

Forecasting and Structural Analysis of Vietnam's Import-Export of Goods and Services, 2015-2024: A Multi-Model Time-Series Approach

Tran Ngoc Dan^{1*}

¹Faculty of Information Technology, Thuyloi University. 175 Tay Son - Kim Lien Ward - Hanoi City, Vietnam

*Correspondence: Tran Ngoc Dan email: tranngocdan@tlu.edu.vn

Abstract-Purpose: This study analyzes and forecasts the structure and trends of Vietnam's imports and exports of goods and services during 2015-2024, with an emphasis on industry, economic sector, SITC group, country/territory markets, major commodity groups, trade-to-GDP ratio, and services trade. **Method:** The study uses descriptive statistics, structural comparison, time-series visualization, and six Minitab forecasting models: Trend Analysis, Decomposition, Moving Average, Single Exponential Smoothing, Double Exponential Smoothing, and Winter's Method. MAPE is used as the main selection criterion; if MAPE is equal, MSD is used as a tie-breaker. **Dataset:** The dataset consists of official-style annual trade tables ID277-ID291, including total import-export turnover, export and import value by sector, SITC classification, partner markets, major commodities, trade-to-GDP ratio, and services trade. **Results:** For total import-export turnover (ID277), Trend Analysis is optimal with MAPE = 3, MAD = 9,345, and MSD = 145,029,092. For exports by market (ID282), Decomposition is optimal with MAPE = 4 and MSD = 256,030,069. For major exports (ID283), Winter's Method is optimal with MAPE = 97 and MSD = 9,087,557. For major imports (ID289), Single Exponential Smoothing is optimal with MAPE = 35 and MSD = 3,374,986. **Value:** The research provides a quantitative framework for identifying export-import dynamics, trade concentration risks, FDI dependence, and model-based policy implications for Vietnam's trade strategy.

Keywords-Import-export forecasting; Vietnam trade; Minitab; MAPE; SITC; FTA; time-series analysis.

I. INTRODUCTION

The research object is Vietnam's import-export structure and trend from 2015 to 2024. The source material defines the research scope through tables ID277-ID291, covering total merchandise trade, exports and imports by economic sector, SITC classification, country group, country and territory, major commodity groups, the export-import-to-GDP ratio, and services trade. This multidimensional design is important because Vietnam's trade performance is shaped simultaneously by commodity structure, partner-market dependence, foreign-invested enterprises, logistics quality, and free trade agreements.

Vietnam's trade growth after 2015 occurred in a volatile environment. Recent literature confirms that foreign trade forecasting is difficult because exports and imports have nonlinear relationships with income, exchange rates, supply-chain disruptions, and international policy shocks [1]-[4]. Studies on Vietnam also show that FTAs such as EVFTA, CPTPP, and RCEP affect trade flows, market access, and export efficiency [5]-[10]. These findings support the need for a structured, model-based analysis rather than descriptive reporting only.

Previous studies have three limitations. First, many studies analyze only one dimension of trade, such as export value, import value, FTA effects, or logistics efficiency, without integrating goods, services, SITC structure and markets in one framework. Second, several macro trade reports are descriptive and do not compare alternative forecasting models by error indicators. Third, advanced AI models improve forecasting but often require high-frequency and multivariate data, while many Vietnamese enterprises and students work with annual public datasets.

The research questions are: RQ1. How did Vietnam's import-export structure change during 2015-2024? RQ2. Which time-series forecasting model gives the lowest error for each ID277-ID291 trade

dataset? RQ3. What do the results imply for Vietnam's export-market diversification, import dependence, FDI dependence and services-trade development?

The value of this study is that it converts a classroom trade dataset into an IJETMS-style research article with a transparent forecasting rule. The output supports enterprises in market planning, inventory preparation, import-material planning and partner-market selection. It also supports policy agencies in evaluating trade integration, logistics bottlenecks, trade concentration, and the role of FDI in export upgrading.

The objectives are to describe the ID277-ID291 data architecture, propose a quantitative research model, compare six time-series methods, select optimal models using MAPE and MSD, present concrete results, and discuss the results using the S.W.O.T principle. The paper is organized as follows: Section II reviews related literature; Section III proposes the research model and mathematical formulas; Section IV reports the empirical results; Section V discusses S.W.O.T implications; and Section VI concludes with contributions, limitations, and future research directions.

