The 6th PCSHS Science Symposium 2026
"Driving the Future with Scientific Innovation"

# Forecasting Rubber Prices and Their Impact on Farmers' Income Using Mathematical Models

Phawinkit Jevprasertpan[1] and Jirawat Onngam[1*]
Manat Pronmanee[1*]
[1]Princess Chulabhorn Science High School Phitsanulok, Makham-Sung, Mueang, Phitsanulok 65000, Thailand
*e-mail: jirawat.onn@pcshstrg.ac.th

## Abstract

This study aims to develop and evaluate mathematical forecasting models for forecasting cup lump rubber prices one year in advance. The predictive performance of three models including SARIMAX, Facebook Prophet, and the Bayesian Dynamic Linear Model (BDLM) was analyzed under three temporal forecasting schemes such as Lagging, Slice, and Two-step. Additionally, independent external factors influencing cup lump rubber prices from January 1, 2014, to December 31, 2024, were investigated and compared. Feature selection was conducted using the Cross-Correlation Function (CCF), LASSO Regression, and Random Forest algorithms to identify key predictors. The results indicate that the Bayesian Dynamic Linear Model (BDLM) outperformed the other candidates. The optimal model configuration incorporated the crop price index ($r = -0.536$), STR20 rubber block export volume ($r = 0.111$), and STR rubber block export volume ($r = 0.137$) as co-exogenous variables. The forecast yielded an mean forecast error of 2.33 Baht, The standard deviation of the forecast error was approximately 3.06% of the forecast price, a Mean Absolute Percentage Error (MAPE) of 4.57%, and a coefficient of determination (R2) of 0.63 Furthermore, integrating these price forecasts into a Vensim-based income simulation model alongside internal farm-level variables obtained from farmer questionnaires (specifically fertilizer usage, pesticide application, and irrigation systems) demonstrated that could help farmers increase net profit by 14.7% and reduce production costs by 37.2% compared to decision-making without forecasting information.

*Keywords: Rubber Price Forecasting, Feature Selection, Income Simulation Model, Cup Lump Rubber, Bayesian Dynamic Linear Model (BDLM)*

## Introduction

Cup lump rubber constitutes a vital economic crop widely produced and commercialized by the majority of rubber smallholders in Thailand due to its straightforward and uncomplicated primary processing requirements. Beyond its raw sale, cup lump rubber serves as the principal precursor for manufacturing Standard Thai Rubber 20 (STR20) block rubber, a critical raw material for the tire manufacturing industry and one of the nation's leading export commodities. Nonetheless, rubber farmers persistently confront severe economic challenges, notably prices falling below production costs. This price depression is driven by volatile global and domestic macroeconomic conditions, including a 53.90% contraction in the demand for truck and bus tires across various countries. Conversely, a 14.27% growth in the demand for medical gloves has periodically exerted positive pressure on rubber prices (Office of Industrial Economics and Plastics Institute of Thailand, 2024). Recognizing these issues, the authors identified the critical need to develop mathematical models for rubber price forecasting to accurately project future price trends. Such predictive tools are essential to empower farmers in production planning and efficient cost management. Furthermore, to analyze the financial dynamics of smallholder households, we integrated a system dynamics simulation model using Vensim. This simulation aims to establish structured guidelines for farm-level operational adjustments, mitigate economic risks associated with price volatility, and enhance household financial stability.

## Methodology

### 1). Collection of Rubber Prices and Influencing Factors

Monthly cup lump rubber prices and structural exogenous factors influencing rubber price dynamics were compiled from reputable government agencies and national databases, including the Office of Agricultural Economics, the Department of Agriculture, and the Rubber Authority of Thailand. The compiled dataset encompasses rubber price indices, agricultural production statistics, block rubber export volumes, substitute synthetic product prices, foreign exchange rates, and various macroeconomic indicators from 2014 to 2025, compiled as monthly time-series data. Concurrently, digital questionnaires were administered to collect primary financial data, socio-economic characteristics, and price risk perceptions from a representative sample of local rubber smallholders.

