Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2024.Doi Number

Forecasting of Dementia Patient Numbers by Integrating Support Vector Regression with the Fly Goose Optimization Algorithm

結合雁行最佳化演算法與支援向量迴歸之失智症患者人數預測

Forecasting of Dementia Patient Numbers Based on FGOA–SVR Parameter Optimization

基於 FGOA - SVR 參數優化於失智症患者人數預測

Support Vector Regression Optimized by Flying Geese Optimization Algorithm for Number of patients with Dementia forecasting 基於雁行最佳化演算法的支援向量迴歸用於失智症患者人數預測

Parameter Optimization of Support Vector Regression Using Flying Geese Optimization Algorithm

使用雁行最佳化演算法進行支援向量迴歸的參數最佳化

PO-HUNG CHEN[1], CHENG-HONG YANG[1,2,3,4], INDRA SATRIAWAN SAKTI[1]
[1] Department of Electronic Engineering, National Kaohsiung University of Science and Technology, Kaohsiung, Taiwan

[2] Department of Information Management, Tainan University of Technology, Tainan, Taiwan
[3]Ph. D. Program in Biomedical Engineering, Kaohsiung Medical University, Kaohsiung, Taiwan
[4]Drug Development and Value Creation Research Center, Kaohsiung Medical University, Kaohsiung, Taiwan
[5]Department of Plastic Surgery, Chia-Yi Christian Hospital, Chia-Yi, Taiwan
[6]Department of Chemical Engineering & Institute of Biotechnology and Chemical Engineering, I-Shou University, Kaohsiung, Taiwan

Corresponding author: Cheng-Hong Yang (chyang@nkust.edu.tw).

ABSTRACT 在全球人口結構迅速高齡化的浪潮下，全球失智症患者人數正以驚人的速度攀升。失智症演變為衝擊全球公共衛生、社會經濟與人類福祉的重大挑戰。失智症的影響認知功能（記憶力、判斷力、理解力、語言能力等）持續退化，最終喪失生活自理能力。可能出現各種精神行為問題，嚴重影響生活品質。本研究提出，有效導入一個最佳化機器學習 Flying Geese Optimization Algoritm Support Vector Regression(FGOASVR)模型來預測失智症患者人數的變化，可提早制定建構預防與風險管理的公共衛生策略。本研究利用台灣衛生福利部提供之全民健康保險資料庫中 1998 年至 2023 年之失智症患者診斷人數年度資料進行實證分析。為了進一步確認 FGOASVR 模型之有效性，本研究亦與三類模型進行比較，包含：統計預測模型：自我迴歸整合移動平均（ARIMA）與 Holt-Winters 指數平滑法（HWETS）；（二）深度學習模型：Long Short-Term Memory（LSTM）；（三）混合式模型：支援向量迴歸（SVR）、Particle Swarm Optimization Support Vector Regression（PSOSVR）、Differential Evolution Support Vector Regression (DE SVR)、Whale Optimization Algorithm Support Vector Regression (WOASVR)與 Harris Hawk Optimization Support Vector Regression (HHOSVR)。模型效能評估則分別透過平均絕對百分比誤差（Mean Absolute Percentage Error, MAPE）、均方根誤差（Root Mean Square Error, RMSE）等常見預測評估指標進行衡量。FGOASVR 的表現明顯優於其他所有模型，男性平均 MAPE 為 3.17，RMSE 則為 0.69，女性平均 MAPE 為 3.42，RMSE 則為 0.96。本研究所提出之 FGOASVR 模型在預測台灣失智症患者數量的年際變化方面表現出更佳之準確度與穩定性，證實其在公共衛生議題之預測應用中具備較優的效能與發展潛力。

INDEX TERMS Dementia forecasting, Flying Geese Optimization Algorithm, Support Vector Regression.

I. INTRODUCTION

The literature indicates that dementia is an acquired, chronic, and progressive cognitive dysfunction in multiple domains, including memory, language, visuospatial, and executive function [3]. Symptoms such as severe memory impairment, disorientation and confusion, mood instability, and behavioral and psychological changes (e.g., hallucinations and delusions) develop as the disease progresses [4]. These symptoms may contribute to the daily social or professional dysfunction of patients, causing emotional exhaustion among caregivers and increasing the socioeconomic burden [5].

Past studies indicated that dementia has a substantial effect on the social care system and societal costs [6]. As an acquired disabling syndrome, characterized by a progressive deterioration in multiple cognitive domains that interferes with daily functioning, several conditions cause dementia symptoms, including Alzheimer's disease, vascular disorders, and Parkinson's disease. The prevalence of dementia among older adults increases as a population ages, significantly affecting the lives of an increasingly large number of older adults globally. Because dementia is the main cause of hospitalization among older adults, the increasing prevalence of dementia places tremendous pressure on the health insurance system. The number of people diagnosed as having dementia worldwide is growing with the increase in the global average living age [7]. The World Alzheimer's Disease Report indicated that the number of people with dementia worldwide exceeded 50 million in 2019 and is expected to increase to 15.2 billion by 2050; moreover, dementia currently incurs a cost of US$1 trillion per year, and this figure is expected to double by 2030 [7]. Based on the rapid increase in dementia patients and financial demands, dementia is a global problem that requires urgent attention. A time series approach could provide a thorough and accurate understanding of the prevalence rate, which will assist in actions taken to address this disease.

A significant portion of the societal cost of dementia is patient care. Studies reported that the average total costs for the last 5 years of life of patients with dementia are higher than those of patients with heart disease or cancer, among other causes of mortality [8]. In 2015, the overall cost of dementia was approximately US $818 billion; 40.4% of this cost was attributed to caregivers. Dementia is often associated with disorientation, confusion, mood instability, and behavioral psychological symptoms; care is thus demanding. Care for patients with dementia is generally more time consuming than care for patients with other diseases [9]. The informal caregivers of patients with dementia often develop depression, anxiety, and physical symptoms and even have a relatively high mortality rate [5,10,11,12]. Therefore, the care of patients with dementia is one of the major sources of socioeconomic burden that should be emphasized in policies on expanding the medical allowance for this population. Suitable social welfare and public health policies necessitate a precise model for predicting the prevalence rate of dementia.

時間序列分析是統計學中歷史悠久且基礎性的研究領域，旨在探索隨著時間推移收集到多變量數據；然而，時間序列數據經常受到結構性突變和隨機擾動的影響，使其呈現出趨勢、周期性、季節性以及非系統性成分高度交織的複雜狀態[1]。為了應對這種高度變異和非平穩的特性，已經提出了許多方法，如自迴歸整體移動平均法(ARIMA)、Holt-Winters 指數平滑法 (HWETS)等。雖然這兩種方法已被應用在不同領域，但仍有其限制。ARIMA 在時間序列分析中被廣泛採用，但其基於線性假設的架構，先天無法非線性異質變異數、閾值效應與相依結構，使模型在面對複雜動態時預測性能明顯不足，這限制了 ARIMA [2]。另一

方面，HWETS 僅容許單一、固定長度的季節循環，難以處理週期及趨勢等資料，且其誤差方差被假設為恆定；趨勢與季節交互作用會擴散誤差，增加預測風險，這限制了 HWETS [3]。

隨著人工智慧技術發展，以機器學習和深度學習演算法為代表，如支援向量迴歸（SVR）及長短期記憶（LSTM）等，演算架構與訓練策略的漸進式優化，其整體預測效能已呈現階段性躍升[4]。其中，SVR 對高維特徵空間之映射能力與對非線性結構的精確擬合，時間序列分析領域中處理複雜動態樣態的核心方法；因此，SVR 在非線性序列預測上的表現顯著優於傳統統計模型所能達成之準確度[5]。Maragatham 與 Devi （2019）藉由導入 LSTM 神經網路門控機制，構建出一套可追蹤序列依賴並預測心理韌性衰竭的模型，驗證了 LSTM 在非線性心理生理訊號解析中的優勢[6]。此外，Wang 等人 （2019）結合 LSTM 與縱向電子健康紀錄，建構死亡風險預測框架，從而識別最可能受惠於安寧療護之失智症患者，LSTM 模型於臨床照護路徑規劃與資源分配上的實證價值 [7]。

深度學習的長短期記憶網路（LSTM）於參數量龐大且結構複雜[4]，機器學習支援向量迴歸（SVR）等演算法在需求預測領域兼具精簡的模型參數與優越的非線性函數逼近能力，因而能夠在維持高度預測精確度的同時，減少過度擬合風險並提高泛化效率[8]。然而，未能精確調校關鍵超參數，支持向量迴歸（SVR）之泛化能力與預測效能將顯著衰減。已研究實證，發展兼具全域搜尋與高效率之參數優化策略，以鑑別最適超參數組合，確保 SVR 模型性能之先決條件 [9]。因此，諸多混合模型因應而生，以解決此最佳化難題，如 PSOSVR、DESVR、WOASVR 與 HHOSVR 等[20,22]。

本研究提出 Flying Geese Optimization Algoritm（FGOA）以解決 SVR 之參數優化策略，以鑑別最適超參數組合。該演算法藉由嵌入嶄新的機制，有效避免陷入區域最佳解，因而特別適用於高維特徵空間之最佳化問題。透過結合生物啟發式優化演算法，FGOA 在高維資料集中有效識別關鍵特徵、提升穩定性，並尋得更卓越之特徵子集合。FGOA 在大規模資料中展現出迅速萃取關鍵特徵的能力，能夠全面探索搜尋空間並發掘具判別力之最佳超參數組合，為後續模型開發奠定可信之基礎。FGOASVR 建構失智症患者人數預測系統；此方法不僅提升了 SVR 參數優化，亦增強了預測模型之穩定性，於時間序列預測等領域達成突破性進展。

II. LITERATURE REVIEW

時間序列分析已廣泛應用於疾病患者人數的預測研究。W.-C. Juang 等人使用 ARIMA 模型來預測南台灣某醫學中心急診室的就診人數，建構適合的統計模型以協助資源調度與避免擁擠。研究 2009 到 2016 年的月度資料，嘗試不同的 ARIMA 組合，最終選定 ARIMA(0,0,1)為最佳模型，其平均絕對百分比誤差（MAPE）為 8.91%，說明預測準確度頗高。研究強調未來 ARIMA 提供了一個有效的基準模型，有助於醫療決策與急診運營規劃[10]。HWETS 方法也已成功應用於疾病患者人數預測。該研究，Holt-Winters 模型用來預測重慶南岸區 2017 年至 2022 年期間 foodborne disease 每月的患者人數，並與 SARIMA 與 ETS 模型進行比較。指出，研究在較短時間序列模型中，Holt-Winters 模型表現出了卓越的預測結果，此模型的均方誤差（MSE）、平均絕對誤差（MAE）與均方根誤差（RMSE）分為 8.78、2.33 與 2.96，皆低於 SARIMA 與 ETS 模型。雖然平均絕對百分比誤差（MAPE）略高，為 39.29%，但整體預測值曲線更接近實際觀察值。Holt-WintersETS 模型持續表現優異，針對趨勢項與季節項的平滑處理，能夠動態地調整時間序列中不同時點的資料權重，從而增強強化模型對具有週期性疾病資料的預測能力[11]。

機器學習相較於傳統統計模型，可有效提升非線性預測的預測能力[12]。該研究成功使用支援向量回歸（SVR）模型，預測印度COVID-19疫情中，每日及累積病例、死亡與康復人數。資料涵蓋2020年3月1日至4月30日的61天，並使用徑向基底函數（RBF）作為核函數。SVR模型在累積病例與死亡數的預測上表現極佳，準確率超過97%，而在每日新增病例的預測上則稍遜，準確率為87%，受到資料中不規則波動造成資料非平穩性，進而降低模型的預測能力。SVR模型展現出高於線性與多項式回歸的準確度與穩定性，在面對變異性高的疫情資料時，其預測結果具備實用價值，有助於政策制定與防疫決策[13]。