II. REVIEW OF RELATED LITERATURE

A. Foreign trade forecasting and model comparison

Recent trade-forecasting literature emphasizes model comparison and hybridization. Yang et al. [1] combine structural economic variables and deep learning for global exports and imports and show that ensemble learning can improve accuracy. Mohamed [2] applies hierarchical time-series forecasting to international trade and shows that ARIMA and exponential smoothing can be combined to improve import-export forecasts. Zhang [3] combines ARIMA with composite quantile regression to exploit both time-series dynamics and explanatory variables. Rincon-Yanez et al. [4] use knowledge graphs and embeddings to predict international trade flows and improve interpretability.

These studies reveal the weakness of purely traditional methods: a single method can miss structural changes, nonlinear patterns, or external shocks. However, they also show that baseline models such as ARIMA, exponential smoothing, decomposition, and moving average remain useful for small datasets because they are interpretable, easy to implement, and suitable for decision-making when data are annual rather than monthly or daily.

B. Vietnam trade structure, FTAs, logistics and export efficiency

Vietnam-related studies stress the importance of FTAs and market structure. Duong and Phan [5] evaluate the trade and investment effects of EVFTA and indicate that Vietnam-EU bilateral trade can grow substantially under tariff and policy scenarios. Nguyen [6] evaluates CPTPP and EVFTA effects on Vietnam's trade flows using gravity-model evidence. Estrades et al. [7] estimate the economic impact of RCEP and show that trade-cost reduction and rules of origin can enlarge regional trade effects.

Recent sectoral and market studies add more detail. Tian et al. [8] analyze China-Vietnam agricultural trade competitiveness and complementarity, showing that bilateral commodity relations need market intelligence and competitiveness indicators. Bui and Le [9] examine logistics quality, FTAs and Vietnam's export efficiency using SFA-GMM and show that logistics and FTAs interact in shaping export performance. Nguyen and Choi [10] evaluate Vietnam's export efficiency after Doi Moi reforms, highlighting the continuing need to improve institutions, infrastructure and partner-specific efficiency.

The literature gap is therefore clear. Existing studies provide strong evidence about FTAs, logistics, bilateral trade and AI forecasting, but they rarely transform the full ID277-ID291 data architecture into a compact model-selection framework for classroom and policy use. This study addresses that gap by combining structural description with Minitab-based time-series evaluation.

III. PROPOSED RESEARCH MODEL

The proposed model follows four stages: data construction, structural feature analysis, multi-model forecasting, and decision interpretation (Figure 1). The input data are the annual series from the ID277-ID291 trade architecture. The output is an optimal model for each dataset and an interpretation of structural trade implications.

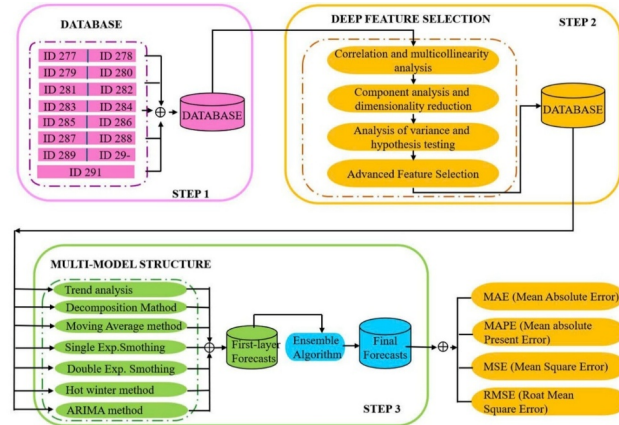


Fig. 1. Proposed research pipeline

A. Step 1: Data construction

Let $Y_{i,t}$ denote the annual trade value of dataset i in year t , where i represents one of the trade tables ranging from ID277 to ID291, and t spans the period from 2015 to 2024. The data are organized into a unified time-series matrix. For each dataset i , the annual trade value is expressed as: $Y_i = \{Y_{i,2015}, Y_{i,2016}, \dots, Y_{i,2024}\}$.