### 2). Screening of Exogenous Factors Influencing Rubber Prices

To mitigate the curse of dimensionality and optimize the predictive accuracy of the models, statistical and machine learning algorithms were implemented to evaluate and screen the structural exogenous factors. Feature selection was conducted using a consensus approach across three distinct methodologies: the Cross-Correlation Function (CCF) for temporal relationships, LASSO Regression for regularization and sparsity, and the Random Forest algorithm for non-linear feature importance ranking. Exogenous variables were selected based on their statistical significance and high feature importance scores.

### 3). Data Preparation of Influencing Factors

The selected exogenous variables were subsequently aligned and structured using three distinct temporal data preparation schemes to overcome the predictive limitations of regression models that require future exogenous values. The three temporal alignment schemes are defined as follows:

- Slice Scheme: Incorporates the most recently observed value of the exogenous variables, holding them constant throughout the Forecastinghorizon
- Lagging Scheme: Applies a 12-month lag to the exogenous variables, allowing historical indicators to project future prices directly.
- Two-step Scheme: Forecasts the future values of the exogenous variables in the first step, and then inputs these forecasted trajectories into the primary cup lump rubber price forecasting models.

### 4). Construction of Mathematical Models

Three linear and dynamic timeseries mathematical forecasting models—SARIMAX, Facebook Prophet, and the Bayesian Dynamic Linear Model (BDLM)—were developed and implemented in Python utilizing the statsmodels, prophet, and pydlm libraries. Hyperparameter optimization was conducted via a systematic Grid Search technique, with model selection determined by minimizing the Akaike Information Criterion (AIC) to ensure parsimonious configurations.

**4.1). Parameters of the SARIMAX Model**

The SARIMAX model incorporates parameters (p, d, q) and (P, D, Q, s). The non-seasonal parameters consist of p (autoregressive order for past values dependency), d (differencing order to achieve stationarity), and q (moving average order of past forecast errors). The seasonal parameters P, D, and Q capture seasonal counterparts, with s set to 12 representing the monthly seasonal cycle.

**4.2). Parameters of the Facebook Prophet Model**

The Facebook Prophet model incorporates parameters (mode)(cps, sps). The Changepoint Prior Scale (cps) controls the flexibility of the trend, the Seasonality Prior Scale (sps) adjusts the magnitude of seasonal fluctuations, and the Seasonality Mode (mode) determines whether seasonal patterns are additive or multiplicative to adapt to rubber price variations.

**4.3). Parameters of the BDLM Model**

The Bayesian Dynamic Linear Model (BDLM) is structured as a state-space formulation, which consists of parameters (w_trend, w_season). The w_trend parameter controls the trend variation rate, while w_season regulates the adaptation of seasonal components to dynamically update states in response to internal trends and exogenous shocks.

**5). Evaluation and Comparison of Models**

A 12-month-ahead out-of-sample forecasting simulation was conducted for the year 2025 to evaluate the predictive accuracy of the three models. The forecasted results were validated against actual historical market prices using standard evaluation metrics: Root Mean Squared Error (RMSE), Coefficient of Determination (R2), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). To prevent model overfitting and maintain mathematical stability, the number of exogenous predictors was constrained to a maximum of 1 to 3 variables.

**6). Income Simulation**

A system dynamics simulation model was constructed using Vensim to analyze and compare the economic outcomes of rubber smallholders. Under the hypothesis that prior knowledge of rubber price trends allows farmers to adopt proactive agricultural practices (such as optimizing fertilizer application, managing water usage on the farm, and managing pesticide use), the simulation evaluates how price forecasting information enables farmers to optimize resource allocation, reduce production costs, and maximize net returns.

## Results and Discussion

**Screening Results of Exogenous Factors**

Following the collection of 23 initial structural exogenous variables, statistical correlation and machine learning feature importance analyses were conducted to identify the most significant drivers of cup lump rubber prices. The top 5 exogenous predictors are summarized in Table 1

**Table 1** Correlation coefficients (r) between cup lump rubber prices and top 5 exogenous variables

| Exogenous Variables | LASSO | Random Forest | CCF (Lag 0) | CCF (Lag 12) |
|---|---|---|---|---|
| INR/THB Exchange Rate | -0.70 | 0.07 | -0.39 | -0.66 |
| Crop Price Index | -0.30 | 0.16 | -0.54 | -0.66 |
| API (Season) | -0.14 | 0.50 | -0.16 | 0.51 |
| Export Volume of STR20 | 0.48 | 0.10 | 0.71 | -0.28 |
| Export Volume of TSR | 0.36 | 0.08 | 0.67 | -0.24 |

**Note:** Let API represent the major agricultural production index (seasonally adjusted). A positive value indicates a direct relation with the cup lump rubber price, while a negative value indicates an inverse relation.