在過去的時間序列預測研究中，支持向量回歸（SVR）被廣泛應用於處理非線性預測問題，然而其預測效果高度依賴於三個超參數（C、ϵ 與σ）的設定。為克服這一限制，引入了基於群體智能的優化方法以提升模型性能。根據 Yang et al.的研究結果[14]，PSOSVR 顯著優於傳統 SVR 模型。在男性失智症患者預測中，SVR 的平均絕對百分比誤差（MAPE）高達 17.53%，而 PSOSVR 僅為 6.01%，誤差減少幅度高達 65.7%。在女

性部分，SVR 的 MAPE 為 12.67%，PSOSVR 為 6.78%，誤差同樣降低了 46.5%。這顯示 PSOSVR 在不同性別樣本中皆能有效減少預測誤差，並提升模型穩定性與泛化能力[14]。Hamdi 等人提出一種DESVR方法[15]，將支援向量迴歸（SVR）與差分演化演算法（DE）結合，藉由DE自動優化SVR模型中的超參數（C、ε、γ），進而在不需人工干預或外部資料的情況下完成血糖預測。該研究在使用CGM資料的條件下，於多個預測時間（Prediction Horizon, PH）內皆展現出穩定且優異的預測表現。相較於使用其他超參數優化方法如遺傳演算法（GA）與粒子群優化（PSO），DE所導出的SVR模型在MAPE與RMSE兩個指標上皆明顯優越，特別是對於15分鐘PH，其平均絕對百分比誤差（Mean Absolute Percentage Error, MAPE）達僅3.74%，顯示MAPE值遠低於多數同類研究的報告值，並且在延長PH至30、45與60分鐘後，預測誤差的增加幅度仍可控制於合理範圍，突顯模型的時間勢，參數優化在提升傳統機器學習模型準確性方面的重要性[5]。該研究HHOSVR來預測每日河川流量，並與其他五種基於元啟發式演算法的SVR混合模型進行比較。研究地點為印度北阿坎德邦的Naula集水區，資料包含2000年至2004年雨季期間的日降雨與河川流量。結果顯示，HHO SVR模型在校正與驗證期間的預測表現均優於其他SVR混合模型，其在誤差指標上取得最小值，相關性與精確度指標亦表現最佳。圖表視覺化也支持這一點，在預測高流量期間尤其明顯。此結果表明，HHO在捕捉非線性與非平穩性的水文資料方面具有優勢，展現其應用於水資源管理的潛力[16]。

III. METHOD

本研究資料來源取自臺灣衛生福利部之健保資料庫，包含 60 歲以上男女失智症患者之年度人數，建立預測失智症病患人數的模型。在資料的預處理過程中，我們會檢查並移除異常的資料之後，將資料分為訓練集和測試集，兩者的比例為 80:20，並使用10倍交叉驗證。時間序列方法用於訓練和調整超參數最佳化獲得最佳模型，以最小化誤差並提

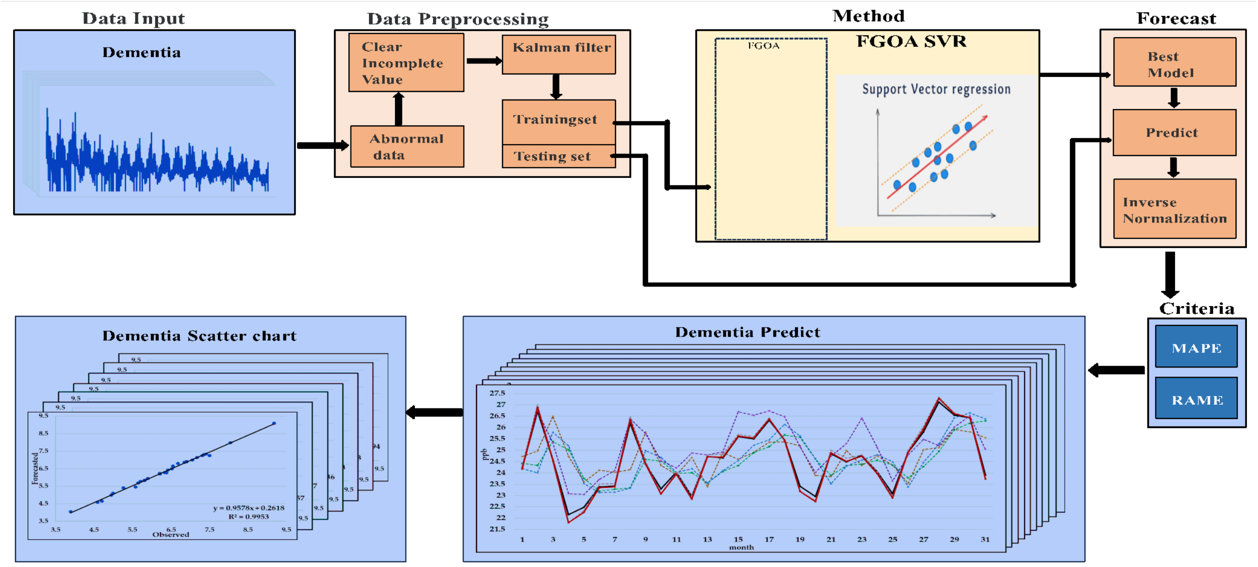

FIGURE 1. System Architecture Diagram for forecasting the Number of Dementia Patients.

穩定性與實用性[15]。Whale Optimization Algorithm (WOA)以座頭鯨的群體狩獵行為為靈感，透過螺旋搜尋與全域最優解探索，有效避免傳統參數調整容易陷入區域最小值的困境。根據 Yang 等人的實證研究[5]，WOASVR 在預測台灣1991至2020年之月電力發電數據時，顯著優於 SVR 模型與統計型預測方法如 ARIMA、ETS 及 HWETS。在風力發電的預測上，WOASVR 的平均絕對百分比誤差（MAPE）為37.28，遠低於 SVR 的 69.9，在太陽能與常規水力發電的預測中亦表現出相對穩定與準確的結果，整體平均 MAPE 為 21.82%，優於 SVR 的 37.55%。在處理波動性強的再生能源數據上具有明顯優昇男女失智症患者預測的準確度。預測失智症病患人數的系統架構如圖 1 所示。

A. TIME SERIES METHOD

1) AUTOREGRESSIVE INTEGRATED MOVING AVERAGE

ARIMA 模型，由 Box 與 Jenkins 於 1976 年所提出，稱為 Box–Jenkins 模型[17]。模型核心在於三種統計特徵以捕捉資料的動態結構，自我回歸（AR）用以描述當前值與其過去觀察值之間的線性依賴關係為 p；差分（I）則為處理時間序列的非平穩性所設計，透過將資料進行一階或多階差分轉換，移除趨勢成分以達致平穩性為 d；而移動平

均（MA）則藉由建模當前值與歷史誤差項（即殘差）之間的關聯，以反映隨機干擾對系統輸出的潛在影響為 q [18]。ARIMA 模型的數學公式如下：

$$\left(1-\sum_{i=1}^{p}\varphi_i L^i\right)(1-L)^D x_t=\left(1+\sum_{i=1}^{q}\theta_i L^i\right)\varepsilon_t \quad (1)$$

L 被用作標記延遲運算子的引入，簡化表示時間序列方程式的複雜度。φ_i 所代表的是自我回歸部分的係數。θ_i 則對應於移動平均成分的參數。ε_t，則代表誤差項，並為模型評估提供了關鍵指標，用以檢驗殘差是否滿足隨機性假設。

2) HOLT WINTERS EXPONENTIAL SMOOTHING

Holt-Winters 指數平滑模型（HWETS），由 Charles Holt 與 Peter Winters 所提出，時間序列預測領域中具代表性之經典方法[19]。其預測機制建立於對時間序列中三項結構性成分，基準水準（level）、趨勢（trend）與季節性（seasonality），進行加權指數平滑。HWETS 模型尤以能同時考量長期趨勢變化與短期週期性震盪的特性見長，使其對於含有明顯季節性循環之資料表現出高度適配性。HWETS 具備建模趨勢與季節成分的雙重能力，因此能有效擷取時間序列潛藏的季節性結構變化，進而提供穩健且具高度準確性的未來數值推估，廣泛應用於流行病監測、銷售預測與公共衛生領域的時序性資料分析[20]。

3) SUPPORT VECTOR REGRESSION

支持向量迴歸（Support Vector Regression，SVR）模型由 Vapnik 等學者於 1997 年正式提出，作為支持向量機（Support Vector Machine, SVM）理論在監督式學習中應用於迴歸分析領域的一項重要擴展[21]。SVR 具有同時處理線性與非線性迴歸問題的能力，特別適用於面對高維度與資料結構複雜性的預測任務[22]。SVR 的核心思想是在經過核技巧處理後所建構的高維特徵空間中，進行線性迴歸分析。透過將原始輸入資料映射至此特徵空間，模型得以在高維空間中建立一條能夠精準擬合資料點並具備預測能力的線性超平面[23]。強化模型對未知資料的泛化能力，SVR 引入了一項關鍵控制參數為正則化係數 C，其目的是在最小化訓練誤差與控制模型複雜度之間取得平衡，防止過度擬合現象[24]。ε 代表對預測值與真實值之誤差的容忍程度；只有當預測誤差超出 ε 時，模型才會對該誤差進行懲罰性處理，有助於忽略微小且非關鍵的偏差，從而提高模型的魯棒性[25]。在處理非線性資料時，SVR 由不同類型的核函數 如線性核 多項式核及高斯徑向基（RBF）核等，進行隱式映射，以捕捉輸入特徵之間潛在而複雜的非線性關係[23]。SVR 模型的數學公式如下：

$$f(x)=\omega^T\varphi(x)+b \quad (6)$$

$$R=C\sum_{i=1}^{N}(\xi_i+\xi_i^*)+\frac{1}{2}\|\omega\|^2 \quad (7)$$

Equation (6) yields the predicted result $f(x)$, which is obtained by taking the inner product between the weight vector ω and the feature vector $\varphi(x)$ and adding the bias term b. Equation (7), a regularization term is introduced to control the model's complexity; C is the regularization constant, and ξ_i and ξ_i^* are slack variables used to quantify the degree of difference between the data and the ε tube. The regularization term is used to minimize prediction errors while controlling the complexity of the model to prevent overfitting.

B. OPTIMIZATION ALGORITHM

1) PARTICLE SWARM OPTIMIZATION (PSO)

The PSO proposed by Kennedy and Eberhart in 1995, is a population-based stochastic optimization algorithm [26]. 演算法的起始階段為粒子群的初始化。首先，建立一個粒子的群體數量，每個粒子均代表一組候選的 SVR 超參數解(C,γ,ε)。這些粒子的初始位置是在預先定義的搜索範圍內隨機生成。同時，為每個粒子賦予一個隨機的初始速度，其大小通常被限制在一個較小的範圍內，以確保搜索過程的穩定性。接下來是適應度評估階段。為了量化每組超參數的性能，定義了一個適應度函數（Fitness Function）。在此框架下，適應度函數為交叉驗證策略計算出的 Mean Absolute Percentage Error (MAPE)。對於群體中的每一個粒子，利用其對應的（C,γ,ε）參數組合來訓練 SVR 模型，並透過交叉驗證計算其 MAPE 值作為適應度分數。在個粒子評估後，更新每個粒子自身的歷史最佳位置（pbest）以及整個群體迄今為止發現的全域最佳位置（gbest）。此後，演算法進入核心的迭代尋優過程。在每一次迭代中，每個粒子的速度和位置都會根據其 pbest 和全域 gbest 進行更新。此更新機制引導粒子向搜索空間中更有潛力的區域移動，從而平衡了個體探索（Exploration）與群體利用（Exploitation）。此迭代過程將持續進行，直到滿足預設的最大迭代次數停止條件。

2) DIFFERENTIAL EVOLUTION (DE)

