P-ISSN 2586-9000
E-ISSN 2586-9027
Homepage : https://tci-thaijo.org/index.php/SciTechAsia
Science & Technology Asia
Vol.28 No.4 October - December 2023
Page: [38-53]
Original research article

Incorporating Decomposition and the Holt-Winters Method into the Whale Optimization Algorithm for Forecasting Monthly Government Revenue in Thailand

Watha Minsan[1,*], Pradthana Minsan[2]

[1]*Data Science Research Center, Department of Statistics, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand*
[2]*Department of Mathematics and Statistics, Faculty of Science and Technology, Chiang Mai Rajabhat University, Chiang Mai 50300, Thailand*

Received 24 July 2023; Received in revised form 15 October 2023
Accepted 31 October 2023; Available online 27 December 2023

ABSTRACT

This study focuses on the forecasting of government revenue in Thailand across four primary sectors: the Revenue Department, Excise Department, Customs Department, and Other Agencies. Acknowledging the critical role of precise and efficient forecasting in policymaking, we proposed two models: the Whale Optimization Algorithm with Holt-Winters (WOA-HW) and the Whale Optimization Algorithm with Decomposition (WOA-D), comparing their performance with two classical models: Classical Decomposition (Classic-D) and Box-Jenkins. The model performances were evaluated using both a training dataset and a test dataset, with Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) serving as key metrics. The results demonstrate that the WOA-D model generally outperformed the other models during the training phase, showcasing its significant potential in time series forecasting. During the testing phase, the WOA-HW model exhibited commendable performance across three datasets: the Revenue Department, Excise Department, and Other Agencies. For the Customs Department dataset, the Box-Jenkins model emerged as the top performer, employing a $SARIMA(2,1,0)(0,1,1)_{12}$ model. This study concludes by emphasizing the effectiveness of these models not only for forecasting government revenue but also for broader applicability in forecasting other time series data.

Keywords: Decomposition; Forecasting; Government revenue; Holt-Winters; Whale optimization algorithm

***Corresponding author:** wathaminsan@gmail.com

1. Introduction

Government revenue in Thailand is derived from the tax revenues collected by the Revenue Department, the Excise Department, and the Customs Department. Additionally, revenue is obtained from other official entities such as the Treasury Department and various state enterprises. Each of these components holds significant importance in the management and administration of the country’s finances. In the fiscal year 2022, the Revenue Department reported a total revenue of approximately 2.166430 trillion baht, with Value-Added Tax (VAT), corporation taxes, and personal income taxes ranking as the primary sources of income. Similarly, the Excise Department generated a revenue of around 503.465 billion baht, with oil taxes, car taxes, and beer taxes serving as the predominant sources. The Customs Department is responsible for collecting import and export duties on goods transported to and from Thailand. Import duties stand out as the main income source, contributing approximately 108.896 billion baht in the fiscal year 2022. Other government agencies derive their primary income from two sources: state enterprises and other government agents. Together, these agencies amassed a total revenue of approximately 290.924 billion baht. Collectively, the government revenue for the fiscal year 2022 amounted to an impressive 3.071272 trillion baht, underscoring its profound significance to the country. Given the critical role of accurate revenue forecasting in economic planning and policymaking, the utilization of statistical forecasting models becomes imperative. These models provide precise estimations, enabling informed decisions for strategic storage or investment plans based on the analysis and forecasted values [1, 2].

Time series analysis is a key statistical approach that examines data characteristics and patterns that change over time. It leverages statistical tools to explore trends, seasonality, patterns, and relationships between data values at different time intervals. Decomposition, proposed by Persons [3] in their 1919 article Indices of Business Conditions, is a renowned method within time series analysis. Persons introduced the concept of dividing a time series into trend, seasonal, and irregular components, a technique still employed by economists and businesses today. The decomposition method dissects a time series into three components: trend, seasonality, and residual. The trend reflects the long-term direction of the data, seasonality represents repeating cyclical patterns, and residual encapsulates random data fluctuations. By projecting the trend and seasonality into the future and adding the residual component, decomposition can forecast future values of a time series. However, this study did not take into account the cyclical and irregular components. This method has found successful application in various domains, including energy forecasting, trend prediction, and seasonal pattern identification. For instance, Mbuli et al. [4] referenced 29 studies that employed the decomposition method in power system forecasting. These studies used both additive and multiplicative decomposition models on various types of data, including hourly, daily, and monthly measurements. Of these, 15 papers solely utilized the multiplicative decomposition method, while 6 papers exclusively applied the additive decomposition method. Other research papers used the decomposition method as the initial forecasting technique, followed by further prediction methods like Artificial Neural Network (ANN) and Autoregressive Integrated Moving Average (ARIMA). In economic research, Koirala [5] examined Nepal's government revenue using five statistical techniques: the Holt method, Winters method, decomposition method, Seasonal Autoregressive Integrated Moving Average (SARIMA) method, and growth rate method. Similarly, Hansen and Nelson [6] investigated the State of Utah's revenue, using machine learning methods like ANN, and traditional time series techniques such as classical decomposition, exponential smoothing, ARIMA, and causal models.

The smoothing technique has found numerous applications in forecasting. Two pioneers, Brown and Holt significantly contributed to the development and refinement of exponential smoothing methods [7]. In the 1950s, Brown introduced two models: simple exponential smoothing, suitable for data without a clear trend or seasonal pattern, and double exponential smoothing, appropriate for data with a trend but no seasonality, also known as Brown's linear exponential smoothing [8]. Holt further developed this field by introducing a method to handle data with a trend [9, 10]. Although termed double exponential smoothing like Brown's model, Holt's method differed by incorporating parameters for both the level of the series and the trend. Collaborating with Winters [11], Holt later adapted his model to accommodate seasonality, leading to the Holt-Winters method (HW) or triple exponential smoothing. This extension added a third parameter for seasonality, enhancing its flexibility for various types of time series data. Numerous practical applications have employed these techniques. For instance, Frank [12] applied moving averages and smoothing techniques to forecast income from crucial sources in eight Florida cities. Likewise, Gianakis and Frank [13] emphasized the significance of forecasting for local governments by employing moving averages, smoothing techniques, and the Box-Jenkins method to inform decisions regarding budgets, staffing, and resource allocation based on expected revenues. In a more recent study, Williams [14] utilized various techniques, including exponential smoothing methods, to forecast 55 revenue data series for 18 local governments.

Numerous studies have explored the hybrid application of the HW method and various optimization techniques to optimize its smoothing parameters. Techniques include particle swarm optimization [15], golden section method [16, 17], fruit fly optimization [18], genetic algorithms [19], optimized fractional grey [20], and cuckoo algorithms [21]. In this study, we focus on employing an optimization method to identify the optimum parameters for the HW method, requiring approximately three parameters, as well as the decomposition method, which involves approximately 14 parameters related to trend (2 parameters) and seasonal components (12 parameters). To estimate these parameters, we utilize the Whale Optimization Algorithm (WOA), previously applied in predicting housing prices in China as demonstrated by Liu and Wu [22]. In their study, they compared the modified Holt's exponential smoothing method incorporating WOA (mH-WOA) with the Box-Jenkins ARIMA model, the gray model, and the backpropagation neural network, reporting favorable results. The mH-WOA yielded positive outcomes for all four datasets analyzed in the study. In this study, we present a model that incorporates WOA with the HW method (WOA-HW) and the decomposition method (WOA-D) for monthly government revenue forecasting in Thailand. WOA facilitates the generation of multiple parameters required for both HW and decomposition. We compare forecasting results obtained from WOA-HW and WOA-D with those derived from the Classical Decomposition (Classic-D) method and the Box-Jenkins model.

