Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v10i3.16045 p-ISSN : 2541-044X

A Multi-Model Time Series Framework for Forecasting Vietnam's Tourism Revenue in the Post-COVID Recovery Era

Ho Nhat Hiep[1)], Nguyen Ngoc Xuan Quynh[2)], Nguyen Thi Van Anh[3)], Minh Ly Duc[4),*]
[1,2,3,4)]Faculty of Commerce, Van Lang University, Ho Chi Minh City, Vietnam
[1)]hiep.2373401210025@vanlanguni.vn, [2)]quynh.2373401210541@vanlanguni.vn,
[3)]anh.2373401210004@vanlanguni.vn, [4)]minh.ld@vlu.edu.vn

Submitted :April 2, 2026 | **Accepted** : May 1, 2026 | **Published** : July 5, 2026

Abstract: This study forecasts Vietnam's tourism revenue in the post-COVID-19 recovery period using a multi-model time series framework. The dataset includes three groups (ID 292–294) covering tourism business performance, economic sectors, and regional revenue. Six forecasting models are applied and evaluated using MAPE, MAD, and MSD. Results show that decomposition and Holt–Winters achieve the best accuracy (e.g., MAPE as low as 18%), while moving average performs well in specific cases (MAPE ≈ 28%). Forecasts indicate that tourism revenue may nearly double by 2030, driven mainly by domestic demand and the non-state sector, although international tourism recovers more slowly.

Keywords: tourism forecasting, time series analysis, multi-model framework, Holt–Winters method, Vietnam tourism.

INTRODUCTION

The global tourism industry has undergone unprecedented disruption due to the COVID-19 pandemic, leading to a dramatic decline in international travel demand and tourism-related revenues (Chen et al., 2023). In Vietnam, this impact was particularly severe, as international tourism, accounting for more than half of total tourism expenditure, nearly collapsed during 2020, resulting in a substantial revenue loss exceeding USD 12 billion (Kumar et al., 2025). Although domestic tourism temporarily played a compensatory role, its relatively low spending capacity proved insufficient to fully offset the decline in international arrivals (Tran et al., 2023).

In the post-pandemic recovery phase, the Vietnamese tourism sector is entering a critical transition period characterized by both opportunities and strategic challenges (Lin et al., 2023). The gradual recovery of global travel demand, especially from 2025 onward, opens prospects for attracting international tourists. However, the industry faces a fundamental dilemma: whether to continue relying on high-volume, low-spending domestic markets or to shift toward higher-value segments that prioritize service quality and premium experiences (Nguyen & Nguyen, 2024). This transformation requires robust analytical tools capable of capturing complex dynamics and supporting evidence-based decision-making in an increasingly volatile environment.

Tourism demand forecasting has long been recognized as a crucial component for strategic planning, investment decisions, and policy formulation (Li et al., 2024). Traditional forecasting approaches, including descriptive statistics and single-model time series techniques, have been widely applied due to their simplicity and interpretability. However, these methods often focus on isolated indicators such as tourist arrivals or aggregate revenue and fail to capture the multidimensional interactions between key components of the tourism system (Fatima & Rahimi, 2024). Furthermore, their predictive performance tends to decline in contexts characterized by structural changes, high volatility, and complex seasonal patterns.

To address these limitations, recent studies have emphasized the use of advanced time series models and hybrid forecasting approaches that integrate multiple techniques to improve accuracy and robustness. Traditional statistical methods such as trend analysis, decomposition, moving averages, and exponential smoothing remain widely used due to their effectiveness in capturing temporal patterns, including trend, seasonality, and irregular fluctuations (Kahraman & Akay, 2023; Iftikhar et al., 2023; Mulla et al., 2024). In parallel, hybrid and machine learning-based approaches have been introduced to enhance forecasting performance by integrating both linear and nonlinear dynamics (Grandón et al., 2024; Li et al., 2024; Nguyen-Da et al., 2023).

However, despite these advancements, several critical limitations persist in the literature. First, many existing studies focus on single-model or hybrid approaches without systematically comparing multiple forecasting

*name of corresponding author

techniques under varying data conditions, which limits the generalizability of their findings (Fatima & Rahimi, 2024; Montero-Manso, 2023). Second, most research assumes relatively stable demand patterns and does not adequately account for structural disruptions caused by external shocks such as the COVID-19 pandemic, which has significantly altered tourism dynamics and recovery trajectories (Lin et al., 2023; Kumar et al., 2025). Third, there is a lack of comprehensive frameworks that integrate multi-source tourism data—such as business performance, economic sectors, and regional characteristics—to capture the heterogeneous and uneven recovery patterns observed across emerging markets like Vietnam (Nguyen-Da et al., 2023; Tran et al., 2023).

