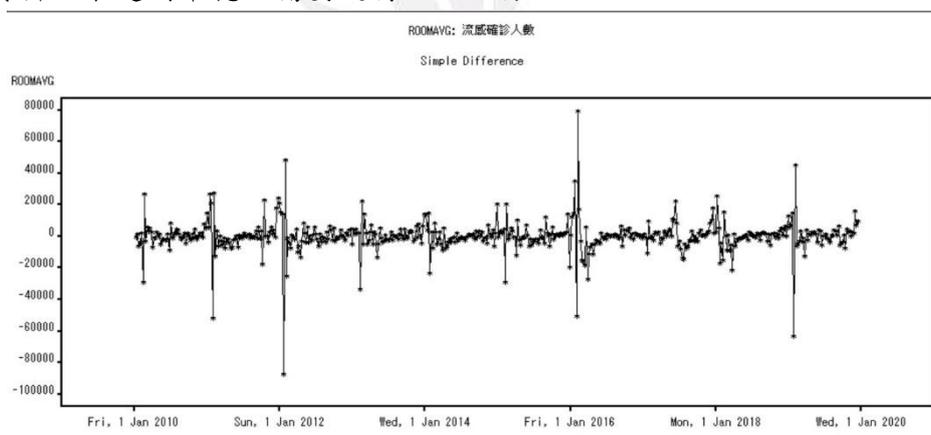


圖 7 原始資料一次差分後時間序列圖

此序列資料經一次差分後之時間序列圖可從圖 7 看出，經過一次差分後資料的平均數已呈現平穩狀態，且進一步是用 ACF 圖、PACF 圖進行驗證，如下圖 8。

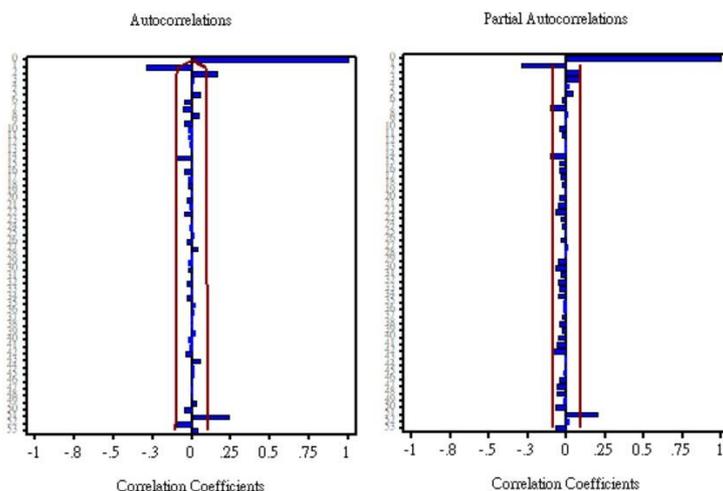


圖 8 原始資料一次差分後 ACF 圖、PACF 圖

從圖 8 的 ACF 圖為 dies down quickly after lag 2，可驗證資料的平均數已達平穩狀態。且透過圖 8 的 ACF 圖為 dies down quickly after lag 2，PACF 圖為 dies down quickly after lag 1，以及在 lag51 的位置 ACF 及 PACF 皆有顯著凸出的現象，本研究判斷 ARIMA 模型應配適為 ARIMA(2, 1, 1)，其中 AR 是在 lag1 和 lag51，MA 是在 lag2。雖然從圖 1 的原始資料時間序列圖中，發現資料具有季節性，但在經過嘗試後，發現加入季節性並不會使模型變更好，因此，最終決定使用 ARIMA(2, 1, 1)進行分析。

2. 參數估計

首先，從表 7 的參數估計結果中可發現，截距項的 p-value 大於 0.05，代表截距項的估計值不顯著，而其餘的模型參數則皆為顯著。因此，考慮將截距項去除後重新估計模型，重新估計模型後的參數估計值如表 8 所示，此時所有的參數估計皆達到顯著，該模型應是可以考慮採用的最終模型。

表 7 ARIMA(2, 1, 1)參數估計結果(含截距項)

Model Parameter	Est imate	Std. Error	T	Prob>ITI
Intercept	80.96529	444.2938	0.1822	0.8555
MA factor 1 lag 2	-0.11275	0.0452	-2.4951	0.0129
AR factor 1 lag 1	-0.25559	0.0423	-6.0475	<.0001
AR factor 1 lag 51	0.24639	0.0441	5.5882	<.0001
Model Variance (sigma squared)	88896203			

