JASET KMUTNB
Research Article

Journal of Applied Science and Emerging Technology (JASET)
Published by Faculty of Applied Science, KMUTNB, Thailand
doi: 10.14416/JASET.KMUTNB.2025.03.004
Vol. 24, No. 3 [2025]: e265260

Forecasting the Elderly Population in Songkhla Province Using the Box-Jenkins Method

Jutamas Boonradsamee[1]*, Kulteera Thongyai[1], and Rugkita Ieamwijarn[2]

[1]Department of Management, Faculty of Business Administration, Rajamangala University of Technology Srivijaya, Songkhla, 90000 Thailand
[2]Department of Accounting, Faculty of Business Administration, Rajamangala University of Technology Srivijaya, Songkhla, 90000 Thailand

* Corresponding Author, E-mail: jutamas.r@rmutsv.ac.th DOI: 10.14416/JASET.KMUTNB.2025.03.004
Received 13 December 2025; **Revised** 25 December 2025; **Accepted** 26 December 2025

ABSTRACT

This research aimed to quantitatively forecast the elderly population in Songkhla province using the Box-Jenkins (Autoregressive Integrated Moving Average - ARIMA) method. This method is based on the assumption that the past behavior of a time series is sufficient to predict future behavior. This study used historical monthly time series data of the elderly population in Songkhla province from 2017 to 2024 to determine the most appropriate ARIMA model. The results showed that the most appropriate model was ARIMA(4,2,0)$(2,0,0)_{12}$, which exhibits a clear and consistent seasonal pattern (seasonal data). The AIC value was 1181.419, and the estimated variance of error (white noise) was 14188. Projections indicate a significant and continuous increase in the elderly population over the next five years, with an expected rise of more than 50,000 people between 2025 and 2029. The resulting forecast provides a quantitative estimate of the future elderly population size in Songkhla province. The research findings aim to provide valuable insights for local and provincial policymakers in Songkhla for effective resource allocation, health care planning, and development of comprehensive social welfare strategies to address the impacts of an aging society.

KEYWORDS: Elderly, Box-jenkins method, Forecasting, Time series model.

Please cite this article as: J. Boonradsamee, K. Thongyai, and R. Ieamwijarn, "Forecasting the Elderly Population in Songkhla Province Using the Box-Jenkins Method," *Journal of Applied Science and Emerging Technology,* vol 24, no 3, pp. 1-13, ID. 265260, December 2025

1. INTRODUCTION

The number of elderly people worldwide is currently growing at a steady and quick pace. According to the World Health Organization, by 2030, one in six people worldwide will be 60 years of age or older. At that time, the proportion of the population aged 60 and over will increase from 1 billion in 2020 to 1.4 billion. By 2050, the global population aged 60 and over will double (2.1 billion), and the number of people aged 80 and over is projected to triple between 2020 and 2050 to 426 million (World Health Organization, 2025). After Nakhon Si Thammarat province, Songkhla province has the second-highest population in Southern Thailand and the Gulf of Thailand. Songkhla province's population has been growing since 2013; however, at an average annual rate of only 0.45%, this increase is not very rapid. The number of elderly people has increased in tandem with changes in the country's demographics. The percentages of people aged 0–14 and 15–59 have steadily declined between 2013 and 2018 (Songkhla Provincial Development Plan, 2021), while the percentage of those over 60 has increased, mirroring the national trend. According to this data, the old population is growing at a lower birth rate, while the child. In order to guarantee a high standard of living for the aged in Songkhla province, pertinent agencies must create policies in the social, economic, and health domains. In the future, there will be a lack of young caretakers and an increase in health issues for the elderly, such as diabetes, hypertension, or age-related degenerative diseases. The provincial labor force will decline at the same time.

This corresponds with information from the Human Resource Development Plan for the old in Songkhla province, which indicates that the province is encountering difficulties due to its increasing elderly population. Consequently, it is imperative to cultivate and refine multifaceted foundational elements within the province to augment the capabilities of the aged. Preliminary data from the official registration system in 2021 indicated that Songkhla province had an elderly population of 229,680 as of June 30, 2021, or 16.27% of the total population. The elderly population can be categorized by their condition: 159,693 are socially engaged, 137,495 have chronic illnesses, 7,301 are home-bound, and 1,249 are bedridden. Almost all elderly people can live freely and carry out everyday tasks because they are socially active. Targeting this group is essential for driving positive societal aging in Songkhla province. According to the Songkhla Provincial Office of Social Development and Human Security (July 27, 2021), the most pressing issue for this demographic is financial distress, characterized by an insufficiency of income to sustain daily livelihoods. They expressed a need for comprehensive support for career and income generation. Health issues were also a major concern, highlighting the desire for a better quality of life among the elderly in Songkhla. Numerous past studies have focused on the quality of life of the elderly, both nationally and within Songkhla province. For example, Noknoi and Boripunt (2017) studied the quality of life of the elderly in Songkhla, Thongyai (2019) examined factors affecting the happiness of the elderly in the Kaoseng community, and Suk-anan (2020) investigated strategies for promoting employment among the Thai elderly. These studies generally focused on the overall quality of life. However, the recent global COVID-19 pandemic has caused changes in all dimensions. Considering the increasing elderly population and their current problems in Songkhla, the research team recognizes the importance of these societal changes.