### Results of Data Alignment Schemes

To determine the optimal temporal scheme for incorporating exogenous inputs, the performance of the three data alignment configurations (Slice, Lagging, and Two-step) was evaluated across all forecasting models. The comparative RMSE results are detailed in Table 2.

**Table 2** Comparison of RMSE values across different data alignment schemes

| RMSE | SARIMAX | Facebook | BDLM |
|---|---|---|---|
| slice | 3.32 | 3.74 | 3.57 |
| lagging | 3.33 | 4.28 | 3.94 |
| two-step | 3.25 | 3.66 | 3.06 |

**Note:** The metrics presented in the table demonstrate significant statistical differences among the forecasting models. This is further supported by other model evaluation metrics not explicitly displayed in this table, which also exhibit notable variations in their predictive performance.

The comparative analysis confirms that the Two-step data alignment scheme yields the highest forecasting accuracy. Dynamically updating the exogenous variables based on their predicted future trends significantly outperformed the Slice scheme (holding values constant) and the Lagging scheme (12-month historical shift). Time-series trend analysis indicated that the underlying trend component of rubber prices underwent structural shifts due to global macroeconomic changes and technological integration. Consequently, the dynamic Two-step scheme successfully captures these shifting trend dynamics, resulting in superior predictive robustness. This dynamic updating mechanism ensures that the forecasting model remains sensitive to external market changes, providing rubber smallholders with more reliable and actionable price projections.

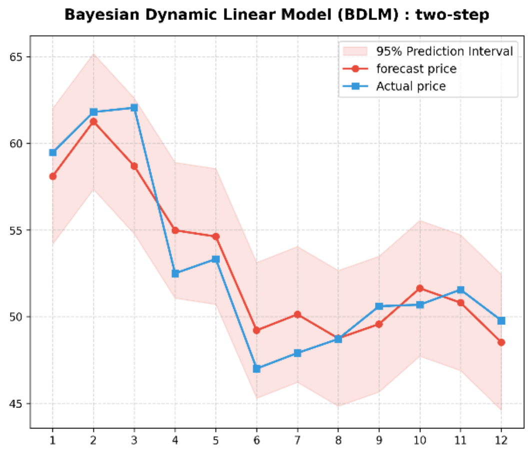
### Comparison of the Most Accurate Models

**Table 3** Overall predictive accuracy metrics of the forecasting models Based on the comprehensive experimental evaluations:

| Metrics | RMSE (Baht) | $R^2$ | MAE (Baht) | MAPE % |
|---|---|---|---|---|
| SARIMAX | 3.25 | 0.59 | 2.68 | 4.98 |
| Facebook | 3.66 | 0.48 | 2.98 | 5.06 |
| BDLM | 3.06 | 0.63 | 2.33 | 4.57 |

The empirical results demonstrate that the Bayesian Dynamic Linear Model (BDLM) combined with the Two-step data alignment scheme achieved the highest predictive performance. The optimal hyperparameter settings were a trend variance (w_trend) of 0.01 and a seasonal variance (w_season) of 0.001. These low transition variances indicate that the internal stochastic fluctuations of the price trend and seasonal components are minimal. Consequently, major price fluctuations are strongly driven by the exogenous factors—specifically, STR20 rubber block exports, total block rubber exports, and the crop price index—rather than sudden internal shifts in the underlying time-series structure. This finding underscores the importance of monitoring global trade volumes and broader agricultural indices when predicting local cup lump rubber prices.

**Fig. 1** 12-month cup lump rubber price forecast using the Two-step BDLM model




### Income Simulation Results

The system dynamics model was populated using survey responses from 64 local rubber farmers, capturing key farm-level characteristics such as average expenditures, fertilizer formulas, tapping frequency, and financial risks. Using Vensim, a dynamic simulation compared the financial outcomes of farmers utilizing rubber price forecasts (proactive scenario) against those who operate without forecasts (reactive scenario).

