

基於 LSTM 遞歸神經網路建構船舶耗材需求預測模型

郭俊良^{1*} 蔡宗憲² 洪培倫³

¹銘傳大學管理學院風險管理與保險學系

²國防大學理工學院資訊工程學系

³國防大學管理學院運籌管理學系

摘 要

軍事後勤管理具高度機敏性，致使導入人工智慧的數位轉型的革新工作進程緩慢，但是商業領域卻已經提出許多創新性的應用，因此軍事後勤管理導入人工智慧進行零附件存貨需求的預測分析已成為急迫性的研究議題。為解決傳統方法造成資產管理效能不張的問題，本研究提出植基於深度學習之 LSTM 長短期遞歸神經網路的零附件需求預測模型，經實證研究結果顯示，LSTM 演算法可以提高料件需求預測的準確率，未來可做為實現精準後勤運籌管理的方法論。

關鍵詞：時間序列分析、需求預測、存貨管理、長短期遞歸網路

Construct a Demand Forecasting Model for Maritime Spare Part Inventory Management based on LSTM

Jiunn-Liang Guo^{1*}, Chung-Hsien Tsai², Pei-Lun Hong³

¹*Department of Risk Management and Insurance, Management College, Ming Chung University*

²*Department of Computer Science and Information Engineering, Chung Cheng Institute of Technology, National Defense University*

³*Department of Logistics Management, Management College, National Defense University*

ABSTRACT

Military logistics management is highly sensitive, resulting in slow progress in the innovation of digital transformation while conducting artificial intelligence. However, many innovative applications have been proposed in the commercial field. Therefore, the issues related to that military logistics management introduces deep learning to predict and analyze spare parts inventory demand are and necessary. In order to solve the problem of asset management inefficiency caused by traditional methods, this study proposes a spare parts demand forecasting model based on LSTM recurrent neural networks. The empirical research results reveal promising performance that LSTM algorithm can improve efficiency for demand forecasting. Accordingly, the proposed scheme provides feasible methodology to achieve precise logistics management in the future.

Keywords: Time Series Analysis, Demand Forecasting, Inventory Management, LSTM

一、前言

隨著人工智慧技術的快速發展，近年來世界各國在各領域均投入大量的資源致力於智能化的數位轉型，特別是國防產業更積極導入人工智慧及大數據分析等先進資通訊科技以進行軍事事務革新，期能藉由國防科技數位轉型提升國家安全層次。台灣囿於複雜的國際地緣政治等外部因素的限制，先進武器獲得不易，往往無法立即獲得符合國軍當前建軍規劃所需的最新裝備，使得國軍裝備經常面臨軍備壽期後勤籌補的問題。雖然透過各種延壽方案仍然能夠維持戰力，但是隨著時間的遞延，許多武器裝備的零附件經常產生消失性商源，間接影響裝備妥善率的嚴峻問題。若是屬關鍵耗材類則籌補更為費時而獲得困難，若屬必要品項，則需耗費更高成本購置（廠商開模、重啟生產線等），或辦理裝備構型變更等作業，以求裝備能符合零附件適用。目前雖已配合政府國防自主政策提高裝備自主性，但關鍵零組件尚須仰賴國外廠商供應，使得零附件存貨管理系統的效能致關重要。

國軍組織調整歷經精實案、精進案迄至精粹案以來，各單位縮編、簡併、裁員，業務簡化未提出有效配套作法，使人力縮編後工作負荷量大增，後勤資訊系統並未同步整合，致使許多資料統計分析功能支援度不佳。為提升補保工作執行效率，並確保零附件供應無虞，運用資料探勘技術分析各式後勤系統裝備及零附件管理運作資訊，提升後勤系統運用效益，推動系統整合及精進，構建共同作業平臺是必要推行之方向。目前國軍各類後勤系統歷年均編列相關預算辦理系統研改以精進內容，故與當初規劃架構與功能略有出入，因各後勤系統由各單位自行開發建置，時常造成資訊獲得時效落差，致使無法有效管制庫存最新資訊，並且也未提供需求預測功能供備料規劃人員參考[1]。

國軍在各式裝備的後勤管理部分主要是以全壽期系統管理（Total Life Cycle Systems Management, TLCSM）方法規劃管理作業流程，涵蓋從武器裝備系統的初始規劃到研發、採購、生產、維護和處置的整個生命週期過程。由於維持作業佔全壽期成本的 60%~70%，因此在武器裝備系統購置後，要確保維護妥善，相關零附件的補充對於武器裝備系統的高效運行至關重要。為確保零附件的供應無虞，需

要對零附件領用量進行高準確度的預測。由於低準確度可能會造成因維護延遲導致使用率降低、庫存不足將影響戰備任務遂行、或庫存過剩將形成庫儲資產積壓，進而衍生預算浪費。因此，精準預測武器裝備系統零附件需求不僅可以降低武器裝備系統全生命週期的運行和維護成本，而且對提高武器裝備系統的戰備水平具有重要意義[2]。

隨著工業 4.0 智慧製造概念的倡導，目前主要有九種科技驅動著各產業的智能化轉型，即大數據分析、深度機器學習、雲端網路、機器人自動化、智慧系統整合、物聯網、網路安全、擴增實境以及數位對映模擬。當中所運用的智慧型數據分析與預測方法更是數位轉型時代提高資料處理效能洞察商業智慧的利器，同時可透過建立預測模型或因素分析改善組織管理效能，創造競爭優勢。

近年來智慧型數據分析於軍事後勤的相應用已受到高度重視，亦成為世界各國未來進行軍事事務革新之重點。軍事後勤乃整體聯合戰力發揮之基礎，而人工智慧預測模型將成為未來作戰指揮決策支援之重要基礎。其中就後勤保修及零附件管理而言，裝備維護之零附件獲得與存貨水準需求規劃則被視為軍事後勤管理之關鍵項目。由於武器裝備獲得困難，國軍大部份裝備經常遭遇中壽期維護成本高漲的問題，主要武器裝備因服役時間長，故為維持這些武器系統及裝備的妥善率及可支援性，以當前的系統支援不足的環境下，已投入相當多的人力資源，在裝備持續老舊的前提下，零附件更換頻率亦逐年提高。由於軍品獲得管道在外購的部分，其交貨期往往動輒在 1 年以上，若未及早進行前置精準規劃，後續在零附件供補上將因緩不濟急而造成裝備無法及時修復等情形，將嚴重影響戰力。

零附件管理在後勤運籌管理之庫存料件需求預測是屬於挑戰性的工作，多樣多量的零附件致使倉儲成本難以精準掌控，而庫存不足亦嚴重延誤裝備維護效率，同時產生因緊急採購的較高成本代價。零附件的年度總需求主要包含計畫性需求與非計畫性需求等兩部分，計畫性需求是以年度定期維護計畫作為編列基礎，並配合裝備維修週期，實施相對應需求估算。而非計畫性需求，往往是突發性之維修保養作業，常使獲得端及預算管制人員不易預判，因此備料時主要以經驗法則執行推估備料量，容易造成需求落差，也就是缺料及備而未

用等情形，進而產生庫儲資產積壓而低效運用的問題，嚴重則易產生久儲未耗資產持續增加之趨勢。為解決此一問題，雖然後勤單位會針對物料採購前建立庫存量審查機制，並於採購案建立前先行完成資格審查，但就存貨管理效能來看，將無法解決久儲未耗問題。因此急需建立精準預測機制，實際檢視需求面資料異動的樣態，並建立完善零附件備料需求預測模型，才能提升後勤存管作業效能。

由於國軍後勤單位現行資訊系統均由各業務單位獨立開發建置，使各資訊系統間無法直接相互構連，系統資料庫的資料無法共享，常需反覆以人工作業方式進行資料傳遞及轉換，不但效能不張還造成人力資源浪費，失去原本資訊化的目的。另外，在資料庫部分，因各資訊系統分由業務主管單位個別發展，其所使用資料庫形態不同，資料庫成立後雖持續十餘年之系統研改，然僅著重在支援業務之執行面，系統各自發展之結果已造成系統內部資料欄位定義、編碼規則等內容差異甚大，無法呈現整合性及共享資訊。因為各資訊系統資料庫整合困難，增加各項資料分析的難度，亦無建立整合備料預測系統，造成需求預測停留在各單位人工紙本，易產生備料上誤差。

另外，國防部在後勤相關維持費的預算時常遭受立法院凍結及刪減，造成預算運用的排擠，整體預算無法滿足年度需求等困境。為使採購預算運用發揮最大效益，需建立一套有效的預測推估機制，用以準確預測需求量並改善零附件獲得的前置規劃，進而可間接提升預算編列的精確度，以確保預算使用效率。若能導入大數據及人工智慧方法，建立具體可行且量化之高司決策模型，未來可優化整體後勤運籌決策流程。