The DE algorithm, proposed by Storn and Price in 1997, is a population-based stochastic optimization technique that incorporates core operations such as mutation, crossover, and selection [27]. DE 的運作過程始於初始化一個包含 NP 個體的群體，這些個體的參數值在預定義範圍內隨機生成。這些範圍通常基於對數尺度，以涵蓋參數的廣泛變化範圍並避免數值溢出。隨後，DE 通過突變、交叉和選擇三個核心操作進行迭代優化。突變操作通過隨機選擇三個不同個體，並利用差異向量生成突變向量，其中縮放因子 F （通

常取 0.5 到 1.0）控制差異向量的大小，從而影響探索的步長。交叉操作將突變向量與目標向量結合，生成試驗向量 ui ，通過交叉概率 CR （通常取 0.8 到 1.0）決定哪些維度從突變向量繼承，確保新解的多樣性。選擇操作則比較試驗向量和目標向量的適應值，保留適應值更優的個體進入下一代。適應值通常定義為 SVR 模型在驗證集上的性能指標，例如 Mean Absolute Percentage Error (MAPE）。為了提高泛化能力，適應值計算通常採用 k 折交叉驗證，通過平均多折的 MAPE 來評估參數組合的穩健性。DE 迭代這些步驟，直到滿足終止條件，例如達到最大迭代次數或適應值變化小於某閾值，最終輸出最優參數組合。

3) WHALE OPTIMIZATION ALGORITHM (WOA)

The WOA introduced by Mirjalili and Lewis in 2016, is a metaheuristic optimization algorithm inspired by the foraging behavior of humpback whales [28]．在初始化階段，我們先於參數空間內隨機生成多個鯨魚個體，每一個個體代表一組 (C,γ,ε)值。這些值通常會先經過對數尺度的映射，以因應正則化參數與核係數在數量級上的巨大差異。在每一次迭代中，首先評估所有鯨魚個體在驗證集上的 SVR 預測誤差（如 MAPE），並將誤差最小者標記為「領頭鯨魚」（即目前的最佳解）。在「包圍捕獵」階段，其餘鯨魚會朝向領頭鯨魚的方向進行位置更新，更新幅度由一組隨機係數與迭代遞減參數所決定，其意義在於逐步收斂至領頭解; 該機制有如真實鯨魚利用氣泡環逐漸靠近獵物的策略.當迭代進入「氣泡網捕食」模式時，每頭鯨魚不再採用線性向量遷移，而是沿著一條對數螺旋曲線圍繞領頭鯨魚進行跳動，這個過程可以幫助它們跳出局部區域，探索更多可能的解空間。螺旋模型的引入不僅加強了探索能力，也避免整體早熟收斂。隨著迭代次數的增加，隨機探索強度逐漸降低，而粒子之間的競爭與協同則導致整體解集中於誤差最低的區域。最終，當演算法達到預定的最大迭代次或驗證誤差的改進幅度趨近於零時，整個過程結束，並回傳領頭鯨魚所代表的最優參數值組合。

4) HARRIS HAWK OPTIMIZATION (HHO)

The HHO algorithm proposed by Heidari in 2019, is inspired by the cooperative hunting behavior of Harris hawks [29]. 在算法初始階段，HHO 生成一個由多隻鷹組成的種群，每隻鷹的三維位置對應一組（C, γ, ε）。這些初始位置通常均勻分布在預先確定的參數範圍內。每當鷹群都需先在整個種群上評估 MAPE。最小的 MAPE 為當前“獵物”位置，代表迄今最優超參數組合。之後進入圍捕階段，HHO 根據“獵物”的逃逸能量決定使用不同的策略。當“獵物”逃逸能量較大時，算法側重全局探索，類似鷹群在廣闊空域中環視獵物，此時各個候選解會以較大步幅在參數空間內隨機跳躍，規避早期收斂到局部最優。反之，當逃逸能量衰減到一定水平，HHO 轉入局部開採，對參數值進行精細調整：部分鷹會在獵物周圍做出“軟包圍”動作，以較小的步幅逼近最優解, 而其他鷹則可能採用“硬包圍”策略, 快速朝獵物逼近以鎖定最優點。在這些過程中，能量指標和隨機性共同作用，確保算法既能深入搜索最有希望的區域，又能保留足夠的多樣性以避免早期收斂。隨著迭代次數不斷增長，“獵物”的逃逸能量逐步降低，鷹群的行動也從粗略搜索轉為更聚焦的微調。此時，參數更新速度變得緩慢且精確，每一次位置更新都意味著在 C、γ、ε 空間中對候選解進行最後的收斂探尋。最終，當算法滿足停止條件（如最大迭代或適應度無明顯改進），所獲得的全局最優“獵物”位置即為最終選定的 SVR 超參數組合，其對應的 MAPE 最小，代表在驗證數據上達到最佳預測精度。

5) FLYING GEESE OPTIMIZATION ALGORITM (FGOA)

The algorithm of the proposed Flying Geese Optimization algorithm (FGOA) is explained in this section

5.1.1 Geese : social behavior

Geese are renowned for their loyalty and strong sense of community. Geese typically ,mate for life, forming strong pair bonds that aid in raising their young and defending nesting territories, this life long partnership provides social advantages,such as higher status and better access to resources(Horton 2024).

5.1.2 dynamic behavior

Perhaps you have seen the distinctive V-formation of a flock of Canada Geese while migrating (Figure 2). This technique helps them reduce wind resistance, making their journey more efficient and energy-saving. Migratory animals must optimise their energetic and cognitive performances during the yearly long-distance travels . Brent geese migrate from northeast Canada to Ireland, a non-stop journey of over 3400 miles(5.500km). Canada geese fly from the northernmost regions of Canada and the artic circle into the USA, a journey of some 1500 miles(2400km). if they find a strong tailwind, they can fly some 1000 to 1500 miles in just oen day (birdfact 2023). Canada Geese migrate in large flocks. Perhaps you have seen a distinct V-formation of a flock of Canada Geese. Few things are as unique as the remarkably intelligent system the Canada geese use for migrating. To reduce wind resistance the strongest and freshest birds lead the group at the head of the V. when they get tired, they drop back and the next ones take their place. This system can allow for over fifteen thousand miles covered in twenty-four hours. Canada geese have established paths for travel and the flock will have a set route with a designated rest stop for food and water. Ducks are outgoing and social animals that are at their most comfortable when traveling in big groups, which they refer to as “paddlings” when they are on the water. They spend their days foraging for food in the grass or in the shallow water, and at night, they sleep with the other individuals with whom they paddle. (GGO, 2024)

FIGURE 2. Flying Geese formation in real life.

encourages the best-performing geese to provide support to the colony, fostering a more cooperative and efficient search process.

5.2.2 Algorithm mechanism

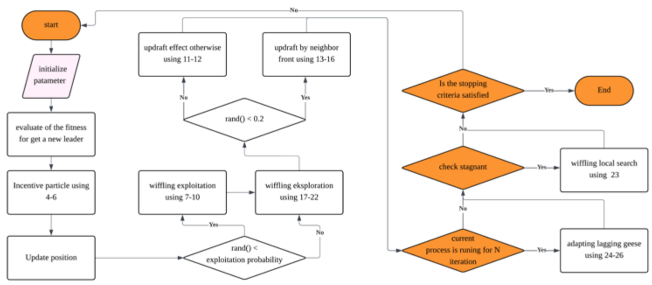

FIGURE 4. Flowchart of the FGOA algorithm

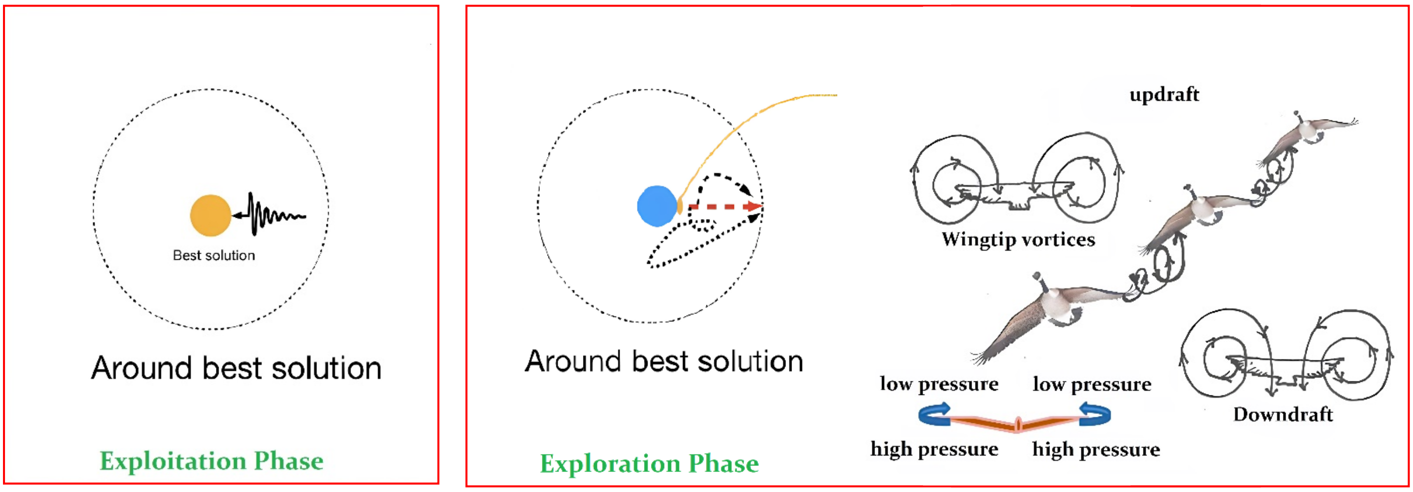

FIGURE 3. Proposed Goose Optimization exploitation and exploration with updraft system.

5.2 Flying Geese Optimization Algorithm (FGOA)

In this section, we will introduce the algorithm of FGOA, which is concisely given as:

5.2.1 Novelty of FGOA

From the previous section, where we discussed the social behavior of geese, this section introduces the unique algorithmic inspirations of FGOA, derived from the migratory behavior of flying geese. The first novelty is wiffling motion utilizes a zig-zag pattern to enhance exploration flexibility. To prevent stagnation and premature convergence, the fatigue reset mechanism replaces the leader when necessary. The adaptive updraft strategy expands the search space, allowing for broader exploration and avoiding local optima. Additionally, the algorithm includes assistance for weaker individuals, ensuring that geese that fall behind still contribute to the optimization process. Lastly, an incentive mechanism

In this section, the exploration search role of FGOA is presented. The FGOA algorithm begins by randomly initializing the positions of individual geese (each representing a candidate solution). Start with position vector $X_i (i = 1,2,3, \dots n)$ where n is representing total number of geese. X_i Uniformly between $minx$ and $maxx$ .for each dimension and V_i(velocity) uniformly with [-0.1,0.1]; each X_i will records their best position (X_{best}) and who is get the best position (g_{best}) will be considered for the next leader.

5.2.1 Incentive phase

To maintain a balance between exploration and exploitation in FGOA, we propose an incentive system that modifies the best-performing solution based on the performance distribution of the population. This prevents early convergence by allowing underperforming particles to contribute to the global best update, making the search process more comprehensive. The incentive threshold is computed as follows:

$$\omega_{treshold} = \left| \frac{1}{N} \sum_{i=1}^{N} f(x_i) - f(g_{best\_score}) \right| \tag{4}$$

Where N represent the number of geese in the population. $f(X_i)$ denotes the fitness score of the i-th goose, and $f(g_{best})$ is the global best fitness score. Each particle get a evaluated again by :

$$f(X_i) < \omega_{treshold} \times f(g_{best\_score}) \tag{5}$$

If the condition is met, a particle that may exhibit below-average behavior in comparison to the global best will be selected to guide the global best update. To incorporate information from the underperforming geese, the global geese has modifies its position using *probabilistic interpolation* between its own goose position (X_i) and current position:

$$X_{i+1} = r \times g_{best} + (1 - r) \times X_i \tag{6}$$

Where r is random value in [0,1] for each position. This algorithm control g_{best} toward best position.

