Songklanakarin J. Sci. Technol.
46 (2), 102–109, Mar. – Apr. 2024

http://www.sjst.psu.ac.th

*Original Article*

# Forecasting the number of Thai overseas workers through a better model selection (assuming no pandemic)

Wipada Phonduang[1], Phontita Thiuthad[1*], Nabendu Pal[2], and Pattaraporn Tusto[3]

[1] *Statistics and Applications Research Unit, Division of Computational Science, Faculty of Science, Prince of Songkla University, Hat Yai, Songkhla, 90110 Thailand*

[2] *Department of Mathematics, University of Louisiana at Lafayette, Lafayette, Louisiana, 70504 United States of America*

[3] *Mathematics Curriculum, Faculty of Science and Technology, Phranakhon Rajabhat University, Bang Khen, Bangkok, 10220 Thailand*

Received: 21 July 2023; Revised: 23 February 2024; Accepted: 29 February 2024

## Abstract

Thailand's economy, like many other ASEAN countries, heavily relies on remittances from Thai overseas workers. Predicting the number of such workers in the coming years is crucial for stakeholders engaged in economic planning and growth. This study aimed to find a reliable statistical model for this purpose, utilizing data from pre-pandemic years to compare two predictive methods. The dataset spans 120 months, from January 2010 to December 2019, and was obtained from the Library System, Department of Employment of the Royal Thai Government. The chosen forecasting techniques are the Box-Jenkins method and Holt-Winter's Exponential Smoothing method. Primary parameters were first estimated based on the given data, and then evaluated through the Bayesian Information Criterion (BIC). To assess model performance, the dataset was split into two subsets. The primary parameters were applied to the first 60 months' data (January 2010 to December 2014) to estimate secondary parameters. The remaining 60 months' data were then used to evaluate how well each model predicts the response variable, measured by the prediction mean absolute error (PMAE) and prediction root mean squared error (PRMSE). Based on the findings, the Box-Jenkins method outperformed Holt-Winter's method in forecasting the monthly number of fresh Thai overseas workers. This study's insights can serve as a valuable template for predicting overseas workers' numbers in other ASEAN countries with socio-economic-cultural contexts similar to Thailand. Accurate predictions can aid decision-making and planning for sustainable economic development in the region.

**Keywords**: Box-Jenkins' SARIMA method, Holt-Winter's Exponential Smoothing method, Thai overseas workers, prediction mean absolute error (PMAE), prediction root mean squared error (PRMSE)

## 1. Introduction

Back in the 1960s, Thailand implemented a successful family planning program to tame the population explosion by which the population growth rate went down from greater than 3% to below 1% in just 20 years. The success of controlling the population growth rate can be gauged by looking at the latest available figures: the population growth rate in the years 2017 through 2020 had been 0.35%, 0.32%, 0.28% and 0.25% respectively. The population growth rate is expected to come down even lower in 2021 and 2022. Yet, the absolute size of the population is still growing, and by the end of 2022 the population is expected to cross 70 million. The burgeoning population brings many challenges including continued deforestation, worsening of pollution, increasing traffic accidents, and above all - a steady rise in unemployment rate.

*Corresponding author
Email address: phontita.t@psu.ac.th

Rapid urbanization in Thailand has been witnessing a large pool of people entering the organized labor force, and since 2013 the country is experiencing a rising unemployment rate. The unemployment rate, as a percentage of the total labor force, was seen decreasing from a high of approximately 3.4% in 1998 to 0.3% in 2013, but then suddenly it reversed the course, and since 2013 it has risen steadily to 1.4% in 2021 (Overseas Employment Administration Division of Thailand's Ministry of Labour, 2021). As a result, a large number of Thai workers are looking for job opportunities abroad. This not only helps the job seekers earn a respectable living, but also it helps them acquire extra technical knowhow, which can be channelized to help improve the domestic economy of Thailand (Overseas Employment Administration Division of Thailand's Ministry of Labour, 2017). That is why it is one of the primary strategies of the Royal Thai Government's Department of Employment to help the labor force look for overseas employment opportunities, and thereby alleviate the current domestic unemployment rate.

Overseas employment of Thai workers has resulted in a positive change in the overall economy; not only does it help the country earn a good amount of foreign exchange, but it also promotes developing the human resources. At a personal level, many Thai workers are eager to take up overseas employment due to various factors, which range from a lack of domestic work opportunities that suits their skills and demands, to helping the family's social standing (Sansiri, 2008).

The aforementioned prevailing socio-economic condition in Thailand has prompted us to study the recent historical data on overseas employment, model it using a suitable method, and then forecast the number of Thai overseas workers, which in turn can help the country's policy makers take necessary steps from a logistical point of view.

In this research we have used two available model types to fit the monthly Thai overseas employment data, namely - (a) Box-Jenkins' 'Seasonal Autoregressive Integrated Moving Average' (SARIMA) model, and (b) Holt - Winter's 'Exponential Smoothing' (ES) model. The objective of this study was to determine which of these two models can provide the most accurate prediction of the number of Thai overseas workers (in the pre-pandemic setting) based on existing data.