本研究針對公部門某單位後勤庫儲系統有關船舶零附件的歷史領用紀錄進行分析，運用深度學習之長短期遞歸神經網路（LSTM）建構零附件需求量預測模式，用以改善存貨管理人員備料之準確性，並提供後續後勤運籌管理人員之決策參考依據。透過歷史撥發異動紀錄，建立常模分析評估，以即時掌握單位後勤作業現況及需求，具體整合後勤管理資源，提升後勤運籌管理效能。因此本研究目的為：建立通用零附件備料之科學預測模型，提供後勤備料之參考、驗證與比較傳統時間序列與深度學習的預測方法最後將應用 LSTM 深度學習方法建構最佳化之零附件時間序列需求預測

模型。

二、文獻探討

2.1 存貨管理

存貨管理中有關物料整備的支援度係裝備保養維修達成高妥善率之關鍵因素，為確保各式裝備發揮應有的功能，需依據完善的計畫備料規劃，運用需求分析科學方法，參考技術文件及庫儲歷史資料，藉審查、稽核機制，精實推估保修計畫備料需求量，同時根據需求、獲得、儲存、分配等面向，滿足裝備保修各期程的管制目標。Wild[3]提出存貨管理是一項在企業公司中越來越重要的任務，並且通常是公司成功的因素，在庫存的控制中，包括存儲、發貨和運輸，運用良善的庫儲管理系統可以穩定成本和間接費用，並提供更好的預算控制，良好的系統可以改善供應鏈中各個階層的問題。存貨管理是一個非常複雜的問題領域，正是這種複雜性使得開發一種系統成為輔助決策的必要要件[4]。

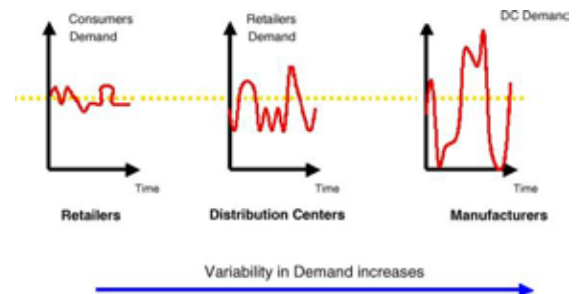


圖 1. 長鞭效應圖

此外供應鏈管理（Supply Chain Management, SCM）是協調貨物、服務、訊息和資金從原材料到零件供應商、製造商、批發商、零售商到消費者的過程，鏈中的每個節點都必須使用來自其相應上游的需求訊息來決定要採購產品之購買時間及數量，而往往在供應鏈中缺乏協調和共享的訊息，導致真正的需求無法正確獲得，就如同長鞭效應（bullwhip effect），如圖 1 所示，需求量往往因為無法獲得精確的資訊，而導致一連串推估量值偏誤的放大現象[5]。為維持裝備妥善及滿足客戶需求，企業在供應鏈中的各階層都必須嚴密掌控相關資訊，才能確保預估的正確性。

蘇雄義[6]提到，制定存貨政策時，物流管理策略的良窳對於企業帶來風險及影響具有

關鍵性影響。而存貨管理最重要的一環在於平衡存貨的供需，也就是說應具有足夠的存貨，來滿足顧客對產品的需求，不會因為存貨短缺而喪失銷售機會，然而在持有存貨時會產生管理、資金、保險等成本，使得企業不願意持有過多的存貨。Boylan and Syntetos[4]指出有效的存貨管理對許多公司相當重要，如航空業、電信產業等，在這類企業公司中，存在著許多庫存，而這些存貨對於公司營運產生重大影響。因此，如何有效管理存貨，確保在低存貨水準下，滿足所有顧客的需求，這是存貨管理中最具挑戰性的工作。有越來越多的企業導入先進的供應鏈管理模式，運用協同規劃、預測與補貨（Collaborative Planning, Forecasting, and Replenishment, CPFR）來改善存貨周轉率。

備料與存貨管理是密不可分的，而存貨即泛指所有可留用於未來、具經濟價值而目前暫時處於閒置狀態的資源，是一種包含各式型態（自原料至成品）之集合體。存貨管理主要目的在於以最低總成本維持最低原物料存量，並能適時、適量、適質地滿足生產及銷售所需，提高週轉能力[7]，故如何降低庫儲成本，在需求預測的部分，除了在各企業佔了重要的一環，亦為零附件管理重要課題，Wild[3]指出對歷史和預期需求的詳細分析提供了更好的預測，可以改善庫存過剩及預算浪費的情形。

存貨管理在客觀方法中，最常見的是經濟訂單數量理論（Economic Ordering Quantity, EOQ），該理論為每項物料確定兩個數量：再訂貨點和訂貨數量。當現有庫存到達再訂購點時，一個訂購量等於訂購數量以補充庫存。計算訂單數量通常只需要預測每段期間的平均需求。相反，計算正確的再訂貨點需要估計在補貨訂單生成到其到達庫存之間的時間間隔內的整個需求分佈，也就是提前期[8]。然而在裝備武器採購費用昂貴，後續後勤維修能量之維持，相關零附件之管理佔了相當大的比重，而備料是以最低總成本維持最低原物料存量適時、適量支援維修任務，確保各式武器裝備妥善，並積極運用各單位專業人力，參考技術文件及後勤支援分析資料，精實檢討保修所需備料，避免因需求失真產生「待料」、「呆料」，並建立一套方法儲存適當之備份元件是備料作業主要的目的。因此，以庫儲系統歷史領耗用紀錄來分析預測相關裝備之未來修期備料狀況，可提高備料精度，減少預算浪費。

2.2 需求預測

需求預測往往與資料量大小具高相關性，Mayer-Schönberger and Cukier[9]等學者的研究指出巨量資料分析的核心在於預測模型，而該模型需以大量資料做為分析運算的基礎，結合機器學習等人工智慧方法進行訓練並判斷最佳參數組合及訊息。Hua and Zhang[10]則提到零附件需求預測的準確性對庫存控制具高度重要性。良好的預測對可以降低持有成本，最大限度地減少浪費，並減少所需的營運資金，來增加現金流以及提高履行客戶訂單的能力[11]。目前在各產業中均積極地以機器學習建立需求預測來降低額外費用所導致的物流成本及產品短缺所產生的機會成本[12]。Chen and Chen[13]的研究提到若能正確預測實際需求量則可有效控制庫存水準並可直接訂定正確的採購量。由此可知，如何分辨雜訊準確判斷數據訊號做出誤差最小的預測，是需求預測方法所需關注的思考面向。

需求預測的方法主要可分為量化方法（定量法）與質化方法（定性法）兩種。余桂霖[14]指出量化預測是一種分析與時間序列有關之歷史資料的研究策略，著重在檢視歷史資料的時間樣態與所關注議題之間的關聯性，也就是利用過往時間之歷史資料來預測未來的結果。而質化預測是在當歷史資料無法利用時，由專家學者的判斷來進行預測。量化預測又可分為兩種類型：時間序列與因果分析，而時間序列原理係指按照時間發生先後順序進行排列紀錄，透過觀察歷史的時序資料來分析未來需求值。時間序列的預測方法可分為時間序列統計分析及機器學習兩類別相關預測方法分類如圖 2，較常見的傳統時間序列方法有移動平均法、指數平滑法(Exponential Smoothing, ETS)及自迴歸積分移動平均(Autoregressive Integrated Moving Average Model, ARIMA)等，而機器學習則是基於類神經網路及其衍伸的方法，例 LSTM 等。本研究資料係基於時間因子樣態，屬性為序列資料，將探討資料間前後關係，所採用之預測方式為歷史資料來實施未來需求值之預測，故將運用時間序列相關演算法進行分析研究，期建立預測模型供後勤人員決策使用。



圖 2. 推估預測方法分類

2.3 時間序列分析

存貨管理之需求分析主要針對歷史資料進行時間脈絡的研究，資料屬性可視為時間序列分析的研究領域。余桂霖[14]提出時間序列是由經濟或物理現象在連續的均等時間內所產生一連串觀察值，而分析是一種透過已發（歷史）的時間序列資料的特性，來預測或推估未來數值的過程。表示對某一現象時間序列的成分加上分解，分析各成分的變化情況，往往會發現它們呈現出某種形態的變化，而未來該成分亦將順從同一形態或呈現另一形態而變化。

時間序列分析的目的是在研究過去時間點所發生的歷史事件，來更能夠了解資料的基組成架構，進而給予預測未來事件的工具及方法，它是一種使用跨越某個已發生時間的序列去進行預測，時間序列亦屬數據探勘的重要領域，而數據探勘已廣泛用於各個領域的需求預測，諸如能源[15]、旅遊[16]、交通[17]、水管理[18]、製造[19]、自行車共享[20]、零售藥店[21]、醫院[22]、物流[23]等，這些領域的實證研究結果可證明時間序列分析具有方法的泛用性。至於庫儲需求預測已有學者採用相關方法來解決產業中的特定問題，而相關問題大多具有時間序列的特性，因此可透過分析已發生（歷史）的時間序列資料來預測或推估未來需求。