表 8 ARIMA(2, 1, 1)去除截距項的參數估計結果

Model Parameter	Estimate	Std. Error	T	Prob>ITI
MA factor 1 lag 2	-0.11279	0.0451	-2.4984	0.0128
AR factor 1 lag 1	-0.25554	0.0422	-6.0523	<.0001
AR factor 1 lag 51	0.2464	0.0441	5.5935	<.0001

Model Variance (sigma squared)	88729596			
--------------------------------	----------	--	--	--

從表 8 的參數估計結果，此 ARIMA 模型可表示為：

$$(1 - 0.25554B + 0.2464B^{51})(y_t - y_{t-1}) = a_t + 0.11279a_{t-2}$$

$$\hat{\sigma}^2 = 88729596$$

從模型的參數估計結果來看，可推測前一期的就診人數成長率對本期就診人數成長率呈現負相關，參數估計值為-0.25554，而去年同期的後一期的就診人數成長率對本期就診人數成長率則呈現正相關，參數估計值為 0.2464。

3. 殘差檢定

在模型殘差中，希望殘差能夠為隨機且平穩的，因此透過 ACF、PACF 及白噪音檢定殘差是否存在自相關，以及使用單根檢定檢測殘差是否平穩。從圖 9 的殘差圖中可看出大部分殘差為隨機的，但在部分時間段似乎存在自相關，因此進一步進行檢定來做為判斷。透過圖 10 的 ACF 圖及 PACF 圖結果發現殘差檢定通過，雖然在 lag15 的位置，超出兩倍標準差一些，但對整體來說並無太大影響。透過圖 11 的白噪音檢定(左圖)結果，發現 p-value 皆大於 0.05，代表我們沒有足夠的證據拒絕殘差不隨機，因此推斷殘差在 ARIMA(2, 1, 1)模型中為隨機的，不存在自我相關。從單根檢定(右圖)結果，發現 p-value 皆小於 0.05，代表有足夠證據拒絕殘差為不平穩的，顯示才差已達到平穩狀態。因此，在 ARIMA(2, 1, 1)模型的設置之下，殘差診斷皆通過，顯示該模型應是值得進一步做為預測分析應用的模型。