Therefore, this research aims to forecast the upward trajectory of the aging population in Songkhla through a time-series analysis. By utilizing the Box-Jenkins method, a predictive model is constructed to provide empirical estimations of future elderly population trends. This study employs quantitative forecasting, a systematic approach that leverages historical time-series data to extrapolate future trends through the application of rigorous mathematical models. The Box-Jenkins method is distinguished by its high degree of accuracy and precision in quantitative forecasting. Its robustness stems from the prerequisite verification of time-series stationarity before estimation. Notably, this approach remains data-driven, as it does not necessitate a priori model specification, allowing the model structure to emerge from the empirical data itself. In this study, the forecasting model was manually specified, yielding a specialized model for the low-income demographic that offers granular insights. This approach is particularly efficacious for time-series data lacking discernible cyclical or seasonal patterns. The robustness of the Box-Jenkins method is well-documented across diverse forecasting domains. For instance, Anvari et al. (2016) developed an automated Box-Jenkins tool to analyze passenger demand in urban rail systems, while Idrus and Mohamed (2020) conducted a comparative study between Box-Jenkins and artificial neural networks to forecast air travel volume in Malaysia. Furthermore, in the context of demographic projections, Ayele and Zewdie (2017) utilized ARIMA modeling on annual Ethiopian population data (1961–2009), identifying the ARIMA (2,1,2) as the optimal fit. Parallel to these global applications, this research employs the Box-Jenkins ARIMA technique, utilizing a longitudinal dataset spanning 1960 to 2017 to project population trends.

The remainder of this article is structured as follows. Section 2 elucidates the theoretical framework of the Box-Jenkins forecasting methodology. Section 3 details the data collection process and analytical procedures, followed by Section 4, which presents the conclusion and discussion. In Section 5, numerical results from a performance comparison across various scenarios are provided. Finally, Section 6 outlines policy implications, offering strategic recommendations for future research.

2. METHODS

2.1 Box-Jenkins Method

The Box-Jenkins method is a quantitative forecasting technique based on the idea that the past behavior of a time series is sufficient to predict its future behavior. Unlike other forecasting methods, which require the forecaster to pre-specify the form of the relationship, especially when the time series lacks a clear trend, cycle, or seasonality, the Box-Jenkins method does not require a fixed model form before analysis. The model is identified during the analysis process itself, which involves the following steps:

Step 1. Calculate the value of the self-correlation function. (Autocorrelation Function: ACF) and Partial Autocorrelation Function (PACF) are the first step for stationary time situations, namely the time requirement for which the forecast value is to be calculated, to find the ACF and PACF values to be used for processing in the model, or to select the model which will be mentioned, or the amount of data to be examined again, as observed. These functions help in identifying the appropriate model by indicating the number of past observations (lags) to be considered. Calculate the ACF value using the following equation:

$$r_j = \frac{\sum_{t=1}^{N-j}(X_t - \bar{X})(X_{t+j} - \bar{X})}{\sum_{t=1}^{N}(X_t - \bar{X})^2}, \quad (2.1)$$

where X_t is the data or observations at time

j is the number of time intervals between the data $j = 1, 2, 3, ..., k$,

N is the total number of data, and

$\bar{X}$ is the average of all the data where

$$\bar{X} = \frac{\sum_{t=1}^{N} X_t}{N}.$$

Calculate the PACF value using the following equation:

$$\hat{\phi}_{kk} = \begin{cases} r_1 & ; k = 1 \\ \dfrac{r_k - \sum_{j=1}^{k-1}(\hat{\phi}_{(k-1)j} r_{k-j})}{1 - \sum_{j=1}^{k-1}(\hat{\phi}_{(k-1)j} r_j)} & ; k = 2, 3, 4, ..., \end{cases} \quad (2.2)$$

and $\hat{\phi}_{kk} = \hat{\phi}_{(k-1)j} - \hat{\phi}_{kk}\hat{\phi}_{(k-1)(k-j)}$; $j = 1, 2, 3, ..., k-1.$

Step 2. The appropriate model is determined by examining the patterns of the ACF and PACF plots, as summarized in Table 1.

The quantitative model used in this research's forecasting is the ARIMA(p,d,q) process. p is the number of regression terms, d is the rank of the difference that keeps the data stable, and q is the number of moving average terms. For example, the ARIMA (2,1,2) process has a rank 1 difference (d=1) that maintains stability, along with two regression terms and two moving average terms. If d=0, the ARIMA (p,d=0,q) process means ARMA (p,q). Note that the ARIMA(p,0,0) process means the AR(p) process, and the ARIMA(0,0,q) process means the MA(q) process. The ARIMA (p,d,q) model is derived from the ARMA(p,q) process, which has the following general form.