傳統的時間序列模型雖然有不錯的成果，惟近期在人工智慧的進步下，有更多深度學習的方式可進一步提高預測的準確度。Wang, Chien, and Trappey[24]主張 ARIMA 方法的缺點為時間序列數據必須具備定態(stationary)，而且只能處理線性關係，該研究透過實際半導體生產廠訂單資料，比較 ARIMA 和長短期遞歸神經網路(Long Short-Term Memory, LSTM)模型，運用均方根和平均絕對百分比誤差進行評估和驗證，發現 LSTM 在短期預測方面表現出更高的準確性，明顯提高了需求和庫存預測的準確性和效率。另外 Dali and Chengcheng[25]說明由於零附件需求的不確定性，傳統的時間序列模型難以獲得良好的預測結果，然而 LSTM 模型已可以在裝備的零附件得到很好的預測結果，為了提高準確性，基於 LSTM 模型架構下提出了一種整合經驗模態分解 (Ensemble Empirical Mode Decomposition, EEMD)和 LSTM 的組合預測模型，透過均方根和平均絕對百分比誤差實驗評估證明，它的準確性遠超過傳統 ARIMA 時間序列預測模型。

Abbasimehr 等學者[26]提到時間序列預測技術以深度學習 LSTM 的方法進行需求預測，可得到不錯的結果。另外由於仍有部份文獻顯示傳統方法在時間需求預測方面也獲得不錯的表現，因此本研究將運用 LSTM 與 ARIMA 模型進行比較，以確認較佳之零附件預測模型方法。

2.4 長短期遞歸神經網路 LSTM

遞歸神經網路 (Recurrent Neural Network, RNN) 是屬於深度神經網路的一種架構，是利用分析資料發生的前後順序樣態，並從前期階段學習的模型來預測未來的趨勢。該方法在預測下一步的輸出值時需要記錄前期階段的數據，藉此模擬人類大腦記憶過去資訊的機制。在 RNN 中，隱藏層做為內部儲存器，用於儲存在讀取順序數據的前期階段獲得的資訊，此特點是利用前期分析的資訊來預測未來隱藏的重要序列資料。典型 RNN 有先天的缺點，因該方法不適合記住較長的序列資訊[27]。並且在執行倒傳遞運算時，若每個時間點之權重均小於 1 時，傳遞到初始值其乘積將趨近於 0，則預測結果的誤差就會消失，因而產生梯度消失情形。反之若權重值大於 1 時，則會造成誤

差值放大效果而產生梯度爆炸的問題[28]。為解決此問題，Hochreiter and Schmidhuber[29]提出長短期神經網路(LSTM)，透過加入遺忘閘的設計，如此可以在較長的時間區間內儲存訊息，並記住較長的短時間區段資訊。

RNN 和 LSTM 之間的主要區別在於 LSTM 可以儲存長期時間序列資訊，因為 LSTM 網路結構包含控制訊息的閘門、輸入閘 (Input Gate)、遺忘閘 (Forget Gate) 和輸出閘 (Output Gate) 的組成架構[30]。其中遺忘閘依據前次資訊及當時輸入的資訊來計算，可以選擇記住的數據和遺忘的數據，來防止內部單元值無限制地增長，以彌補 RNN 因梯度消失和爆炸而無法學習的狀況。Siami-Namini et al. [27] 運用金融數據比較 ARIMA 和 LSTM 作為預測時間序列數據的代表性技術的準確性，實驗結果證明 LSTM 明顯優於 ARIMA。

從上述文獻得知，深度學習演算法在需求預測上較傳統方法準確，其中長短期遞歸神經網路(LSTM)為各領域在做需求預測之基礎，並且在實證上的預測效果的準確性較高，惟在零附件需求預測之領域文獻著墨較少，因此，本研究提出基於 LSTM 演算法的模型架構，以庫儲系統之零附件歷史耗用紀錄資料來訓練模型，期能獲得準確之預測結果。誠如前述文獻探討顯示傳統方法與神經網路之運用各有所長，目前深度學習為最有效及運用最廣之預測方式，特別針對時間序列資料預測效果最為準確。因此，本研究將與傳統 ARIMA 方式進行比較，對於提高勘估備料的精準度之預測，關係著裝之妥善率目標之達成、降低庫儲資產及擲節年度預算，以達預算管控及維持裝備妥善之目的。

三、研究方法

3.1 研究架構

本研究將蒐集某單位歷史零附件耗用資料，並依據資料特性進行預處理使本研究數據更臻完備。研究程序透過資料蒐集、資料預處理、建模分析、研究評估機器學習相關演算法之方式，找到最佳預測模型，最後產生實證結果，研究架構圖如圖 3。

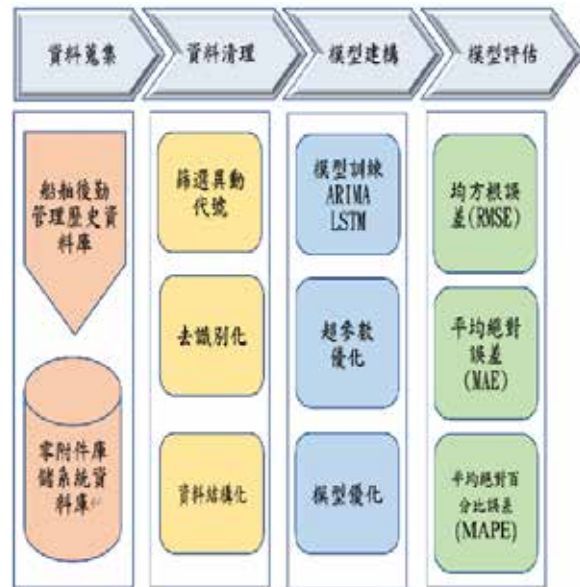


圖 3. 研究架構圖

研究對象為公部門某單位裝備零附件自 2010 年至 2020 年止之 10 年歷史領用紀錄。資料蒐集時間頻率則考量各裝備維修屬性，將以通用料件之週資料為彙整時間單位。由於資料屬性符合時間序列資料類型，因此在資料清理階段將檢視系統中之歷史領用紀錄，篩選相關資料異動代號，並彙整每週耗用情形。為訓練模型，結構化資料完成處理後將分割為訓練及測試集，以訓練集資料來建立模型，而測試集資料則用來評估訓練集所建立之模型之表現，同時為測試模型之穩固性，對於數據進行交叉驗證，最後產生備料需求預測模型。

3.2 長短期遞歸模型 LSTM

LSTM 可以預測具有未知持續時間的時間間隔的時間序列，與傳統 RNN 相比具顯著優勢[31]。LSTM 網路架構包三種神經網路層：(1) 輸入層，其神經元數量等於輸入變量的數量，(2) 單個或多個隱藏層和 (3) 輸出層的神經元數量等於輸出變量的數量。記憶單元的結構具有三種類型的閘門：(1) 遺忘閘 f_t ，(2) 一個輸入閘 i_t ，以及 (3) 一個輸出閘 o_t ，模型架構如圖 4。在記憶單元中，在每個時間 t ，輸入由來自輸入序列 X_t 的一個元素和前一期的輸出組成時間步長 h_{t-1} 。(a) 遺忘閘接受這些輸入並決定將刪除哪些資訊，(b) 輸入閘決定將哪些資訊添加到記憶中，(c) 輸出閘決定資料分析後的輸出值。

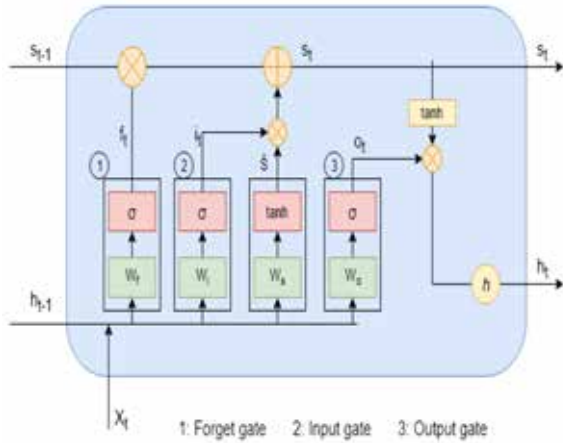


圖 4. LSTM 神經網路架構圖

LSTM 主要運用三個神經元閘門運算單元來架構整體模型訓練流程。首先，遺忘閘 f_t 負責將當次輸入值 X_t 和前次輸出值 h_{t-1} 結合(分別乘上權重值 W_f)作為當次(t)運算的自變項，並加上偏差項 b_f ，之後再以 sigmoid(σ)激活函數(Activation Function)進行非線轉換的計算。神經元閘門單元的設計主要是用來訓練並決定各神經元的最適權重 W_f 的參數組合，遺忘閘 f_t 的方程式如公式 (1)。

$$f_t = \sigma(W_f X_t + W_f h_{t-1} + b_f) \quad (1)$$

第二個閘門結構是輸入資訊與記憶篩選閘門，主要是用來分析當期輸入資訊及獲得更新記憶狀態的特徵資訊。這部份包含有兩個子神經元閘門：第一個子神經元閘門 i_t ，對當前的輸入做選擇性記憶，同樣取當次輸入 X_t 和前次輸出 h_{t-1} 相結合作為輸入(分別乘上權重值 W_i)，而 b_i 為偏差項，並使用 σ 來決定是否保留記憶資訊，如公式 (2)。第二個子神經元閘門 $\tilde{s}_t$ ，則是用來計算前次輸出的資訊(前期記憶值)與本次輸入值的相關性，因為相關程度的值域是介於-1 至+1 的區間，因此閘門是使用 $Tanh$ 激活函數進行非線性轉換，如公式 (3)。

$$i_t = \sigma(W_i X_t + W_i h_{t-1} + b_i) \quad (2)$$

$$\tilde{s}_t = \tanh(W_s X_t + W_s h_{t-1} + b_s) \quad (3)$$

當完成前述兩個閘門單元運算後，隨後會將 f_t 與前一次狀態 s_{t-1} 進行乘積運算，輸出值則代表是否保留前一次訊息或丟棄不要得資訊的決定，最後再加上 i_t 與 $\tilde{s}_t$ 乘積的結果，用以