5.2.2 update velocity velocity and position

In the Metaheuristic algorithm, the balance between exploration and exploitation is a crucial factor that determines the effectiveness of obtaining an optimal solution. Exploitation refers to the utilization information from the g_{best} to improve existing solution, meanwhile, exploration is designed to find new solution in the space to avoid getting trapped in local optima. The Dynamic Groups behavior of the FGOA divides all the individuals into exploitation (n_1) and exploration(n_2). The FGOA initiates $n\%$ for exploitation (n_1) and $m\%$ for exploration, as shown in Figure 3.

5.2.3.1 whiffling exploitation phase

In FGOA, we refer to exploitation as whiffling exploitation. Exploitation get formulated to help goose increase convergence with directing particle to g_{best} while adding small random disturbances to avoid premature convergence and local optima. First, the direction is generated experience level $(\gamma)$with random 0 until 1 and calculated (sometimes reversed with 20% probability to allow flexibility):

$$\vec{d} = g_{best} - X_i \tag{7}$$

Afterward, the goose will experience some disturbance to avoid premature convergence by adding random number :

$$\vec{r}f = r_{number} \times |\vec{d}| \tag{8}$$

Where rf is a random scaling factor and r_{number} is a random generated number. Next, the direction vector is normalized as $\hat{d} = \frac{\vec{d}}{||\vec{d}||+\epsilon}$, where ϵ is small value (10^{-10}) prevent division with zero, and give some random noise $r_s \in (0, \sigma)$ where $\sigma = 0.5 \times (1 - \gamma)$. Then, a small angular deviation is introduced using a uniformly distributed random angle $\theta \in [0, \frac{\pi}{6}(1 - \gamma)]$ together with a randomly generated, normalized unit direction vector $\hat{r}_d \in [0,1]^D$.

$$\Delta X = \|\vec{d}\| \times (\hat{d} \times cos\theta + \hat{r}_d \times sin\theta) + \vec{r}_f + \vec{r}_s \tag{9}$$

the meaning is that the goose move toward to g_{best} with some random modification. However, to ensure that the goose remains within the valid search space, we apply normalization using the clip function :

$$X_{i+1} = clip(X_i + \Delta X, X_{min}, X_{\max}) \tag{10}$$

Where X_{min} and X_{max} is a minimum and maximum area of dimension. This prevents the goose from going out of bounds in the search dimension. The strength of the updraft is randomly initialized within range of 0.1 to 2.0.

5.2.3.2 updraft phase

At each iteration, the particle either applies the updraft influence based on its front neighbor with a 50% probability or applies a general updraft effect otherwise.

a. Updraft by best position

The updraft phase simulates the aerodynamic advantage geese gain from rising air currents, which increases their movement efficiency. This mechanism enhances exploration by introducing random updraft effects. As shown in lines 80 through 91 in Algorithm 1, the new velocity is updated considering the updraft effect and the parameters for inertia, cognitive, and social influence.

In every iteration, the number of updrafts is randomly selected within the range of 1 to 3. For each generated updraft, the center is randomly determined within the search space based on the predefined limits, $minx$ and $maxx$.

Furthermore, the distance between agent and the center of updraft will calculate with Euclidean distance. If the agent is within the range of influence of the specified updraft, then the influence force is defined Gaussian influence (https://doi.org/10.1016/B978-044452701-1.00060-0):

$$influence = S \times \exp\left(-\frac{d^2}{2\sigma^2}\right) \tag{11}$$

Where S represent the strength of updraft, d is the distance between agent and center of updraft, and σ is deviation standard set to 1.0 .

After we set the influence of updraft, vector acceleration giving to the agent with direction to the center of updraft, which formula :

$$V_{new} = V_i + influence \times \frac{c-p}{d+\epsilon} \tag{12}$$

Where V_i is the current velocity, c is position of center updraft, p is agent is position. This implementation of this mechanism to improve exploration in the search space by providing additional impetus toward the center of dynamically generated updraft.

b. updraft by neighbor front

In addition to random environmental updrafts, geese benefit from the aerodynamic upwash generated by the wings of leading birds in a V-formation. This phenomenon improves energy efficiency and flight alignment within the formation. Inspired by this behavior, the front neighbor updraft influence mechanism is introduced to improve local guidance and exploitation in the population.

As shown in lines 58-78 in Algorithm 1, this mechanism identifies a front-aligned neighbor that is ahead of the current particle in the formation direction (from X_i *to* g_{best}). The procedure begins by computing the formation direction vector:

$$\hat{d}_{form} = \frac{g_{best}-X_{best}}{||g_{bes}-X_{best}||+\varepsilon} \quad (13)$$

Where $\varepsilon = 1 \times 10^{-8}$ to avoid division by zero. For each other agent in the population, a normalized relative direction vector is computed:

$$\hat{d}_{pq} = \frac{n-X_i}{||n-X_i||+\varepsilon} \quad (14)$$

Where n denotes the position of the selected front neighbor, after that calculate dot product :

$$\propto_a = \hat{d}_{pq} \times \hat{d}_{form} \quad (15)$$

Only neighbors with $\propto_a > 0.5$ are considered . The neighbor with the highest $\propto_a$ is selected as the leading front neighbor. If such a neighbor exists, and the distance between its position(r) and the particle's best position is within a moderate range (less than 1.0), then the particle's velocity is influenced according to the following equation:

$$V_{new} = V_{new} + \frac{0.5}{r+\varepsilon} \times \frac{n-g_{best}}{||n-g_{best}||+\varepsilon} + \delta \quad (16)$$

where r is the distance between the neighbor and the particle's best position, and $\delta \sim U(-0.1,0.1)^D$ is a small random perturbation vector that prevents strict deterministic alignment.

5.2.3.2 exploration phase

After get a updraft effect, the goose do it the exploration. exploration plays a role in searching for new solutions across the entire search space to avoid getting trapped in local optima. Exploration ensure that the algorithm can caught all the area in domain dimension before the goose do the exploitation. The value of V we make a reference for next formulate its aim for renew the velocity which :

$$V_{new} = V_{new} + A|C(X_{best} - X_i)| \quad (17)$$

Where $A = 2 * (r_1 * az) - az$ amnd $C = 2 * r_2$. With r_1 and r_2 parameter change linearly from random dimension. And az parameter changed linearly from 2 to 0. The velocity update equation is:

$$V_{new} = inertia \times V_i + cognitive \times r_1 \times (X_{best} - X_i) + social \times r_2 \times (g_{best} - X_i) \quad (18)$$

the goose performs a random perturbation, which mix the directions and ensure the goose does not concentrate solely on the

$$X_{new} = V_{new} + X_i + 0.1 \times r_n(0,1) \quad (19)$$

Where r_n represent Gaussian random noise to ensure the goose can move to the new direction with different move for every goose. Exploration also led by g_{best} :

$$direction\ to\ g_{best} = g_{best} - X_i \quad (20)$$

After that :

$$rf = mutation\ probability * a|direction\ to\ g_{best}| \quad (30)$$

Where a is random factor in the dimension of space. Momentum is introduced using :

$$momentum = m_f \times V_i \quad (21)$$

Where m_f represent momentum factor. The final position update is:

$$X_{i+1} = \alpha \times direction\ to\ g_{best} \times (1-\alpha) * rf + momentum \quad (27)$$

To maintain valid search boundaries :

$$X_{i+1} = clip(X_i, X_{minx}, X_{maxx}) \quad (22)$$

4.2.4 whiffling local search

If the goose is stagnant or cannot find the optimal solution in another iteration, whiffling local search will activate to help the goose by giving a minor disturbance to each goose in the population to help find the optimal solution:

$$new_{position} = X_i + rand\ pertubation \quad (23)$$

Where the perturbation is sampled from a uniform distribution in the rang [-0.001,0.001], allowing for fine grained local exploration. If the $new_{position}$ finds a new best position, then the position will be updated with the new position. This approach avoids *premature convergence* (https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=ON+STAGNATION+OF+THE+DIFFERENTIAL+EVOLUTION+ALGORITHM&btnG=)by ensuring that each goose in the population not only exploits the best solution for now but also explores.

4.2.5 assist lagging geese

This formulation is designed to help geese with low performance stay within the dynamic population and not fall behind the colony during optimization. This strategy aims to improve the efficiency of exploration and exploitation and ensure the stability of the convergence algorithm. The algorithm includes a mechanism for assisting lagging geese to ensure that all individuals in the population remain competitive. As shown in lines 28 through 30 of Algorithm 1, geese whose fitness falls below the average are moved toward the global best, ensuring that no individual is left behind and promoting diversity in the population. The first step to assist lagging geese is identifying lagging geese with:

$$\bar{S} = \frac{1}{N}\sum_{i=1}^{N} S_i \tag{24}$$

Where s_i represent objective goose function value to i −th and N is total goose in population. If the objective value worse than threshold, then:

$$S_i = 1.5 \times \bar{S} \tag{25}$$

For the goose get a lagging goose criteria, the position will renew with push them to global best or average goose with good performance with formula :

$$g_i^{new} = X_i + \lambda(X_g - g_{best}) \tag{26}$$

Where λ which control how much the individuals is moved toward the g_{best}.

Algorithm 1 Flying Geese Optimization Algorithm (FGOA)
1: **Initialize parameters**: population of geese X_i, velocities V_i, global best as g_{best}
2: **Evaluate initial fitness** for all geese
3: **Set the** g_{best} based on the best fitness
4: **while** stopping criteria are not met **do**
5: **for** each geese i **do** ▷ // incentive threshold:
Compute $\omega_{treshold}$ using Eq. 4
6: **if** $f(X_i) < (\omega_{threshold} \cdot g_{best})$ **then**
7: Compute using Eq. 5
8: **end if**
9: **end for**
10: **for** each geese i **do** ▷ // Whiffling exploitation
11: Compute Whiffling exploitation using Eq. ?-?
▷ // Whiffling exploration
12: Compute Whiffling exploration using Eq.?-?
13: **if** $rand < 0.5$ **then** ▷ // Updraft mechanism
14: Find front neighbor using Eq. ?
15: **else**
16: Generate random updraft centers using Eq.
17: **end if**
18: **end if**
19: **end for**
▷ // Assist lagging geese
20: **if** i mod 100 == 0 **then**
21: Calculate average score of all geese
22: **for all** $geese\ y$ **do**
23: **if** $X_y\ score > 1.9 \times average\ score$ **then**
24: Compute using Eq. ?
25: **end if**
26: **end for**
27: **end if**
28: **if** g_{best} score unchanged **then**
29: Increment fatigue and stagnant counters
30: **else**
31: Reset counters to zero
32: **end if**
33: **for all** $geese\ y$ **do** ▷ // Local whiffling search
34: $X_{i+1} = X_y + rand\ uniform()$
35: Evaluate X_i
36: **if** $f(X_i) < f(X_y)$ **then**
37: Update X_{best} and $f(X_{best})$
38: **end if**
39: **end for**
▷ // Change leader if fatigued
40: **if** fatigue counter ≥ fatigue threshold **then**
41: Set new leader as current g_{best}
43: **end if**
44: **end while**
45: return g_{best}

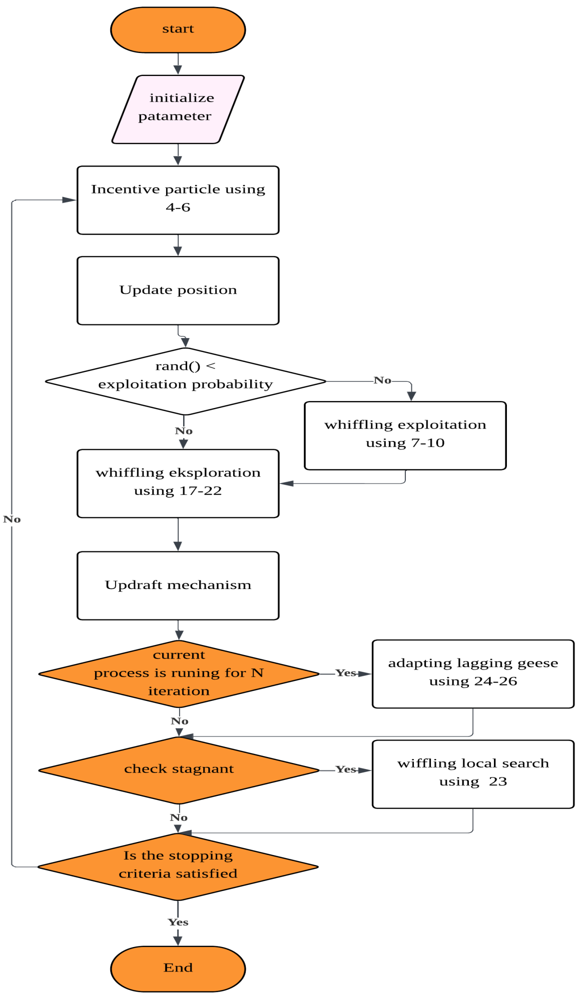

FIGURE 2. The flow chart of the FGOA.