## 2. Materials and Methods

In this research, the monthly Thai overseas employment data were gathered from the Library System, Department of Employment of the Royal Thai Government (Overseas Employment Administration Division of Thailand's Ministry of Labour, 2021) from Jan 2010 to Dec 2019, covering a period of 120 months, say $\{Y_1, Y_2, \ldots, Y_{120}\}$, *i.e.*, $Y_t$ represents the number of fresh (or first time) Thai overseas workers in month-$t$, with $t$=1 representing Jan 2010, and $t$=120 standing for Dec 2019. Figure 1 illustrates the time series plot of monthly first time Thai overseas workers from Jan 2010 to Dec 2019. The data can be seen to have a volatile time series pattern.

This is due to the fact that it has a trend component and changes (i.e., variations) in the same way as it climbs up or down over a given period. In other words, there is a seasonal variation with period of twelve units of time (i.e., 12

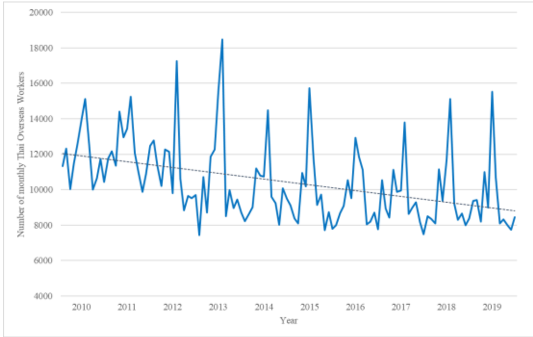


Figure 1. Time series plot of monthly Thai overseas workers

months) because the series data is per month. This period can also be seen from the correlograms in Figure 2. There appear to be annual or 12-month spikes in the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) correlograms. Therefore, two forecasting models were chosen to fit the data: SARIMA model (Box-Jenkins method), and Holt-Winter's ES model.

### 2.1 Box-Jenkins' SARIMA model

The Seasonal Autoregressive Integrated Moving Average model or SARIMA model, developed by Box and Jenkins in 1976, is based on the behavior of historical data influencing the present and helps explain future patterns or events (Cowpertwait & Metcalfe, 2009). It has been used as a popular model for time series data.

SARIMA $(p, d, q)(P, D, Q)_s$ model is an ARIMA model with a seasonal component where $(p, d, q)$ is the non-seasonal part of the model, and $(P, D, Q)$ is the seasonal part of the model, while $s$ is a period of the time series in each season. The Box-Jenkins SARIMA method is a forecasting technique that begins by identifying a stationary time series, checking to see if the time series has a constant mean $E(Y_t)$ and variance $V(Y_t)$ by performing the Augmented Dickey Fuller (ADF) test. If the series is not stationary, differencing and/or logarithmic transformation can be used to make it stationary (Manayaga & Ceballos, 2019). Once the time series has reached stationary behavior, i.e., the preliminary values of $D$ and $d$ have been fixed where the parameter $d$ is the order of difference frequency from non-stationary time series to stationary series, then the tentative models are established by determining the values of $p$, $q$, $P$, and $Q$ using the ACF and PACF plots of the stationary series. The ACF measures the amount of linear dependence between observations in the time series that are separated by a lag $q$. The PACF helps to determine how many autoregressive terms $p$ are necessary. The SARIMA$(p, d, q)(P, D, Q)_s$ model can be stated as follows (Bowerman & O'Connell, 1993; Box, Jenkins, & Reinsel, 1994):

$$\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D Y_t = \delta + \theta_q(B)\Theta_Q(B^s)\varepsilon_t \quad (1)$$

where $Y_t$ is an observation of the original series at time $t$, $\varepsilon_t \overset{iid}{\sim} N(0,\sigma^2)$ is an error term at time $t$, $d$, and $D$ are degrees of non-seasonal and seasonal differencing, $B$ is the

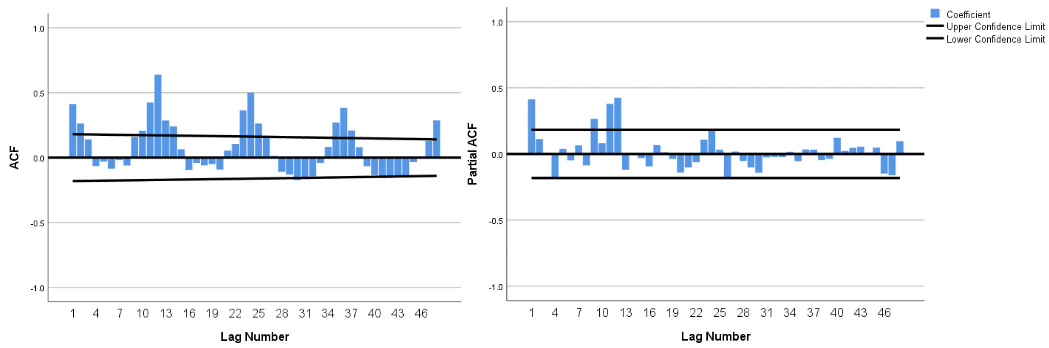


Figure 2. The ACF and PACF plots of the time series

backshift operator such that

$$B^s Y_t = Y_{t-s}; \tag{2}$$

non-seasonal autoregressive operator of order $p$ $(AR(p))$ is

$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p; \tag{3}$$

seasonal autoregressive operator of order $P$ $(SAR(P))$ is

$$\Phi_P(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \ldots - \Phi_P B^{Ps}; \tag{4}$$