更新記憶狀態，並得到新的記憶狀態單元值 s_t ，如公式 (4)：

$$s_t = f_t \cdot s_{t-1} + i_t \cdot \tilde{s}_t \quad (4)$$

最後是輸出閘門單元則是用來決定輸出值，結構同樣是取當前輸入 X_t 和前次輸出 h_{t-1} 相結合，並分別乘上權重值 W_o ，而 b_o 為偏差項，並使用 σ 來決定輸出 o_t ，如公式 (5)。本期輸出的隱藏記憶資訊 h_t 為公式 (6)，此運算會將記憶狀態單元值 s_t 透過 $Tanh$ 激活函數計算，再與輸出 o_t 進行乘積運算而得到最後 h_t 。

$$o_t = \sigma(W_o X_t + W_o h_{t-1} + b_o) \quad (5)$$

$$h_t = \tanh(s_t) * o_t \quad (6)$$

3.3 模型訓練超參數與優化

LSTM 神經網路模型的訓練過程中隱含超參數的設定，其目的是為調校及優化訓練模型以產生更準確的結果，藉由調整這些超參數可使模型更佳地適配因變項。超參數主要可區分為激活函數、損失函數、優化器、丟棄率和學習率等。

3.3.1 激活函數

激活函數是深度學習神經網路實現非線性輸出的關鍵，其功能如同感知器的開關，用來決定輸出值的大小。由於每一個神經元運算都必需經過此過程，可知激活函數是深度學習模型中最重要的超參數[32]。由於神經網路是透過調整神經元和資料流之間的權重來使訓練與學習，如果沒有轉換，則深度學習的效果與一般神經網路的效果沒有區別。在 LSTM 網路架構中使用到的非線性激活函數主要為 Sigmoid 及 Tanh 函數。Sigmoid 函數的值域介於 0 到 1 之間，由於此值域的特性可模擬控制神經元節點是否被激發並進行資訊傳導的功能，如同神經運算單元是否傳遞資訊的開關，因而成為神經網路運算節點主要的非線性轉換函式選擇。然而 sigmoid 函數在倒傳遞梯度運算的極值上常會變得非常小，所以在深度神經網路架構中易導致梯度消失 (vanishing gradient) 的問題而間接影響訓練的效能，詳如方程式(7)及圖 5 所示。

$$f(x) = \frac{1}{1+e^{-x}} \quad (7)$$

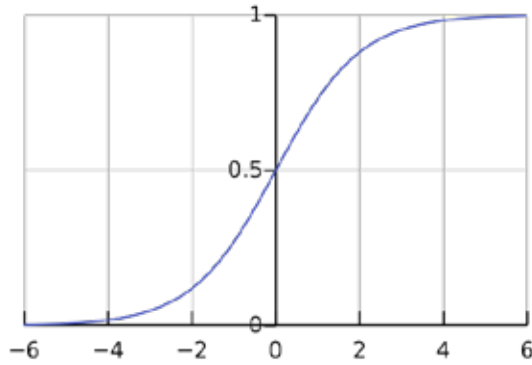


圖 5. Sigmoid 函數

Tanh 函數則是類似 Sigmoid 的一種非線性函數，值域介於 -1 到 1 之間，常運用在輸出範圍需要以負值表示，或是偏好參數平均值為 0 時使用，同樣地與 Sigmoid 函數一樣在倒傳遞運算時其極值亦可能有梯度消失的問題，詳如方程式(8)及圖 6 所示。

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (8)$$

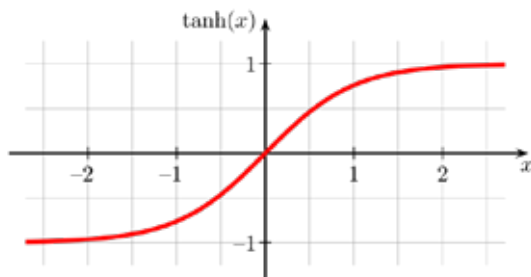


圖 6. Tanh 函數

3.3.2 損失函數

LSTM 網路架構中的損失函數(Loss Function)則是在訓練過程中用來計算誤差的目標式，藉以評估每一次訓練結果的預測值與真實值的差異程度，輸出值越小越好，基本上可分別針對迴歸和分類問題設計不同的目標函數。若運用 LSTM 在時間序列或迴歸的問題時，其主要的損失函數有均方誤差(Mean Square Error, MSE)、平均絕對值誤差(Mean Absolute Error, MAE)，而若運用在分類的問題則常見的函數為交叉熵(cross-entropy)。

在迴歸的問題中，通常期望模型預測值能夠接近真實的值，但實務上預測值基本上與實際值之間一定存在差距，此差距即是統計上的「殘差(residual)」，而損失函數中的 Loss 就是

計算「實際值和預測值的殘差」，而迴歸常用的損失函數為：均方誤差(Mean Square Error, MSE)、平均絕對值誤差(Mean Absolute Error, MAE)。

均方誤差(MSE)是用來計算預測值與真實值之間差異的均方值，其設計是因為計算所有資料的平均值，而誤差值包含正負數，所以加上平方即去除正負值在加總過程中相互抵消的效果，藉此可評估訓練模型的預測值與實際值的差異程度，如方程式(9)；另外，除了均方誤差外，亦可使用平均絕對值誤差(MAE)，以絕對值修正誤差值使其輸出為正數，如方程式(10)。

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (9)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (10)$$

本研究的零附件需求預測模型將以均方誤差(MSE)及平均絕對值誤差(MAE)作為模型訓練的損失函數。

3.3.3 優化器

隨著深度學習領域的快速發展，優化器與神經網路模型的表現有著高度相關性[33]。神經網路中每一個神經元有屬於自己的權重，而優化器(Optimizer)是用來調整權重值參數，使模型達到最佳值的重要方法，常見的優化器有 SGD、Momentum、Adagrad、Adadelta、RMSprop 及 Adam 等。

隨機梯度下降法(Stochastic Gradient Descent, SGD)是深度神經網路最常使用的優化方法，是基於隨機值進行梯度計算來修正權重，雖然訓練速度快，但由於更新頻繁，容易導致準確率下降或不易找到全域最佳解[34]。Momentum 是「動量」的意思，此優化器為模擬物理動量的概念，在同方向的維度上學習速度會變快，方向改變的時候學習速度會變慢。Adagrad 則是在學習過程中不斷調整學習率，依照梯度去調整學習率。優點是能加快訓練速度，在前期梯度較小時(坡度較平坦)能夠放大梯度，後期梯度較大時(坡度較陡峭)能約束梯度，但缺點是在訓練中後段時有可能梯度趨近於 0，而過早結束學習過程。Adadelta 為 Adagrad 的更新版本，可改善 Adagrad 的缺點，使用梯度均方根的指數衰減平均值作為第二梯度時刻，僅依賴前一時刻的平均和當前梯度

的訊息來計算，根據轉換更新的移動窗口來調整學習速度。RMSprop 是 Hinton [35] 所提出，可改善 Adagrad 的缺點。此外 RMSprop 方法比 Adagrad 增加一個衰減計算，用於衡量前次梯度變化的程度來更新學習率以減緩 Adagrad 學習率下降過快的問題。Adam 則是近期在各種模型表現較佳的方法，使用取代傳統隨機梯度下降過程的一階優化算法，基於訓練數據迭代地更新神經網路權重。由於上述優化器方法各有優缺點，因此本研究後續將透過實驗檢視並比較各優化器於本研究模型中的表現以確定最適配之優化器方法。

3.3.4 丟棄率

深度神經網路結構複雜，加上往往訓練數據量大，因而容易產生模型過度配適 (overfitting) 的問題。丟棄率 Dropout 是一種修正過度配適的方法，可在訓練時隨機將神經元的激活值預先設置為零 [36]。在訓練時每一次的迭代 (epoch) 皆以一定的機率丟棄隱藏層神經元使得被丟棄的神經元不會傳遞訊息，此外在倒傳遞運算 (backward propagation) 時，被丟棄的神經元因為其梯度是 0，所以在訓練時不會過度依賴某一些神經元，藉此達到降低過配適的效應 [37]，如圖 7。

