

# Markov Switching Autoregressive Model with Time Varying Parameter and Model Error Compensator for Thailand GDP Forecast

Chuyuttaphon Lueangsuwan<sup>1†</sup> David Banjerdpongchai<sup>2</sup> Wantanee Poonvaralak<sup>3</sup>

<sup>1,2</sup>Center of Excellence in Intelligent Control Automation of Process Systems  
Department of Electrical Engineering, Faculty of Engineering, Chulalongkorn University,

<sup>3</sup>Sasin School of Management, Chulalongkorn University  
254 Phayathai Road, Pathumwan, Bangkok 10330, Thailand

(Email: <sup>1</sup>6772016821@student.chula.ac.th, <sup>2</sup>bdavid@chula.ac.th (corresp. author), <sup>3</sup>wantanee.poonvoralak@sasin.edu)

**Abstract:** Time series modeling plays a crucial role in addressing complex problems across interdisciplinary field, including macro-economic, financial, and engineering. Most of studies primarily concentrate on handle volatility and improve predictive accuracy. The objective of this study is to enhance the time series model, namely, Markov switching autoregressive model with time varying parameter (MS-AR-TVP), to capture complex patterns for prediction. A model error compensator is a key to improve the accuracy of the model by adding the model's residuals and Kim collapsing procedure to find the model parameter. We integrate MS-AR-TVP model with the model error compensator and name the new model MS-AR-TVP-ME. We apply the proposed model to Thailand GDP dataset from 1993 to 2024 to illustrate the effectiveness of the proposed model. We compare the results of MS-AR-TVP and MS-AR-TVP-ME with that of Long Short-Term Memory (LSTM). MS-AR-TVP-ME outperforms MS-AR-TVP and LSTM in terms of mean square error and mean absolute percentage error.

**Keywords:** Markov switching autoregressive model, prediction, time varying parameter, model error compensator, long short-term memory.

## 1. INTRODUCTION

The time series modeling plays a crucial role as an analytical tool across disciplines, including financial [1], macro-economic [2], and engineering. In recent years, both of statistical model and machine learning technique have been major research topics to analyze the sequence of sampling data collected over time constant. With the primary objectives, the model understands the underlying structure and forecast the future data. To enhance accuracy, the volatility of time series data is a challenge for linear time series model, such as autoregressive integrated moving average (ARIMA) [3]. This model was heavily used amongst the econometrics disciplines to solve macro-economic problem.

To begin with the machine learning model, particularly Long Short-Term Memory (LSTM) is one of the famous models to solve the time series data [4]. Unlike the traditional time series model, the assumption such as a stationary is not necessary in machine learning models to capture the fluctuation [5]. Other models, deep leaning architectures like LSTM have indicated the high performance in forecasting high-dimension and noisy time series [6]. However, under the law of large numbers, they require a large volume of dataset for efficiency in training.

The linear time series model such as Markov switching autoregression (MS-AR) model is a novel model to forecast the economic field [2]. The model integrated the Markov switching to the model for handling the fluctuation by separated the model to several regimes [7]. Moreover, the extended MS-AR model called Markov switching autoregression with time varying parameter (MS-AR-TVP). The model adds time varying parameter terms to traditional models able to capture the more fluctuation [9]. The parameter estimation method utilized the Kim collapsing procedure and log-likelihood function

to find the time varying parameters and AR coefficients, respectively [8, 9]. Nevertheless, the MS-AR-TVP models do not have ability to capture the outlier of the data in the long-term [9] due to the volatility effect of data set. Technically, they could not rapidly adjust the model parameters to respond to large residuals.

Our work proposes a novel method named Markov switching autoregressive model with time-varying parameters and model error compensator (MS-AR-TVP-ME). To contribute to the forecasting enhancement, we apply the model error compensator which takes into account the residual of the MS-AR-TVP model to correct significant errors and improve overall prediction.

This paper is organized as follows. Section 2 presents mathematical modeling of MS-AR-TVP. Section 3 describes the parameter estimation. Section 4 describes MS-AR-TVP-ME. Section 5 gives the numerical results of Thailand GDP prediction and compares the performance between three models: LSTM, MS-AR-TVP, and MS-AR-TVP-ME. Finally, Section 6 gives conclusions.