non-seasonal moving average operator of order $q$ $(MA(q))$ is

$$\theta_q(B) = 1 + \theta_1 B + \theta_2 B^2 + \ldots + \theta_q B^q; \tag{5}$$

seasonal moving average operator of order $Q$ $(SAM(Q))$ is

$$\Theta_Q(B^s) = 1 + \Theta_1 B^s + \Theta_2 B^{2s} + \ldots + \Theta_Q B^{Qs}; \tag{6}$$

and $\delta$ is a nonzero constant at which $\mu$ is the mean of series or constant estimate of this model as

$$\delta = \mu \phi_p(B) \Phi_P(B^s). \tag{7}$$

We also call $p$, $d$, $q$, $P$, $D$, $Q$ the primary (or structural) parameters and the $\phi_i$'s, $\Phi_i$'s, $\theta_i$'s, and $\Theta_i$'s the secondary parameters. Let $\tau = (p, d, q, P, D, Q)$ be a set of the primary parameters of the model.

## 2.2 Holt-Winter's exponential smoothing model

Like SARIMA, the Exponential Smoothing (ES) model is also widely used to model time series. The formulation of exponential smoothing forecasting methods arose in the 1950s. Holt (Holt, 2004) and Winters (Winters, 1960) extended simple exponential smoothing to allow the forecasting of data with a trend and capture seasonality. There are two variations to this method that differ in the nature of the seasonal component. The additive method is preferred when the seasonal variations are roughly constant through the series, while the multiplicative method is preferred when the seasonal variations are changing proportional to the level of the series. With the additive method, the seasonal component is expressed in absolute terms in the scale of the observed series, and in the level equation, the series is seasonally adjusted by subtracting the seasonal component. The Holt-Winter's ES method comprises the forecast equation and three smoothing equations - one for the level $a_t$, one for the trend $b_t$, and one for the seasonal component $S_t$, with corresponding smoothing parameters $\alpha$, $\gamma$, and $\delta$, respectively. We use $s$ to denote the frequency of the seasonality, i.e., the number of seasons in a year. So, the additive Holt-Winter's ES method (Holt, 2004) involves a forecast equation and the three smoothing equations for $m = 1, 2, \ldots$, as follows:

forecast equation

$$\widehat{Y}_{t+m} = (a_t + m b_t) + S_{t-s+m}; \tag{8}$$

level equation

$$a_t = \alpha(Y_t - S_{t-s}) + (1-\alpha)(a_{t-1} + b_{t-1}); \tag{9}$$

trend equation

$$b_t = \gamma(a_t - a_{t-1}) + (1-\gamma) b_{t-1}; \tag{10}$$

seasonal equation

$$S_t = \delta(Y_t - a_t) + (1-\delta) S_{t-s}; \tag{11}$$

where $a_t$ denotes an estimate of the level of the series at time $t$, $b_t$ denotes an estimate of the trend (slope) of the series at time $t$, $S_t$ denotes an estimate of the seasonal component of the series at time $t$, when $\alpha$, $\gamma$, and $\delta$ are the smoothing parameters for the level, trend, and seasonal components, respectively, at which $0 \le \alpha, \gamma, \delta \le 1$ (Sanguanrungsirikul, Chiewananta vanich, & Sangkasem, 2015). We also call $\alpha$, $\gamma$, $\delta$ the primary (or structural) parameters. Let $\tau = (\alpha, \gamma, \delta)$ be a set of the primary parameters of the model.

## 2.3 Model fitting and validation steps

This study is an application of the above two models to see which one fits the data best. The main objective of the paper is to compare the two forecasting methods and to predict future values of the monthly Thai overseas workers by using SPSS version 25 and Minitab version 20 software packages. Moreover, the model parameters are estimated by using maximum likelihood estimation algorithm in SPSS software. The best model for each method will be selected for the lowest Bayesian information criterion (BIC) as provided by the software. The BIC is a criterion for model selection among a finite set of models; models with lower BIC are generally preferred. It will measure the goodness of fit and is used to decide which model has the best fit to the data, in a comparison of estimated models.

Step-1: Use the whole 120 periods of data $\{Y_1, Y_2, \ldots, Y_{120}\}$ to estimate the primary (or structural) parameters $\tau$ and select the best model for each method with the lowest BIC value as provided by the software. This also gives the optimal values of the primary parameters, say $\tau_*$.