i. p^{th}-order autoregressive model: AR(p)

$$X_t = \delta + \phi_1 X_{t-1} + \phi_2 X_{t-2} + ... + \phi_p X_{t-p} + u_t, \quad (2.3)$$

where X_t is the data or observation at the time t,

δ is the process constant,

u_t is the random error at time t,

$\phi_1, \phi_2, ..., \phi_p$ is the coefficient of the regression term, and $X_{t-1}, X_{t-2}, ..., X_{t-p}$ is the response variable at the lag, $t-1, t-2, ..., t-p.$

ii. q^{th}-order moving average model: MA(q)

$$X_t = \delta + u_t - \theta_1 u_{t-1} - \theta_2 u_{t-2} - ... - \theta_q u_{t-q}, \quad (2.4)$$

where X_t is the data or observation at time t,

δ is the constant mean,

u_t is the white noise process at time t,

$\theta_1, \theta_2, ..., \theta_q$ is the coefficient of the moving average term, and $u_{t-1}, u_{t-2}, ..., u_{t-q}$ is the moving average term q.

iii. p^{th} and q^{th}-order autoregressive moving average model: ARMA (p,q)

$$X_t = \delta + \phi_1 X_{t-1} + \phi_2 X_{t-2} + ... + \phi_p X_{t-p} + u_t - \theta_1 u_{t-1} - \theta_2 u_{t-2} - ... - \theta_q u_{t-q}, \quad (2.5)$$

where X_t is the data or observation at time t,

δ is the constant mean,

u_t is the white noise process at time,

$\theta_1, \theta_2, ..., \theta_q$ is the coefficient of the moving average term,

$u_{t-1}, u_{t-2}, ..., u_{t-q}$ is the moving average term q.

$\phi_1, \phi_2, ..., \phi_p$ is the coefficient of the regression term, and

$X_{t-1}, X_{t-2}, ..., X_{t-p}$ is the response variable at the lag, $t-1, t-2, ..., t-p.$

Table 1 Identifying Models using ACF and PACF.

Model	ACF	PACF
AR(p)	Decays exponentially	Cuts off after lag p
MA(q)	Cuts off after lag q	Decays exponentially
ARMA(p,q)	Decays exponentially	Decays exponentially

Table 2 The results of testing the stationarity of the data through the Unit Root Test method, the results before and after data adjustment.

Augmented Dickey-Fuller test					
At level		1st Difference		2nd Difference	
t - statistics	p - value	t - statistics	p - value	t - statistics	p - value
-1.3975	0.8254	-2.8841	0.2113	-7.7183	0.0100

Table 3 Comparison of best–fit ARIMA models through Akaike Info Criterion (AIC) requirements.

MODEL	AIC
$ARIMA(2,2,2)(1,0,1)_{12}$	Inf
ARIMA(0,2,0)	1287.986
$ARIMA(1,2,0)(1,0,0)_{12}$	1205.391
$ARIMA(0,2,1)(0,0,1)_{12}$	1219.348
ARIMA(1,2,0)	1278.459
$ARIMA(1,2,0)(2,0,0)_{12}$	Inf
$ARIMA(1,2,0)(1,0,1)_{12}$	Inf
$ARIMA(1,2,0)(0,0,1)_{12}$	1239.545
$ARIMA(1,2,0)(2,0,1)_{12}$	Inf
$ARIMA(0,2,0)(1,0,0)_{12}$	1208.09
$ARIMA(2,2,0)(1,0,0)_{12}$	1197.401
ARIMA(2,2,0)	1261.342
$ARIMA(2,2,0)(2,0,0)_{12}$	1189.355
$ARIMA(2,2,0)(2,0,1)_{12}$	Inf
$ARIMA(2,2,0)(1,0,1)_{12}$	Inf
$ARIMA(3,2,0)(2,0,0)_{12}$	1188.844
$ARIMA(3,2,0)(1,0,0)_{12}$	1194.276
$ARIMA(3,2,0)(2,0,1)_{12}$	Inf
$ARIMA(3,2,0)(1,0,1)_{12}$	Inf
$ARIMA(4,2,0)(2,0,0)_{12}$	**1181.419**
$ARIMA(4,2,0)(1,0,0)_{12}$	1188.255
$ARIMA(4,2,0)(2,0,1)_{12}$	Inf
$ARIMA(4,2,0)(1,0,1)_{12}$	Inf
$ARIMA(5,2,0)(2,0,0)_{12}$	Inf
$ARIMA(4,2,1)(2,0,0)_{12}$	1183.713
$ARIMA(3,2,1)(2,0,0)_{12}$	Inf
$ARIMA(5,2,1)(2,0,0)_{12}$	1185.765

Step 3. Parameter estimation is the process of estimating the parameters contained in a time series model using the maximum likelihood method $L(\underline{\phi}, \underline{\theta}, \delta, \sigma_u^2 | X_t, t = 1, 2, 3, \ldots, N)$.

The estimated values of ϕ, θ, and δ can be calculated by minimizing the sum of squared errors, that is

$$\text{Minimize}\sum_{t=1}^{n}\varepsilon_t^2 \quad (2.6)$$

$$\varepsilon_t = X_t - \hat{\phi}_1 X_t - \ldots - \hat{\phi}_p X_{t-p} - \hat{\delta} + \hat{\theta}_1 \varepsilon_{t-1} + \ldots + \hat{\theta}_q \varepsilon_{t-q}. \quad (2.7)$$

It is an approximate value u_t, which is considered from the equation.

$$u_t = X_t - \hat{\phi}_1 X_t - \ldots - \hat{\phi}_p X_{t-p} - \hat{\delta} + \hat{\theta}_1 u_{t-1} + \ldots + \hat{\theta}_q u_{t-q}, \quad (2.8)$$

when the approximate values of ϕ, θ, and δ are found, the approximate values of σ_u^2 are obtained by

$$\sigma_u^2 = \frac{1}{N}\sum_{t=1}^{N}\varepsilon_t^2. \quad (2.9)$$

Let $\hat{\beta}$ the factors of the various parameters be replaced by the statistic used to test the t estimator given by

$$t_{\hat{\beta}} = \frac{\hat{\beta}}{SE(\hat{\beta})}, \quad (2.10)$$

where $SE(\hat{\beta})$ is the standard error of $\hat{\beta}$ and degrees of freedom is the number of terms N minus the parameter to be estimated.