2. Research Methods

2.1 Data preparation

Thailand's government revenue is derived from the Revenue Department, the Excise Department, the Customs Department, and Other Agencies. Secondary data for this study was gathered from the Ministry of Finance [1]. This data, organized monthly, spans from October 2012 to April 2023, yielding a total of 127 data points. The dataset was subdivided into two distinct subsets for the purposes of this study. The first subset, known as the training dataset, consists of 115 data points from October 2012 to April 2022. This subset was employed to develop an appropriate forecasting model for each method being considered. The second subset, referred to as the test dataset, includes 12 data points from May 2022 to April 2023.

2.2 The components of a time series

Understanding the components of a time series is crucial for the selection of an appropriate forecasting model. In this study, statistical tests and thorough examinations of the time series plot are employed alongside an analysis of the inherent data characteristics to facilitate a comprehensive understanding of these components.

2.2.1 The Runs Test

The Runs Test is a statistical test used to assess the presence of a trend in a time series data. It examines whether consecutive values in the time series tend to be consistently increasing or decreasing. Deseasonalizing a time series using centered moving averages involves calculating the average of observations within a defined window centered around each data point. The time series data (Y^T) was first processed through a central 12-month moving average to eliminate seasonal influences, resulting in a deseasonalized dataset. Subsequently, the Runs test was applied to this deseasonalized data to further examine its characteristics.

2.2.2 The Kruskal-Wallis Test

The Kruskal-Wallis Test is a non-parametric statistical test employed to assess the presence of significant differences between groups. Although it is typically applied to compare multiple independent groups, it can also be adapted to examine seasonal variation in time series data. To isolate the seasonal component, it is necessary to eliminate the trend by calculating $Y^S = Y - Y^T$, where Y^S represents the detrended data, obtained by subtracting the trend component (Y^T) from the original time series (Y). The detrended data Y^S is subsequently utilized in the Kruskal-Wallis testing.

2.2.3 Levene's Test

Levene's Test is a statistical test used to assess the equality of variances across different groups. While it is commonly used for comparing variances in independent samples, it can also be adapted to check for homoscedasticity in a time series context.

2.3 Forecasting model

In conducting this study, the following symbols have been defined:

Y_t represents the time series data at time t.

$\hat{Y}_t$ represents the forecasting data at time t.

ε_t represents the residual, which is assumed to adhere to a normal distribution with constant variance (homoscedasticity), independence, and a mean of zero.

t represents the time period, ranging from 1 to n_1 or n_2 or n, depending on the case. In the case of model construction, $t = 1, 2, ... n_1$ represents the number of data points in the training dataset. In the case of model testing, $t = 1, 2, ... n_2$ represents the number of data points in the test dataset. In the case of future forecasting, $t = 1, 2, ... n$ represents the total number of data points. And s represents the time interval of a season, with a value of 12.

2.3.1 Decomposition Method

The Classic-D method is a time series forecasting technique that separates a time series into its different components, namely trend, seasonal, and residual components. Here are the general steps involved in the classical decomposition forecasting:

Step 1: Data Preparation: Collect the historical time series data intended for forecasting. Ensure that the data is in a suitable format and covers a sufficient time period.

Step 2: Visualize the Time Series: Plot the time series data to visualize its overall pattern, trends, and seasonal fluctuations. This step assists in understanding the characteristics of the data.

Step 3: Determine the Seasonal Period: Identify the length of the seasonal cycle in the data. It could be daily, weekly, monthly, or any other pattern that repeats over a fixed time interval.

Step 4: Detrend the Data: Remove the trend component from the time series to isolate the seasonal and residual components. This can be done using techniques such as center the moving averages and subtract them from the original time series (for additive models).

Step 5: Estimate the Seasonal Component: Calculate the seasonal component by averaging the values for each season or period within the seasonal cycle. Adjust the seasonal by subtracting the overall average from each of these to give you the seasonal component (for additive models).

Step 6: Remove the Seasonal Component: Subtract the seasonal component from the original time series (for an additive model) to obtain the deseasonalized time series.

Step 7: Estimate the trend Component: Calculate the trend component by linear regression.

Step 8: Reconstruct the Forecast: Combine the forecasted trend and seasonal components to obtain the final forecast for the original time series.

Step 9: Evaluate and Refine: Assess the accuracy of the forecast by comparing it with actual values. If necessary, refine the forecasting model or methodology based on the evaluation results.

In this study, we focus solely on the additive decomposition based on the test results presented in Table 1, discussed further in section 3. The model represented by Eqs. (2.1)-(2.2), is utilized for both modeling and forecasting purposes.

$$Y_t = \beta_0 + \beta_1 t + S_t + \varepsilon_t, \tag{2.1}$$

$$\hat{Y}_t = \hat{\beta}_0 + \hat{\beta}_1 t + \hat{S}_t, \tag{2.2}$$

where β_0, β_1 are the y-intercept, and the slope coefficient, respectively. $\hat{\beta}_0, \hat{\beta}_1$ are the estimated coefficients of β_0, β_1, respectively. S_t is seasonal component at time $t, \hat{S}_t$ estimated seasonal component of S_t.

2.3.2 Smoothing Method

The HW approach considers both additive and multiplicative seasonal patterns. In this study, our focus is exclusively on the additive seasonal pattern based on the findings presented in Table 1, which will be further elaborated upon in section 3. Eqs. (2.3)-(2.6) represent the calculations for the additive HW method, which captures the additive seasonal component.

$$\hat{Y}_{t+p} = \hat{T}_{t+p} + \hat{S}_{t+p-s} \quad \text{for} \quad p = 1, 2, \ldots, \tag{2.3}$$

where $\hat{T}_{t+p} = \hat{T}_t + p\hat{\beta}_t$, $\hat{Y}_{t+p}$ is the forecasted value at time $t+p$, $\hat{T}_t$ is the level of the time series, $\hat{\beta}_t$ is the trend and $\hat{S}_t$ is the seasonality.

$$\hat{T}_t = \alpha(Y_t - \hat{S}_{t-s}) + (1-\alpha)(\hat{T}_{t-1} + \hat{\beta}_{t-1}), \tag{2.4}$$

$$\hat{\beta}_t = \gamma(\hat{T}_t - \hat{T}_{t-1}) + (1-\gamma)\hat{\beta}_{t-1}, \tag{2.5}$$

$$\hat{S}_t = \delta(Y_t - \hat{T}_t) + (1-\delta)\hat{S}_{t-s}, \tag{2.6}$$

where α is the smoothing parameter for level, γ is the smoothing parameter for trend, and δ is the smoothing parameter for seasonal.

The parameters α, γ, and δ are smoothing parameters that range between 0 and 1. These parameters determine the weight given to the current observation and the previous smoothed values when updating the level, trend, and seasonal components. Values closer to 0 result in more smoothing, while values closer to 1 give more weight to recent observations.