Step-2: Diagnostic check with normality and residual white noise of the models:

(a) use the normal Q-Q plot and Kolmogorov-Smirnov test to look for normality,

(b) use Ljung-Box test to confirm that the residuals have the white noise property.

Step-3: Once we obtain the appropriate model for each method, to see how well they perform, use the $\tau_*$ (from Step-1) on the first 60 months' data $\{Y_1, Y_2, \ldots, Y_{60}\}$ and fit the model to predict $Y_{61}$. This automatically estimates the secondary parameters and in turn gives the predicted value for Jan 2015 which is $\hat{Y}_{61}$ (set the number of periods to forecast to be one). The error in prediction is $(Y_{61}-\hat{Y}_{61})$.

Step-4: Use the same $\tau_*$ (from Step-1) on the data $\{Y_2, Y_3, \ldots, Y_{61}\}$ to predict $Y_{62}$ (*i.e.*, shifting the test dataset by 1 unit of time). This automatically estimates the secondary parameters and in turn gives the predicted value for Feb 2015 which is $\hat{Y}_{62}$. The error in prediction is $(Y_{62}-\hat{Y}_{62})$.

Step-5: Keep on sliding by one time point until the predicted value for Dec 2019 will be obtained by using the same $\tau_*$ (from Step-1) on the data $\{Y_{60}, Y_{61}, \ldots, Y_{119}\}$ to get $\hat{Y}_{120}$ and the error in prediction will be $(Y_{120}-\hat{Y}_{120})$.

Step-6: Compare SARIMA model against Holt-Winter's ES model by using the criteria of the prediction mean absolute error (PMAE) and the prediction root mean squared error (PRMSE) defined as:

$$PMAE = \frac{1}{60}\sum_{t=61}^{120}\left|Y_t - \hat{Y}_t\right| \tag{13}$$

$$PRMSE = \sqrt{\frac{1}{60}\sum_{t=61}^{120}(Y_t - \hat{Y}_t)^2} \tag{14}$$

The process is summarized in Figure 3 to depict a model's structural parameter selection, followed by secondary parameter estimation and finally the model validation.

## 3. Results

### 3.1 The results of SARIMA model

The Augmented Dickey-Fuller (ADF) test is used to determine whether the time series is stationary with the null hypothesis $H_0$: the data is nonstationary. The results

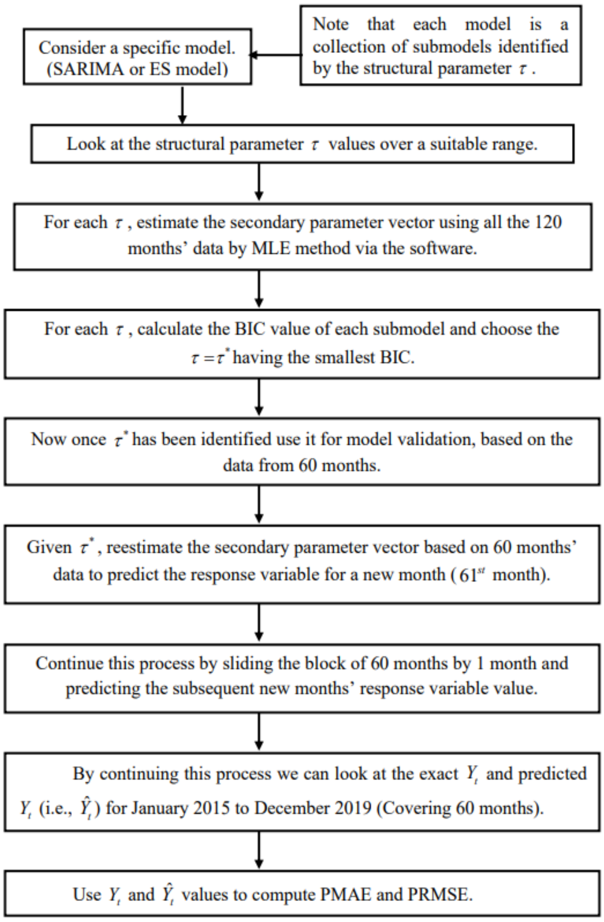


Figure 3. A summary flowchart about model fitting and evaluation

show that the ADF value is (-5.993) and its $p$-value is 0.01 (less than 0.05) which implies the null hypothesis is rejected. Therefore, the value of parameter $d$ will be zero as the series is stationary.

The ACF and PACF plots of the series in Figure 2 illustrate that the series is nonstationary in seasonal terms, as seen by the slow attenuation of the seasonal peaks (*i.e.*, at lag 12, 24, 36, ...) of the ACF. To achieve stationarity, the first differencing for the seasonal term is applied to the series. The first differenced series (in seasonal term) seems to be stationary since the autocorrelation of the data is reduced and there is no obvious pattern as seen in Figure 4, which means $D = 1$.