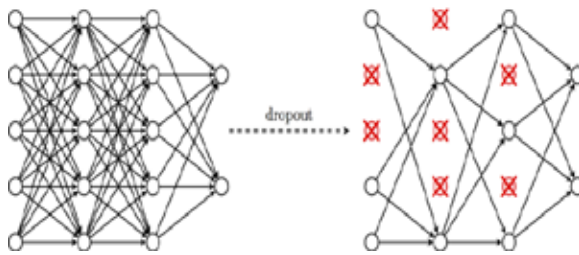


圖 7. Dropout 示意圖

3.3.5 學習率

學習率 (Learning Rate) 主要是用來決定損失函數梯度調整神經網路權值運算的速度，也就是在梯度下降的過程中控制每次權重更新的步幅大小，最終影響模型的收斂速度與表現。若學習率太小，雖可確定神經網路之收斂，但卻影響整體模型的訓練速度，可能會產生學習緩慢及落於局部區域最佳解的情形。相反的，如果學習率過大，雖可加快學習速度，但很有可能導致神經網路運算過於發散而錯過全域最佳解。由於此參數可以提高神經網路表

現透過較佳的學習率可改善神經網路模型 [38]，因此本研究亦將透過實驗組合決定相關較適的學習率值。

3.4 模型評估

為使模型評估以量化方式顯示其表現，並且計算出預測值與實際值之差異，在各項文獻中已提出許多模型評估方法來計算誤差，考量使用單一評估指標無法正確顯示模型差異，故本研究將以平均絕對誤差 (Mean Absolute Error, MAE)、平均絕對百分比誤差 (Mean Absolute Percentage Error, MAPE) 以及均方根誤差 (Root-Mean-Square Error, RMSE) 等方式來進階評估模型，當誤差值越小，則代表模型準確越高。

MAPE 是評估預測準確度最常用的指標之一種。MAPE 為每項絕對誤差的和除以實際值。實際上，它是誤差百分率的平均值。使用 MAPE，當百分比越小，則代表模型越佳 [39]，方程式如 (11)。

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \times 100\% \quad (11)$$

以平均絕對誤差 (MAE) 方式來進階評估模型，平均絕對誤差是一個很好的預測評估指標，它是所有單一觀測值與算術平均值偏差的絕對值的平均。與平均誤差相比，MAE 由於離差被絕對值化，不會出現正負相抵消的情況，因而，平均絕對誤差能更好地反映預測值誤差的實際情況。當誤差值越小，則代表模型準確越高 [40]，方程式如 (12)。

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (12)$$

RMSE 是一種經常用於評估模型而獲得預測準確性的度量，它是預測值與實際值之間的平方差平均值的平方根，測量實際值和預測值之間的差異或殘差，該指標比較不同模型對特定數據的預測誤差，而不是數據集之間的預測誤差 [27]，公式如 (13)：

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (13)$$

四、實驗與研究發現

本研究對象為某單位零附件管理系統，主要篩選 2010 年至 2020 年之間船舶零附件歷史耗用紀錄，針對高消耗量且不具機敏性之料件項目作為資料收集標的。零附件存貨水準與領用頻率與裝備維修保養作業效率高度相關，同時由於資料具時間序列特性，資料的質量將影響模型的訓練效率與預測準確性，故僅挑選具有高頻的領用品項作為研究對象。經篩選每年供補數大於 500 次以上品項後再以週為時間單位彙整歷史耗用資料，資料筆數共計 515 筆。考量裝備零附件資料之機敏性，原始資料已先行完成去識別化，並比對實際耗用之異動紀錄，刪除相同品項，以免重複計算，最後清理後的資料集主要以零附件耗用及供補數量等變項來分析時間序列資料之變化樣態(pattern)。

4.1 時間序列之需求預測模型

為檢驗深度學習模型的效益，本研究將 ARIMA 傳統時間序列方法建立需求預測模型的比對基準(benchmark)。考量後續比對的一致性，資料的分割均以隨機選取 70% 資料作為訓練集來訓練模型，其餘的 30% 則為測試集來測試模型的效能。初步以敘述統計分析來檢視時間序列資料集之變化，經時間序列折線圖可得到週零附件領用量的異動變化，如圖 8。

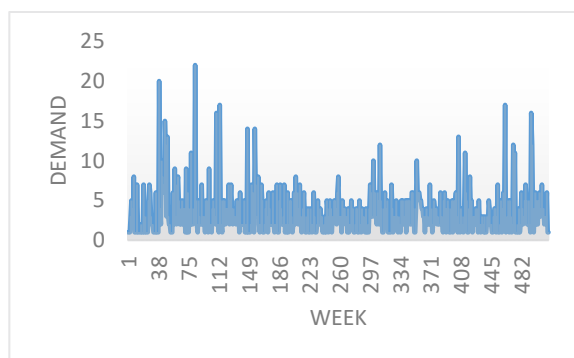


圖 8. 某零附件 2010 至 2020 年歷史需求異動

4.2 ARIMA 模型

為建立研究比較基準，首先將運用自迴歸綜合移動平均模型 (ARIMA) 建構需求預測模型，方法原理是基於迴歸移動平均(ARMA)廣義模型，是一種結合自迴歸(AR)過程和移動

平均(MA)以構建時間序列的複合模型，可應用在分析非平穩時間序列資料。由於 ARIMA 模型必須先將序列轉換為平穩的隨機序列，也就是需先行透過差分運算的過程將非平穩資料轉換為平穩資料。若考量週期性則可以季節差分與 ARIMA 模型相結合的 SARIMA 建立模型，因為 SARIMA 更適用於具有週期性特徵的時間序列數據建模。在本實驗中，使用 ARIMA 演算法分析資料集時間序列的變化，觀察原始數據在時間序列圖的變動，圖中呈現零附件每週撥發量在過去的異動情形，如圖 9。

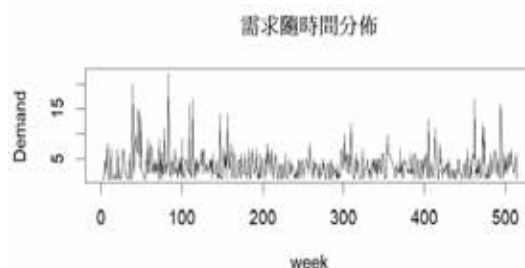


圖 9. 原始資料之時間序列分佈

從原始時間序列圖中初步可發現每週的領用量具有穩定的趨勢，後續經單根檢定可進一步確認均數及方差是否在一定範圍內，如果檢定後 p -value 未小於 0.05，則需要將資料進行轉換及差分。在對非定態資料之處理一般會利用差分或採用分解法修正資料的趨勢性與季節性，藉以消除趨勢和季節波動的影響，使序列達到平穩化。

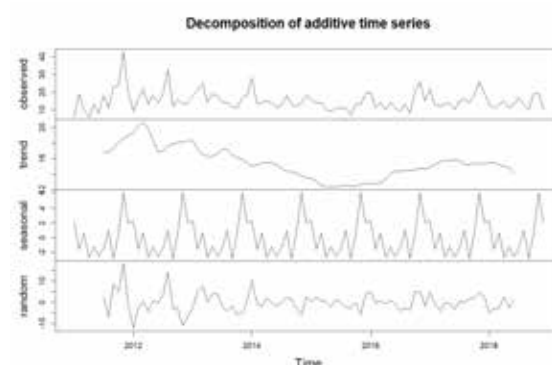


圖 10. 原始資料時間序列分解圖

本實驗初步假設研究資料帶有季節及趨勢性，因此透過分解法來分解時序圖(如圖 10)，圖中共有四個子圖，第一部分是顯示原始資料集的序列分佈，第二部分為資料均值的趨勢線，第三部分則萃取並視覺化資料的季節性樣態，第四部份則是將第一部份的實際值減去

第二部分的趨勢，再減去第三部分的季節性則可以獲得資料的隨機性。

此外序列是否穩定(定態)將影響序列資料的分析結果，所以在進行時間序列分析之前本研究將先檢定資料是否符合定態的假設。除了以單根檢定(Augmented Dickey Fuller, ADF)外，也可以使用殘差檢定分析。另外自相關係數(Auto Correlation Function, ACF)及樣本偏自我相關函數(Partial Autocorrelation Function, PACF)圖形也可用來驗證資料的定態與否。在單根檢定(ADF)部分，為了確定資料其平穩性，其 p -value 要小於 0.05。經分析本研究資料集 ADF 的 p -value 為 0.01，即可初步確認資料集屬於定態。而殘差檢定結果可從殘差圖(如圖 11)發現資料均分佈於信賴區間(藍色虛線間)，故也可再次確認資料集符合序列的平穩性。