Step 4. Model Suitability Checking: Once a time series model has been selected and its parameters have been estimated, the next step is to test its suitability. The principle for model suitability checks is based on the random probability properties of the error values u_t, particularly regarding the lack of self-correlation. That is, if the selected model is suitable and its parameter values are known, we will use Q the Chi-squared distribution, which is calculated using the following:

$$Q(k) = \{(N-d)[(n-d)+2]\}\sum_{j=1}^{k}\frac{r_j^2}{[(N-d)-j]}. \quad (2.11)$$

The Q statistic is used to test for non-self-correlation of error values u_t. It is the degree of freedom $Q = k$ of the statistic minus the number of parameters to be estimated for the chosen model.

where k is the number of time intervals (lags) between the data

N is the total number of observations of the time series,

d is the rank of the time series difference,

and r_j is the autocorrelation at the lag j.

Step 5. Forecasting. Once the model assigned to the time series has been verified to be suitable, the next step is to apply the model equations to the forecast to find the forecast value.

2.2 Describing the Real Data

This research focuses on the quantitative analysis of a forecasting model for the elderly population in Songkhla Province. Statistical time series data from 2017 to 2024 were collected and analyzed using the Box-Jenkins method (or ARIMA model) to accurately forecast future trends and growth rates of the elderly population, providing important insights for provincial-level care planning. The collected data are presented in Figure 1.

Figure 1 shows that the number of elderly people in Songkhla Province has been steadily increasing each year. When considering each district, Hat Yai District has the highest number of elderly people, followed by Mueang Songkhla District, Sadao District, Chana District, and Singhanakhon District, which are the top five districts in Songkhla Province with the highest number of elderly people.

As shown in Figure 2, the average elderly population in Songkhla Province has been increasing annually and is likely to continue to increase. This

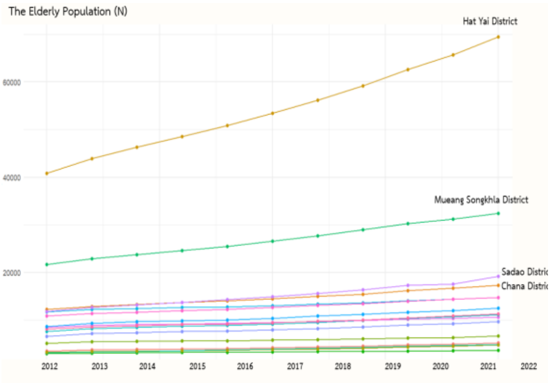

Figure 1 Annual Number of Elderly Population by District in Songkhla Province.

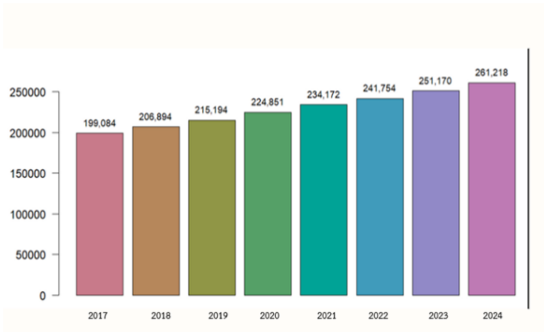

Figure 2 The Average number of elderly people per year, 2017 – 2024, of Songkhla Province.

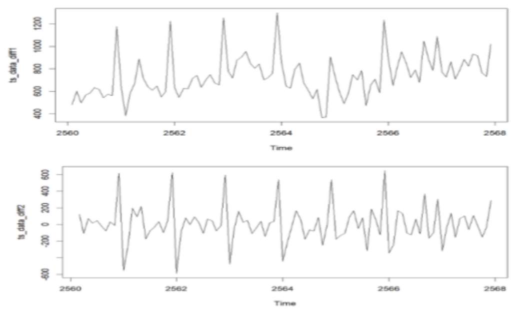

Figure 3 The Stationary Characteristics of Elderly Population Data by Finding 1st Difference and 2nd Difference.

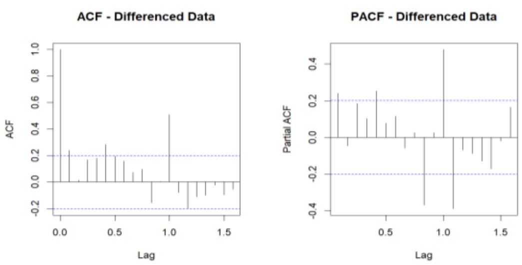

Figure 4 ACF and PACF diagrams.

increase in the elderly population impacts management in many areas, including the economy, society, and health. This is particularly the case with a decline in the workforce, increased financial burdens, and the increasing demand for health and welfare services. This has led government agencies to plan and adjust policies to accommodate the growing elderly population. Furthermore, in today's society, the birth rate has significantly decreased. According to the Department of Older Persons Affairs, an "aging society" is defined as a society in which the proportion of the elderly or the population aged 60 years and over has steadily increased, while the proportion of the birth rate and the working-age population has decreased. The results of the quantitative analysis of the elderly population forecasting model in Songkhla Province, using the Box-Jenkins method, were used to create a forecast model for the future elderly population. Historical data on the elderly population in Songkhla Province were collected from 2017 to 2024.