2.4 Whale Optimization Algorithm

The WOA method is introduced as a powerful optimization technique for parameter estimation in time series forecasting. Inspired by the hunting behavior of humpback whales, the WOA algorithm has gained recognition for its effectiveness in solving complex optimization problems across various domains. Mirjalili and Lewis [23] proposed the WOA algorithm, drawing significant interest as a nature-inspired optimization algorithm capable of tackling real-world challenges. The number of WOA citations has shown remarkable growth, with 37 citations in 2016 and reaching 7,410 citations by the end of March 2023, reflecting its popularity and impact in solving diverse optimization problems [24]. In this study, we employ the WOA method to estimate the parameters in the HW method and the decomposition technique, leveraging its exploration and exploitation capabilities. The application of the WOA algorithm offers a promising approach to enhance the accuracy and efficiency of forecasting models in time series analysis. We provide a comprehensive overview of the WOA method, including its formulation and implementation, in this section.

Assuming there are N whales and m parameters, each whale's position is represented as $x_i = (x_i^1, x_i^2, ..., x_i^m), i \in \{1, 2, ..., N\}$. During the optimization search process, whales perform three types of movements: encircling prey, bubble-net attacking, and searching for prey. At any given time, a whale updates its position by performing one of these three actions.

2.4.1 Encircling prey

The encircling prey movement in the WOA is inspired by the hunting behavior of humpback whales. It involves the whales surrounding the target in a coordinated manner, maximizing the chances of capturing the prey. In the algorithm, this movement is simulated by updating the positions of the whales based on their current positions and the best solution found so far. The position update equation for the encircling prey movement can be represented as follows:

$$\vec{X}(t+1) = \vec{X}^*(t) - \vec{A} \cdot \vec{D}, \tag{2.7}$$

where $\vec{X}(t+1)$ is the new position of the whale, $\vec{X}^*(t)$ is the best position found by the whale so far, $\vec{A}$ is the amplitude coefficient, and $\vec{D}$ is a randomly generated vector. This movement allows the whales to explore the search space effectively and converge towards promising solutions. The vectors $\vec{A}$ and $\vec{D}$ are calculated as follows:

$$\vec{A} = 2 \cdot \vec{a} \cdot \vec{r} - \vec{a},$$
$$\vec{D} = \left| \vec{C} \cdot \vec{X}^*(t) - \vec{X}(t) \right| \; ; \vec{C} = 2 \cdot \vec{r},$$

where $\vec{a}$ is linearly decreased from 2 to 0 over the course of iterations (in both exploitation and exploration phases), $\vec{r}$ is a random vector in [0, 1], $\vec{C}$ is a coefficient vector, and $\cdot$ is an element-by-element multiplication.

2.4.2 Bubble-net attacking (exploitation phase)

Bubble-net attacking is characterized by the humpback whales encircling the prey within a shrinking circle while following a spiral-shaped path. To capture this simultaneous behavior, a probabilistic approach is incorporated where there is a 50% chance of selecting either the shrinking encircling mechanism or the spiral model to update the positions of the whales during the optimization process. The mathematical model is defined as follows:

$$\vec{X}(t+1) = \vec{X}^*(t) - \vec{A} \cdot \vec{D} \quad \text{if } p < 0.5,$$
$$\vec{X}(t+1) = \vec{D}' \cdot e^{bl} \cdot \cos(2\pi l) + \vec{X}^*(t) \quad \text{if } p \geq 0.5, \tag{2.8}$$

where p is a random number in [0,1], $\vec{D}' = \left| \vec{X}^*(t) - \vec{X}(t) \right|$ and indicates the distance of the i^{th} whale to the prey (best solution obtained so far), b is a constant for defining the shape of the logarithmic spiral, l is a random number in [-1, 1]. Besides the bubble-net technique, humpback whales also undertake random searches for prey. The mathematical model for this search process is as follows.

2.4.3 Search for prey (Exploration phare)

This exploration mechanism allows the whales to venture into different areas of the search space, increasing the chances of discovering better solutions. The mathematical model for the exploration phase involves randomly updating the position of the whales based on the following equations:

$$\vec{D} = \left| \vec{C} \cdot \vec{X}_{rand} - \vec{X}(t) \right|,$$

$$\vec{X}(t+1) = \vec{X}_{rand} - \vec{A} \cdot \vec{D}, \quad (2.9)$$

where $\vec{X}_{rand}$ is a random position vector (a random whale) chosen from the current population. This mechanism, coupled with the emphasis on exploration where $|\vec{A}| \geq 1$, enhances the algorithm's capability to explore a diverse range of solutions and prevent being trapped in local optima. The three movements are illustrated in the pseudo-code provided in Fig. 1., which details this procedure.

```
Input: the number of whales: N, maximum iterations: T_max, time limit:
MaxTime, the fitness value fails to improve after a specified: T_improve, the
bound of search area: Range Initialize X_i = (x_i^1, x_i^2, ..., x_i^m), X*
While (t < T_max or time < MaxTime or the fitness value fails to improve
after a specified T_improve)
  For i = 1 to N
    Check if any search agent goes beyond the search space and amend it
    For m = 1 to 3
      p = rand[0,1]
      Update a, r, A, C, D, D', b, l, X_rand
      If p ≥ 0.5 then
        Update x_i^m the position of the current search agent by the Eq. (2.8)
      Elseif p < 0.5 and |A| < 1
        Update x_i^m the position of the current search agent by the Eq. (2.7)
      Elseif p < 0.5 and |A| ≥ 1
        Update x_i^m the position of the current search agent by the Eq. (2.9)
      Endif
    End for
  End for
  Calculate fitness(X_i) using HW/Decomposition
  Update X* if there is a better solution
  t = t + 1
End while
Return X*
```

Fig. 1. Pseudo-code of WOA.

In this research, all three types of whale movements were utilized to enhance the forecasting model's performance. The 'Encircling prey' movement was employed for its effectiveness in local search optimization, focusing on areas of the solution space with high fitness values. The Bubble-net attacking' technique was incorporated due to its prowess in balancing exploration and exploitation, aiding in a more diverse search. Lastly, the 'Search for prey' movement was included for its broad exploration capabilities, allowing the algorithm to scan a larger solution space efficiently.

2.5 Hybrid the Whale Optimization Algorithm with Holt-Winters (WOA-HW)

The selection of parameters has a great impact on the performance of the WOA-HW model. In this study, WOA was used to solve the α, γ, and δ parameters of HW. The computational procedure is described as follows:

Step 1: Generate initialization parameters: The position of each whale represented the three parameters $\vec{X}_i = (\alpha_i, \gamma_i, \delta_i)$ within the bound of [0, 1].

Step 2: Select movement: Each whale chooses an action for its next move by selecting one of the three movements. This selection determines whether the whale moves closer to the prey or expands its search. Refer to pseudo-code Fig. 1. for a detailed illustration of this procedure. By input of whale are $N = 30$, $T_{\max} = 1{,}000$, $MaxTime = 180$ sec., $T_{improve} = 50$, and $m = 3$.