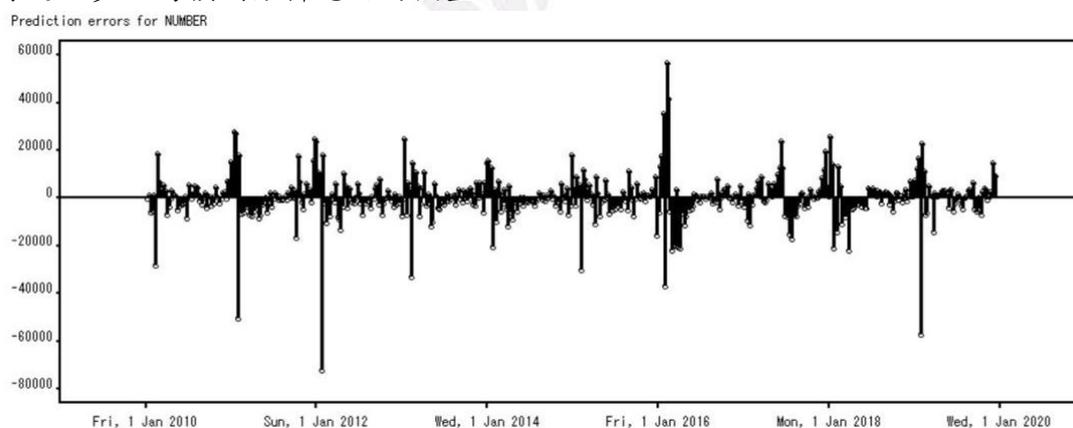


圖 9 ARIMA(2, 1, 1)殘差圖

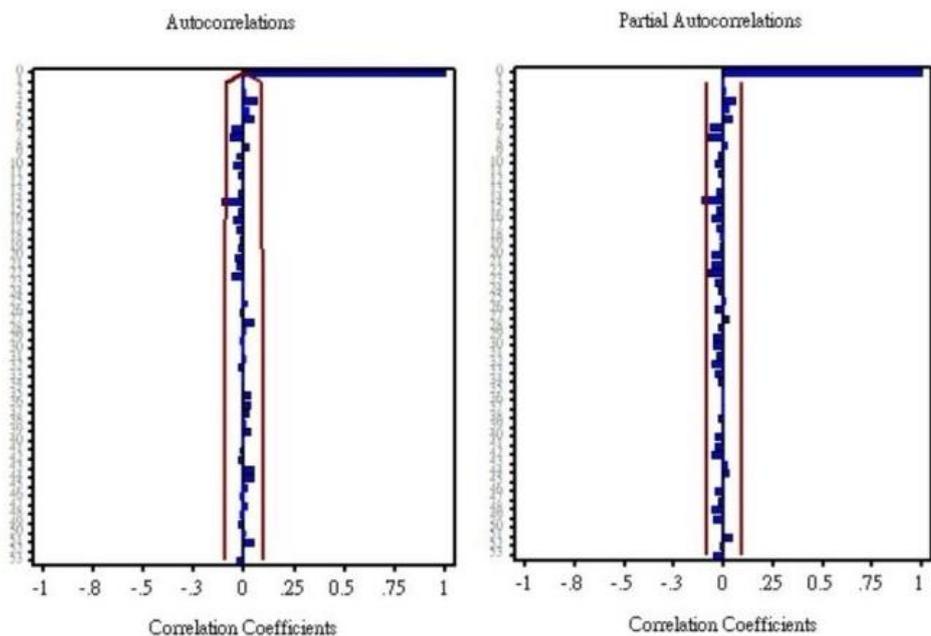


圖 10 ARIMA(2, 1, 1)殘差 ACF(左)、PACF(右)圖

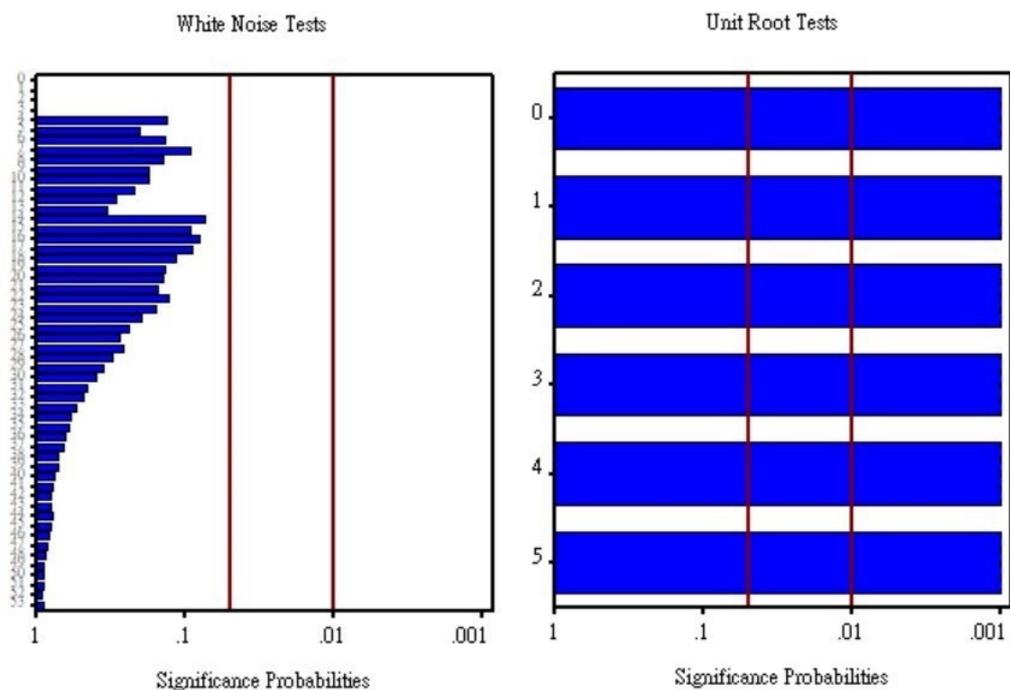


圖 11 ARIMA(2, 1, 1)白噪音檢定(左)及單根檢定(右)

4. 預測結果

本研究最後利用所找出來的 ARIMA(2, 1, 1)模型進行下一年度(2020 年)的流感人數預測，並在圖 12 中呈現出預測值(紫色線)，以及標示出 95%預測區間(上界為黃色線，下界為灰色線)。最後，在表 9 中列出預測值及 95%預測區間。