## 2. MATHEMATICAL MODELING

The MS-AR-TVP model was initially developed by Inayati et al. [9]. There are time-varying parameters to capture the dynamic over time. The MS order  $m$  autoregression with time-varying parameter is defined by:

$$y_t = \mu_{S_t}^* + \mathbf{h}_t \boldsymbol{\beta}_{t,S_t} + \varepsilon_t, \quad (1)$$

$\boldsymbol{\beta}_{t,S_t} = \boldsymbol{\delta}_{S_t}^* + \mathbf{F}_{S_t} \boldsymbol{\beta}_{t-1,S_t} + \mathbf{v}_t$ , where  $y_t$  is data at time  $t$ . Let  $\mathbf{h}_t$  and  $\boldsymbol{\beta}_{t,S_t}$  be

$$\mathbf{h}_t = [y_{t-1} \quad y_{t-2} \quad \cdots \quad y_{t-p}]$$

$$\boldsymbol{\beta}_{t,S_t} = [\beta_{t,1,S_t} \quad \beta_{t,2,S_t} \quad \cdots \quad \beta_{t,p,S_t}]^T$$

---

<sup>†</sup> Chuyuttaphon Lueangsuwan is the presenter of this paper.

where  $\beta_{t,1,S_t}$  is a time varying parameter at state  $S_t$  and time  $t$  of autoregression order 1,  $\mathbf{v}_t$  is a variance vector of state equation,  $\mathbf{F}_{S_t}$  is a diagonal matrix which constructed the AR coefficients order 1 to  $p$ .

$$\mathbf{F}_{S_t} = \begin{bmatrix} \phi_{1,S_t} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \phi_{p,S_t} \end{bmatrix}$$

$\mu_{S_t}^* = \mu_{S_t} - \beta_{t,1,S_t}\mu_{S_{t-1}} - \beta_{t,2,S_t}\mu_{S_{t-2}} - \cdots - \beta_{t,p,S_t}\mu_{S_{t-p}}$  where  $\mu_{S_{t-p}}$  denotes the mean of data at state  $S_t$ .

$$\boldsymbol{\delta}_{S_t}^* = [\delta_{1,S_t}^* \quad \delta_{2,S_t}^* \quad \cdots \quad \delta_{p,S_t}^*]^T$$

$$\boldsymbol{\delta}_{p,S_t}^* = \boldsymbol{\delta}_{p,S_t} - \phi_{p,S_t}\boldsymbol{\delta}_{p,S_{t-1}}$$

where  $\boldsymbol{\delta}_{p,S_t}$  means the mean of time varying parameter of autoregression order  $p$  at state  $S_t$  and  $P$  is the transition probability matrix.

$$P = \begin{bmatrix} P_{11} & \cdots & P_{1m} \\ \vdots & \ddots & \vdots \\ P_{m1} & \cdots & P_{mm} \end{bmatrix} \quad (2)$$

The summation of each row in  $P$  is equal to 1.  $\sum_{i=1}^m P_{ji} = 1$  when  $i \in \{1, 2, \dots, m\}$ .  $\varepsilon_t$  is a variance of measurement equation.

### 3. PARAMETER ESTIMATION

This section focuses on parameter approximation for MS-AR-TVP. To approximate the time-varying parameters and AR coefficients in each regime, the Kalman Filtering [9], Kim collapsing Procedure [8], and Hamilton Filtering [7] are applied to estimate the time-varying parameters. The Kalman filtering is employed for estimate the next state of time-varying parameter and its variance. Then, Kim collapsing procedure is applied to determine the parameter in the current state. Finally, Hamilton filtering is utilized to calculate variables from the Kim collapsing procedure, that is  $P[S_{t-1} = i, S_t = j | \gamma_t]$  and  $P[S_t = j | \gamma_t]$ . AR coefficients are calculated by minimizing the negative log-likelihood function. In this paper, we consider the MS(2)-AR(1)-TVP model as follows.

$$y_t = \mu_{S_t}^* + y_{t-1}\beta_{t,1,S_t}, \quad (3)$$

$$\beta_{t,S_t} = \delta_{1,S_t}^* + \phi_{1,S_t}\beta_{t-1,S_t} \quad (4)$$

On each regime of MS(2)-AR(1)-TVP, there is an only one-time varying parameter on the random state at  $S_t = j, S_{t-1} = i$  and  $\gamma_{t-1}$ . The time-varying parameter is determined by

$$\beta_{t|t-1}^{(i,j)} = E[\beta_t | S_t = j, S_{t-1} = i, \gamma_{t-1}] \quad (5)$$

where  $\boldsymbol{\gamma}_{t-1} = [y_{t-1} \quad y_{t-2} \quad \cdots \quad y_1]^T$ .