Step 3: Calculate fitness: The fitness is computed by evaluating the Root Mean Square Error (RMSE) of the dataset using the position of each whale in HW method. Therefore, we could judge the best position $(\vec{X}^*)$.

Step 4: Establish stopping criteria: The algorithm terminates when the fitness value fails to improve after a specified number of iterations, reaches a time limit, or reaches the maximum number of iterations $T_{\max}$. At this point, the optimal parameters are output.

Step 5: Evaluation: Evaluate HW performance with the optimal parameters. We assess the forecast accuracy by measuring the error between the forecasted data and the dataset.

The objective function of WOA-HW uses the following equation:

Objective Minimize RMSE(α,γ,δ),

$$\text{Variable range} \begin{cases} 0 \le \alpha \le 1 \\ 0 \le \gamma \le 1, \\ 0 \le \delta \le 1 \end{cases}$$

where $\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(Y_t - \hat{Y}_t\right)^2}$, n represents for the length of dataset; Y_t and $\hat{Y}_t$ refer to the actual value and forecasting value of WOA-HW, respectively.

2.6 Hybrid the Whale Optimization Algorithm with Decomposition (WOA-D)

The selection of parameters has a great impact on the performance of the WOA-D model. In this study, WOA was used to solve the $\hat{\beta}_0, \hat{\beta}_1, \hat{S}_1, \hat{S}_2, \hat{S}_3, \hat{S}_4, \hat{S}_5, \hat{S}_6, \hat{S}_7, \hat{S}_8, \hat{S}_9, \hat{S}_{10}, \hat{S}_{11}$ and $\hat{S}_{12}$ parameters of decomposition. The computational procedure is described as follows:

Step 1: Set Constraint upper bound and lower bound of parameters: Calculate the trend component by linear regression to get $\hat{\beta}'_0$, and $\hat{\beta}'_1$. Set the constraint upper bound and lower bound of parameters of $\hat{\beta}_0$, and $\hat{\beta}_1$ using the following equations:

Constraint upper bound

$$\hat{\beta}_0 = 1.2\hat{\beta}'_0,\ \hat{\beta}_1 = 1.2\hat{\beta}'_1.$$

Constraint lower bound

$$\hat{\beta}_0 = 0.8\hat{\beta}'_0,\ \hat{\beta}_1 = 0.8\hat{\beta}'_1.$$

These constraints apply when the parameter is positive. Conversely, if the parameter is negative, the upper and lower bounds should be switched.

Remove the trend component from the time series using differencing ($\Delta Y_t = Y_t - Y_{t-1}$). Set the constraint upper bound and lower bound of parameters of $\hat{S}_1, \hat{S}_2, \ldots, \hat{S}_{12}$ by using the following equations:

Constraint upper bound (US) is +(extreme value) of amplitude of ΔY_t.

Constraint lower bound (LS) is −(extreme value) of amplitude of ΔY_t.

Step 2: Generate initialization parameters: The position of each whale is represented by the 14 parameters $\vec{X}_i = (\hat{\beta}_{0i}, \hat{\beta}_{1i}, \hat{S}_{1i}, \hat{S}_{2i}, \hat{S}_{3i}, \hat{S}_{4i}, \hat{S}_{5i}, \hat{S}_{6i}, \hat{S}_{7i}, \hat{S}_{8i}, \hat{S}_{9i}, \hat{S}_{10i}, \hat{S}_{11i}, \hat{S}_{12i})$ within the bound of [0, 1]. These parameters can be adjusted to the appropriate units before calculating the fitness, alleviating any concerns about their values.

Step 3: Select movement: Each whale chooses an action for its next move by selecting one of the three movements.

Step 4: Scaling parameters: We configure the whale algorithm to search for parameters within the boundary of [0, 1]. Therefore, it is necessary to adjust the units of the parameters before calculating the fitness value. The following equation is employed for this purpose:

Original Value = *Scaled Value* × (*Constraint Upper Bound* − *Constraint Lower Bound*) + *Constraint Lower Bound*.

If the original values are denoted as $(\hat{S}_i)$, the seasonal adjustment can be calculated using the formula:

Adjust $\hat{S}_i = \hat{S}_i - \sum_{i=1}^{12} \hat{S}_i / 12$, then $\sum_{i=1}^{12} \hat{S}_i = 0$.

(Note: Multiplicative decomposition Adjust $\hat{S}_i = 12 \times \hat{S}_i / \sum_{i=1}^{12} \hat{S}_i$, then $\sum_{i=1}^{12} \hat{S}_i = 12$)

In this equation, the original value represents the parameter value based on the original data unit, while the scaled value is the value obtained by the whale algorithm within the range of [0, 1]. This step holds significant importance, especially when dealing with parameters of different units and a considerable number of parameters.

We will continue to refer to pseudo-code Fig. 1. for a detailed illustration of this procedure. However, it is important to note that we have added the "Scaling parameters" line before the step "Calculate $fitness(X_i)$ using decomposition". The input values for the whale algorithm are as follows: $N = 30$, $T_{\max} = 1{,}000$, $MaxTime = 180$ sec., $T_{improve} = 50$, and $m = 14$.

Step 5: Calculate fitness: The fitness is computed by evaluating the RMSE of the dataset using the position of each whale in decomposition method. Therefore, we could judge the best position $(\vec{X}^*)$.

Steps 6-7: Establish stopping criteria: and *Evaluation*: Follow the same procedure as in Steps 4-5 of the WOA-HW process. The objective function of WOA-D uses the following equation:

Objective Minimize RMSE $(\hat{\beta}_0, \hat{\beta}_1, \hat{S}_1, \hat{S}_2, \hat{S}_3, \hat{S}_4, \hat{S}_5, \hat{S}_6, \hat{S}_7, \hat{S}_8, \hat{S}_9, \hat{S}_{10}, \hat{S}_{11}, \hat{S}_{12})$,

Variable range $\begin{cases} 0.8\hat{\beta}_0' \le \hat{\beta}_0 \le 1.2\hat{\beta}_0' \\ 0.8\hat{\beta}_1' \le \hat{\beta}_1 \le 1.2\hat{\beta}_1' \\ \text{LS} \le \hat{S}_i \le \text{US}) \text{ for } i = 1,2,...,12 \end{cases}$,

where $\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(Y_t - \hat{Y}_t\right)^2}$, n represents the length of dataset; Y_t and $\hat{Y}_t$ refer to the actual value and forecasting value of WOA-D, respectively.

2.7 Evaluation criteria

To compare the performance of the model more clearly, we adopt the following three metrics to evaluate the effect of the model: RMSE, Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). The formulas are as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n_2}\sum_{t=1}^{n_2}\left(Y_t - \hat{Y}_t\right)^2},$$

$$\text{MAE} = \frac{1}{n_2}\sum_{t=1}^{n_2}\left|Y_t - \hat{Y}_t\right|,$$

$$\text{MAPE} = 100 \times \frac{1}{n_2}\sum_{t=1}^{n_2}\left|Y_t - \hat{Y}_t\right| \Big/ Y_t,$$

where n_2 represents the length of the test dataset; Y_t and $\hat{Y}_t$ refer to the actual value and forecasting value of the test dataset, respectively. If n_2 is replaced with n_1, all three metrics will be adjusted to evaluate the performance during the training phase of the dataset.