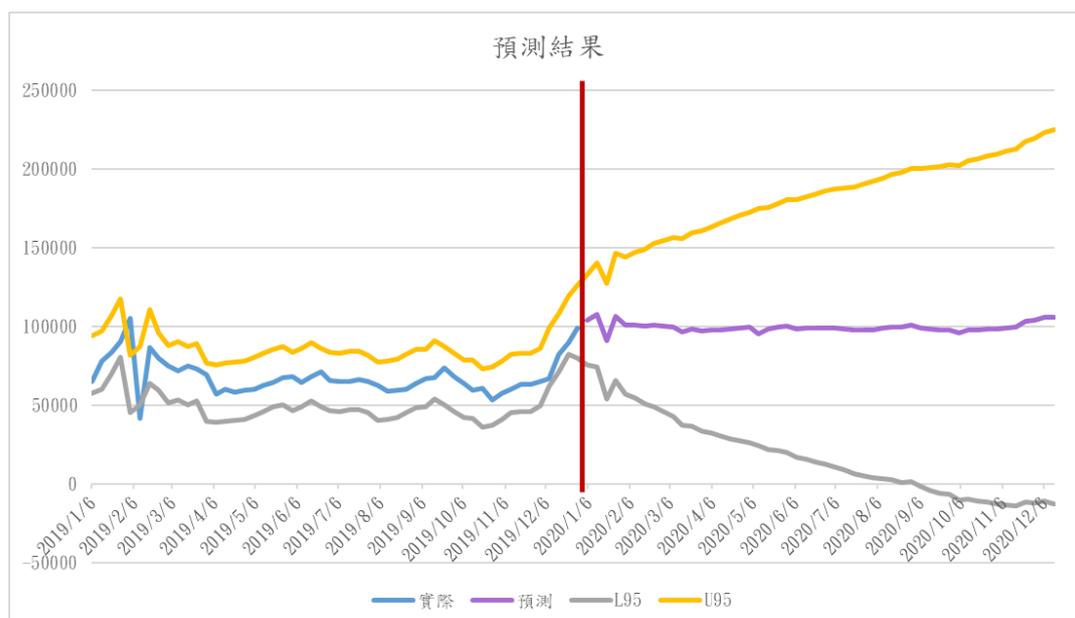


圖 12 ARIMA(2, 1, 1)預測圖

表 9 ARIMA(2, 1, 1)預測值及 95%信賴區間

時間	預測	L95	U95
2020/1/5	104262	75629	132895
2020/1/12	107554	74660	140449
2020/1/19	91043	54284	127802
2020/1/26	106346	66115	146578
2020/2/2	100834	57401	144267
2020/2/9	100988	54575	147401
2020/2/16	100166	50953	149379
2020/2/23	101077	49215	152939
2020/3/1	100507	46125	154889
2020/3/8	99698	42908	156489
2020/3/15	96788	37687	155888
2020/3/22	98418	37094	159742
2020/3/29	97429	33959	160898
2020/4/5	97995	32450	163540
2020/4/12	98040	30483	165596
2020/4/19	98573	29063	168083
2020/4/26	99035	27624	170445
2020/5/3	99583	26322	172844
2020/5/10	95665	24599	174731
2020/5/17	98688	21859	175516
2020/5/24	99793	21241	178345

2020/5/31	100286	20048	180524
2020/6/7	98787	16898	180677
2020/6/14	99085	15577	182594
2020/6/21	98931	13835	184028
2020/6/28	99287	12632	185943
2020/7/5	98901	10715	187088
2020/7/12	98486	8794	188177
2020/7/19	97637	6465	188809
2020/7/26	97982	5353	190611
2020/8/2	98064	4001	192127
2020/8/9	98967	3491	194443
2020/8/16	99499	2631	196367
2020/8/23	99553	1313	197793
2020/8/30	101010	1416	200603
2020/9/6	99265	-1663	200194
2020/9/13	98556	-3691	200802
2020/9/20	97813	-5735	201360
2020/9/27	98175	-6658	203007
2020/10/4	96265	-9837	202367
2020/10/11	97829	-9527	205185
2020/10/18	98086	-10511	206682
2020/10/25	98769	-11052	208591
2020/11/1	98543	-12492	209577
2020/11/8	99137	-13096	211371
2020/11/15	99461	-13959	212881
2020/11/22	103189	-11405	217783
2020/11/29	104001	-11755	219758
2020/12/6	106090	-10817	222998
2020/12/13	106012	-12751	224775

(三)指數平滑法模型(Exponential smoothing method)

1. 模型配適

在指數平滑法中，我們比較了 Linear (Holt's) Exponential Model、Winters Model-Additive、Winters Model-Multiplicative 和 Simple Exponential Smoothing 四種模型的 Root Mean Square Error，選擇 Root Mean Square Error 最小的模型進行分析。