3. DATA AND ANALYSIS

Data analysis revealed a steadily increasing trend of the elderly population each year, resulting in a shrinking workforce, financial burdens, and demand for health and welfare services. To accommodate the future expansion of the elderly population in Songkhla Province, researchers developed a quantitative forecasting model of the elderly population in Songkhla Province using the Box-Jenkins method. This model was developed by collecting historical statistics of the elderly population in Songkhla Province from 2017 to 2024. The steps can be summarized as follows:

3.1 Testing for Stationarity

The stationarity of the time series data was tested using the Augmented Dickey-Fuller (ADF) Unit Root

Test (Dickey and Fuller, 1979). For time series analysis, the researchers used the R statistical software (v4.3.2; R Core Team 2023) and core functions from the forecasting package v8.21 (Hyndman et al., 2023) for ARIMA modeling and the tseries package v0.10-54 (Trapletti and Hornik, 2024) for stationarity testing.

The hypotheses are given as:

H_0 : The data is non-stationary

H_1 : The data is stationary

From the Table 2 the data testing through the Unit Root Test method, the p-value results are as follows: The data on the number of elderly people in Songkhla Province has a p-value of 0.8254, which is higher than the level of significance (Level of Significant) set at 0.05, indicating that the data on the number of elderly people used in the study are not stationary. Therefore, the researcher calculated the first difference (1^{st} Difference) and got a p-value of 0.2113, indicating that the data on the number of elderly people used in the study are not stationary. When calculating the first difference, the researcher calculated the second difference (2^{nd} Difference) and got a p-value of 0.01, which is lower than the significance level of 0.05, resulting in data that is stationary at the significance level of 0.05. The study results can explain the differences in data, as shown in Figure 3.

3.2 Identifying the Model Order (p, q)

Consider the ACF (Autocorrelation Function) and PACF (Partial Autocorrelation Function) plots of the differentiated time series data used to identify the order (p and q) of the ARIMA (Autoregressive Integrated Moving Average) model, as shown in Figure 4

Figure 4 illustrates the ACF - Differenced Data diagram showing the relationship between the time series and historical values (lags). The results show that the ACF value rapidly decreases (cuts off) after the first lag. At lag 1, the first bar exceeds the dotted blue line (value 0.2), indicating a significant relationship. At lags 2 and 3, the bar falls within the range of the dotted blue line, indicating a non-significant relationship. Finally, at lag 12, the relationship is again high (approximately 0.5), indicating seasonality at lag 12 (since it is monthly data). In other words, for MA (q), the rapid "cut off" of the ACF after lag 1 suggests that the model may have a sequential MA (moving average) component, or perhaps a seasonal MA.

The PACF - Differenced Data diagram shows the relationship between the time series and historical values. By eliminating the effect of the intermediate values, at Lag 1, the first bar exceeds the blue dotted line (value 0.2), indicating a significant relationship. At Lag 2 and Lag 3, the bars are in the range of the blue dotted line. At Lag 10 and Lag 12, the relationship is again significant, which may be related to the seasonal structure, i.e., for AR (p): the PACF "Decays", i.e., it declines slowly or "Cuts off". From the figure, the PACF also "Cuts off" after Lag 1, because Lag 2 and Lag 3 are not significant, which may indicate that the model may have an AR (Autoregressive) component at the order of magnitude.

The ACF and PACF plots for the differenced series. The ACF (Autocorrelation Function) exhibits a distinct cut-off after lag 1, with the first lag significantly exceeding the 95% confidence threshold. Additionally, a prominent seasonal spike is observed at lag 12 (lag 1.0), suggesting a seasonal moving average (SMA) component. Simultaneously, the PACF (Partial Autocorrelation Function) shows significant spikes up to lag 4 before cutting off, which typically characterizes an AR(4) process. Following the Box-Jenkins identification logic, these patterns justify the inclusion of both non-seasonal and seasonal parameters in the candidate SARIMA model.

3.3 Finding the Best ARIMA Model

The best ARIMA model was automatically found to fit the time series data of the elderly population in

Table 4 Goodness-of-Fit and Forecasting Accuracy of the SARIMA(4, 2, 0)(2, 0, 0)$_{12}$ model.

Parameter	Lag	Coefficient	Standard Error
AR1	1	-0.4733	0.1005
AR2	2	-0.5194	0.1086
AR3	3	-0.3151	0.1081
AR4	4	-0.3249	0.1008
SAR1	12	0.4596	0.101
SAR2	24	0.3599	0.1112
$\sigma^2 = 14188$	AIC = 1180.12		
RMSE = 114.0422	MAPE = 0.03790666	ACF1 = -0.02079397	

Table 5 The increasing trend of the number of elderly people in Songkhla Province in 2025.