#### 3.1 Kalman Filtering

Let  $\beta_{t|t-1}^{(i,j)}$  be an estimation of  $\beta_t$  based on the information at  $t-1$  and the previous state is  $i$  and the current state is  $j$ . Form  $B_{t|t-1}$  of dimension  $m \times m$  collecting estimation  $i \in \{1, 2, \dots, m\}$  and  $j \in \{1, 2, \dots, m\}$ .

$$B_{t|t-1} = \begin{bmatrix} \beta_{t|t-1}^{(1,1)} & \cdots & \beta_{t|t-1}^{(1,m)} \\ \vdots & \ddots & \vdots \\ \beta_{t|t-1}^{(m,1)} & \cdots & \beta_{t|t-1}^{(m,m)} \end{bmatrix}$$

In addition,  $b_{t|t-1}$  is given by

$$b_{t|t-1} = [b_{t|t-1}^1 \quad b_{t|t-1}^2 \quad \cdots \quad b_{t|t-1}^m]^T$$

where  $b_{t|t-1}^j = E[\beta_t | S_t = j, \gamma_{t-1}]$ .

Kalman filtering plays a crucial role in prediction. The algorithm of Kalman Filtering is given as follows [9, 10]:

Prediction:

$$\begin{aligned} \beta_{t|t-1}^{(i,j)} &= \delta_{1,j}^* + \phi_{1,j}\mathcal{B}_{t-1|t-1}^i, \\ \boldsymbol{\omega}_{t|t-1}^{(i,j)} &= \phi_{1,j}\boldsymbol{\omega}_{t-1|t-1}^i\phi_{1,j} + \sigma_{1,j}^2, \\ \eta_{t|t-1}^{(i,j)} &= y_{t-1} - \mu_j^* - y_{t-1}\beta_{t|t-1}^{(i,j)}, \\ V_{t|t-1}^{(i,j)} &= y_{t-1} + \boldsymbol{\omega}_{t|t-1}^{(i,j)}y_{t-1} + s_j^2, \end{aligned} \quad (6)$$

Updating:

$$\begin{aligned} \beta_{t|t}^{(i,j)} &= \beta_{t|t-1}^{(i,j)} + \boldsymbol{\omega}_{t|t-1}^{(i,j)}y_{t-1}[V_{t|t-1}^{(i,j)}]^{-1}\eta_{t|t-1}^{(i,j)}, \\ \boldsymbol{\omega}_{t|t}^{(i,j)} &= \left(1 - \boldsymbol{\omega}_{t|t-1}^{(i,j)}y_{t-1}[V_{t|t-1}^{(i,j)}]^{-1}y_{t-1}\right)\boldsymbol{\omega}_{t|t-1}^{(i,j)}, \end{aligned} \quad (7)$$

where  $\boldsymbol{\omega}_{t|t-1}^{(i,j)}$  is an estimation  $\boldsymbol{\omega}_t$  based on the information at  $t-1$  and the previous state is  $i$  and the current state is  $j$ .

$$\boldsymbol{\omega}_{t|t-1}^{(i,j)} = E[(\beta_t - \beta_{t-1})^2 | S_t = j, S_{t-1} = i, \gamma_{t-1}]$$

Let  $W_{t|t-1}$  be a  $m \times m$  matrix consisting of state  $i \in \{1, 2, \dots, m\}$  and  $j \in \{1, 2, \dots, m\}$ .

$$W_{t|t-1} = \begin{bmatrix} \boldsymbol{\omega}_{t|t-1}^{(1,1)} & \cdots & \boldsymbol{\omega}_{t|t-1}^{(1,m)} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\omega}_{t|t-1}^{(m,1)} & \cdots & \boldsymbol{\omega}_{t|t-1}^{(m,m)} \end{bmatrix}$$

where

$$\boldsymbol{\omega}_{t|t-1}^{(i,j)} = E[(\beta_t - \beta_{t-1}^{(i,j)})^2 | S_t = j, S_{t-1} = i, \gamma_{t-1}]$$