表 10 Exponential smoothing method 的 Root Mean Square Error

Exponential smoothing method	Root Mean Square Error
Linear (Holt's) Exponential Model	9859.1
Winters Model-Additive	10025.8
Winters Model-Multiplicative	11198
Simple Exponential Smoothing	9854.5

經過比較 Simple Exponential Smoothing 的 Root Mean Square Error 最小且資料呈現無趨勢，因此選擇 Simple Exponential Smoothing 進行參數估計與分析。

2. 參數估計

從表 11 的參數估計結果來看，模型為：

表 11 Simple Exponential Smoothing 的參數估計

ROOMAVG：流感確診人數				
Simple Exponential Smoothing				
Model Parameter	Estimate	Std. Error	T	Prob> T
LEVEL Smoothing Weight	0.77067	0.0303	25.4515	<.0001
Residual Variance (sigma squared)	97298542			
Smoothed Level	96460			

$$\hat{y}_{T+1} = l_T, \quad l_T = 0.77067y_T + 0.22933l_{T-1}$$

3. 殘差檢定

透過 ACF、PACF 及白噪音檢定殘差是否存在自相關，以及使用單根檢定檢測殘差是否平穩。從圖 13 的殘差圖中可看出大部分殘差為隨機的，但在 2016 年至 2019 年間存在殘差自相關。從圖 14 的 ACF 及 PACF 檢定結果中，發現在 lag2 和 lag51 的地方檢定未通過，代表，殘差存在自我相關性。從圖 15 的白噪音檢定(左圖)結果中，發現 p-value 小於 0.05，代表殘差並非隨機的。

由於此模型的 RMSE 已是我們所配適的指數平滑法模型中最小的，但殘差檢定結果仍存在自相關，因此受模型限制，這份資料並不適合使用指數平滑法模型(Exponential Smoothing Method)。