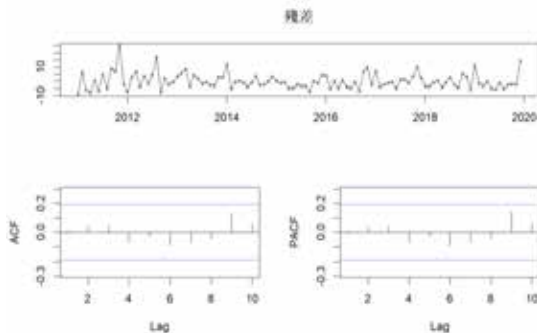


圖 11. ARIMA 模型檢定

最後以 ARIMA(1,0,0)參數進行模型訓練，並將預測結果以三項模型評估指標(MAPE、MAE、RMSE)分別加以計算，相關指標值將做為後續實驗的比較基準，模型如圖 12。

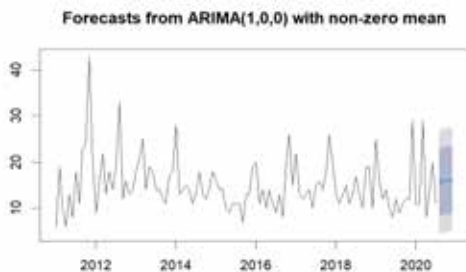


圖 12. ARIMA 模型預測圖

4.3 LSTM 模型

誠如文獻探討的研析，許多領域的研究已證明深度學習演算法在需求預測可以得到較佳的結果，因此本研究將進一步運用深度學習

LSTM 遞歸神經網路進行建模及預測，最後再與 ARIAM 模型進行比較。本實驗的 LSTM 模型預設的激活函數是使用 sigmoid，超參數實驗之優化器比較則分別採用 Adadelata、RMSprop 及 Adam，並搭配不同神經元數及學習率進行組合實驗，藉此獲得最佳模型參數組合，如圖 13。



圖 13. 超參數實驗組合

4.3.1 模型建構

本資料集經整理後共計為 515 筆序列資料並將其分割為訓練集及測試集，隨後進行模型的預測與評估，LSTM 模型初始參數如表 1。

表 1. LSTM 模型初始參數

Layer (type)	Output Shape	Param
lstm (LSTM)	(1, 1, 10)	480
lstm_1 (LSTM)	(1, 10)	840
dense (Dense)	(1, 1)	11

4.3.2 模型優化組合實驗

由於學習率對於模型好壞有著一定程度影響，藉由微調學習率來查看資料集變化，本研究將檢視學習率參數 0.003 至 0.03 區間對模型的影響。另外，神經元的數量亦會影響訓練結果，故實驗後續將分別以神經元 20 及 30 搭配學習率的組合進行測試用以觀察不同神經元與學習率組合之變化。至於優化器是用來調整參數使模型訓練達到最佳結果，根據文獻分析結果，本研究選擇三種優化器(Adadelata、RMSprop、Adam)並搭配神經元 20 及 30 的設定，再以不同學習率進行比較分析。

實驗一：優化器 Adadelata 搭配神經元 20 之學習率優化。實驗結果發現，當學習率為 0.007 時可得到最小的 MAPE 值(0.0025)，若從圖 14 可以觀察學習率與模型表現的相關

性，初期隨學習率遞增，模型表現未有顯著變化，但是隨著學習率大於 0.02 時，MAPE 值則快速增加，由此可知在此實驗情境下，較佳的學習率約在 0.01 至 0.02 之間。

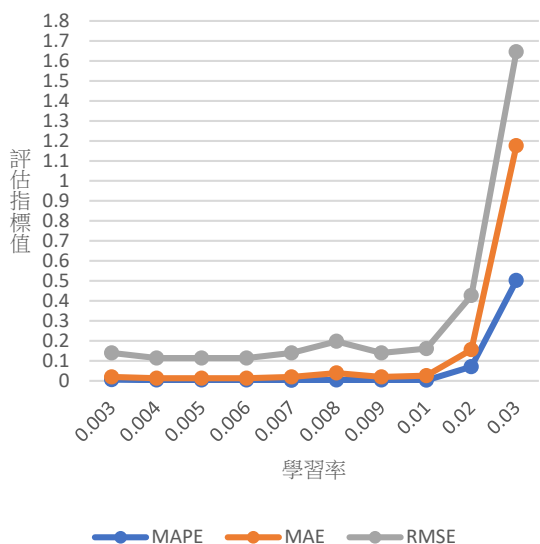


圖 14. 學習率優化(Adadelta+神經元 20)

實驗二：優化器 Adadelta 搭配神經元 30 之學習率優化。實驗結果發現，當學習率高於 0.01 之後，其模型就逐漸表現不佳；另外在學習率低於 0.007 狀況下，其 MAPE 值等三種評估指標則無變化；另外，與神經元 20 相同，在隨著學習率遞減，相關評估指標之績效亦同時遞增，如圖 15。在本研究初始設定之學習率區間內，依 MAPE 值等三種評估指標，可得知當學習率為 0.008 為模型最佳結果，其數值與神經元 20 時，其學習率 0.007 相同。

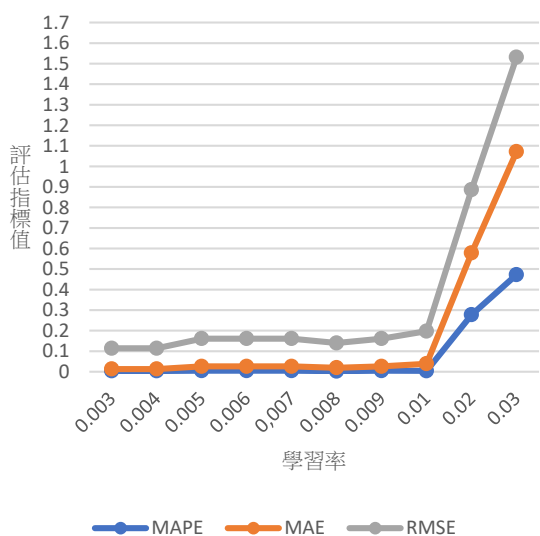


圖 15. 學習率優化(Adadelta+神經元 30)

實驗三：優化器 RMSprop 搭配神經元 20 之學習率優化。實驗結果發現，當學習率為 0.003 及 0.007 時的模型表現較佳，再由圖 16 發現，當學習率變化時，對於模型優化程度較不明顯。

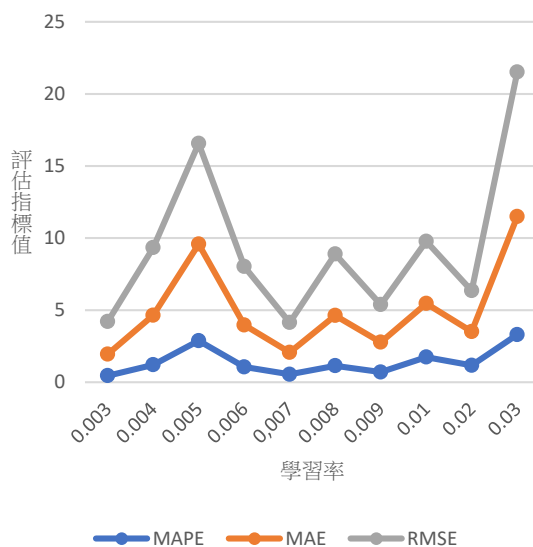


圖 16. 學習率優化(RMSprop+神經元 20)

實驗四：優化器 RMSprop 搭配神經元 30 之學習率優化。實驗結果發現，當學習率為 0.006 時模型較佳，惟與神經元 20 時相同，學習率調整對模型優化並無改善，如圖 17。

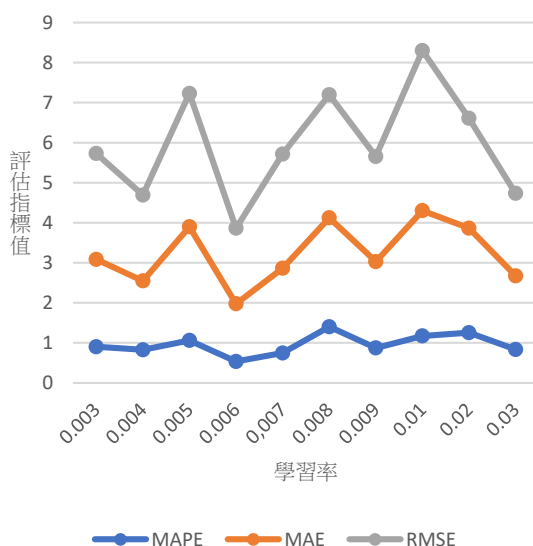


圖 17. 學習率優化(RMSprop+神經元 30)