The experiments were conducted in a Google Colab environment [25], which automatically provides access to Python programming capabilities. The hardware specifications for the Colab environment are equivalent to a machine with an Intel(R) Core(TM) i5-9400 CPU @ 2.90GHz, 16GB RAM, running on a Windows 11 operating system.

3. Experiment and Discussion

3.1 Results of time series motion characteristics analysis

The trend, seasonal, and variance characteristics of the data were assessed using the Runs test, Kruskal-Wallis test, and Levene's test, respectively, as outlined in Table 1. Further examination of the time series plot and the inherent nature of the data enhanced the analysis. The Runs test indicated a trend in all four sectors of government income, supported by a P-value less than the predetermined significance level. Similarly, the Kruskal-Wallis test revealed seasonal variations for all four sectors, also supported by a *P*-value less than the predetermined significance level. Additionally, we employed the Autocorrelation function (ACF) validation in the Box-Jenkins model construction process, which helped us identify trends and seasonality as detailed in Tables 2-5. In fitting the Box-Jenkins model, we set d = 1 for all data sets and D = 1 for data sets corresponding to the Revenue Department, Customs Department, and Other Agencies. Even though the ACF did not support that the Excise Department has a seasonal component, when considered in conjunction with the Kruskal-Wallis test and the nature of the data, we decided that the Excise Department also has seasonality. Lastly, all four sectors exhibited constant variance, corroborated by a P-value not less than the predetermined significance level. Consequently, the additive decomposition approach and the additive HW method were considered appropriate for this analysis.

The additive decomposition and HW model were selected for this study based on its suitability for handling linear seasonal patterns and consistent trend components in the time series data. In comparison with multiplicative models, additive decomposition offers greater simplicity and is particularly effective for data where variations around the trend do not significantly vary with the level of the time series. Given these strengths and the specific characteristics of the dataset, both the additive decomposition and the HW models stand out as highly reliable frameworks for accurate forecasting. These methods are well-suited to handle all four components of the government revenue data.

3.2 Results of the experiment

In order to evaluate the performance of WOA-HW and WOA-D in forecasting government revenue, we compared these models with two classical forecasting models: Classic-D, and the Box and Jenkins model. These models, when applied to the same data, yielded different results, as shown in Tables 2-5. Moreover, Figs. 2-5 illustrate the fitting curves for both the train and test datasets across the four government revenue sectors.

Table 1. The test statistics and P-values from Runs, Kruskal-Wallis, and Levene's tests for each of the four revenue sectors.

Government Revenue	Runs Test		Kruskal Wallis Test		Levene's Test	
	Test Statistic	*P*-value	Test Statistic	*P*-value	Test Statistic	*P*-value
Revenue	-2.2578	0.000*	87.09	0.000*	0.15	0.701
Excise	-7.5261	0.000*	37.39	0.000*	0.34	0.562
Customs	-1.8815	0.000*	38.88	0.000*	0.22	0.643
Other	1.8815	0.000*	44.57	0.000*	2.78	0.098

Note: * Significant Level 0.05.

Table 2. Evaluation criteria of the Revenue Department.

Method	Parameter	Train Dataset			Test Dataset		
		MAPE	RMSE	MAE	MAPE	RMSE	MAE
WOA-HW	$(\alpha,\gamma,\delta) =$ 0.09, 0, 0.53	7.34	19,657	12,127	8.08*	22,389*	16,951*
WOA-D	$(\hat{\beta}_0,\hat{\beta}_1,\hat{S}_1,\hat{S}_2,\hat{S}_3,\hat{S}_4,\hat{S}_5,\hat{S}_6,\hat{S}_7,\hat{S}_8,\hat{S}_9,\hat{S}_{10},\hat{S}_{11},\hat{S}_{12}) =$ 141359.1, 202.1, -37785.9, -26859.5, -21172.8, -18396, -31731.6, -24039.6, -29764.6, 74336.5, 57410.3, -35741.9, 33141.7, 60603.4	6.09*	16,383*	10,208*	9.61	29,813	20,303
Classic-D	$(\hat{\beta}_0,\hat{\beta}_1,\hat{S}_1,\hat{S}_2,\hat{S}_3,\hat{S}_4,\hat{S}_5,\hat{S}_6,\hat{S}_7,\hat{S}_8,\hat{S}_9,\hat{S}_{10},\hat{S}_{11},\hat{S}_{12}) =$ 141447, 199.3, -34434.5, -25808.6, -20902.9, -17318.7, -32068.4, -21981.7, -28067.9, 70216.5, 56629.8, -33776.8, 32937.6, 54575.7	6.19	16,573	10,274	9.43	31,084	20,483
Box-Jenkins	$SARIMA(p,d,q)(P,D,Q)_{12} = SARIMA(0,1,1)(1,1,0)_{12}$	17.96	54,857	27,198	8.73	26,234	19,093

Note: * Minimum Value

Table 3. Evaluation criteria of the Excise Department.

Method	Parameter	Train Dataset			Test Dataset		
		MAPE	RMSE	MAE	MAPE	RMSE	MAE
WOA-HW	$(\alpha,\gamma,\delta) =$ 0.18, 0, 0.13	10.80*	6,765	4,327*	14.73*	5,819*	5,601*
WOA-D	$(\hat{\beta}_0,\hat{\beta}_1,\hat{S}_1,\hat{S}_2,\hat{S}_3,\hat{S}_4,\hat{S}_5,\hat{S}_6,\hat{S}_7,\hat{S}_8,\hat{S}_9,\hat{S}_{10},\hat{S}_{11},\hat{S}_{12}) =$ 35267.6, 125.1, -2380.3, -485.7, 3197.7, 3524.5, -243.8, 3912.5, 433, -2806.9, -2890.7, 136.7, -1597.4, -799.5	13.22	6,560*	4,968	31.91	12,288	12,160
Classic-D	$(\hat{\beta}_0,\hat{\beta}_1,\hat{S}_1,\hat{S}_2,\hat{S}_3,\hat{S}_4,\hat{S}_5,\hat{S}_6,\hat{S}_7,\hat{S}_8,\hat{S}_9,\hat{S}_{10},\hat{S}_{11},\hat{S}_{12}) =$ 35235, 125.8, -1688.0, 738.5, 2280.7, 1976.8, -204.7, 4423.2, 2664.6, -948.1, -1891.4, -3125.6, -1806.2, -2419.9	12.44	6,728	4,954	32.01	12,325	12,207
Box-Jenkins	$ARIMA(p,d,q) = ARIMA(0,1,1)$	11.45	7,252	4,521	18.04	7,141	6,750

Note: * Minimum Value.

Table 4. Evaluation criteria of the Customs Department.