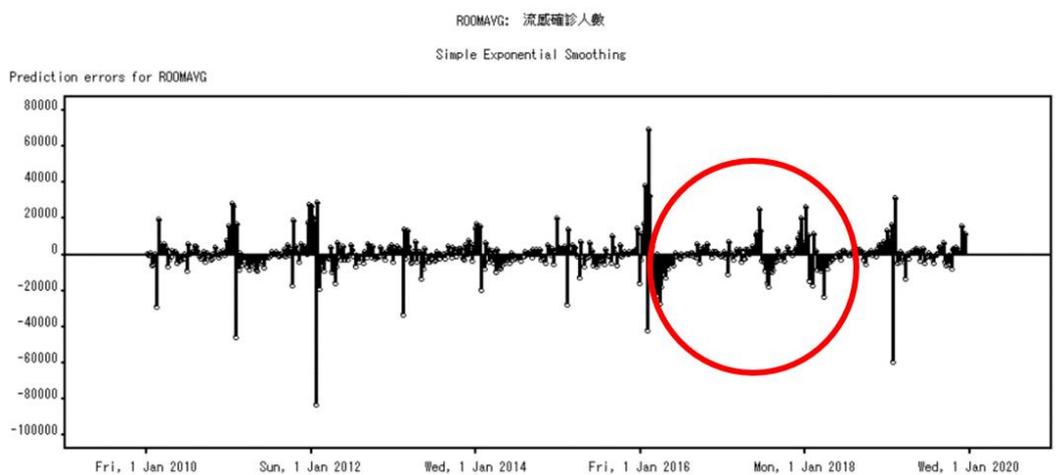


圖 13 指數平滑法殘差圖

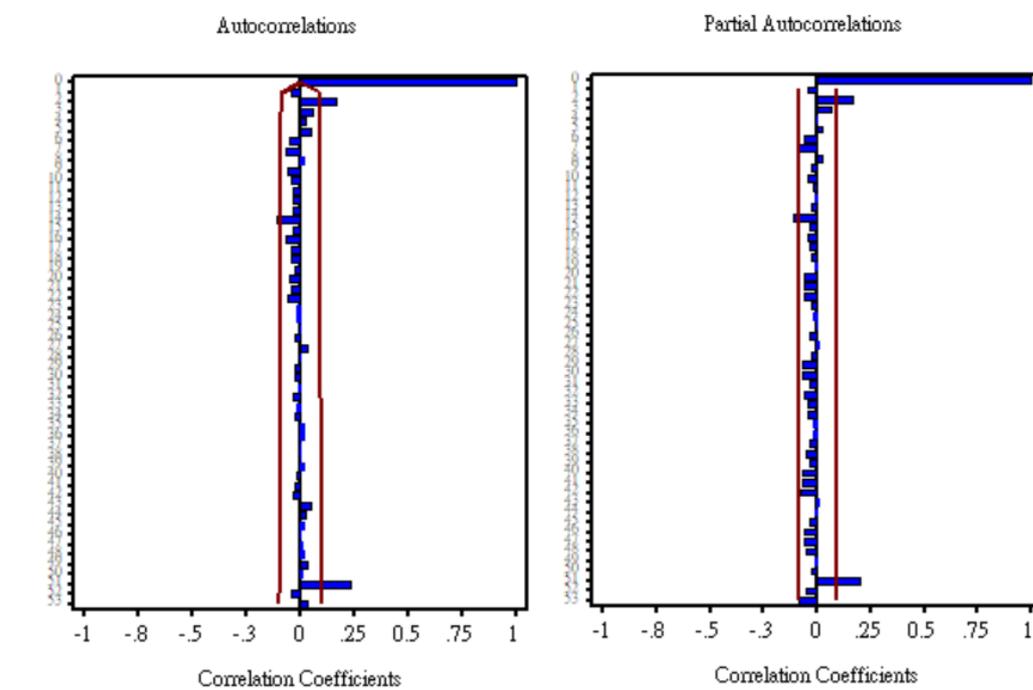


圖 14 指數平滑法的 ACF、PACF 圖

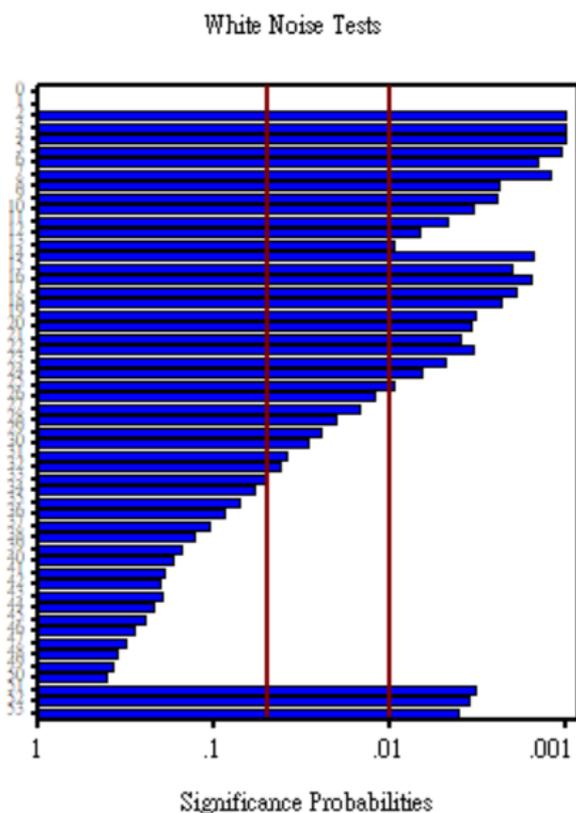


圖 15 指數平滑法的白噪音檢定

4. 預測結果

利用 Simple Exponential Smoothing 模型預測下一年(2020 年)的流感人數，並在圖 16 中呈現出來，並標示 95%信賴區間。最後在表 12 列出預測值。