C. PERFORMANCE CRITERIA

評估時間序列預測模型的效能時，常見的評估指標包括平均絕對百分比誤差（Mean Absolute Percentage Error, MAPE）以及均方根誤差（Root Mean Squared Error, RMSE），評估指標能夠從不同層面反映模型預測的準確性與穩定性。

其中，MAPE 為常用的相對誤差衡量指標之一，其透過計算預測值與實際觀測值之間誤差的相對百分比，進而求得其平均值，以評估模型在整體預測過程中的相對準確性。若模型之 MAPE 值愈低，則代表其預測結果與實際值之差距愈小，亦即模型具備更高的預測精確度。MAPE 的數學定義詳見公式（20）。RMSE 為衡量預測誤差幅度的另一常用指標，其透過計算預測誤差平方之平均後開根號，以反映模型預測值與實際觀測值之間偏差的平均程度。RMSE 數值愈低，則代表模型預測結果之誤差變異愈小，顯示其在預測任務中表現越精準穩定。RMSE 的計算公式可參閱公式(21)，作為定量評估預測效能的標準依據[33]。

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_i-f_i}{y_i}\right| \times 100\% \tag{20}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i-f_i)^2} \tag{21}$$

Equations (20), for each observation i, compute the prediction error as $y_i - f_i$, and then divide this prediction error by the actual value y_i to obtain the relative error. Next, take the absolute value of the relative error for all observations and calculate the mean. Finally, multiply by 100% to present the error as a percentage. Equations (21), calculate the squared prediction error for each observation, $(y_i - f_i)^2$. Then, take the average of all squared errors, i.e., $\frac{1}{N}\sum_{i=1}^{N}(y_i-f_i)^2$.

IV. RESULTS

A. DATA DESCRIPTION

在本研究中，我們從臺灣衛生福利部之健保資料庫收集了從 1998 年到 2023 年共 26 年的 60 歲以上男女失智症患者之年度人數. 研究涵蓋台灣各區域, 包括 60~64 歲, 65~69 歲, 70~74 歲, 75~79 歲, 80~84 歲, 85 歲 and above, 詳細資料如表 1、2。每個區域的資料分為兩個子集：由 1998 年至 2015 年的男女失智症患者之年度人數資料組成的訓練集，以及由 2016 年至 2023 年的男女失智症患者之年度人數資料組成的測試集。

TABLE I.
NITROGEN DIOXIDE AND SULFUR DIOXIDE, AIR POLLUTION IN TAIWAN, 2005-2021.

Years Old	SD	Min	Max	Mean	COV(%)
60~64	0.35	1.01	2.12	1.49	0.24
65~69	2.64	3.50	10.59	6.07	0.44
70~74	3.04	5.89	17.19	9.34	0.33
75~79	3.75	5.27	17.74	12.60	0.30
80~84	6.14	3.71	22.53	14.03	0.44
85 and above	12.57	1.99	35.17	17.80	0.71

TABLE II.
NITROGEN DIOXIDE AND SULFUR DIOXIDE, AIR POLLUTION IN TAIWAN, 2005-2021.

Years Old	SD	Min	Max	Mean	COV(%)
60~64	0.38	0.93	2.34	1.46	0.26
65~69	3.15	3.08	11.80	6.94	0.45
70~74	5.28	4.65	22.99	11.68	0.45
75~79	8.22	4.71	28.44	17.11	0.48
80~84	12.37	3.90	42.26	20.36	0.61
85 and above	21.32	2.57	68.17	27.10	0.79

B. MACHINE LEARNING PARAMETER SETTINGS

支持向量迴歸（SVR）已應用於時間序列預測領域，並取得了優異的效果. SVR 是一種很有前景的時間序列預測方法，它具有參數少、預測能力強、訓練速度快等優勢[23]。因此相較於其他方法，SVR 在對高維度、稀疏或含雜訊的數據時，仍能保持穩健的預測性能[25]。 在本研究中, 我們採用 SVR 模型與超參數最佳化. 我們指定了 SVR 指數參數值範圍，包括 $C = (2^0 \sim 2^{30})$, $\sigma = (2^{-30} \sim 2^1)$, and $\varepsilon = (2^{-30} \sim 2^1)$ [5] 。 表 2、3 列出 SVR 模型的超參數最佳化訓練結果。

C. ANALYSIS OF FORECAST RESULTS FOR THE NUMBER OF DEMENTIA PATIENTS

本研究探討了多種時間序列與人工智慧模型，包括傳統的自回歸移動平均模型（ARIMA）、Holt-Winters 指數平滑（HWETS），及整合支持向量回歸（SVR）的優化算法如粒子群優化（PSO）、差分演化（DE）、鯨魚優化（WOA）、哈里斯鷹優化 (HHO)、長短期記憶網路 (LSTM)，以及新提出的雁行最佳化演算法（FGOA）結合 SVR 之性能表現。研究透過平均絕對百分比誤差 (MAPE) 與均方根誤差（RMSE）衡量模型預測的準確性與穩定性。

從男性失智症患者的預測結果表 5， FGOASVR 的表現明顯優於其他所有模型，平均 MAPE 僅為 3.17，RMSE 則為 0.69。與傳統的 ARIMA 和 HWETS 相較之下，其 MAPE 分別為 6.90、5.24 與 RMSE 分別為 1.13、1.02 仍較高。而女性失智症患者的預測分析結果表 6 呈現相似趨勢。FGOA-SVR 模型同樣顯著優於其他方法，整體平均 MAPE 僅 3.42%，RMSE 為 0.96，表現一致穩定。DE-SVR 與 HHO-SVR 模型則顯示明顯的性能波動，尤其在 70~44 歲族群，其 MAPE 分別達 37.07%與 13.29%，凸顯這些方法在處理特定年齡區段數據的局限性。

TABLE III.
MALE SVR MODEL OBTAINS HYPERPARAMETRIC OPTIMIZED TRAINING RESULTS.

	SVR			PSOSVR			DESVR		
	C	σ	ε	C	σ	ε	C	σ	ε
60~64	1.024E+03	6.250E-02	2.500E-01	4.564E+08	1.288E+00	6.559E-02	1.024E+03	6.058E-01	9.313E-10
65~69	3.277E+04	6.104E-05	4.000E-03	4.916E+08	4.882E-01	1.386E-01	1.024E+03	4.491E-02	1.937E-01
70~74	1.024E+03	1.221E-04	1.250E-01	5.680E+08	3.899E-01	1.995E-01	1.024E+03	7.617E-02	2.895E-01
75~79	1.024E+03	2.000E-03	1.250E-01	7.394E+08	8.517E-07	3.920E-01	1.074E+09	1.454E-03	2.599E-01
80~84	1.024E+03	9.000E-04	1.250E-01	5.190E+08	2.735E-05	3.739E-01	2.621E+05	9.766E-04	9.766E-04
85 and above	1.024E+03	2.000E-03	6.250E-02	8.839E+08	8.926E-03	1.656E-01	1.024E+03	1.447E-04	9.313E-10
	WOASVR			HHOSVR			FGOASVR		
	C	σ	ε	C	σ	ε	C	σ	ε
60~64	1.024E+03	6.056E-01	9.537E-07	1.068E+09	1.989E+00	6.609E-02	5.282E+05	2.298E-05	1.816E-02
65~69	1.059E+03	3.090E-03	1.058E-01	1.366E+03	2.428E-03	1.059E-01	1.638E+04	5.000E-01	3.283E-02
70~74	1.028E+03	1.630E-02	9.578E-07	1.024E+03	1.628E-02	1.064E-06	1.353E+06	8.225E-07	4.605E-04
75~79	4.963E+08	4.170E-03	2.579E-01	2.023E+08	1.968E-03	2.573E-01	4.569E+06	6.076E-07	7.947E-04
80~84	7.428E+07	3.401E-03	2.338E-01	4.765E+06	2.135E-04	5.363E-04	4.189E+06	5.493E-07	3.564E-06
85 and above	1.024E+03	1.445E-04	9.537E-07	1.024E+03	1.450E-04	9.537E-07	2.889E+06	1.527E-06	3.056E-05

TABLE IV.
FEMALE SVR MODEL OBTAINS HYPERPARAMETRIC OPTIMIZED TRAINING RESULTS.

	SVR			PSOSVR			DESVR		
	C	σ	ε	C	σ	ε	C	σ	ε
60~64	2.048E+03	2.000E-04	3.000E-03	7.318E+08	9.099E-01	8.622E-02	1.024E+03	1.007E+00	7.224E-04
65~69	8.192E+03	3.125E-02	1.250E-01	2.657E+08	1.979E-04	2.829E-01	1.024E+03	1.111E-03	1.412E-01
70~74	1.024E+03	2.000E-03	2.500E-01	3.367E+08	3.110E-06	4.156E-01	6.711E+07	7.629E-06	2.441E-04
75~79	4.096E+03	2.000E-03	1.250E-01	3.581E+07	9.313E-10	1.352E+00	1.024E+03	8.019E-05	2.770E-02
80~84	8.192E+03	3.000E-05	6.250E-02	5.356E+08	9.313E-10	1.288E+00	1.024E+03	4.001E-04	1.439E-02
85 and above	8.192E+03	4.883E-04	6.250E-02	2.048E+03	1.526E-05	1.250E-01	1.092E+07	9.313E-10	9.313E-10
	WOASVR			HHOSVR			FGOASVR		
	C	σ	ε	C	σ	ε	C	ε	σ
60~64	1.394E+03	9.294E-01	1.301E-06	1.079E+03	9.971E-01	2.150E-05	5.282E+05	2.298E-05	1.816E-02
65~69	8.952E+08	2.987E-04	2.426E-01	8.652E+05	4.991E-05	8.198E-04	1.638E+04	5.000E-01	3.283E-02
70~74	1.519E+03	3.111E-02	1.374E-01	3.708E+06	1.818E-03	6.935E-03	3.277E+04	1.221E-04	7.451E-09
75~79	3.823E+03	2.828E-05	1.528E-06	4.433E+03	2.629E-05	2.933E-05	4.569E+06	6.076E-07	7.947E-04
80~84	4.041E+04	1.825E-04	1.343E-03	7.131E+03	2.387E-04	6.018E-03	4.189E+06	5.493E-07	3.564E-06
85 and above	1.067E+07	9.313E-10	9.537E-07	1.024E+03	8.669E-06	1.550E-02	2.889E+06	1.527E-06	3.056E-05

TABLE V
PREDICT THE NUMBER OF MALE DEMENTIA CASES FROM 2016 TO 2023 USING DIFFERENT METHODS.