As an illustration in Figure 4, the ACF shows significant positive spike at lag 18, a significant negative spike at lag 13, and a large negative spike at lag 12 and the PACF reveals significant negative spike at lag 13, and a large negative spike at lag 12. Therefore, there are several interpretations that can be drawn from Figure 4 ($d = 0$, $D = 1$) since it has a large number of parameters and combinations of terms, but it is appropriate to try out a wide range of models to choose a best-fitting model. So, one possible parameter set for determining the tentative SARIMA models will be $p = 0, 1, 12, 13$, $q = 0, 1, 12, 13, 18$, $P = 0, 1, 3$, and $Q = 0, 1$. Table 1 represents only some of the tentative models with their corresponding BIC values. However, from all such

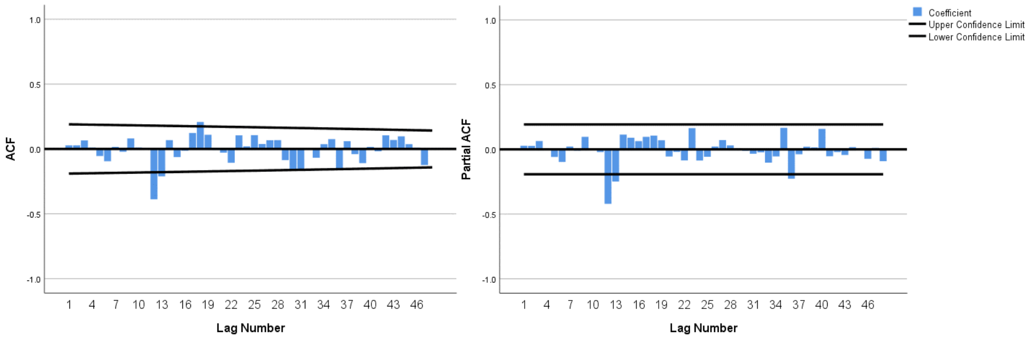


Figure 4. The ACF and PACF plots of the differenced (in seasonal term) series

Table 1. The tentative SARIMA models for the number of Thai overseas worker data

| SARIMA$(p,d,q)(P,D,Q)_s$ | BIC | SARIMA$(p,d,q)(P,D,Q)_s$ | BIC |
|---|---|---|---|
| SARIMA$(0,0,0)(0,1,0)_{12}$ | 14.923 | SARIMA$(1,0,0)(0,1,0)_{12}$ | 14.975 |
| SARIMA$(0,0,0)(0,1,1)_{12}$ | 14.624* | SARIMA$(1,0,0)(0,1,1)_{12}$ | 14.676 |
| SARIMA$(0,0,0)(1,1,0)_{12}$ | 14.786 | SARIMA$(1,0,0)(1,1,0)_{12}$ | 14.833 |
| SARIMA$(0,0,0)(1,1,1)_{12}$ | 14.673 | SARIMA$(1,0,0)(1,1,1)_{12}$ | 14.725 |
| SARIMA$(0,0,0)(3,1,0)_{12}$ | 14.775 | SARIMA$(1,0,0)(3,1,0)_{12}$ | 14.827 |
| SARIMA$(0,0,0)(3,1,1)_{12}$ | 14.765 | SARIMA$(1,0,0)(3,1,1)_{12}$ | 14.818 |
| SARIMA$(0,0,1)(0,1,0)_{12}$ | 14.975 | SARIMA$(1,0,1)(0,1,0)_{12}$ | 15.018 |
| SARIMA$(0,0,1)(0,1,1)_{12}$ | 14.777 | SARIMA$(1,0,1)(0,1,1)_{12}$ | 14.711 |
| SARIMA$(0,0,1)(1,1,0)_{12}$ | 14.834 | SARIMA$(1,0,1)(1,1,0)_{12}$ | 14.891 |
| SARIMA$(0,0,1)(1,1,1)_{12}$ | 14.725 | SARIMA$(1,0,1)(1,1,1)_{12}$ | 14.763 |
| SARIMA$(0,0,1)(3,1,0)_{12}$ | 14.827 | SARIMA$(1,0,1)(3,1,0)_{12}$ | 14.863 |
| SARIMA$(0,0,1)(3,1,1)_{12}$ | 14.818 | SARIMA$(1,0,1)(3,1,1)_{12}$ | 14.853 |

*model with the lowest BIC value

possible models, the SARIMA$(0,0,0)(0,1,1)_{12}$ model gave the smallest BIC.

Checking that the residuals of SARIMA$(0, 0, 0)(0, 1, 1)_{12}$ model are white noise was done by testing the Ljung-Box $Q$ statistic, which was found to be insignificant at $\alpha = 0.05$ (Ljung-Box $Q$ (lag 18) = 18.263, $p$-value = 0.372), *i.e.*, the residuals are uncorrelated. Moreover, the residuals of the SARIMA$(0, 0, 0)(0, 1, 1)_{12}$ model are normally distributed, which can be seen from the normality plot in Figure 5 and was confirmed by Kolmogorov-Smirnov test (Kolmogorov-Smirnov test statistic = 0.085, $p$-value = 0.051). Therefore, the SARIMA$(0, 0, 0)(0, 1, 1)_{12}$ model is appropriate and can be used to forecast the number of Thai overseas workers.