Method	Parameter	Train Dataset			Test Dataset		
		MAPE	RMSE	MAE	MAPE	RMSE	MAE
WOA-HW	$(\alpha,\gamma,\delta) =$ 0.28, 0, 0.43	6.54	719	583	13.09	2,612	1,599
WOA-D	$(\hat{\beta}_0,\hat{\beta}_1,\hat{S}_1,\hat{S}_2,\hat{S}_3,\hat{S}_4,\hat{S}_5,\hat{S}_6,\hat{S}_7,\hat{S}_8,\hat{S}_9,\hat{S}_{10},\hat{S}_{11},\hat{S}_{12}) =$ 9817.3, -13.7, 116.3, 484.1, 534.4, 191.9, -818.8, 419.4, -421.9, -378.6, -263.3, -89.1, 316.1, -90.5	5.93	646*	518*	20.22	3,045	2,317
Classic-D	$(\hat{\beta}_0,\hat{\beta}_1,\hat{S}_1,\hat{S}_2,\hat{S}_3,\hat{S}_4,\hat{S}_5,\hat{S}_6,\hat{S}_7,\hat{S}_8,\hat{S}_9,\hat{S}_{10},\hat{S}_{11},\hat{S}_{12}) =$ 9806, -13.5, 147.9, 411, 550.1, 223, -795.9, 189.5, -436.1, -297.8, -329.4, 5.7, 322.3, 9.6	5.92*	652	520	20.04	3,039	2,301
Box-Jenkins	$SARIMA(p,d,q)(P,D,Q)_{12} = SARIMA(2,1,0)(0,1,1)_{12}$	16.31	3,267	1,521	10.33*	2,388*	1,310*

Note: * Minimum Value.

Table 5. Evaluation criteria of Other Agencies.

Method	Parameter	Train Dataset			Test Dataset		
		MAPE	RMSE	MAE	MAPE	RMSE	MAE
WOA-HW	$(\alpha,\gamma,\delta) =$ 0.03, 0, 0.41	38.79	14,198	10,019	27.46*	11,424	8,108*
WOA-D	$(\hat{\beta}_0,\hat{\beta}_1,\hat{S}_1,\hat{S}_2,\hat{S}_3,\hat{S}_4,\hat{S}_5,\hat{S}_6,\hat{S}_7,\hat{S}_8,\hat{S}_9,\hat{S}_{10},\hat{S}_{11},\hat{S}_{12}) =$ 26040.8, 33.8, 22778.3, -11072.9, 3559.4, -2935.8, -5429.2, -7380, 17467.3, 2661.3, -10242.6, 5652.5, -9046, -6012.4	37.59*	12,560*	9,234	35.34	9,443*	8,404
Classic-D	$(\hat{\beta}_0,\hat{\beta}_1,\hat{S}_1,\hat{S}_2,\hat{S}_3,\hat{S}_4,\hat{S}_5,\hat{S}_6,\hat{S}_7,\hat{S}_8,\hat{S}_9,\hat{S}_{10},\hat{S}_{11},\hat{S}_{12}) =$ 26041, 33.8, 24847.2, -14700.6, -1988.9, -50.4, -4797, -7909.4, 18351.2, 1299.3, -8192.2, 9096.2, -8176.9, -7778.6	37.89	12,826	9,130*	34.63	9,579	8,263
Box-Jenkins	$SARIMA(p,d,q)(P,D,Q)_{12} = SARIMA(0,1,2)(2,1,0)_{12}$	47.22	15,171	11,239	28.34	12,725	8,711

Note: * Minimum Value.

Table 6. Percentage change of WOA-D parameters relative to Classic-D.

Revenue of Government	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{S}_1$	$\hat{S}_2$	$\hat{S}_3$	$\hat{S}_4$	$\hat{S}_5$	$\hat{S}_6$	$\hat{S}_7$	$\hat{S}_8$	$\hat{S}_9$	$\hat{S}_{10}$	$\hat{S}_{11}$	$\hat{S}_{12}$
Revenue Department	0	1	10	4	1	6	1	9	6	6	1	6	1	11
Excise Department	0	1	41	166	40	78	19	12	84	196	53	104	12	67
Customs Department	0	2	21	18	3	14	3	121	3	27	20	1,655	2	1,042
Other Agencies	0	0	8	25	279	5,725	13	7	5	105	25	38	11	23

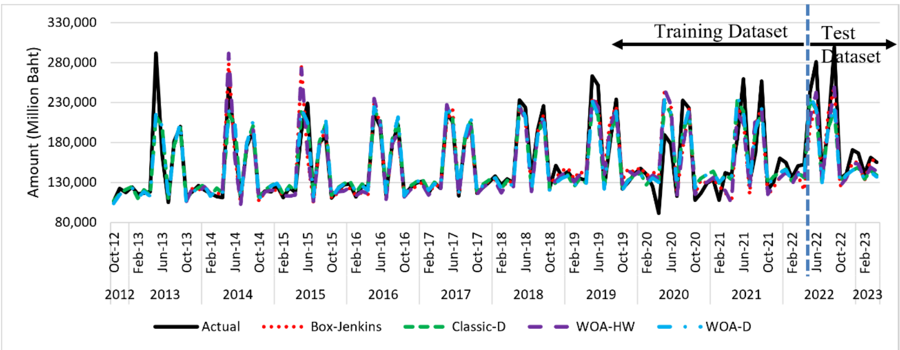

Fig. 2. Fitting curves of training dataset and test dataset for the Revenue Department.

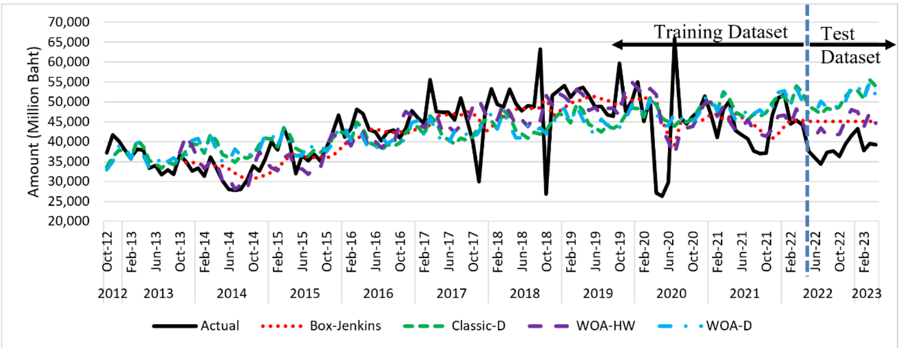

Fig. 3. Fitting curves of training dataset and test dataset for the Excise Department.

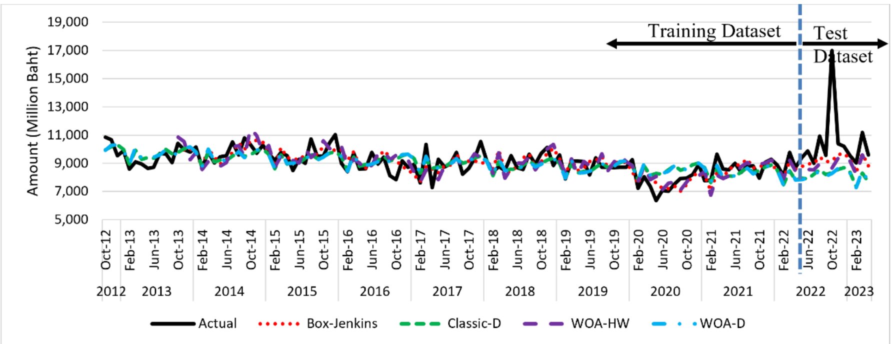

Fig. 4. Fitting curves of training dataset and test dataset for the Customs Department.