Years Old	Criteria	ARIMA	HWETS	SVR	PSO SVR	DE SVR	WOA SVR	HHO SVR	LSTM	FGOA SVR
60~64	MAPE	11.67	4.61	10.00	9.52	8.45	8.45	11.12	17.99	**3.37**
	RMSE	0.31	0.09	0.21	0.20	0.17	0.17	0.23	0.08	**0.07**
65~69	MAPE	11.10	7.86	5.59	37.36	37.27	4.54	4.64	15.57	**3.95**
	RMSE	1.12	0.90	0.74	3.85	4.09	0.65	0.67	1.84	**0.51**
70~74	MAPE	1.93	6.48	5.02	31.71	29.01	36.14	36.14	2.09	**1.49**
	RMSE	0.29	1.00	0.83	5.52	5.48	7.53	7.53	0.29	**0.22**
75~79	MAPE	5.91	4.76	9.94	3.40	12.86	14.85	13.48	6.90	**3.35**
	RMSE	1.10	0.85	1.72	0.67	2.56	2.87	2.66	1.46	**0.67**
80~84	MAPE	2.77	3.15	7.66	3.89	9.55	11.33	4.68	6.03	**2.73**
	RMSE	0.65	0.69	1.89	0.90	2.41	2.97	1.08	1.31	**0.61**
85 and above	MAPE	8.01	4.57	11.85	7.63	10.65	10.65	10.66	9.79	**4.10**
	RMSE	3.31	2.58	4.06	3.81	3.66	3.66	3.67	3.33	**2.03**
Average	MAPE	6.90	5.24	8.34	15.59	17.97	14.33	13.45	9.73	**3.17**
	RMSE	1.13	1.02	1.58	2.49	3.06	2.98	2.64	1.39	**0.69**

MAPE, mean absolute percentage error; RMSE, root mean square error; boldface, the optimal values in each row. ARIMA, Autoregressive Integrated Moving Average; HWETS, holt winters exponential smoothing; SVR, supports vector regression; PSOSVR, Particle Swarm Optimization supports vector regression; DESVR, differential evolution supports vector regression; WOASVR, whale optimization algorithm supports vector regression; HHOSVR, Harris Hawks Optimization supports vector regression; LSTM, long short-term memory; FGOASVR, Flying Geese Optimization Algoritm supports vector regression.

TABLE VI
PREDICT THE NUMBER OF FEMALE DEMENTIA CASES FROM 2016 TO 2023 USING DIFFERENT METHODS.

Years Old	Criteria	ARIMA	HWETS	SVR	PSO SVR	DE SVR	WOA SVR	HHO SVR	LSTM	FGOA SVR
60~64	MAPE	12.35	5.90	14.13	8.76	9.87	9.86	9.84	11.32	**5.30**
	RMSE	0.24	0.11	0.26	0.16	0.19	0.19	0.19	0.22	**0.11**
65~69	MAPE	7.52	5.90	10.74	19.94	6.24	17.12	17.67	18.88	**3.44**
	RMSE	0.94	0.84	1.45	2.45	0.81	2.09	2.19	2.15	**0.46**
70~74	MAPE	3.41	6.68	12.73	5.51	37.07	26.38	13.29	5.47	**2.94**
	RMSE	0.72	1.30	3.39	1.11	11.77	7.44	2.76	1.08	**0.62**
75~79	MAPE	4.75	4.40	8.92	6.18	3.74	4.52	4.53	13.13	**2.85**
	RMSE	1.64	1.27	2.64	2.17	1.43	1.68	1.69	3.98	**0.97**
80~84	MAPE	2.41	6.13	9.48	3.621	16.81	14.61	14.86	3.55	**1.94**
	RMSE	1.08	2.204	3.65	1.40	6.96	6.04	6.14	1.45	**0.73**
85 and above	MAPE	7.36	8.69	9.25	7.01	5.06	5.25	4.46	4.15	**4.04**
	RMSE	4.79	5.01	5.91	4.04	3.42	3.49	3.06	3.04	**2.84**
Average	MAPE	6.30	6.28	10.88	8.50	13.13	12.96	10.78	9.42	**3.42**
	RMSE	1.57	1.79	2.88	1.89	4.10	3.49	2.67	1.99	**0.96**

MAPE, mean absolute percentage error; RMSE, root mean square error; boldface, the optimal values in each row. ARIMA, Autoregressive Integrated Moving Average; HWETS, holt winters exponential smoothing; SVR, supports vector regression; PSOSVR, Particle Swarm Optimization supports vector regression; DESVR, differential evolution supports vector regression; WOASVR, whale optimization algorithm supports vector regression; HHOSVR, Harris Hawks Optimization supports vector regression; LSTM, long short-term memory; FGOASVR, Flying Geese Optimization Algoritm supports vector regression.

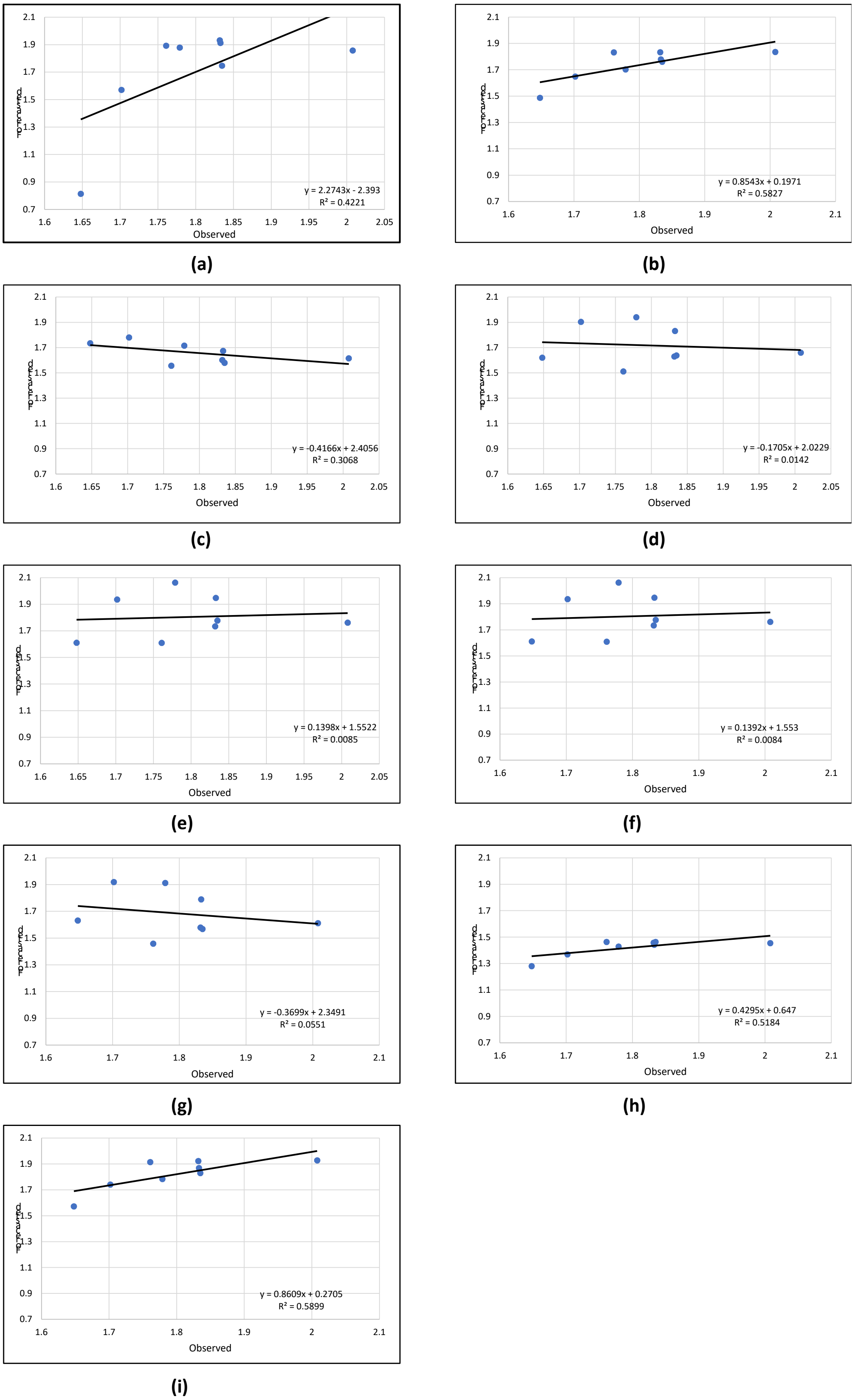

FIGURE 3. Scatter chart of Dementia forecasts (1-month forecasts) from 2016 to 2023, using different methods: (a) ARIMA, autoregressive composite moving average; (b) HWETS, holt winters exponential smoothing; (c) SVR, exponential smoothing; (d) PSOSVR, Particle Swarm Optimization Support Vector Regression; (e) DESVR, Differential Evolution Support Vector Regression; (e) WOASVR, Whale Optimization Algorithm Support Vector Regression; (e) HHOSVR, Harris Hawk Optimization Support Vector Regression; (f) LSTM, long short-term memory; (g) FGOASVR, Flying Geese Optimization Algoritm Support Vector Regression.

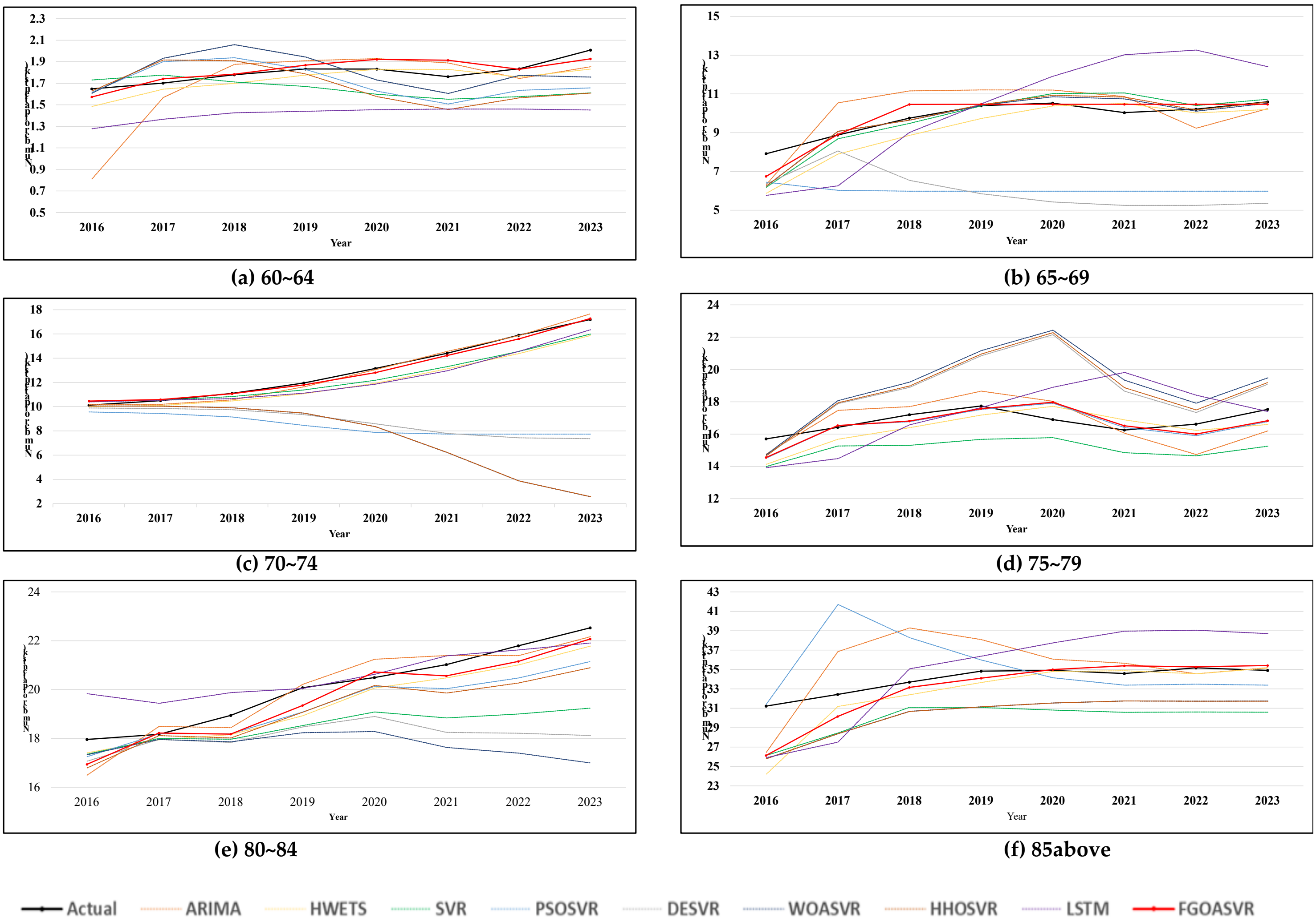

FIGURE 4. Dementia uses seven methods to predict the outcome.