Month	Point Forecast	95% Confidence Interval	
		Lower	Upper
Jan	266,790.70	266,557.30	267,024.20
Feb	267,556.90	267,130.80	267,983.00
Mar	268,391.00	267,799.40	268,982.60
Apr	269,217.00	268,454.60	269,979.40
May	270,021.90	269,088.60	270,955.20
Jun	270,846.50	269,708.60	271,984.50
Jul	271,680.90	270,311.60	273,050.20
Aug	272,509.40	270,901.30	274,117.50
Sep	273,465.30	271,609.40	275,321.20
Oct	274,289.80	272,178.40	276,401.10
Nov	275,063.40	272,683.90	277,442.80
Dec	276,085.20	273,423.20	278,747.20

Songkhla Province using the Akaike Information Criterion (AIC). The results are shown in Table 3.

From Table 3, it was found that the best Seasonal ARIMA (SARIMA) model is based on the Akaike Information Criterion (AIC). The model with the lowest AIC was found to be SARIMA(4,2,0)(2,0,0)$_{12}$, which is the model with clear and consistent seasonal data, with an AIC of 1181.419. Several models resulted in infinite AIC values. This phenomenon occurs when the model structure is incompatible with the data's stochastic properties, primarily due to a lack of stationarity or invertibility. In such cases, the iterative estimation procedure fails to converge, often as a consequence of including redundant parameters beyond the information capacity of the time series. Although several candidate models, such as SARIMA(4,2,1)(2,0,0)$_{12}$ exhibited AIC values comparable to the selected model, SARIMA(4,2,0)(2,0,0)$_{12}$ was preferred based on the principle of parsimony. The marginal improvement in model fit provided by the additional moving average (MA) parameter in the more complex configurations did not justify the increased model complexity. Furthermore, SARIMA(4,2,0)(2,0,0)$_{12}$ achieved the absolute minimum AIC (1181.419), suggesting an optimal balance between goodness-of-fit and model simplicity, thereby minimizing the risk of over-parameterization and enhancing forecasting robustness.

3.4 Model Estimation and Diagnostics

The residuals of the model were found to be white noise, meaning that there was no autocorrelation remaining, confirming that the model accurately

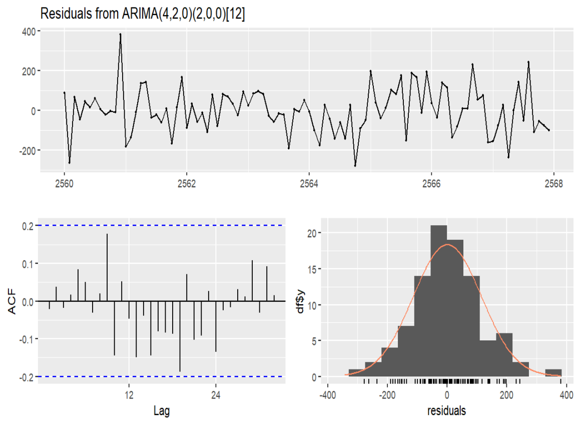

Figure 5 Residual Diagnostics.

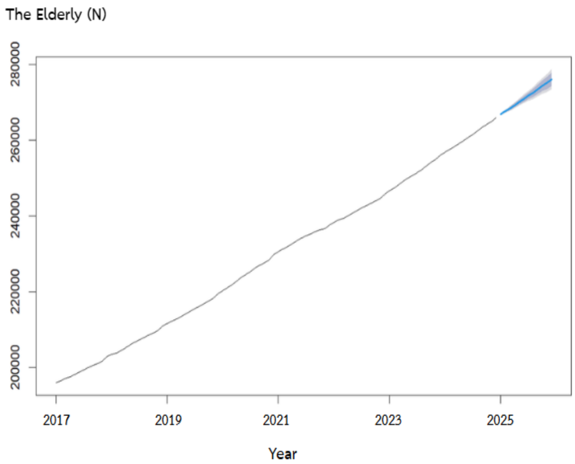

Figure 6 The Number of elderly people in Songkhla Province in 2025.

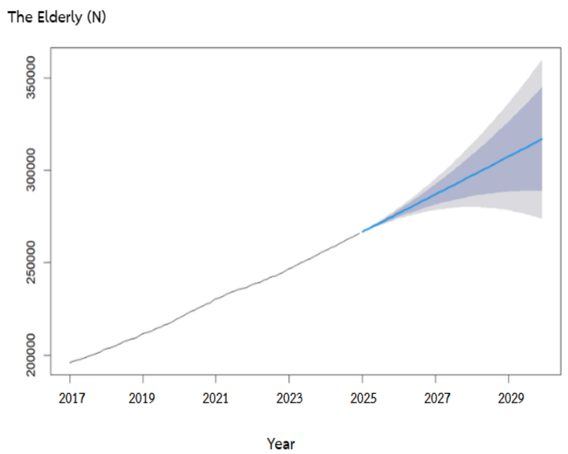

Figure 7 The Number of elderly people in Songkhla Province from 2025

captured the underlying patterns in the data. The results are described in Table 4

From Table 4, the results of the formally set and orthogonal systems for SARIMA(4,2,0)(2,0,0)$_{12}$ are obtained. Technical data are obtained by taking the difference between 2 values. The stationary data are softened by the Autoregressive (AR) component, the non-seasonal lag of 4 periods, and the lag of 2 parts at the lag of 12 periods.