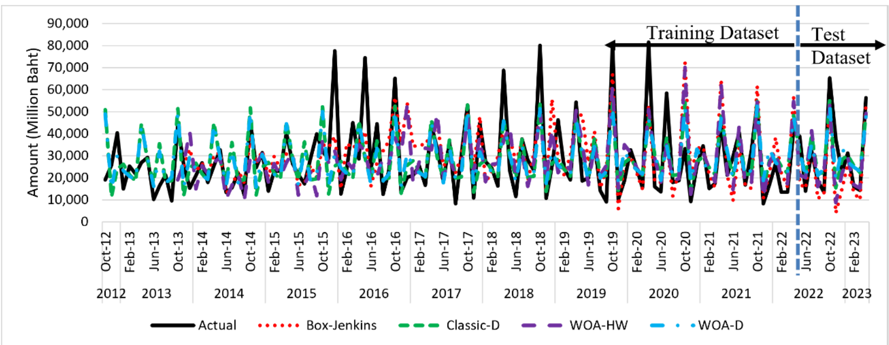

Fig. 5. Fitting curves of training dataset and test dataset for the Other Agencies.

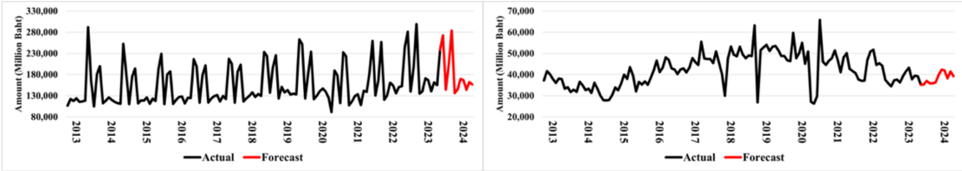

Fig. 6. Forecast curves for the Revenue Department (a), and the Excise Department (b)

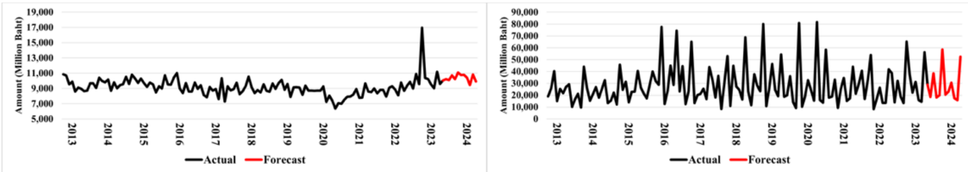

Fig. 7. Forecast curves for the Customs Department (a), and the Other Agencies (b).

Box-Jenkins is a statistical model commonly used in forecasting or for comparison with new models that are being introduced, such as in [4-6, 13, 22]. In this research, the Box-Jenkins model construction process is briefly summarized as follows: The parameters for the Box-Jenkins ARIMA model were computed using a multi-step approach. First, the consistency of the time series was examined through the analysis of the time series curve, the Autocorrelation Function, and the Partial Autocorrelation Function (PACF) curves. The Dickey-Fuller (DF) test was utilized for unit root checking, with the maximum order of Lag set at 12. The Bayesian Information Criterion (BIC) was employed to determine the lowest value. If the time series was found to be non-stationary, it was transformed to become stationary in order to obtain d and D values. A bounding set of p, q, P, and Q values ranging from 0 to 3 was employed. To find suitable parameters, a grid search method, a machine learning technique for hyperparameter tuning, was utilized. This approach proved more effective than the traditional method of considering ACF and PACF curves, which can be challenging to determine when multiple possibilities exist. Therefore, instead of dealing with 256 possible models, the grid search method streamlined the process.

Tables 2-5 illustrate that the WOA-D method outperformed the other three models in three datasets during the training phase. Specifically, the Revenue Department showed the lowest values across all three evaluation criteria. In terms of the Customs Department and Other Agencies, WOA-D yielded the lowest values in two out of three evaluation criteria. However, for the Excise Department, the WOA-HW method recorded the lowest values in two evaluation criteria. Notwithstanding, the WOA method demonstrated strong performance across all datasets when applied to parameter value searches for model fitting. During the training phase, the only model that surpassed the WOA method in performance was the Classic-D method applied to the Customs Department data, which resulted in the lowest MAPE. Also, when the Classic-D method was utilized for Other Agencies, it produced the lowest MAE. Thus, it can be concluded that the WOA method excelled in 10 out of 12 evaluation metrics across the four datasets, with three metrics evaluated per dataset.

The WOA-D method outperforms the Classic-D in parameter tuning, primarily due to its capability to search for all parameters simultaneously. In contrast, the Classic-D method conducts a search for each parameter sequentially, which might lead to a local optimization situation. However, the simultaneous parameter search approach of the WOA-D allows for a broader exploration of the solution space, thus increasing the likelihood of achieving global optimization. The importance of creating parameter boundaries is highlighted in section 2.6, Hybrid Whale Optimization Algorithm with Decomposition (WOA-D), where Scaling Parameters is identified as a crucial step.

Examining the parameters of the HW forecasting model, one can see it is capable of estimating the parameters α, γ and δ for the time series of the four government revenue sectors, as demonstrated in Tables 2-5. Particularly notable is the relatively large δ (the seasonal smoothing factor) in the Revenue Department, Customs Department, and Other Agencies, standing at 0.53, 0.43, and 0.41, respectively. This indicates that the influence of the previous year's season carries a weight of approximately 40-50%, diminishing exponentially for each preceding year. However, $\gamma = 0$ across all four government revenue sectors indicates that the model will not update the trend component based on new data. It should not be interpreted as an absence of trend within the time series but rather as an assumption by the model that the trend will remain constant throughout the forecast period. The level adjuster also plays a role in fitting the trend to the dataset, as expressed in Eq. (2.3) and $\hat{T}_{t+p} = \hat{T}_t + p\hat{\beta}_t$.

This study examines two methods of forecasting with decomposition: Classic-D and WOA-D. Given the model's 14 parameters, distinct outcomes were observed from each method. The percentage change of WOA-D parameters relative to Classic-D is presented in Table 6. The Revenue Department exhibited the smallest divergence in parameters, with the most considerable discrepancy of 11% occurring in December. The Excise Department displayed a moderate variation in parameters, peaking at a 196% difference in August. Similarly, the Customs Department showed moderate parameter disparity, with the maximum divergence reaching 1,655% in October. Other Agencies demonstrated the largest parameter differences, with a whopping 5,725% discrepancy in April. The significant parameter differences underscore the importance of the method used to derive the model parameters. The superior performance of WOA-D in optimizing the fitness objective lends credence to the effectiveness of the WOA-D and WOA-HW techniques.

3.3 Forecasting government revenue for the next 12 months

In pursuit of the most suitable model for forecasting government revenue, the performance of different models was evaluated using the test dataset. The WOA-HW model showed commendable performance across three datasets: Revenue Department, Excise Department, and Other Agencies. These results indicate that the WOA-HW model can effectively be used to forecast the subsequent 12 months, outperforming all other models in the comparison. For the Customs Department dataset, the Box-Jenkins model emerged as the top performer, satisfying the assumptions of residual normal distribution, constant variance (homoscedasticity), independence, and a mean of zero. As a result, both the WOA-HW and Box-Jenkins models were utilized to forecast future government revenue across the four sectors. The forecasting results are visually represented in Figs. 6-7.