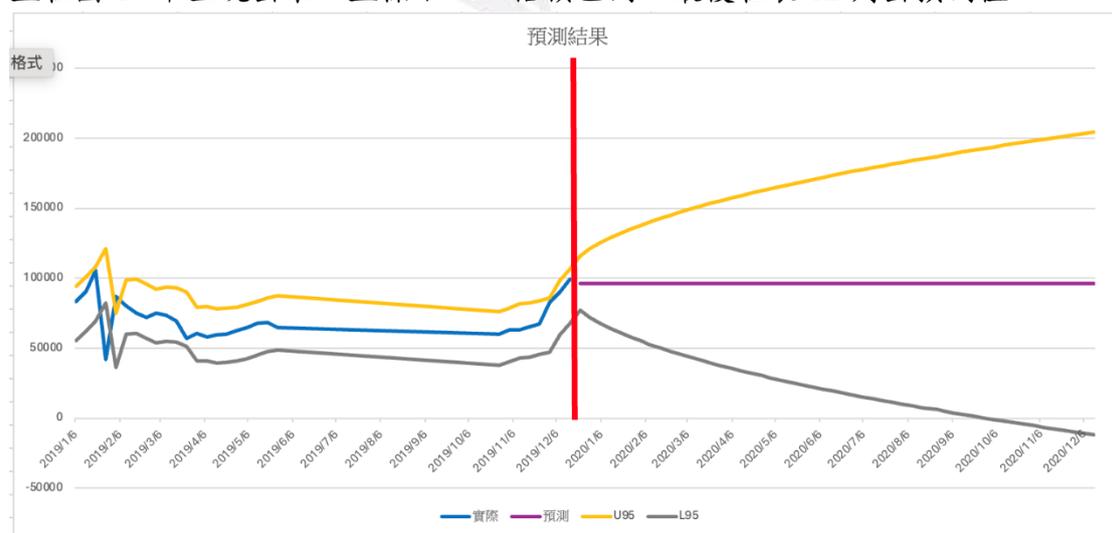


圖 16 Simple Exponential Smoothing 預測圖

表 12 Simple Exponential Smoothing 預測值及 95%信賴區間

時間	預測	L95	U95
2020/1/5	96460	67864	125057

2020/1/12	96460	64215	128705
2020/1/19	96460	60939	131981
2020/1/26	96460	57941	134979
2020/2/2	96460	55160	137761
2020/2/9	96460	52555	140366
2020/2/16	96460	50096	142825
2020/2/23	96460	47760	145160
2020/3/1	96460	45532	147388
2020/3/8	96460	43397	149523
2020/3/15	96460	41345	151575
2020/3/22	96460	39367	153554
2020/3/29	96460	37455	155466
2020/4/5	96460	35603	157318
2020/4/12	96460	33806	159115
2020/4/19	96460	32058	160862
2020/4/26	96460	30357	162563
2020/5/3	96460	28699	164221
2020/5/10	96460	27080	165840
2020/5/17	96460	25499	167422
2020/5/24	96460	23951	168969
2020/5/31	96460	22436	170484
2020/6/7	96460	20952	171969
2020/6/14	96460	19496	173425
2020/6/21	96460	18067	174854
2020/6/28	96460	16664	176257
2020/7/5	96460	15285	177636
2020/7/12	96460	13929	178992
2020/7/19	96460	12594	180326
2020/7/26	96460	11281	181639
2020/8/2	96460	9988	182933
2020/8/9	96460	8714	184207
2020/8/16	96460	7458	185463
2020/8/26	96460	6219	186701
2020/8/30	96460	4998	187923
2020/9/6	96460	3792	189129
2020/9/13	96460	2602	190319
2020/9/20	96460	1427	191494

2020/9/27	96460	265.7931	192655
2020/10/4	96460	-881.2299	193802
2020/10/11	96460	-2015	194935
2020/10/18	96460	-3136	196056
2020/10/25	96460	-4244	197164
2020/11/1	96460	-5340	198261
2020/11/8	96460	-6425	199345
2020/11/15	96460	-7498	200418
2020/11/22	96460	-8560	201481
2020/11/29	96460	-9612	202532
2020/12/6	96460	-10653	203574
2020/12/13	96460	-11684	204605

第三節 預測結果比較

透過比較 Mean Square Error (MSE)、Root Mean Square Error (RMSE)、Mean Absolute Percent Error (MAPE)、Mean Absolute Error (MAE) 和 Correlation Coefficient (R)來找出時間序列迴歸模型、ARIMA 和指數平滑法三種方法所配適的模型何者最佳。檢定標準為，MSE、RMSE、MAPE 及 MAE 最小以及 R^2 最大的方法。

表 113 預測結果比較

模型	MSE	RMSE	MAPE	MAE	R^2
時間序列迴歸	75253760	8674.9	7.84008	4923.3	0.822
ARIMA(2, 1, 1)	88441034	9404.3	8.49660	5454.5	0.791
指數平滑法	97111439	9854.5	8.39985	5450.3	0.77

經過比較後，發現時間序列迴歸模型所配適的模型 Seasonal Dummies+AR(2) 的誤差最小，但是從圖 3 的 ACF、PACF 檢定結果及圖 4 的白噪音檢定結果發現，殘差存在自相關，推測時間序列迴歸模型並不適用於此資料，所以，我們選擇 ARIMA(2, 1, 1)模型作為最終的模型。此模型為：

$$(1 - 0.25554B + 0.2464B^{51})(y_t - y_{t-1}) = a_t + 0.11279a_{t-2}$$

第四章 結論與建議

第一節 結論

在原始時間序列圖中，觀察到資料具有季節性，每年流感的高峰期大約在每年年初，但在進行模型配適時，發現加入季節差分並不會使我們所配適的模型變好，因此模型中未加入季節差分。從最終模型 $ARIMA(2, 1, 1)$ 中，發現本週類流感確診人數成長率與上週呈負相關，本週類流感確診人數成長率會與去年的下週呈正相關，衛生福利部疾病管制署表示流感疫苗的保護力通常不超過一年，因此，推測若去年同時期流感成長率上升，許多人去打疫苗，到今年疫苗失去保護力，造成本期確診人數成長率上升。而本週類流感確診人數成長率與上週呈負相關，推測為民眾發現上週確診人數較高則會開始有一些預防動作，例如：勤洗手、戴口罩等，或是較嚴重時，政府也會介入進行一些預防措施，進行環境周遭的消毒，因此本週類流感確診人數成長率呈現負相關。