Good-fit test (Goodness-of-Fit) found that the AIC value of 1180.12, the lowest value used to test this, provided the best performance between that and simplicity (penalized for complexity). The recency factor for the deviation (white noise) was 14188. In forecasting (forecast accuracy), the RMSE value was 114.04, and the MAE value was 87.07, indicating the average size of the motherboard. The MAPE value was 0.0379. The deviation rate is very low, clearly indicating that the components can explain and predict the data to a high degree. An investigation of the residual diagnostics found an ACF1 value of −0.02079, and the autocorrected residual at lag 1 was above zero. For the review, the residual diagnostics were adjusted to white noise without any structure or manipulation.

To evaluate the statistical significance of the SARIMA(4, 2, 0)(2, 0, 0)$_{12}$ model, the t-test was applied to each estimated parameter. As shown in Table 4, all coefficients (AR1 to AR4 and SAR1 to SAR2) yielded absolute t-values greater than 1.96. These results demonstrate that both the non-seasonal and seasonal components are statistically significant, providing strong evidence for the model's structural validity. Furthermore, the model demonstrates high forecasting precision with a MAPE of 3.79% and a near-zero ACF1 for residuals, indicating that the model has effectively captured both the non-seasonal trends and the 12-month seasonal cycles present in the time series.

For the first time, SARIMA(4,2,0)(2,0,0)$_{12}$ was able to capture the pattrn and relationship of the data over time in the elderly population and was ready for formal forecasting. The results of the Residual Diagnostics examination are shown in Figure 5.

The SARIMA(4, 2, 0)(2, 0, 0)$_{12}$ model equation can be written as the following equation (12). Stationary data is defined because $d = 2$ and, $S = 12$ which z_t are stationary variables, are defined as follows:

$$z_t = (1-B)^2 y_t = y_t - 2y_{t-1} + y_{t-2}, \quad (3.1)$$

where y_t is the time series data at time t, and

B is the backward shift operator.

The equation in the form of AR operator is the product equation of Non-Seasonal AR(4) and Seasonal AR(2) acting on z_t without MA term because of $q = 0$ and $Q = 0$ as follows equation

$$(1-0.4596B^{12} - 0.3599B^{24})(1-(-0.4733)B-(-0.5194)B^2 \\ -(-0.3151)B^3 - (-0.3249)B^4)z_t = w_t, \quad (3.2)$$

when w_t (White Noise Error Term) with variance σ^2 equal to 14188, the total AR coefficient (α_i) is given by polynomial expansion as follows.

$$z_t = \sum_{i=1}^{28} \alpha_i z_{t-1} + w_t$$

$$\begin{aligned} z_t = {} & (-0.4733)z_{t-1} + (-0.5194)z_{t-2} + (-0.3151)z_{t-3} + (-0.3249)z_{t-4} \\ & + (0.4596)z_{t-12} + (0.3599)z_{t-24} \\ & - (0.2175)z_{t-13} - (0.2387)z_{t-14} - (0.1448)z_{t-15} - (0.1493)z_{t-16} \\ & - (0.1703)z_{t-25} - (0.1869)z_{t-26} - (0.1134)z_{t-27} - (0.1170)z_{t-28} \\ & + w_t \end{aligned}$$

3.5 Ljung-Box Q Test

The SARIMA(4, 2, 0)(2, 0, 0)$_{12}$ model was tested using the Ljung-Box Q statistic under the following hypothesis:

H_0 : The model is suitable

H_1 : The model is not suitable

The SARIMA(4, 2, 0)(2, 0, 0)$_{12}$ model is an appropriate model for forecasting the number of elderly people in Songkhla province. The Ljung-Box Q test value is 19.947, and the p-value is 0.1071, which is greater than the significance level of 0.05, meaning that the model is appropriate for forecasting data because the residuals are "White Noise", indicating that there is no significant relationship. The analysis was conducted in R using the forecast and tseries packages.

4. FORECASTING RESULTS AND DISCUSSION

The time series of elderly people in Songkhla province can be forecast using the SARIMA(4, 2, 0)(2, 0, 0)$_{12}$ forecasting model for the elderly population in 2025 can be summarized as follows in Table 5.

From Table 5, the point prediction values show that the number of elderly people has continuously increased throughout 2025, starting in January 2025 with 266,791 people and ending in December 2025 with 276,085 people. This confirms the aging society in which the elderly population will increase steadily each month. The width of the 95% confidence interval has expanded significantly when considering the time period. In January 2025, the 95% confidence interval was approximately 467 people wide, and in December 2025, the 95% confidence interval was approximately 5,324 people wide. The expansion of the number of elderly people is shown in Figure 6

The time series of elderly people in Songkhla province can be forecast using the SARIMA(4, 2, 0)(2, 0, 0)$_{12}$. Trends in the increasing number of elderly people in Songkhla Province from 2025 to 2029.

The forecasting results for the five years from 2025 to 2029 reveal a consistent and statistically significant upward trend in the elderly demographic. According to the model estimates, the number of elderly citizens is projected to rise from 266,791 in January 2025 to 316,940 by December 2029. Over this 60-month horizon, the total increase is expected to reach approximately 50,149 individuals, reflecting a steady expansion of the aging population in the region. This continuous growth trajectory, as illustrated in Figure 7, underscores the intensifying demographic shift toward an aging society in Songkhla province.