實驗五：優化器 Adam 搭配神經元 20 之學習率優化。實驗結果發現學習率的調整對模型優化程度並無明顯幫助，如圖 18。

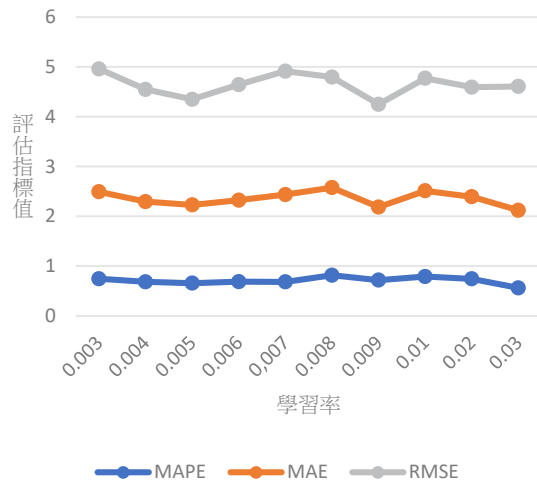


圖 18. 學習率優化(Adam+神經元 20)

實驗五：優化器 Adam 搭配神經元 30 之學習率優化。實驗結果發現學習率的調整對模型優化程度並無明顯幫助，如圖 19。

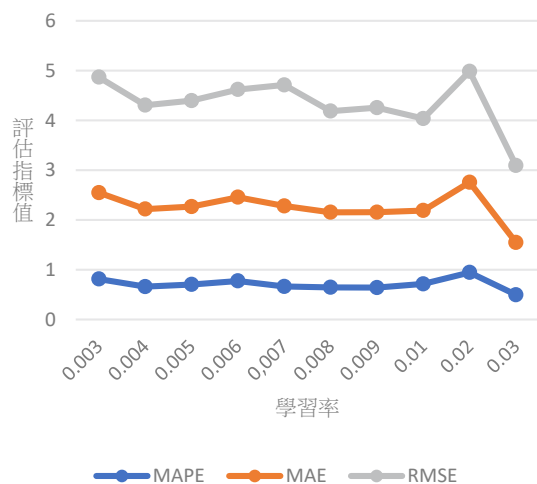
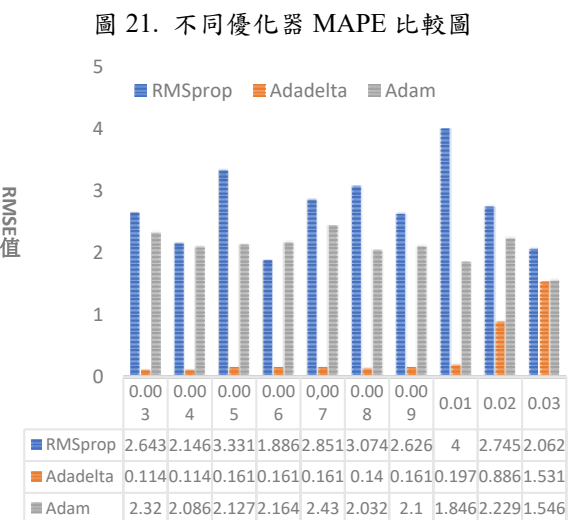
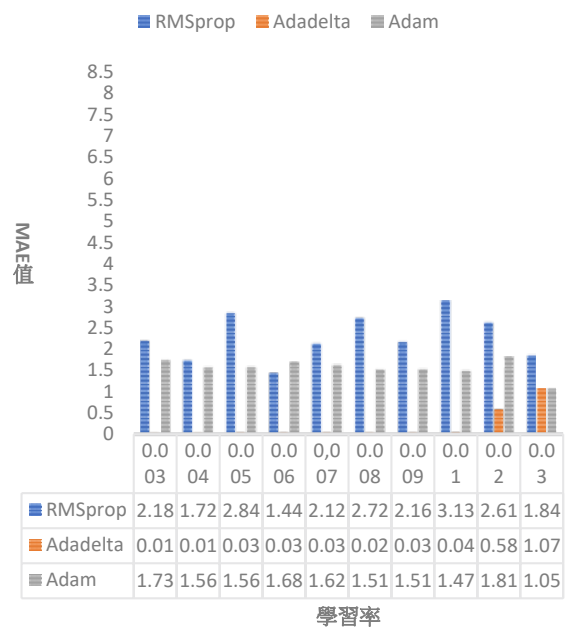
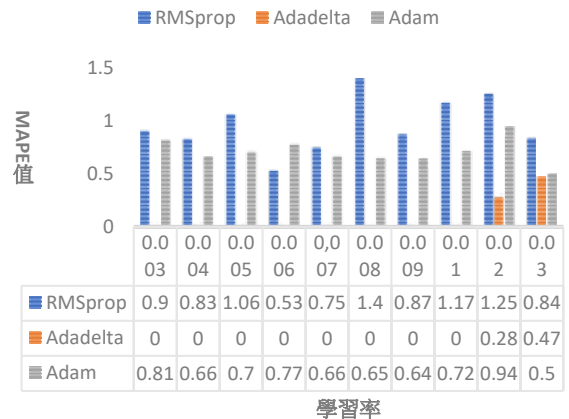


圖 19. 學習率優化(Adam+神經元 30)

優化器 Adam 搭配神經元 20 及 30 的模型設計下，經實驗發現在學習率為 0.03 時，其 MAPE 值有再下降之趨勢，故再以學習率 0.04 及 0.05 實施測試，發現 MAPE 為 0.7942 及 1.1171，依數據顯示並未有更佳之結果。

優化器綜合比較：為了解何種優化器可產生較佳的訓練模型，在神經元 30 時以 10 種學習率值進行統計值的彙整及分析，由 MAPE、MAE 及 RMSE 三種模型評估指標可以明顯發現 Adadelata 優化器對本研究的模型訓練產生較佳的結果，如圖 20、21、22。



從綜合分析中可以瞭解在不同優化器實驗中為了優化及提升預測模型的表現，使用超參數調整來訓練模型，並從中挑選出最佳之超參數。以 MAPE 評估指標(以 MAPE 值為例)觀察，Adadelata 優化器的準確度較佳，在搭配神經元 30 及學習率 0.008 時為模型最佳組合，不同優化器模型表現比較如表 2。

表 2. 不同優化器比較表(以 MAPE 為例)

優化器	MAPE	最佳值
Adam	0.6466	
Adadelata	0.0025	◎
RMSprop	1.4016	

分別藉由優化器、神經元及學習率搭配組合，調教修正模型學習，在各優化器下將各超參數調整至預測最佳組合，將 LSTM 預測模型圖以視覺化方式呈現，發現實際值與預測值的重疊比例呈現高度重疊，如圖 23。

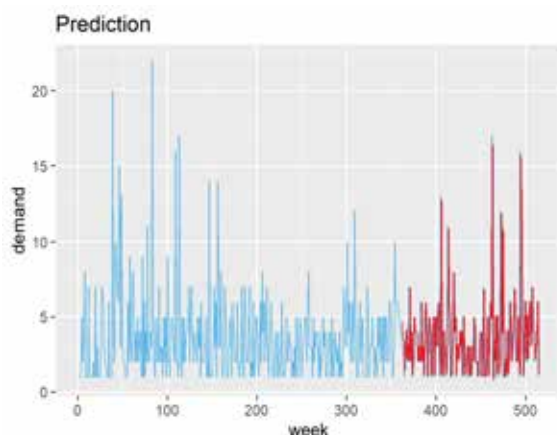


圖 23. LSTM 預測模型圖

最後歸納上述綜合比較的結果來進行 ARIMA 及 LSTM 兩種需求預測模型比較，經研究證實，LSTM 模型的確優於傳統 ARIMA 方法。

表 3. 模型評估指標比較表

模型	MAPE	MAE	RMSE	最佳值
LSTM	0.0025	0.0195	0.1396	◎
ARIMA	0.4420	6.4205	0.5954	

五、結論

本研究探討零附件歷史耗用趨勢，由實驗結果發現運用 LSTM 深度學習演算法建立通用零附件備料之科學預測方法可精準預測未來需求以提供後勤備料決策參考。此外為檢視比較基於不同時間序列需求預測方法：ARIMA 及 LSTM 的差異，以 MAPE、MAE 及 RMSE 等三種評估指標做為模型優劣的依據。各項實驗的比較分析可驗證 LSTM 深度學習演算法優於傳統 ARIMA 時間序列方法。此外為優化 LSTM 模型，在不同超參數(優化器、神經元及學習率)的組合實驗結果發現優化器對模型的影響最大，而神經元及學習率在調校上對模型的影響亦有相當程度影響，對模型在優化後所產出最佳參數組合，可作為零附件需求量的最佳預測模型。

本研究提出導入 LSTM 深度學習模型可提升後勤運籌管理決策，運用歷史資料研析可以精準預測未來零附件需求量，經由模型所產出之需求預測量，可達成改善需求規劃。因此建議在未來公部門船舶後勤管理系統中可運用深度學習演算法為架構，大規模建立零附件需求預測模型可提升後勤備料的效能。

此外囿於研究對象的庫儲系統資料格式不一而影響模型訓練效能，但未來隨著新建立的系統，應可進一步導入用深度學習演算法如卷積神經網路 (CNN)、遞歸神經網路(RNN) 及長短期神經網路 (LSTM) 等，直接整合至系統功能以利於巨量數據的學習與分析，並且在未來研究上可考量更多影響因子，如任務型態、頻次、天候、採購週期等不同變數納入考量，將可以獲得更實務的預測結果，以達精準後勤之目標。

參考文獻

- [1] 黎立珊，“整合型後勤資訊系統執行武器裝備全壽期管理之規劃”，海軍學術雙月刊，第 52 卷，第 2 期，第 68-81 頁，2018。
- [2] Choi, B., & Suh, J. H., “Forecasting Spare Parts Demand of Military Aircraft: Comparisons of Data Mining Techniques and Managerial Features from the Case of South Korea,” Sustainability, Vol. 12, No. 15, pp. 6045-6065, 2020.