B. Step 2: Feature and structural analysis

Correlation analysis is employed to examine the degree of co-movement among key trade indicators, including exports, imports, market concentration, and services trade. The Pearson correlation coefficient is calculated as: $r_{xy} = \frac{\sum (x_t - \bar{x})(y_t - \bar{y})}{\sqrt{\sum (x_t - \bar{x})^2 \sum (y_t - \bar{y})^2}}$, where x_t and y_t represent the observed values of two trade variables at time t , and $\bar{x}$ and $\bar{y}$ denote their respective sample means.

Principal Component Analysis (PCA) is applied to reduce multiple correlated trade variables into a smaller number of uncorrelated principal components representing the major structural dimensions of international trade, such as export scale, import dependence, and market concentration. The contribution of each principal component is measured by its explained variance ratio: $\text{Explained Variance Ratio}_j = \frac{\lambda_j}{\sum \lambda_j}$, where λ_j is the eigenvalue corresponding to the j -th principal component.

Analysis of Variance (ANOVA) is used to determine whether the mean trade values differ significantly across economic sectors or commodity groups. The test statistic is defined as: $F = \frac{MS_{\text{between}}}{MS_{\text{within}}}$, where MS_{between} and MS_{within} denote the mean squares between and within groups, respectively.

Lasso regression is proposed as an extension for future research to identify the most influential explanatory variables when additional macroeconomic indicators, such as gross domestic product (GDP), exchange rates, foreign direct investment (FDI), and logistics performance indices, are incorporated into the analysis. The Lasso optimization problem is formulated as: $\min_{\beta} \left[\frac{1}{2n} \sum_{t=1}^n (y_t - X_t \beta)^2 + \lambda \sum_j |\beta_j| \right]$, where β is the vector of regression coefficients, X_t represents the explanatory variables, y_t is the response variable, n is the sample size, and λ is the regularization parameter controlling the degree of coefficient shrinkage. This approach simultaneously performs

variable selection and regularization, thereby improving model interpretability and reducing the risk of overfitting.

C. Step 3: Multi-model forecasting

Six forecasting methods are implemented in Minitab to analyze historical trade data and generate future projections.

Trend Analysis is used to capture the long-term deterministic movement of a time series and is expressed as: $Y_t = a + bt + e_t$, where Y_t is the observed trade value at time t , a is the intercept, b is the trend coefficient, and e_t is the random error term.

Decomposition Analysis separates the observed time series into its underlying components, including trend, seasonal or cyclical patterns, and irregular fluctuations. The additive decomposition model is defined as: $Y_t = T_t + S_t + C_t + I_t$, where T_t , S_t , C_t , and I_t represent the trend, seasonal, cyclical, and irregular components, respectively.

The Moving Average method reduces short-term fluctuations by averaging the most recent n observations. The one-step-ahead forecast is calculated as: $F_{t+1} = \frac{Y_t + Y_{t-1} + \dots + Y_{t-n+1}}{n}$, where F_{t+1} is the forecast for the next period.

Single Exponential Smoothing assigns greater weight to the most recent observation while retaining information from previous forecasts. The forecasting equation is: $F_{t+1} = \alpha Y_t + (1-\alpha)F_t$, where α is the smoothing parameter ($0 < \alpha < 1$).

Double Exponential Smoothing, also known as Holt's method, extends exponential smoothing by incorporating both the level and the trend of the series. The updating equations are: $L_t = \alpha Y_t + (1-\alpha)(L_{t-1} + B_{t-1})$, $B_t = \beta(L_t - L_{t-1}) + (1-\beta)B_{t-1}$, with the m -step-ahead forecast given by: $F_{t+m} = L_t + mB_t$, where L_t denotes the level, B_t represents the trend, and β is the trend smoothing parameter.

Winter's Method (Holt–Winters exponential smoothing) further extends Holt's method by incorporating seasonal effects. The forecasting equation is: $F_{t+m} = (L_t + mB_t)S_{t-s+m}$, where S_{t-s+m} is the seasonal component and s denotes the seasonal length.

The Autoregressive Integrated Moving Average (ARIMA) model is proposed as a potential extension for future research to model autocorrelated and nonstationary trade series. The general ARIMA model is expressed as: $\phi(B)(1-B)^d Y_t = \theta(B)e_t$, where $\phi(B)$ and $\theta(B)$ are the autoregressive and moving-average polynomials, respectively, d is the order of differencing, B is the backshift operator, and e_t is the white-noise error term. ARIMA provides a flexible framework for forecasting time series that exhibit temporal dependence and nonstationary behavior.