For five years, the trend in elderly population growth for interval planning, health, welfare, and infrastructure planning for 2028-2029 should use a

95% confidence interval as the basis for resource allocation. The model should be reviewed and updated annually with new actual data to reduce the width of the long-term confidence interval. Since the forecast for December 2029 has a 95% chance that the elderly population will not exceed 359,948.6 people, relevant agencies should prepare budgets and personnel to cope with this peak level to avoid service shortages if the elderly population actually increases.

5. POLICY IMPLICATION

This research provides quantitative evidence that there is an urgency for provincial and national agencies to develop and implement comprehensive policies to meet the needs of an aging society. The findings can serve as an important tool for strategic planning of healthcare, social welfare, economic development, and infrastructure to ensure a good quality of life for the growing elderly population in Songkhla Province.

Relevant agencies should allocate budgets to support the development of public infrastructure for the elderly at the community and local levels. That is, it is recommended that government agencies allocate budgets to support the management of public spaces sufficient to accommodate the increasing number of elderly people in the future.

ACKNOWLEDGMENT

This research was funded by Thailand Science Research and Innovation Fund (NSRF), Fundamental Fund (FF-67), and Rajamangala University of Technology Srivijaya project number 4696535

REFERENCES

Ayele, A. W., & Zewdie, M. A. (2017). Modeling and forecasting Ethiopian human population size and its pattern. *International Journal of Social Sciences, Arts and Humanities*, *4*(3), 71-82.

Anvari, S., Tuna, S., Canci, M., & Turkay, M. (2016). Automated Box–Jenkins forecasting tool with an application for passenger demand in urban rail systems. *Journal of Advanced Transportation*, *50*(1), 25-49.

Banchongphatthana, P., & Phantharak, K. (2021). *A study and comparison of ARIMA, ARIMAX, and GARCH forecasting models on the SET50 index before and after the COVID-19 pandemic crisis* [Master's thesis]. Thammasat University. (in Thai)

Chanachok, K. (2010). *Guidelines for promoting the quality of life of the elderly in Plapak subdistrict administrative organization, Nakhon Phanom province* [Master's thesis]. Khon Kaen University.

Department of Provincial Administration, Ministry of Interior. (2024). *Official Registration Statistics System*.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, *74*(366a), 427-431.

Duangjai, K. (2011). *Factors affecting the quality of life of the elderly in Lam Sin subdistrict, Si Nakharin district, Phatthalung province* [Master's thesis]. Thaksin University. (in Thai)

Hoksakul, K. (2008). *The participation of the elderly in organizing health activities*. Chiang Mai University. (in Thai)

Husin, W. Z. W., & Aziz, N. A. (2021). Box-Jenkins model for forecasting malaysia life expectancy. *International Journal of Academic Research in Business and Social Sciences*.

Hyndman, R., Athanasopoulos, G., & Bergmeir, C., et al. (2023). *forecast: Forecasting functions for time series and linearmodels.* Retrieved August 28, 2025, from https://pkg.robjhyndman.com/forecast/

IDRUS, N., & MOHAMED, N. (2020). Forecasting the number of airplane passengers using box-jenkins and artificial neural network In Malaysia. *Universiti Malaysia Terengganu Journal of Undergraduate Research*, *2*(4), 89-100.

Jiandon, C., Suwannapong, N., & Boonshuyar, C., et. al. (2011). Quality of life of rural elderly in Wangnamkheaw District, Nakhonratchasima Province. *Journal of Public Health*, *41*(3), 229-239.

Kasem, T., & Tantiphalachiva, K. (1985). *Medical care for the elderly*. Bangkok: Arun Printing.

Lorchirachoonkul, W., & Jittordwech, J. (2015). *Forecasting Techniques.* Bangkok: Chulalongkorn University. Press. (in Thai)

Noknoi, J., & Boripunt, W. (2005). A study on the quality of life of the elderly in Songkhla province. *Princess of Naradhiwas Univ J. 9*(2), 15–27.

Prompak, C. (2013). *Aging society of Thailand*. Bangkok: The Printing Office of the Secretariat of the Senate.

Rochana, K., & Wongprachan, R. (2025). Statistical models for air quality forecasting: A case study of particulate matter less than 2.5 microns (PM2.5) in Chiang Mai Province. *NKRAFA J Sci Technol. 21*(2), 242–258.

Songkhla Provincial Development Plan. (2021). *Provincial Government Report*. Songkhla, Thailand.

Suk-anan, S. (2020). The Promoting Work Strategy for Thai Aging. *Journal of Social Synergy. 11*(1), 1–17.

Thongyai, K. (2019). Factors affecting the happiness of the elderly in Kaoseng community, Mueang district, Songkhla province. *MUT Journal of Business Administration*, *16*(1), 154-179.

Trapletti, A., & Hornik, K. (2024). *tseries: time series analysis and computational finance*. R package version 0.10–58.

World Health Organization. (2025). *Ageing and health*. Retrieved December 22, 2025, from https://www.who.int/news-room/fact-sheets/detail/ageing-and-health