**Fig. 2** Monthly accumulated profit of rubber farmers

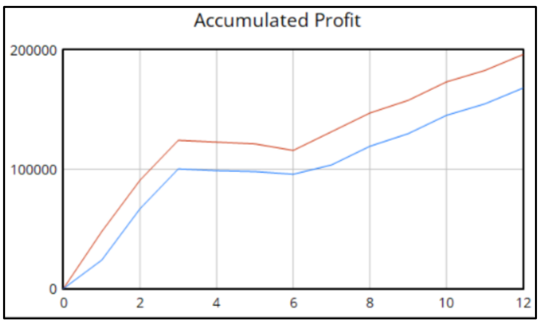
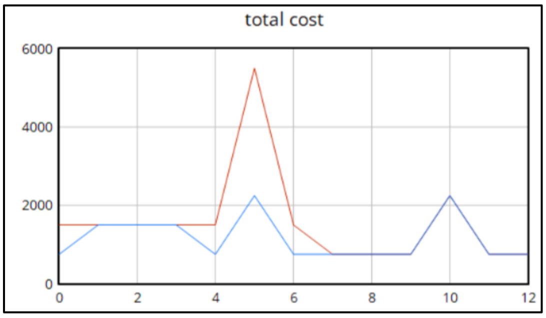



**Fig. 3** Monthly cost of rubber farmers




**Note:** The blue curves represent the baseline scenario without forecasting information.
The red curves represent the forecast-informed scenario.

The simulation results demonstrate that integrating price forecasts into decision-making allows farmers to reduce production costs by 37.2% and increase net profit by 14.7% compared to the baseline scenario without forecast information. This confirms that mathematical price forecasting is a highly effective tool for improving household agricultural economy.By providing smallholders with advanced price information, they can make proactive adjustments to their farming schedules, optimize fertilizer and chemical usage, and mitigate the risks of selling during price troughs, thereby securing sustainable long-term livelihoods.

### Conclusions

This study demonstrates that the Bayesian Dynamic Linear Model (BDLM) coupled with the Two-step data alignment scheme is the most suitable method for forecasting cup lump rubber prices, yielding the lowest error standard deviation of 3.06%. The overall model performance ranks as follows: BDLM, SARIMAX, and Facebook Prophet. The key exogenous drivers influencing price movements are STR20 rubber block exports, total block rubber exports, and the crop price index of Thailand. Ultimately, integrating these mathematical price predictions into agricultural management allows farmers to proactively adapt to price fluctuations, leading to a 14.7% increase in net profit and a 37.2% reduction in production costs, fully achieving the objectives of this research. These findings provide a valuable framework for policy makers and agricultural cooperatives to develop supportive decision systems for smallholders.


## Acknowledgments

The authors wish to express their sincere gratitude to their project advisor, Mr. Manat Pronmanee, for his invaluable guidance, support, and constructive feedback throughout the course of this study.

We also convey our appreciation to Mrs. Preeyaporn

Jevprasertpan for her assistance in coordinating and distributing the surveys to the rubber farmers, which was vital for the empirical analysis.

Finally, we thank Princess Chulabhorn Science High School Phitsanulok and the faculty of the Mathematics Department for providing the resources, administrative assistance, and an encouraging research environment.


## References


Putsorn, P. (2013). Forecasting the Export Volume of Ribbed Smoked Sheet Rubber Grade 1 Using the Box-Jenkins Method. Undergraduate Special Project, Burapha University, Department of Mathematics, Faculty of Science. (in Thai)

Rungruanganunt, W., et al. (2014). A Study on the Impact of Factors Affecting the Price Stabilization of Para Rubber. Undergraduate Special Project, King Mongkut's University of Technology North Bangkok. (in Thai)

Khan, F., et al. (2021). Forecasting daily new infections, deaths and recovery cases due to COVID-19 in Pakistan by using Bayesian Dynamic Linear Models. PLOS ONE, 16(6), e0253367. https://doi.org/10.1371/journal.pone.0253367

Hossain, M. M., et al. (2022). Monthly rainfall prediction at catchment level with the Facebook Prophet model using observed and CMIP5 decadal data. *Hydrology*, *9*(6), 111. https://doi.org/10.3390/hydrology9060111