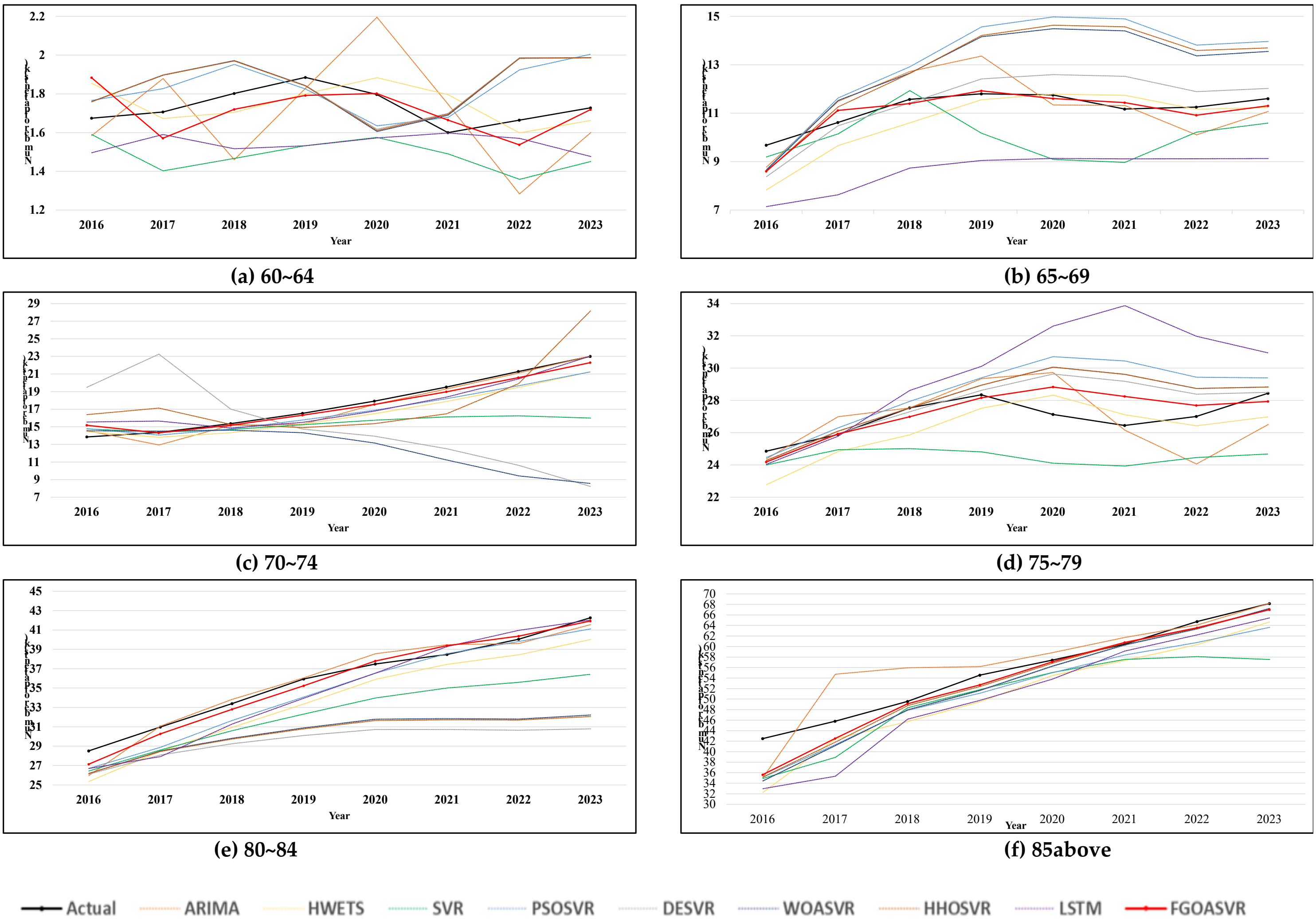

FIGURE 5. Dementia uses seven methods to predict the outcome.

V. DISCUSSION

本研究中，我們從台灣衛生福利部健康保險資料庫中萃取出的失智症患者人數 資料以年為單位分析了自 1998 年至 2023 年間。進一步驗證本研究所提出模型的有效性，我們比較了統計模型(ARIMA、HWETS)、深度學習模型(LSTM)及混合模型(SVR、PSOSVR、DESVR、WOASVR、HHOSVR、FGOASVR) 的預測結果。整體而言，透過 MAPE 與 RMSE 的評估指標，結果顯示 FGOASVR 模型在預測效能上明顯優於其他現有的模型。本節將對上述的統計模型、深度學習模型及混合模型進行深入討論。

A. HYBRID MODELS: COMPARISON OF THE OTHER METAHEURISTIC ALGORITHMS AND FGOASVR MODELS

支持向量迴歸（Support Vector Regression, SVR）能同時處理線性與非線性迴歸問題，已廣泛應用於時間序列預測與曲線擬合。其方法論以懲罰參數 C 控制模型允許的訓練誤差上限 。γ 決定 RBF 核在原始空間中的「感應半徑」。ε-insensitive 損失函數為基礎 當 ε 隨樣本量適當縮放時，支持向量的上下界可以被束縛在一個次線性成長率，從而保證在固定運算資源下維持風險收斂速度，因此具備優異的泛化能力與對高維、小樣本資料的適應性[34]。本研究選擇以 SVR 作為序列預測之基準模型，惟實驗結果顯示：若懲罰參數 C、核函數尺度 σ（或 γ）以及 ε-insensitive 帶寬未經細緻調校，其預測誤差往往高於傳統統計模型。此一現象凸顯超參數設定對 SVR 效能之關鍵性，本研究採用最佳化策略以精確搜尋最適超參數組合，對提升模型預測性能之必要[14]。

由表 V 與表 VI 可見, FGOASVR 在 2016–2023 年男性與女性失智症病例數預測中，平均 MAPE 分別為 3.17% 與 3.42%，明顯優於 ARIMA、HWETS 及各類混合式 SVR 模型(最接近者 HWETS 僅能維持 5.24% / 6.28%)。相同的優勢亦反映在 RMSE（0.69 與 0.96）上，說明 FGOA 在 SVR 超參數搜尋時能更可靠地逼近全域最適，而非停留於侷限的區域最適解。FGOA 之所以具備較強的全域探索能力，首先源自其模擬候鳥飛行時的多重資訊流與能量共享機制。演算法以「領頭—側翼—墊後」三層角色動態輪替為核心，使最佳個體（領鵝）與次佳個體之間維持長距離耦合；當原領鵝於特定迭代內未能改善全域適應值，即觸發「更換領導」與「落後援助」機制，以增加族群穿越高維解空間之機率[35]。 將此 FGOA 演算法用於 SVR 之 C、γ、ε 三參數搜索時，即能以更細膩的全域探索與後期局部剖析，找到使預測誤差（MAPE、RMSE）同時下降的解組合，而不致因過早鎖定次佳支點而產生高變動或欠擬合。

B. DEMENTIA PREVENTION AND INTERVENTIONS

With declining mortality in younger populations, dementia is expected to become one of the greatest global health concerns of the 21st century. Although dementia is not curable, its management and the delay of its manifestation are considered to be theoretically possible. Studies have reported that the course of the disease can be modified with adequate care [53], which supports the focus on the manipulation of modifiable risk factors. In 2017, nine potentially modifiable risk factors were reported, including hypertension, obesity, depression, and low social contact [3]. Three new modifiable risk factors, namely traumatic brain injury, excessive alcohol consumption, and air pollution, were introduced in 2020, with convincing evidence [3]. Approximately 30–50% of dementia cases are attributed to these potentially modifiable risk factors. A reduction of 10–25% of these risk factors can reduce the number of patients with dementia by 1.1 to 3 million worldwide [54]. Furthermore, postponement of the onset of dementia by even 2 years can reduce the burden on public health, society, interpersonal relationships, and the economy [55]. Based on this study's proposed model, policy administrators, medical workers, and stakeholders can implement more effective and extensive societal policies on dementia prevention and care among society. The following suggestions on dementia prevention and care are provided for the aim of a more dementia-friendly society.

Promoting resilience in an aging society is a far-reaching approach to dementia prevention. The maximization of care quality and reduction of dementia incidence should begin at the community level, including through the promotion of dementia awareness and knowledge. According to the UK National Institute of Health and Care Excellence and the US National Institute of Health, social isolation is a potentially modifiable risk factor [56,57]. Aging people may experience loneliness and a lack of social contact and social participation, and the promotion of social engagement opportunities is necessary within the community. Moreover, education and intellectual stimulation alternatives have been demonstrated to enhance cognitive resilience later in life [58]. Therefore, within communities, the establishment of supportive social networks that encourage interaction will alleviate loneliness, hence reduce dementia possibilities.

The cost and burden of dementia care are tremendous and continue to rise as the global population ages. The average total cost incurred by patients with dementia exceeds the total costs of patients with other diseases [8]. Patients with dementia are often elderly people approaching their last years of life; thus, their workforce productivity is naturally weaker. As a group that has relatively low capability of coping with such a household financial crisis, the illness contributes to patients' cognitive and physical burden and hinders the ability of their families to afford future health care [8]. Therefore, especially financially, dementia care often calls for more medical health care support than other illnesses [9]. To actively promote high-quality dementia care, additional medical expenditure on dementia care and prevention is necessary and strongly recommended.

Furthermore, the prevalence of dementia affects not only patients but also their family or the health care workers who must live with these patients and deal with the behavioral and emotional effects of dementia. As mentioned earlier, patients with dementia often also experience disorientation, confusion, mood instability, and behavioral or psychological symptoms. As a result, under high pressure for an extended period of time, studies have reported that informal caregivers of patients with dementia often develop poor mental health, and have a relatively high mortality rate [5,10,11,12], which results in further socioeconomic problems. Additional dementia-care training is necessary for the development of adequate dementia care, which should also include the emphasis on caregivers' mental and physical health. Not only the quality of care of patients with dementia, but also their family caregivers should be emphasized in future policies; an expand in the medical allowance and societal support for this particular population should be considered.

Prevention is more effective than a cure. Proper social welfare and public health policies necessitate a precise model for predicting the prevalence rate of dementia. The purpose of this study was to examine whether an alternative prediction model, namely the proposed LSTM network, could effectively predict the trends among the population of patients with dementia. The results demonstrate that the proposed model was not only applicable, but also significantly more accurate than the other models. This precise model can successfully predict the prevalence of dementia and can thus aid government administrations in the development of relevant strategies. For example, policymakers can manage the budget allocated to dementia care to reduce its occurrence by implementing legislative changes, developing preventive interventions for younger populations, and providing ongoing education and care for elderly adults and their families. Future research is warranted to investigate the performance of the proposed LSTM network for the prediction of trends in other illnesses.