D. Step 4: Error indicators and decision rule

Forecasting performance is evaluated using three widely adopted accuracy measures: Mean Absolute Deviation (MAD), Mean Squared Deviation (MSD), and Mean Absolute Percentage Error (MAPE). These metrics assess forecast errors from different perspectives and provide a comprehensive basis for comparing alternative forecasting models.

Mean Absolute Deviation (MAD) measures the average magnitude of forecast errors without considering their direction and is calculated as: $MAD = \frac{1}{n} \sum_{t=1}^n |Y_t - F_t|$, where Y_t is the observed value, F_t is the forecasted value, and n is the number of observations.

Mean Squared Deviation (MSD) evaluates the average squared forecast error, assigning greater weight to larger deviations. It is defined as: $MSD = \frac{1}{n} \sum_{t=1}^n (Y_t - F_t)^2$.

Mean Absolute Percentage Error (MAPE) measures forecasting accuracy in percentage terms and facilitates comparisons across trade series with different scales. It is computed as:

$$MAPE = \frac{100}{n} \sum_{t=1}^n \left| \frac{Y_t - F_t}{Y_t} \right|.$$

Among the candidate forecasting methods, the optimal model for each trade dataset is selected according to the minimum MAPE value, as MAPE provides an intuitive measure of relative forecasting accuracy. If two or more models produce the same MAPE, the model with the smaller

MSD is chosen because it imposes a greater penalty on large forecasting errors. The model selection criterion is expressed as: $M_i^* = \arg \min_m \text{MAPE}_{i,m}$; if a tie occurs, select $\arg \min_m \text{MSD}_{i,m}$, where M_i^* denotes the selected forecasting model for trade dataset i , and m indexes the competing forecasting methods. This model selection procedure ensures that the chosen forecasting model achieves the highest predictive accuracy while minimizing the impact of large forecast errors.

IV. RESEARCH RESULTS

Table I. ID277-ID291 data architecture

ID	Data-architecture content
ID277	Total circulation of exported and imported goods
ID278	Export value of goods by economic sector
ID279	Export value of goods by SITC classification
ID280	Export value structure by SITC classification
ID281	Export value by economic sector and commodity group
ID282	Export value by country group, country and territory
ID283	Major exported commodities
ID284	Import value of goods by economic sector
ID285	Import value by SITC classification
ID286	Import value structure by SITC classification
ID287	Import value by economic sector and commodity group
ID288	Import value by country group, country and territory
ID289	Major imported commodities
ID290	Export and import value ratio to GDP
ID291	Export and import of services

The data architecture covers merchandise trade, services trade, commodity structure, market structure and trade-to-GDP ratio. In the source file, the complete error-comparison table reports MAPE, MAD and MSD from ID277 to ID290; ID291 appears in the architecture but has no complete error row in the extracted table.

Table II. Optimal model selection by MAPE and MSD

ID	Research object	Optimal model	MAPE	Other reported indicators
ID277	Total import-export turnover	Trend Analysis	3	MAD=9,345; MSD=145,029,092
ID278	Export by economic sector	Double Exp. Smoothing	42,772,500	MAD=8,945,480; MSD=18,235,400,000,000
ID279	Export by SITC	Moving Average	15,475.2	MAD=676.048; MSD=1,246,370,000,000
ID280	Export structure by SITC	Decomposition	14,188.0	MAD=24.0; MSD=954.7
ID281	Export by sector and group	Decomposition	4	MAD=14,553; MSD=256,030,069
ID282	Export by market	Decomposition	4	MAD=14,553; MSD=256,030,069
ID283	Major exports	Winter's Method	97	MAD=1,943; MSD=9,087,557
ID284	Import by economic sector	Moving Average	428,258	MAD=820.824; MSD=16,222,600,000,000
ID285	Import by SITC	Decomposition	2,412	MAD=79,473; MSD=9,975,444,135
ID286	Import structure by SITC	Trend Analysis	2,060.51	MAD=23.63; MSD=899.66
ID287	Import by sector and	Decomposition	4,536	MAD=50,390;

	group			MSD=4,113,735,738
ID288	Import by market	Moving Average	1,662	MAD=28,815; MSD=2,165,144,959
ID289	Major imports	Single Exp. Smoothing	35	MAD=1,208; MSD=3,374,986
ID290	Trade/GDP ratio	Moving Average	8,413.58	MAD=25.13; MSD=1,251.72