As depicted in Fig. 6(a) for the Revenue Department, the WOA-HW model projects a pattern akin to the preceding 12 months. This prediction aligns with the δ value of 0.53, as reported in Table 2. The closer the values are to 1, the more emphasis they place on recent observations. A notably low MAPE of 8.08 in the test dataset implies the high accuracy of this forecasting model. In the case of the Excise Department, portrayed in Fig. 6(b), the model has three estimated parameters outlined in Table 3, all relatively small. Lower values correspond to increased smoothing in the model. Again, a low MAPE value of 14.73 in the test dataset signals the model's high forecasting accuracy. Turning to Fig. 7(a), the Box-Jenkins model maintains high accuracy in forecasting, as suggested by its low MAPE value of 13.09 in the test dataset. Fig. 7(b) showcases the Other Agencies model, which anticipates a pattern similar to the previous 12 months. The δ value of 0.41, as detailed in Table 5, indicates that recent seasonal patterns are somewhat weighted but are also smoothed due to the high variability in the data. The MAPE value in the test dataset stands at 27.46, the highest among all datasets. This is consistent with the inherent nature of the data, which encompasses revenue from various sources and hence exhibits higher variability than the first three revenue sources.

4. Conclusions

The results of this study demonstrate a successful forecasting of government revenue, with different sectors emerging as key factors. Therefore, it is necessary to make informed decisions when formulating policies for government tax collection or projecting future revenue. Notably, seasonal influences are discernible across all sectors of government revenue.

In this study, an approach incorporating the Whale Optimization Algorithm with Holt-Winters (WOA-HW) and Decomposition (WOA-D) was introduced to forecast future government revenue in Thailand. This proposed method demonstrated superior performance over traditional models in forecasting Thailand's government revenue. Regardless of the sources of government revenue represented in the time series data, the WOA-HW and WOA-D models offer superior forecasting compared to Classic-D. Consequently, users can trust in the models' ability to handle level adjustment, trend considerations, and seasonal components of the time series data. However, it remains inconclusive as to which method, either WOA-HW or WOA-D, is superior. Therefore, if the data exhibit both trend and seasonal components, it is recommended to employ both forecasting methods and select the one that yields the most accurate results based on predetermined criteria. The WOA can effectively determine the relevant parameters. In conclusion, this paper contributes an effective tool for studying time series data. In future research, enhanced versions of WOA-HW or WOA-D will be deployed to forecast other time series data.

Acknowledgments

This study is supported by the Data Science Research Center, Department of Statistics, Faculty of Science, Chiang Mai University, Thailand.

References

[1] Ministry of Finance Operation Center. Ministry of Finance. Available from: https://dataservices.mof.go.th/menu3?id=2&page=&freq=month&mf=10&yf=2565&mt=4&yt=2566&sort=asc&search_text=.

[2] Tax Planning Office. The Excise Department. Available from: https://www.excise.go.th/cs/groups/public/documents/document/dwnt/ndcw/~edisp/uatucm470306.pdf.

[3] Persons W. Indices of business conditions. Review of Economics and Statistics, 1919;1:5-110.

[4] Mbuli N, Mathonsi M, Seitshiro M, Pretorius JHC. Decomposition forecasting methods: A review of applications in power systems. Energy Reports 6, 2020;298-306.

[5] Koirala TP. Government revenue forecasting in Nepal. NRB Economic Review, 2012;24(2):47-60.

[6] Hansen JV, Nelson RD. Neural networks and traditional time series methods: A synergistic combination in state economic forecasts. IEEE Transactions on Neural Networks, 1997;8(4):863-73.

[7] Gardner ES. Exponential smoothing: The state of the art. Journal of forecasting, 1985;4(1):1-28.

[8] Brown RG. Statistical forecasting for inventory control. McGraw/Hill; 1959.

[9] Holt CC. Forecasting seasonals and trends by exponentially weighted moving averages. Report to the Office of Naval Research (ONR 52);1957.

[10] Holt CC. Forecasting seasonals and trends by exponentially weighted moving averages. International Journal of Forecasting, 2004;20(1):5-10.

[11] Winters PR. Forecasting Sales by Exponentially Weighted Moving Averages. Management Science, 1960;6(3):324-42.

[12] Frank HA. Municipal revenue forecasting with time series models: A Florida case study. The American review of public administration, 1990;20(1):45-59.

[13] Gianakis GA, Frank HA. Implementing Time Series Forecasting Models: Considerations for Local Governments. State and Local Government Review, 1993;25(2):130-44.

[14] Williams DW. Local government revenue forecasting methods: Competition and comparison. Journal of Public Budgeting, 2016;28(4):488-526.

[15] Ortiz RRL. The accuracy rate of Holt-Winters model with particle swarm optimization in forecasting exchange rates. Journal of Computers, 2016;11(3):216-24.

[16] Tresnani HW, Sihabuddin A, Mustofa K. Parameter optimization in grey Holt-Winter exponential smoothing using golden section. Journal of Mathematics and Natural Science, 2018;25(3):312-25.

[17] Hani'ah M, Putri IK, Ririd ARTH. Parameter optimization of Holt-Winters exponential smoothing using golden section method for predicting indonesian car sales. In Proceedings of IEIT 2021: 1st International Conference on Electrical and Information Technology; 2021. pp. 21-6.

[18] Jiang W, Wu X, Gong Y, Yu W, Zhong X. Holt–Winters smoothing enhanced by fruit fly optimization algorithm to forecast monthly electricity consumption. Energy, 2020;193:116779.

[19] Hani'ah M, Putri IK, Ririd ARTH. Genetic algorithms for Holt Winter exponential smoothing parameter optimization in indonesian car sales forecasting. In Proceedings of the 2022 Annual Technology, Applied Science and Engineering Conference (ATASEC 2022). Atlantis Press; 2022. pp. 159-71.

[20] Zhou W, Tao H, Jiang H. Application of a novel optimized fractional grey Holt-Winters model in energy forecasting. Sustainability, 2022;14(3118):1-18.

[21] Mauricio CC, Ostia CF. Cuckoo search algorithm optimization of Holt-Winter Method for distribution transformer load forecasting. In Proceeding of the 2023 9th International Conference on Control, Automation and Robotics (ICCAR), Beijing, China; 2023. pp. 36-42.

[22] Liu L, Wu L. Predicting housing prices in China based on modified Holt's exponential smoothing incorporating whale optimization algorithm. Socio- Economic Planning Sciences, 2020;72:1-8.

[23] Mirjalili S, Lewis A. The whale optimization algorithm. Advances in Engineering Software, 2016;95:51-67.

[24] Nadimi-Shahraki MH, Zamani H, Varzaneh ZA, Mirjalili S. A systematic review of the whale optimization algorithm: theoretical foundation, improvements, and hybridizations. Archives of Computational Methods in Engineering, 2023;1-47.

[25] Google Colab. Overview of Colaboratory Features. Available form: https://colab.research.google.com/notebooks/intro.ipynb