Moreover, let  $w_{t|t-1}$  be

$$w_{t|t-1} = [\omega_{t|t-1}^1 \quad \omega_{t|t-1}^2 \quad \cdots \quad \omega_{t|t-1}^m]^T$$

where  $\omega_{t|t-1}^j = E[(\beta_t - \mathcal{B}_{t|t-1}^j)^2 | S_t = j, \gamma_{t-1}]$

$\eta_{t|t-1}^{(i,j)}$  is a conditional estimation prediction error based on the information when  $t = t-1$  and the previous state is  $i$  and the current state is  $j$ .  $V_{t|t-1}^{(i,j)}$  is a conditional variance  $v_t$  based on the information when  $t = t-1$  and the previous state is  $i$  and the current state is  $j$ .  $\sigma_{1,j}^2$  is the variance of  $\varepsilon_t$  at the current state  $j$  and  $s_j^2$  is the variance of  $\mathbf{v}_t$  at the current state  $j$ .

#### 3.2 Kim Collapsing Procedure

From Eq. (6), estimation parameters  $\beta_{t-1|t-1}^i$  and  $\boldsymbol{\omega}_{t-1|t-1}^i$  shall be calculated by Kim collapsing procedure. The parameters relate the current state of  $B_{t|t}$  to  $b_{t|t}$  and  $W_{t|t}$  to  $w_{t|t}$  [8, 9]. Kim collapsing procedure is a key to solving the problem by transforming  $m \times m$  matrix to  $m \times 1$  matrix. The process is as follows.

$$\begin{aligned} b_{t|t}^j &= \frac{\sum_{i=1}^m P[S_{t-1}=i, S_t=j | \gamma_t] \beta_{t|t-1}^{(i,j)}}{P[S_t=j | \gamma_t]}, \\ \boldsymbol{\omega}_{t|t}^j &= \frac{\sum_{i=1}^m P[S_{t-1}=i, S_t=j | \gamma_t] (\boldsymbol{\omega}_{t|t-1}^{(i,j)} + (\mathcal{B}_{t|t-1}^i - \beta_{t|t-1}^{(i,j)})^2)}{P[S_t=j | \gamma_t]} \end{aligned} \quad (8)$$

#### 3.3 Hamilton Filtering

From eq. (8) parameters that are  $P[S_{t-1} = i, S_t =$

$j|\gamma_t]$  and  $P[S_t = j|\gamma_t]$  are still under computation, To estimate parameters  $P[S_{t-1} = i, S_t = j|\gamma_t]$  and  $P[S_t = j|\gamma_t]$ , Hamilton filtering [7] is utilized as follows.

$$P[S_t = j, S_{t-1} = i|\gamma_t] = \frac{f(y_t|S_{t-1}=i, S_t=j, \gamma_{t-1}) \times P[S_{t-1} = i|\gamma_{t-1}]}{\sum_{j=1}^2 \sum_{i=1}^2 f(y_t|S_{t-1}=i, S_t=j, \gamma_{t-1}) \times (P[S_{t-1} = j|\gamma_{t-1} = i] \times P[S_{t-1} = i|\gamma_{t-1}])^2} \quad (9)$$

$$P[S_t = j|\gamma_t] = \sum_{i=1}^2 P[S_t = j, S_{t-1} = i|\gamma_t] \quad (10)$$

$$f(y_t|S_{t-1} = i, S_t = j, \gamma_{t-1}) = \frac{1}{\sqrt{2\pi V_{t-1}^2}} e^{-\left[ y_{t-1} - \left( \mu_{S_t}^* + \gamma_{t-2} (\delta_{1, S_t}^* + \phi_{1, S_t} \beta_{t-2, S_t}) \right) \right]^2} \quad (11)$$

### 3.4 Log Likelihood Function

We define the negative log-likelihood of loss function

$$J = -\ln L(\Phi) = -\sum_{t=1}^T \ln [f(y_t|\gamma_{t-1}; \Phi)]$$

The AR coefficients of each regime are computed by minimizing  $J$ . That is

$$\min_{\Phi} J \quad (12)$$

## 4. MODEL ERROR COMPENSATOR

The model error compensator [11] is a model which models the residual at the same input. The model in Eqs. (3)-(4) is transformed to the Eqs. (15)-(20). Furthermore, the MS(2)-AR(1)-TVP with model error compensator (MS-AR-TVP-ME) is defined as follows:

$$\hat{y}_t = \hat{\mu}_{S_t}^* + \hat{\gamma}_{t-1} \hat{\beta}_{t,1,S_t} \quad (13)$$

$$\hat{\beta}_{t,S_t} = \hat{\delta}_{1,S_t}^* + \hat{\phi}_{1,S_t} \hat{\beta}_{t-1,S_t} \quad (14)$$

$$\hat{P} = \begin{bmatrix} \hat{p}_{11} & \hat{p}_{12} \\ \hat{p}_{21} & \hat{p}_{22} \end{bmatrix} \quad (14)$$