### 3.2 The results of Holt-Winter's exponential smoothing model

Based on the least BIC provided by the software, along with the results of diagnostic checking of the models, the model with level smoothing $\alpha = 0.098034$, the slope smoothing $\gamma = 5.67 \times 10^{-6}$, and the seasonal smoothing $\delta = 2.28 \times 10^{-4}$ was the best model of Holt-Winter's ES type.

Checking that the residuals of Holt-Winter's ES model are white noise was done by testing the Ljung-Box $Q$ statistic, which was found to be insignificant at $\alpha = 0.05$ (Ljung-Box $Q$ (lag 18) = 22.702, $p$-value = 0.091), *i.e.*, the residuals are uncorrelated. However, the residuals of the Holt-Winter's ES model are not normally distributed, which can be seen from the normality plot in Figure 6 and this was confirmed by Kolmogorov-Smirnov test (Kolmogorov-Smirnov test statistic = 0.103, $p$-value = 0.003). Therefore, the Holt-Winter's ES model is not appropriate to forecast the number of Thai overseas workers, as the model assumptions do not seem to be valid.

### 3.3 Accuracy of the models

The second subset of the number of Thai overseas workers data from Jan 2015 to Dec 2019 was used to validate the models by looking at the PMAE and PRMSE as shown in Table 2. SARIMA$(0,0,0),(0,1,1)_{12}$ model from Box-Jenkins method gave the lower PMAE as well as PRMSE. It was also found that the PMAE and PRMSE of Holt-Winter's ES model are large, as it is not an appropriate model to forecast the number of Thai overseas workers in the first place.

Moreover, one can see how much SARIMA model is providing improvement over the Holt-Winter's ES model by looking at the Relative Improvement (RI) of SARIMA over ES model as follows.

RI in terms of PMAE

$$= \left(\frac{2399.08 - 921.98}{2399.08}\right) \times 100\% = 61.57\%$$

RI in terms of PRMSE

$$=\left(\frac{2904.08-1361.23}{2904.08}\right)\times 100\% = 53.13\%$$

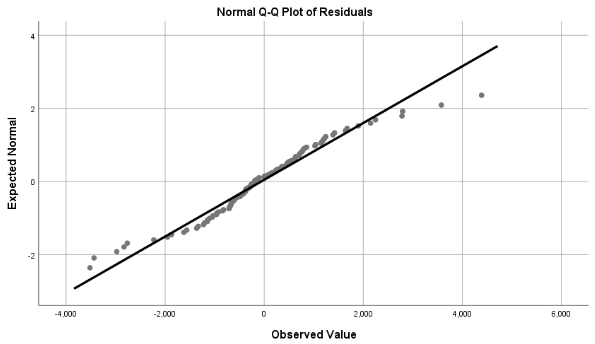


Figure 5. The normality plots of the residuals of SARIMA(0, 0, 0)(0, 1, 1)$_{12}$ model

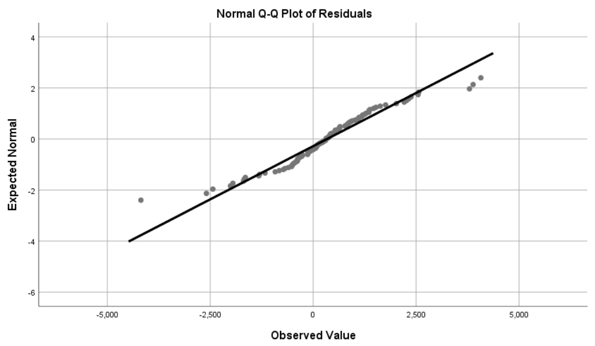


Figure 6. The normality plots of the residuals of Holt-Winter's ES model

Table 2. A performance comparison of the two models

| Forecasting model | Accuracy measures | |
|---|---|---|
| | PMAE | PRMSE |
| SARIMA(0,0,0),(0,1,1)$_{12}$ | 921.98* | 1361.23* |
| Holt-Winter's ES | 2399.08 | 2904.08 |

*The lowest (best) accuracy measure

The comparison graphs between the actual number of Thai overseas workers, fitted values, and predicted values using the SARIMA model and Holt-Winter's ES model during Jan 2010 to Dec 2019, are shown in Figure 7 and Figure 8, respectively.

It is evident from Figure 7 and Figure 8 why Box-Jenkins' SARIMA model is performing better than Holt-Winter's ES model. It appears that Holt-Winter's ES is adhering to the overall general trend but missing the aspect of seasonality, whereas SARIMA is capturing the seasonality better and focusing on the overall general trend also.