C. CONTRIBUTION OF THIS PAPER

本研究針對 1998 至 2023 年間之失智症患者人數資料進行縱深剖析，並以七種模型比較預測性能。其學術貢獻可概述如下：首先，透過時序統計與訊號分解技術精確捕捉患者人數的長期依存結構，奠定後續模型建構基礎；其次，本研究提出以 Flying Geese Optimization Algoritm (FGOA)為核心的預測框架，利用領雁-僚雁隊形調度、上升氣流協同及疲勞檢測重置等生物啟發機制，於高維特徵空間中執行全域搜尋與自適應參數調整；實證結果顯示 FGOA 模型對失智症患者人數之長期趨勢與極端波動皆具備優於傳統統計基準的精準度，證明其在公共衛生情境下之可行性。本研究成果為政府部門制定相關策略提供量化依據，針對年輕族群的預防性介入，以及對高齡者及其家庭之持續教育與照護支援等措施。

D. LIMITATIONS

未來工作可進一步驗證所提 FGOA 架構在其他疾病流行趨勢預測中的表現; 若能融入更多臨床與多模態變項並透過 FGOA 的協同特徵選取篩選與動態調參，將可進一步提升失智症盛行率預測之準確性，亦符「預防重於治療」的研究初衷。未來可應用於船舶航跡推估、潮位預報、金融市場波動預測與即時交通事故風險預警等多元場景，顯示其在序列預測任務上具備高度可遷移的泛化潛力。

VI. CONCLUSION

準確預測不同性別與年齡族群中失智症的病患人數，將有助於提供實證依據，以發展預防或延緩失智症發作的介入措施。本研究所提出的Flying Geese Optimization Algoritm Support Vector Regression (FGOASVR)在預測準確性方面，優於ARIMA、HWETS、SVR、PSOSVR、DESVR、WOASVR及HHOSVR等模型，男性平均MAPE為3.17，RMSE則為0.69，女性平均MAPE為3.42，RMSE則為0.96。全球和台灣由於人口老化，失智症患者人數預計將持續大幅增長。這不僅帶來巨大的醫療和社會照護負擔，也凸顯了在預防、早期診斷、治療和長期照護方面加強投入的迫切性。運用準確的FGOASVR模型發展失智症患者的健康與社會照護策略，改善大眾對失智症的認知，減少污名化，並建立更友善、支持性的社會環境，是未來必須共同努力的重要方向。

ACKNOWLEDGMENT

This work was supported by the National Science and Technology Council, Taiwan (under Grant no. 111-2221-E-165-002-MY3).

REFERENCES

[1] R. J. Hyndman, A. B. Koehler, R. D. Snyder, and S. Grose, "A state space framework for automatic forecasting using exponential smoothing methods," *International Journal of forecasting,* vol. 18, no. 3, pp. 439-454, 2002.

[2] M. Zhang, X. Li, and L. Wang, "An adaptive outlier detection and processing approach towards time series sensor data," *Ieee Access,* vol. 7, pp. 175192-175212, 2019.

[3] A. B. Koehler, R. D. Snyder, and J. K. Ord, "Forecasting models and prediction intervals for the multiplicative Holt–Winters method," *International Journal of Forecasting,* vol. 17, no. 2, pp. 269-286, 2001.

[4] C.-H. Yang, P.-H. Chen, C.-H. Wu, C.-S. Yang, and L.-Y. Chuang, "Deep learning-based air pollution analysis on carbon monoxide in Taiwan," *Ecological Informatics,* vol. 80, p. 102477, 2024.

[5] C.-H. Yang, B.-H. Chen, C.-H. Wu, K.-C. Chen, and L.-Y. Chuang, "Deep learning for forecasting electricity demand in Taiwan," *Mathematics,* vol. 10, no. 14, p. 2547, 2022.

[6] G. Maragatham and S. Devi, "Retracted article: LSTM model for prediction of heart failure in big data," *Journal of medical systems,* vol. 43, no. 5, p. 111, 2019.

[7] L. Wang *et al.*, "Development and validation of a deep learning algorithm for mortality prediction in selecting patients with dementia for earlier palliative care interventions," *JAMA network open,* vol. 2, no. 7, pp. e196972-e196972, 2019.

[8] W. Fan *et al.*, "Support vector regression model for flight demand forecasting," *International Journal of Engineering Business Management,* vol. 15, p. 18479790231174318, 2023.

[9] X. Du, H. Xu, and F. Zhu, "Understanding the effect of hyperparameter optimization on machine learning models for structure design problems," *Computer-Aided Design,* vol. 135, p. 103013, 2021.

[10] W.-C. Juang, S.-J. Huang, F.-D. Huang, P.-W. Cheng, and S.-R. Wann, "Application of time series analysis in modelling and forecasting emergency department visits in a medical centre in Southern Taiwan," *BMJ open,* vol. 7, no. 11, p. e018628, 2017.

[11] X. Xian *et al.*, "Comparison of SARIMA model, Holt-winters model and ETS model in predicting the incidence of foodborne disease," *BMC Infectious Diseases,* vol. 23, no. 1, p. 803, 2023.

[12] K. S. Bohnsack, M. Kaden, J. Abel, and T. Villmann, "Alignment-free sequence comparison: A systematic survey from a machine learning perspective," *IEEE/ACM Transactions on Computational Biology and Bioinformatics,* vol. 20, no. 1, pp. 119-135, 2022.

[13] D. Parbat and M. Chakraborty, "A python based support vector regression model for prediction of COVID19 cases in India," *Chaos, Solitons & Fractals,* vol. 138, p. 109942, 2020.

[14] S. Yang, H.-C. Chen, C.-H. Wu, M.-N. Wu, and C.-H. Yang, "Forecasting of the prevalence of dementia using the lstm neural network in Taiwan," *Mathematics,* vol. 9, no. 5, p. 488, 2021.

[15] T. Hamdi, J. B. Ali, V. Di Costanzo, F. Fnaiech, E. Moreau, and J.-M. Ginoux, "Accurate prediction of continuous blood glucose based on support vector regression and differential evolution algorithm," *Biocybernetics and Biomedical Engineering,* vol. 38, no. 2, pp. 362-372, 2018.

[16] A. Malik, Y. Tikhamarine, D. Souag-Gamane, O. Kisi, and Q. B. Pham, "Support vector regression optimized by meta-heuristic algorithms for daily streamflow prediction," *Stochastic Environmental Research and Risk Assessment,* vol. 34, pp. 1755-1773, 2020.

[17] G. E. Box and G. M. Jenkins, "Time series analysis: Forecasting and control San Francisco," *Calif: Holden-Day,* 1976.
[18] S. Chaudhuri and D. Dutta, "Mann–Kendall trend of pollutants, temperature and humidity over an urban station of India with forecast verification using different ARIMA models," *Environmental monitoring and assessment,* vol. 186, pp. 4719-4742, 2014.
[19] P. R. Winters, "Forecasting sales by exponentially weighted moving averages," *Management science,* vol. 6, no. 3, pp. 324-342, 1960.
[20] D. Nolan and T. Speed, "Teaching statistics theory through applications," *The American Statistician,* vol. 53, no. 4, pp. 370-375, 1999.
[21] V. Vapnik, S. Golowich, and A. Smola, "Support vector method for function approximation, regression estimation and signal processing," *Advances in neural information processing systems,* vol. 9, 1997.
[22] P. G. Nieto, E. F. Combarro, J. del Coz Díaz, and E. Montañés, "A SVM-based regression model to study the air quality at local scale in Oviedo urban area (Northern Spain): A case study," *Applied Mathematics and Computation,* vol. 219, no. 17, pp. 8923-8937, 2013.
[23] M. Castelli, F. M. Clemente, A. Popovič, S. Silva, and L. Vanneschi, "A machine learning approach to predict air quality in California," *Complexity,* vol. 2020, 2020.
[24] C.-H. Yang, C.-F. Lee, and P.-Y. Chang, "Export-and import-based economic models for predicting global trade using deep learning," *Expert Systems with Applications,* vol. 218, p. 119590, 2023.
[25] C. Zhou *et al.*, "A new model transfer strategy among spectrometers based on SVR parameter calibrating," *IEEE Transactions on Instrumentation and Measurement,* vol. 70, pp. 1-13, 2021.
[26] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks*, 1995, vol. 4, pp. 1942-1948 vol.4.
[27] L. Resende and R. H. C. Takahashi, "Contributions to Dynamic Analysis of Differential Evolution Algorithms," (in eng), *Evol Comput,* vol. 31, no. 3, pp. 201-232, Sep 1 2023.
[28] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Advances in engineering software,* vol. 95, pp. 51-67, 2016.
[29] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. Chen, "Harris hawks optimization: Algorithm and applications," *Future generation computer systems,* vol. 97, pp. 849-872, 2019.
[30] K. Akamatsu, "A historical pattern of economic growth in developing countries," *The developing economies,* vol. 1, pp. 3-25, 1962.
[31] R. K. Hamad and T. A. Rashid, "GOOSE algorithm: a powerful optimization tool for real-world engineering challenges and beyond," *Evolving Systems,* vol. 15, no. 4, pp. 1249-1274, 2024.
[32] E.-S. M. El-Kenawy, N. Khodadadi, S. Mirjalili, A. A. Abdelhamid, M. M. Eid, and A. Ibrahim, "Greylag goose optimization: nature-inspired optimization algorithm," *Expert Systems with Applications,* vol. 238, p. 122147, 2024.

[33] X. Liu, Z. Lin, and Z. Feng, "Short-term offshore wind speed forecast by seasonal ARIMA-A comparison against GRU and LSTM," *Energy,* vol. 227, p. 120492, 2021.

[34] V. Cherkassky and Y. Ma, "Practical selection of SVM parameters and noise estimation for SVM regression," *Neural networks,* vol. 17, no. 1, pp. 113-126, 2004.

[35] H. Bian *et al.*, "Improved snow geese algorithm for engineering applications and clustering optimization," *Scientific Reports,* vol. 15, no. 1, p. 4506, 2025.

Po-Hung Chen received the B.S. degree from the Department of Electronic Engineering, Fortune Institute of Technology, in 2013. He is currently working toward the MS, PhD student degree in the Department of Electronic Engineering, National Kaohsiung University of Science and Technology, Taiwan. He has rich experience in machine learning, computer programming, Embedded System, and design. His main areas of research are time series, data mining, machine learning, and deep learning.

Cheng-Hong Yang (M'00-SM'03) received his Ph.D. in computer engineering from North Dakota State University in 1992. He was served as the president of National Kaohsiung University of Applied Science in Taiwan from 2012 to 2016. Currently, he is a chair professor in the Department of Electronic Engineering at National Kaohsiung University of Science and Technology, Taiwan. He has authored/coauthored over 380 refereed publications and several book chapters. He is an Editorial Board Member of multiple international journals. His main areas of research are fuzzy control, evolutionary computation, bioinformatics, and data analysis. Prof. Yang is a Senior Member of the IEEE, fellow of the Institution of Engineering and Technology, and fellow of the American Biographical Institute.

Cheng-San Yang received the Ph.D. degree from National Cheng Kung University, Tainan, Taiwan. He is currently a physician in the Department of Plastic Surgery and serves as the Medical Matters Secretary of Chia-Yi Christian Hospital, Taiwan. He has authored or coauthored numerous refereed publications. His research interests include bioinformatics, plastic and reconstructive surgery, and microscopy surgery.

Li-Yeh Chuang is a professor in the Department of Chemical Engineering and Institute of Biotechnology and Chemical Engineering at I-Shou University, Kaohsiung, Taiwan. She received an M.S. from the Department of Chemistry at the University of North Carolina in 1989 and a Ph.D. from the Department of Biochemistry at North Dakota State University in 1994. She has authored/coauthored over 300 refereed publications. Her main areas of research are bioinformatics, biochemistry, and genetic engineering.

Ting-Jen Hseuh received the M.S. degree in Electrical Engineering from the National Kaohsiung University of Applied Sciences, Kaohsiung, in 2004, and the Ph.D. degree in Microelectronics from the National Cheng Kung University (NCKU), Tainan, Taiwan, in 2008. From 2008 September to 2009 January, he was a postdoctoral researcher with the Institute of Microelectronics, Advanced Optoelectronic Technology Center, Center for Micro/Nano Science and Technology, NCKU; From 2009 January to 20018 January, he was a researcher withthe National Nano Device Laboratories (NDL). Currently, he is a Professor in the Department of Electronic Engineering, National Kaohsiung University of Science and Technology, Kaohsiung. He has published more than 140 contributions in journals. His current research interests nanoscale 1D semiconductors and gas sensor.