Define model error as  $\tilde{y}_t = y_t - \hat{y}_t$ ,  $(15)$

$$\tilde{y}_t = \tilde{\mu}_{S_t}^* + \tilde{\gamma}_{t-1} \tilde{\beta}_{t,1,S_t} \quad (16)$$

$$\tilde{\beta}_{t,S_t} = \tilde{\delta}_{1,S_t}^* + \tilde{\phi}_{1,S_t} \tilde{\beta}_{t-1,S_t} \quad (16)$$

$$\tilde{P} = \begin{bmatrix} \tilde{p}_{11} & \tilde{p}_{12} \\ \tilde{p}_{21} & \tilde{p}_{22} \end{bmatrix} \quad (17)$$

where the  $\hat{\cdot}$  stands for a parameter utilized in the nominal model as MS-AR-TVP and  $\tilde{\cdot}$  stands for a parameter utilized in the residual model of MS-AR-TVP. Thus, the output of MS-AR-TVP-ME denotes as

$$\check{y}_t = \hat{y}_t + \tilde{y}_t \quad (18)$$

### 4.1 Algorithm

The algorithm of MS-AR-TVP-ME is described as follows. Denotes  $T$  the number of the data. The stopping criterion is specified by the iteration reaching Iter\_max.

1. Clean data  $y_1, y_2, \dots, y_T$
- For  $i$  in range (2): #for nominal and error model
2. Change data to stationary time series
3. Find  $P$  matrix
4. Split the data to train validation and test data
5. For iter in range (Iter\_max):
- For  $j$  in range (T):
- Set  $\beta_{0|0}^j, w_{0|0}^j$  for Kalman filter and  $P[S_t = j|\gamma_0]$
- Execute Kalman filter Eqs. (6)-(7)
- Run Hamilton filter Eqs. (9)-(11)
- Utilize Kim collapsing procedure Eq. (8)
- Minimize Eq. (12) by Nelder-Mead method
- Update the optimal AR coefficients in validation
- Construct error data  $\tilde{y}_t$

Find the MS-AR-TVP-ME  $\hat{y}_t + \tilde{y}_t$

## 4.2 Model Evaluation

The model evaluation is based on the mean square error (MSE) and mean absolute percentage error (MAPE). To begin with MSE, MSE measures the average of the squares of error. Then, MAPE indicates the percentage by measuring the error and actual data.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (19)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% \quad (20)$$

where  $y_i$  is an actual data and  $\hat{y}_i$  is a predicted data.

## 5. NUMERICAL RESULTS

The Gross Domestic Product (GDP) of Thailand from 1993 to 2024 [13] is utilized in the experiment. The data which has 128 data points for each quarter is separated into three parts: training data, validation data, and testing data equal to 71, 31, and 26 points, respectively. Fig.1 (top) illustrates the Thailand GDP from 1993 to 2024. Let  $G_t$  denote the GDP at index  $t$ . The original time series is a non-stationary, indicated by the Augmented Dickey-Fuller test. The p-value of 0.9886 is above 0.05 significantly threshold. It implies that the null hypothesis cannot be rejected. Thus, the original time series does not have fixed mean and variance over time. We modify the GPD data to stationary series as

$$Z_t = \ln(G_t) - 0.995 \ln(G_{t-1}) \quad (21)$$

where  $Z_t$  is a modified GDP. This transformation applies the difference of logarithm function with a smoothing coefficient of 0.995. The result of ADF test, which obtained a p-value of 0.01375—below the 0.05 significance threshold. It implies that the modified time series has a constant mean and variance over time and suitable to time series model. Fig. 1 (bottom) shows  $Z_t$ .