The actual and predicted values (predict one value at a time by sliding $t = 1$) of the number of Thai overseas workers by SARIMA(0,0,0),(0,1,1)$_{12}$ model from Jan 2015 to Dec 2019 are given in Table 3 to see how close the two are.

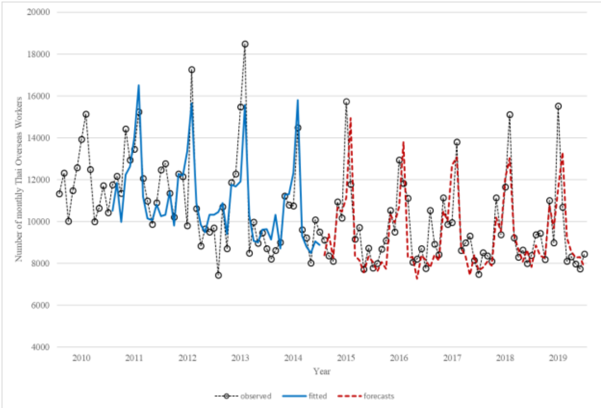


Figure 7. The actual time series data of the number of Thai overseas workers from Jan 2010 to Dec 2019 along with the fitted as well as predicted values by the SARIMA model

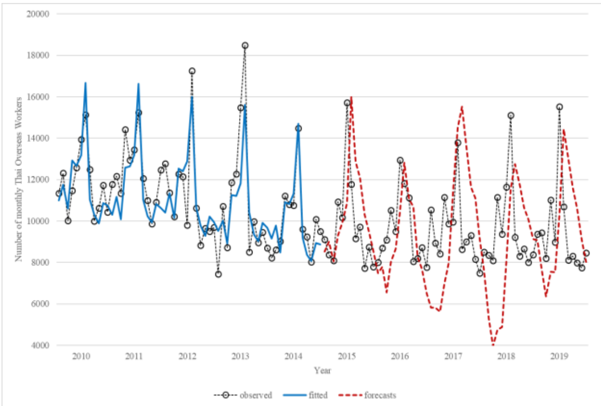


Figure 8. The actual time series data of the number of Thai overseas workers from Jan 2010 to Dec 2019 along with the fitted as well as predicted values by the ES model

Then, using the 60 months' data (Jan 2015 to Dec 2019) to do the prediction (ignoring the pandemic period of Jan 2020 – Dec 2023), one has the forecasted monthly number of Thai overseas workers in Jan 2024 – Dec 2026 as seen in Table 4 and Figure 9.

## 4. Conclusions and Suggestions

In today's globalized world, overseas workers play a crucial role in the economies of many developing countries, including Thailand. To accurately predict the number of Thai overseas workers each month, two types of forecasting models were compared: Box-Jenkins' SARIMA, and Holt-Winter's ES. Using the pre-pandemic data from the Royal Thai Government's Department of Employment, spanning from Jan 2010 to Dec 2019, the study covered a period of 120 months. The findings revealed that SARIMA(0,0,0),(0,1,1)$_{12}$ model outperformed the Holt-Winter's ES model in a significant way in predicting the number of monthly Thai overseas workers, as it is a powerful tool in the analysis of time series since it is capable of modeling a very wide range of time dependent data (Cowpertwait & Metcalfe, 2009). This is a significant finding

Table 3. The actual (A) and predicted (P) values (predicting one step ahead) of the number of Thai overseas workers by SARIMA(0,0,0),$(0,1,1)_{12}$ model

| Month | 2015 | | 2016 | | 2017 | | 2018 | | 2019 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | P | A | P | A | P | A | P | A | P |
| Jan | 9111 | 8386 | 7995 | 7603 | 10539 | 7795 | 8511 | 7803 | 9377 | 8854 |
| Feb | 8367 | 9392 | 8690 | 8057 | 8928 | 8407 | 8355 | 8074 | 9430 | 8353 |
| Mar | 8098 | 8249 | 9098 | 7732 | 8426 | 8136 | 8097 | 7896 | 8194 | 8252 |
| Apr | 10936 | 10656 | 10525 | 10310 | 11135 | 10500 | 11145 | 10193 | 11008 | 10769 |
| May | 10176 | 10463 | 9510 | 9954 | 9878 | 10046 | 9370 | 9518 | 8985 | 9674 |
| Jun | 15722 | 11142 | 12940 | 10708 | 9957 | 12674 | 11651 | 12087 | 15525 | 11561 |
| Jul | 11773 | 14937 | 11825 | 13807 | 13796 | 13071 | 15119 | 13030 | 10684 | 13307 |
| Aug | 9148 | 8358 | 11118 | 8268 | 8627 | 8728 | 9217 | 9341 | 8103 | 9214 |
| Sep | 9713 | 8140 | 8055 | 8289 | 8991 | 8335 | 8298 | 8654 | 8312 | 8566 |
| Oct | 7716 | 7573 | 8204 | 7276 | 9299 | 7429 | 8653 | 8094 | 7979 | 8282 |
| Nov | 8738 | 8314 | 8710 | 8387 | 8158 | 8385 | 8004 | 8646 | 7748 | 8291 |
| Dec | 7793 | 8066 | 7767 | 8154 | 7481 | 7662 | 8381 | 7821 | 8456 | 7813 |