Table III. Selected concrete results for strategic interpretation

ID	Optimal model	MAPE	MAD	MSD
ID277	Trend Analysis	3	9,345	145,029,092
ID281	Decomposition	4	14,553	256,030,069
ID282	Decomposition	4	14,553	256,030,069
ID283	Winter's Method	97	1,943	9,087,557
ID288	Moving Average	1,662	28,815	2,165,144,959
ID289	Single Exp. Smoothing	35	1,208	3,374,986
ID290	Moving Average	8,413.58	25.13	1,251.72

The results show strong model heterogeneity across trade dimensions. Total import-export turnover (ID277) is best described by Trend Analysis, with MAPE = 3 and MSD = 145,029,092, suggesting a strong long-run movement. Export values by market (ID282) and by sector/commodity group (ID281) are best represented by Decomposition, both with MAPE = 4 and MSD = 256,030,069, indicating that export structure contains distinguishable components. Major exports (ID283) fit Winter's Method, while major imports (ID289) fit Single Exponential Smoothing, implying different dynamics between outward and inward commodity flows. The trade-to-GDP ratio (ID290) is best fitted by Moving Average, indicating that smoothing can capture macro-ratio stability better than trend or seasonal methods in the provided table.

V. DISCUSSION

The findings confirm the source document's central interpretation: Vietnam's import-export activity increased over the long term but remains exposed to FDI dependence, imported input dependence, and concentration in key partner markets. Compared with recent studies [1]-[10], the present article is less complex than deep learning, knowledge graph or SFA-GMM models, but it is transparent and appropriate for annual ID277-ID291 data.

A. Strengths

The main strength is the multidimensional structure. Unlike studies that focus only on total exports or one FTA, this study covers total turnover, export/import sectors, SITC structure, partner markets, major commodity groups and trade-to-GDP ratio. The model-comparison rule is transparent because MAPE and MSD are reported for each candidate model. ID277 has a low MAPE of 3, and ID281-ID282 both have MAPE = 4, showing reliable trend and decomposition performance for core trade aggregates.

B. Weaknesses

The weakness is that several datasets show very high MAPE values, especially structure-ratio and commodity-detail series such as ID280, ID285, ID286 and ID290. This indicates that annual data are not sufficient to fully capture shocks, structural breaks, and market reallocation. In addition, the extracted source table does not contain complete forecasting-error results for ID291 services trade, although services trade is included in the data architecture.

C. Opportunities

The main opportunity is to combine annual trade data with external variables such as GDP, exchange rate, FDI inflows, logistics performance, partner-country demand, tariff reductions, non-tariff measures and PMI. Recent studies show that hybrid trade forecasting, knowledge graphs, FTAs and logistics variables can improve trade prediction and interpretation [1], [3], [4], [7], [9].

Vietnam can use these extensions to improve export-market diversification, domestic value added and services-trade upgrading.

D. Threats

The main threats are global demand shocks, geopolitical tension, supply-chain disruption, exchange-rate volatility and dependence on imported intermediate goods. The source file highlights dependence on FDI enterprises and major markets such as the United States for exports and China for imports. These threats can weaken forecast reliability and require regular model updating as new data become available.

Compared with previous studies, the current research provides a practical middle path. It is more quantitative than descriptive trade reports because it compares six models with error metrics. It is more accessible than advanced AI models because it can be replicated in Minitab. However, it should not be considered a replacement for multivariate or AI-based forecasting when longer and richer datasets are available.

VI. CONCLUSION

This study provides a structured IJETMS-style research article on Vietnam's import-export of goods and services during 2015-2024. Its main value is to integrate trade structure, time-series forecasting and model-selection logic into one practical framework for students, enterprises and policy planners. The concrete findings are as follows. ID277 total import-export turnover is best fitted by Trend Analysis with MAPE = 3, MAD = 9,345 and MSD = 145,029,092. ID282 export by market is best fitted by Decomposition with MAPE = 4 and MSD = 256,030,069. ID283 major exports are best fitted by Winter's Method with MAPE = 97 and MSD = 9,087,557. ID289 major imports are best fitted by Single Exponential Smoothing with MAPE = 35 and MSD = 3,374,986. ID290 trade-to-GDP ratio is best fitted by Moving Average with MAPE = 8,413.58, MAD = 25.13 and MSD = 1,251.72.