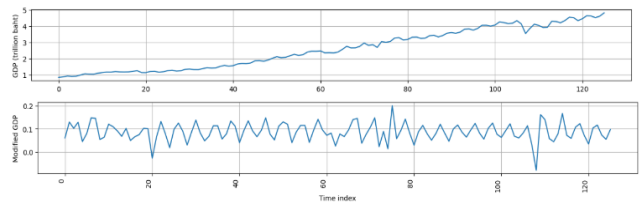


Fig. 1 GDP and modified GDP.

We obtain MS-AR-TVP-ME model for  $Z_t$ .

$$S_t = 1: \begin{aligned} \hat{Z}_t &= \hat{\mu}_1^* + \hat{Z}_{t-1} \hat{\beta}_{t,1,1} \\ \hat{\beta}_{t,1,1} &= \hat{\delta}_{1,1}^* + 0.0592 \hat{\beta}_{t-1,1,1} \end{aligned}$$

$$S_t = 2: \begin{aligned} \hat{Z}_t &= \hat{\mu}_2^* + \hat{Z}_{t-1} \hat{\beta}_{t,1,2} \\ \hat{\beta}_{t,1,2} &= \hat{\delta}_{1,2}^* + 0.0049 \hat{\beta}_{t-1,1,2} \end{aligned}$$

$S_t = 1$  means the economic expansion, and vice versa.

$$\hat{P} = \begin{bmatrix} 0.4421 & 0.5588 \\ 0.5588 & 0.4421 \end{bmatrix}$$

$$S_t = 1: \begin{aligned} \tilde{Z}_t &= \tilde{\mu}_1^* + \tilde{Z}_{t-1} \tilde{\beta}_{t,1,1} \\ \tilde{\beta}_{t,1,1} &= \tilde{\delta}_{1,1}^* - 0.0840 \tilde{\beta}_{t-1,1,1} \end{aligned}$$

$$S_t = 2: \begin{aligned} \tilde{Z}_t &= \tilde{\mu}_2^* + \tilde{Z}_{t-1} \tilde{\beta}_{t,1,2} \\ \tilde{\beta}_{t,1,2} &= \tilde{\delta}_{1,2}^* - 0.02466 \tilde{\beta}_{t-1,1,2} \end{aligned}$$

$$\tilde{P} = \begin{bmatrix} 0.0833 & 0.9167 \\ 0.2000 & 0.8000 \end{bmatrix}$$

$$\check{Z}_t = \hat{Z}_t + \tilde{Z}_t$$

The MS(2)-AR(1)-TVP model utilizes  $\hat{Z}_t$ . The LSTM model was selected by varying the hidden layers and look backsteps. For the optimal model of LSTM, there were 50 hidden layers and 3 look backsteps.

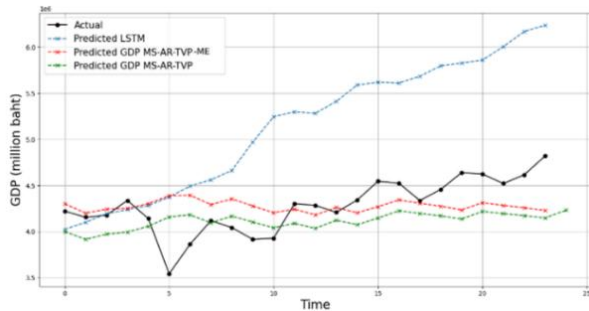


Fig.2 Comparison of three models on the testing data.

Table 1. Performance comparison between three models on the training data.

Method	MSE ( $\times 10^{22} THB^2$ )	MAPE (%)
MS-AR-TVP-ME	4.0277	10.9089
MS-AR-TVP	2.7744	7.6168
LSTM	2.8682	9.4918

Table 2 Performance comparison between three models on the testing data.

Method	MSE ( $\times 10^{22} THB^2$ )	MAPE (%)
MS-AR-TVP-ME	9.82	5.8384
MS-AR-TVP	10.539	6.5552
LSTM	104.60	20.8091

Fig. 2 demonstrates the performance of the MS-AR-TVP-ME comparing with the famous models which are the MS-AR-TVP and LSTM which was modeled with 50 hidden layers and 3 look back steps. Focusing on the performance of training data illustrated in Table 1. To explain the numerical results, the traditional MS-AR-TVP outperforms the proposed model and LSTM. However, turning to the testing data, the LSTM cannot handle the robustness of data shown in Table 2. To sum up, MS-AR-TVP-ME surpasses both traditional MS-AR-TVP and LSTM by integrating the error term to fill the gap of error.