Table 4. The forecasted monthly number of Thai overseas workers by SARIMA(0,0,0),$(0,1,1)_{12}$ model from Jan 2024 to Dec 2026

| Month | 2024 | 2025 | 2026 |
|---|---|---|---|
| Jan | 8904 | 9003 | 8734 |
| Fed | 8581 | 8687 | 8455 |
| Mar | 8143 | 8148 | 8076 |
| Apr | 10772 | 10785 | 10641 |
| May | 9289 | 9163 | 9227 |
| Jun | 12790 | 12795 | 12938 |
| Jul | 12748 | 12316 | 12594 |
| Aug | 9028 | 8486 | 8849 |
| Sep | 8467 | 8210 | 8317 |
| Oct | 8346 | 8011 | 8165 |
| Nov | 8076 | 7684 | 7958 |
| Dec | 7994 | 8259 | 7822 |

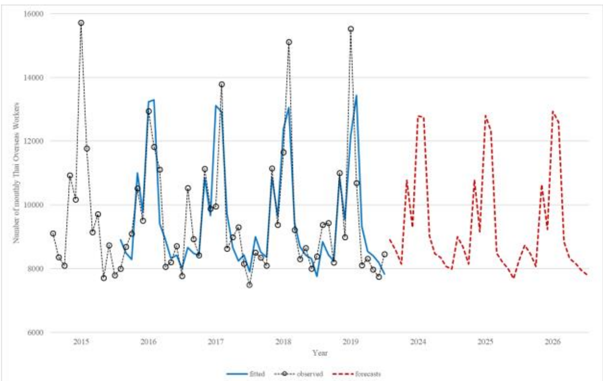


Figure 9. The time series plots of observed (circles) versus the fitted monthly number of Thai overseas workers (blue solid line) from SARIMA model, and the forecasts (red dashed line) for the years 2024 to 2026

which could help policymakers better understand and plan for the future of the Thai workforce abroad, ultimately benefiting the country's economy and its people.

The study recommends two important actions to improve the reliability of forecast results. Firstly, it suggests that forecast results can only be considered reliable when the data used to build the model is less outdated, preferably no more than five years. This is because older data may no longer be relevant or accurate due to the fast-changing economic contexts. Secondly, the study suggests that adding more forecasting techniques or processes can help identify the best model for forecasting, thereby improving the overall accuracy of predictions. It is important to take these recommendations into account in order to improve forecasting reliability and ensure that the predictions are as reliable and accurate as possible.

## Acknowledgements

Authors would like to thank sincerely (i) the Library System, Department of Employment of the Royal Thai Government for providing the online data via their website; (ii) the editors for their valuable time, and (iii) the three anonymous reviewers for their critical comments and useful suggestions which helped in improving the presentation of this work.

## References

Bowerman, B. L., & O'Connell, R. T. (1993). Forecasting and time series: An applied approach (3rd ed.). California, CA: Duxbury Press.

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (1994). Time series analysis forecasting and control (3rd ed.). New Jersey, NJ: Prentice Hall.

Brown, R. G. (1959). Statistical forecasting for inventory control. New York, NY: McGraw Hill.

Cowpertwait, P. S. P., & Metcalfe, A. V. (2009). Introductory time series with R. New York, NY: Springer.

Overseas Employment Administration Division of Thailand's Ministry of Labour. (2017). Strategies for dispatching job seekers to work abroad for 2017-2021. Retrieved from https://www.doe.go.th/prd/main/strategy/param/site/1/cat/1/sub/0/pull/category/view/list-label on 25 Jul 2021

Overseas Employment Administration Division of Thailand's Ministry of Labour. (2021). Travelling to work overseas. Retrieved from https://www.doe.go.th/

prd/overseas/custom/param/site/149/cat/40/sub/0/pull/detail/view/detail/object_id/284 on 27 Jun 2021

Holt, C. C. (2004). Forecasting seasonals and trends by exponentially weighted averages. *International Journal of Forecasting, 20*(1), 5-10.

Manayaga, E. M. C., & Ceballos, F. R. (2019). Forecasting the remittances of the overseas Filipino workers in the Philippines. *International Journal of Statistics and Economics, 3*(20), 36-44.

Sanguanrungsirikul, D., Chiewanantavanich, H., & Sangkasem, M. (2015). A comparative study to determine optimal models for forecasting the number of patients having epidemiological-surveillance diseases in Bangkok. *KMUTT Research and Development Journal, 38*(1), 35-55.

Sansiri, K. (2008). Factors affecting decision making of Thai labors to work abroad. Independent research [Master's thesis, Chiang Mai University, Chiang Mai, Thailand).

Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. *Management Science, 6*(3), 324–342.