第二節 建議

流感好發期在冬季流感季節開始前幾個月，衛生部可以通過強化流感疫苗宣導來提高公眾的預防意識，特別是鼓勵高風險人群提前接種疫苗。此外，通過提供免費流感疫苗接種服務，降低經濟障礙，以提高整體接種率，並監控各地區接種進度，確保高風險地區的接種率達到預期水平。政府也可以在流感高發季節前，加強預防措施宣導，並且在公共場所加強消毒工作，也能舉辦健康教育相關活動，特別針對兒童和老年人講解流感預防知識和方法。

在流感疫情暴發時期，能夠迅速增加醫療資源配適，加強流感病人的醫療與隔離措施，防止病毒快速擴散。平時也能與國際衛生組織或其他國家加強合作，分享流感病毒株的監測數據和疫苗研發進展，一同應對疫情的威脅。也能透過推動長效疫苗的研發，延長疫苗的保護期，讓人民都有更好的抵抗力。

參考文獻

1. 黃彥嘉、林育如、施惟量、方啟泰(2019)。環境低溫對於類流感急診就診之影響：時間序列分析。台灣公共衛生雜誌，38(5)，539-550。
[https://doi.org/10.6288/TJPH.201910_38\(5\).108031](https://doi.org/10.6288/TJPH.201910_38(5).108031)
2. 吳姿瑩、張家榮、吳文祥(2018)。以醫療區為基礎之流感併發重症疫情監控系統。醫務管理期刊，19(3)，227-248。
[https://doi.org/10.6174/JHM.201809_19\(3\).227](https://doi.org/10.6174/JHM.201809_19(3).227)
3. 周書任(2017)。以社群網路分析進行流感趨勢預測〔碩士論文，淡江大學〕。華藝線上圖書館。<https://doi.org/10.6846/TKU.2017.00938>
4. 莊淑如(2012)。臺灣民眾有類流感症狀時戴口罩之行為意圖及其相關因素探討：H1N1 新型流感大流行與後大流行期間之比較〔碩士論文，國立臺灣大學〕。華藝線上圖書館。<https://doi.org/10.6342/NTU.2012.00326>
5. 鄒佳蓁、柯沛程、簡郁紘、湯宛貽、陳怡如、曾瑋慈、陳宛晴(2021)。流感及其併發症病例數之研究與預測。中國統計學報，59(3)，172-225。
<https://www.airitilibrary.com/Article/Detail?DocID=05296528-202109-202109110010-202109110010-172-225>
6. 維基百科(2023年8月13號) 流感大流行。
<https://zh.wikipedia.org/zh-tw/%E6%B5%81%E6%84%9F%E5%A4%A7%E6%B5%81%E8%A1%8C>
7. 衛生福利部疾病管制署(2023年6月)。健保門診及住院就診人次統計-類流感。<https://data.gov.tw/dataset/14593>
8. 衛生福利部疾病管制署 (2024年2月)。台灣流感速訊。
<https://www.cdc.gov.tw/Uploads/files/b05c5172-32f7-4ff0-82cd-3d6519bff0f8.pdf>
9. 衛生福利部疾病管制署(2023年9月) 疫苗保護力篇。
<https://www.cdc.gov.tw/Category/QAPage/RJ6gxWhhbZlqMrTMBIWPIQ>
10. 錢晨嗣、姜晨彥、夏寒、鄭雅旭、劉星航、楊妹、夏天(2023)。上海市流感樣病例就診百分比時間序列分析和預測模型研究。上海預防醫學，35(2)，116-121。DOI: 10.19428/j.cnki.sjpm.2023.22253
11. 查文婷、李渭通、何嘉慧、何嘉琪、呂媛、劉穎、鄒享玉、易尚輝(2021)。湖南省流行性感冒與氣象因素關係及預測。中國公共衛生，37(3)，537-541。DOI: 10.11847/zgggws1127362