The initial conditions of model parameters are crucial. When the initial values of  $\beta$  are close to zero, the model effectively defaults to its long-term mean. Thus, initial  $\beta_0$  was chosen based on the volatility of the data. selecting appropriate initial AR coefficients is also important to avoid local minima of optimization problem Eq. (12). In the same way, initial  $P[S_t = j|\gamma_0]$  is a probability of each regime to predict the next state—if  $P[S_t = j|\gamma_0]$  is near 1, the process will begin at state  $j$ .

At last, the initial  $w_0$  used in Kalman filter affects both the prediction and updating step. It should be selected to reflect the short-term fluctuation of the data.

## 6. CONCLUSIONS

This paper proposes the novel forecasting model, MS-AR-TVP-ME, for predicting Thailand GDP. This proposed approach is designed to improve the forecasting

accuracy by capturing the fluctuation and filling the error with model error compensation in long-term data. The results indicate that the MS-AR-TVP-ME achieves 15% improvement in MAPE and nearly ten times reduction in MSE compared to that of the LSTM. Moreover, the proposed model outperforms MS-AR-TVP model. Although, the LSTM has strong performance on the training data due to overfitting which lacks robustness in testing data but since The GDP collected data is known as limited numbers data domain that consists of 126 sample sizes. In contrast, the prediction from the extended linear model like a MS-AR-TVP-ME is more suitable under data constraints. In conclusion, our proposed method shows that it outperforms both traditional model and LSTM in terms of MSE and MAPE metrics. For the future work, we will explore the different function cost and the higher order model.

## REFERENCES

- [1] Y. Guo, "Applications of time series analysis in Quantitative Finance," *SSRN*, Jan. 2025, doi: 10.2139/ssrn.5140015.
- [2] M. Chauvet and S. Potter, "Forecasting output," in *Handbook of economic forecasting*, 2013, pp. 141–194. doi: 10.1016/b978-0-444-53683-9.00003-7.
- [3] M. G. Kendall, G. E. P. Box, and G. M. Jenkins, "Time series analysis, forecasting and control.," *Journal of the Royal Statistical Society Series a (General)*, vol. 134, no. 3, p. 450, Jan. 1971, doi: 10.2307/2344246.
- [4] X. Ma, Z. Tao, Y. Wang, H. Yu, and Y. Wang, "Long short-term memory neural network for traffic speed prediction using remote microwave sensor data," *Transportation Research Part C Emerging Technologies*, vol. 54, pp. 187–197, Apr. 2015, doi: 10.1016/j.trc.2015.03.014.
- [5] S. Hochreiter and J. Schmidhuber, "Long Short-Term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.
- [6] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and Computing*, vol. 14, no. 3, pp. 199–222, Jul. 2004, doi: 10.1023/b:stco.0000035301.49549.88.
- [7] J. D. Hamilton, "A new approach to the economic analysis of nonstationary time series and the business cycle," *Econometrica*, vol. 57, no. 2, p. 357, Mar. 1989, doi: 10.2307/1912559.
- [8] C.-J. Kim and C. R. Nelson, *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*. 1999. [Online]. Available: <http://ci.nii.ac.jp/ncid/BA41883574>
- [9] S. Inayati, N. Iriawan, and N. Irhamah, "A Markov Switching Autoregressive Model with Time-Varying Parameters," *Forecasting*, vol. 6, no. 3, pp. 568–590, Jul. 2024, doi: 10.3390/forecast6030031.
- [10] A. C. Harvey, *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press, 1990.
- [11] H. Umei, H. Okajima, N. Matsunaga, and T. Asai, "A design method of model error compensator for MIMO systems," *Transactions of the Institute of Systems Control and Information Engineers*, vol. 27, no. 2, pp. 67–72, Jan. 2014, doi: 10.5687/iscie.27.67.
- [12] M. Camacho, G. Perez-Quiros, and P. Poncela, "Markov-Switching dynamic factor models in real time," *SSRN Electronic Journal*, Jan. 2012, doi: 10.2139/ssrn.2002667.
- [13] Office of the National Economic and Social Development Council, "Quarterly Gross Domestic Product of Thailand" <https://www.nesdc.go.th>