- [3] Wild, T., Best practice in inventory management, Routledge, Chap. 2C, pp. 47-59, 2017.
- [4] Boylan, J. E. and Syntetos, A. A., "Spare parts management: a review of forecasting research and extensions," *IMA journal of management mathematics*, Vol. 21, No. 3, pp. 227-237, 2010.
- [5] Aburto, L., and Weber, R., "Improved supply chain management based on hybrid demand forecasts," *Applied Soft Computing*, Vol. 7, No. 1, pp. 136-144, 2007.
- [6] 蘇雄義, 物流與運籌管理-智慧與創新, 第五版, 華泰文化, 第 25-47 頁, 2021。
- [7] 顏憶茹、張淳智 物流管理: 原理'方法與實例: 前程文化, 第 34-45 頁, 2005。
- [8] Willemain, T. R., Smart, C. N., and Schwarz, H. F., "A new approach to forecasting intermittent demand for service parts inventories," *International Journal of forecasting*, Vol. 20, No. 3, pp. 375-387, 2004.
- [9] Mayer-Schönberger, V. and Cukier, K., Big data: A revolution that will transform how we live, work, and think, Houghton Mifflin Harcourt, Chap. 1, pp. 1-20, 2013.
- [10] Hua, Z. and Zhang, B., "A hybrid support vector machines and logistic regression approach for forecasting intermittent demand of spare parts," *Applied Mathematics and Computation*, Vol. 181, No. 2, pp. 1035-1048, 2006.
- [11] Lindsey, M. and Pavur, R., "Prediction intervals for future demand of existing products with an observed demand of zero," *International Journal of Production Economics*, Vol. 119, No. 1, pp. 75-89, 2009.
- [12] Pawar, N. and Tiple, B., "Analysis on Machine Learning Algorithms and Neural Networks for Demand Forecasting of Anti-Aircraft Missile Spare Parts," *Proc. the 2019 International Conference on Communication and Electronics Systems*, 2019.
- [13] Chen, F.-L. and Chen, Y.-C., "An investigation of forecasting critical spare parts requirement," *Proc. the 2009 WRI world congress on computer science and information engineering*, Los Angeles, USA, pp. 225-230, 2009.
- [14] 余桂霖, 時間序列分析, 台灣五南圖書, 第 12-20 頁, 2013。
- [15] Fan, C., Xiao, F. and Wang, S., "Development of prediction models for next-day building energy consumption and peak power demand using data mining techniques," *Applied Energy*, Vol. 127, pp. 1-10, 2014.
- [16] Colladon, A. F., Guardabascio, B., and Innarella, R., "Using social network and semantic analysis to analyze online travel forums and forecast tourism demand," *Decision Support Systems*, Vol. 123, pp. 113075, 2019.
- [17] Li, W., Cao, J., Guan, J., Zhou, S., Liang, G., So, W. K., and Szczecinski, M., "A general framework for unmet demand prediction in on-demand transport services," *IEEE Transactions on Intelligent Transportation Systems*, Vol. 20, No. 8, pp. 2820-2830, 2018.
- [18] Cao, S., Hou, S., Yu, L., and Lu, J., "Predictive control based on occupant behavior prediction for domestic hot water system using data mining algorithm," *Energy Science & Engineering*, Vol. 7, No. 4, pp. 1214-1232, 2019.
- [19] Van Nguyen, T., Zhou, L., Chong, A. Y. L., Li, B., and Pu, X., "Predicting customer demand for remanufactured products: A data-mining approach," *European Journal of Operational Research*, Vol. 281, No. 3, pp. 543-558, 2020.
- [20] Sathishkumar, V., Park, J., and Cho, Y., "Using data mining techniques for bike sharing demand prediction in metropolitan city," *Computer Communications*, Vol. 153, pp. 353-366, 2020.
- [21] Papanagnou, C. I. and Matthews-Amune, O., "Coping with demand volatility in retail pharmacies with the aid of big data exploration," *Computers & Operations Research*, Vol. 98, pp. 343-354, 2018.
- [22] Jiang, S., Chin, K.-S., Wang, L., Qu, G., and Tsui, K. L., "Modified genetic algorithm-based feature selection combined with pre-trained deep neural network for demand forecasting in outpatient department," *Expert systems with applications*, Vol. 82, pp. 216-230, 2017.
- [23] Dombi, J., Jónás, T., and Tóth, Z. E., "Modeling and long-term forecasting demand in spare parts logistics businesses," *International Journal of Production Economics*, Vol. 201, pp. 1-17, 2018.

- [24] Wang, C.-C., Chien, C.-H., and Trappey, A. J. C., "On the Application of ARIMA and LSTM to Predict Order Demand Based on Short Lead Time and On-Time Delivery Requirements," *Processes*, Vol. 9, No. 7, pp. 1157, 2021.
- [25] Dali, H. and Chengcheng, L., "Demand forecast of equipment spare parts based on EEMD-LSTM," *Proc. the 2021 6th International Conference on Intelligent Informatics and Biomedical Sciences*, Oita, Japan, pp. 230-234, 2021.
- [26] Abbasimehr, H., Shabani, M., and Yousefi, M., "An optimized model using LSTM network for demand forecasting," *Computers & industrial engineering*, Vol. 143, pp. 106435, 2020.
- [27] Siami-Namini, S., Tavakoli, N., and Namin, A. S., "A comparison of ARIMA and LSTM in forecasting time series," *Proc. IEEE International Conference on Machine Learning and Applications (ICMLA)*, Florida, USA, pp. 1394-1401, 2018.
- [28] LeCun, Y., Bengio, Y., and Hinton, G., "Deep learning," *Nature*, Vol. 521, No. 7553, pp. 436-444, 2015.
- [29] Hochreiter, S. and Schmidhuber, J., "Long short-term memory," *Neural computation*, Vol. 9, No. 8, pp. 1735-1780, 1997.
- [30] Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., and Schmidhuber, J., "LSTM: A search space odyssey," *IEEE transactions on neural networks and learning systems*, Vol. 28, No. 10, pp. 2222-2232, 2016.
- [31] Ribeiro, A. M. N., do Carmo, P. R. X., Rodrigues, I. R., Sadok, D., Lynn, T., and Endo, P. T., "Short-Term Firm-Level Energy-Consumption Forecasting for Energy-Intensive Manufacturing: A Comparison of Machine Learning and Deep Learning Models," *Algorithms*, Vol. 13, No. 11, pp. 274, 2020.
- [32] Day, M.-Y. and Lin, Y.-D., "Deep learning for sentiment analysis on google play consumer review," *Proc. IEEE international conference on information reuse and integration (IRI)*, San Diego, USA, pp. 382-388, 2017.
- [33] Kobayashi, T., "SCW-SGD: Stochastically Confidence-Weighted SGD," *Proc. IEEE International Conference on Image Processing (ICIP)*, Abu Dhabi, United Arab Emirates, pp. 1746-1750, 2020.
- [34] Zhang, Z., "Improved adam optimizer for deep neural networks," *Proc. IEEE/ACM 26th International Symposium on Quality of Service (IWQoS)*, Banff, Canada, pp. 1-2, 2018.
- [35] Tieleman, T. and Hinton, G., "Lecture 6.5-rmsprop: Divide the Gradient by a Running Average of Its Recent Magnitude," *COURSERA: Neural Networks for Machine Learning*, Vol.4, pp. 26-31, 2012.
- [36] Wu, H. and Gu, X., "Towards dropout training for convolutional neural networks," *Neural Networks*, Vol. 71, pp. 1-10, 2015.
- [37] Li, Z., Gong, B., and Yang, T., "Improved dropout for shallow and deep learning," *Advances in neural information processing systems*, Vol. 29, pp. 2523-2531, 2016.
- [38] Ren, G., Cao, Y., Wen, S., Huang, T., and Zeng, Z., "A modified Elman neural network with a new learning rate scheme," *Neurocomputing*, Vol. 286, pp. 11-18, 2018.
- [39] Wang, S., Wang, X., Wang, S., and Wang, D., "Bi-directional long short-term memory method based on attention mechanism and rolling update for short-term load forecasting," *International Journal of Electrical Power & Energy Systems*, Vol. 109, pp. 470-479, 2019.
- [40] Lu, X., Ma, C., and Qiao, Y., "Short-term demand forecasting for online car-hailing using ConvLSTM networks," *Physica A: Statistical Mechanics and its Applications*, Vol. 570, pp. 125838, 2021.