The limitations are that the analysis relies mainly on annual time-series data, some detailed commodity and structure series have high error, ID291 services trade lacks a complete error row in the extracted table, and the models do not yet include macroeconomic, logistics, FDI, tariff, exchange-rate or partner-demand variables.

Future research should extend the dataset to monthly or quarterly frequency, integrate macroeconomic and trade-policy variables, apply ARIMAX, VAR, Prophet, Random Forest, XGBoost, LSTM and Transformer models, and build a Power BI or dashboard system for real-time monitoring of Vietnam's export-import structure and market risks.

Acknowledgments: The authors thank Thuyloi University, Vietnam, for academic guidance and support.

Data availability: The article is based on the uploaded course document, the accompanying slide deck, and the ID277-ID291 trade-data architecture described in the source files.

REFERENCES

- [1] C.-H. Yang, C.-F. Lee, and P.-Y. Chang, "Export- and import-based economic models for predicting global trade using deep learning," *Expert Systems with Applications*, vol. 218, Art. no. 119590, 2023, doi: 10.1016/j.eswa.2023.119590.
- [2] R. A. H. Mohamed, "Enhancing forecast accuracy using combination methods for the hierarchical time series approach," *PLoS ONE*, vol. 18, no. 7, Art. no. e0287897, 2023, doi: 10.1371/journal.pone.0287897.
- [3] J. Chumnaul and K. Damkliang, "Integrating quantile regression with ARIMA and ANN for interpretable and accurate PM2.5 forecasting in Hat Yai, Thailand," *Scientific Reports*, vol. 15, Art. no. 43352, 2025, doi: 10.1038/s41598-025-27294-1.
- [4] D. Rincon-Yanez, C. Ounoughi, B. Sellami, T. Kalvet, M. Tiits, S. Senatore, and S. B. Yahia, "Accurate prediction of international trade flows: Leveraging knowledge graphs and their

- embeddings," *Journal of King Saud University - Computer and Information Sciences*, vol. 35, no. 10, Art. no. 101789, 2023, doi: 10.1016/j.jksuci.2023.101789.
- [5] D. T. D. My and P. T. Hoan, "Trade and investment effects of the EU-Vietnam Free Trade Agreement," *Hue University Journal of Science: Economics and Development*, vol. 132, no. 5B, 2023, doi: 10.26459/hueunijed.v132i5B.7167.
- [6] Q. H. Nguyen, "Evaluation of CPTPP and EVFTA on Vietnam trade flows," *Vietnam Social Sciences Review*, vol. 3, no. 215, pp. 30-44, 2023, doi: 10.56794/VSSR.3(215).30-44.
- [7] C. Estrades, M. Maliszewska, I. Osorio-Rodarte, and M. Seara e Pereira, "Estimating the economic impacts of the Regional Comprehensive Economic Partnership," *Asia and the Global Economy*, vol. 3, no. 2, Art. no. 100060, 2023, doi: 10.1016/j.aglobe.2023.100060.
- [8] J. Tian, Y. Zhu, T. B. N. Hoang, and B. K. T. Edjah, "Analysis of the competitiveness and complementarity of China-Vietnam bilateral agricultural commodity trade," *PLoS ONE*, vol. 19, no. 4, Art. no. e0302630, 2024, doi: 10.1371/journal.pone.0302630.
- [9] T. Q. Bui and L. H. Co, "Logistics, trade agreements and export efficiency: Evidence from Vietnam using SFA-GMM," *Journal of International Logistics and Trade*, vol. 23, no. 4, pp. 162-178, 2025, doi: 10.1108/JILT-06-2025-0053.
- [10] T. T. T. Nguyen and B.-Y. Choi, "The export efficiency of Vietnam after 40 years of Doi Moi reforms," *Asia & the Pacific Policy Studies*, vol. 13, no. 2, 2026, doi: 10.1002/app5.70081.