Therefore, there remains a clear need for a structured multi-model forecasting framework that not only systematically evaluates different time series approaches but also captures post-COVID structural changes and multi-dimensional tourism dynamics. Addressing this gap is particularly important in the context of Vietnam, where the tourism sector is undergoing a complex recovery process characterized by strong domestic demand, slow international recovery, and increasing regional disparities.

The main contributions of this research are threefold. First, it proposes a comprehensive analytical framework that integrates multiple forecasting to the existing literature by advancing tourism demand forecasting in the context of post-COVID recovery. Rather than relying on the common assumption of stable demand patterns, the analysis explicitly incorporates structural disruptions and evolving recovery dynamics, thereby offering a more realistic representation of tourism behavior in volatile conditions. In addition, the study develops a structured multi-model comparative framework that enables a systematic evaluation of forecasting performance across different time series approaches, addressing the limitation of model-specific bias observed in prior research. Furthermore, by integrating multi-source tourism data including business performance indicators, economic sectors, and regional characteristics, the research provides a comprehensive and multi-dimensional perspective on tourism recovery, generating deeper empirical insights into the heterogeneous development patterns of Vietnam's tourism industry.

LITERATURE REVIEW

Traditional approaches to tourism forecasting

Forecasting has long played a fundamental role in tourism research, supporting policy planning, investment decision-making, and development strategy development (Ma, 2025). Early approaches relied primarily on qualitative assessments and simple statistical techniques (Kucharavy et al., 2023), where forecasting depended heavily on direct observation, historical comparisons, and expert experience. These methods proved suitable in small-scale contexts or when data was limited (Grandón et al., 2024).

However, this very simplicity is also its core weakness. Traditional methods are highly subjective, difficult to replicate, and lack the ability to model the complex relationships within increasingly expansive and dynamic tourism systems. In the context of increasingly abundant data and multifaceted influencing factors, continued reliance on empirical methods is no longer sufficient to guarantee the accuracy and reliability of forecasts.

Time series models in tourism forecasting

With the advancement of statistical methods, time series analysis has become one of the most widely used approaches in tourism demand forecasting. Time series models analyze historical data recorded over time to identify underlying patterns such as trends, seasonality, and irregular fluctuations (Kaur et al., 2023). Among these methods, decomposition techniques separate a time series into its core components, enabling a clearer understanding of long-term and seasonal effects (Iftikhar et al., 2023). Moving average models are commonly used to smooth short-term fluctuations (Mulla et al., 2024), while exponential smoothing techniques (Kahraman & Akay, 2023), including single, double (Holt's method), and triple (Holt–Winters method), are effective in capturing both trend and seasonal dynamics (Lallawmzuali & Muniyan, 2025). In particular, the Holt–Winters method has been widely recognized for its ability to handle both trend and seasonality simultaneously, making it suitable for tourism data characterized by strong seasonal patterns. These models offer improved accuracy compared to traditional methods and are widely applied in both academic research and practical forecasting applications. However, the effectiveness of time series models depends heavily on data quality and availability. Insufficient or noisy data can significantly reduce forecasting accuracy. Furthermore, individual models may perform differently depending on the characteristics of the dataset, suggesting that no single method is universally optimal.

Advanced and hybrid forecasting approaches

To overcome the limitations of single-model approaches, recent research has increasingly focused on hybrid and multi-model forecasting techniques. These approaches combine multiple models to leverage their individual strengths and improve overall predictive performance (Montero, 2023). Hybrid models often integrate statistical methods with machine learning or ensemble techniques. For example, combining ARIMA models with exponential smoothing or neural networks can enhance forecasting accuracy by capturing both linear and nonlinear

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v10i3.16045 p-ISSN : 2541-044X

patterns (Dhankhar et al., 2024). Ensemble approaches further improve performance by aggregating forecasts from multiple models, reducing model-specific biases, and increasing robustness. In addition, feature selection and dimensionality reduction techniques such as correlation analysis, principal component analysis (PCA), and regularization methods like Lasso have been applied to improve model performance by identifying the most relevant variables and eliminating redundant information (Fang et al., 2023). These methods are particularly useful in tourism forecasting, where multiple factors (e.g., economic conditions, tourist behavior, and regional characteristics) interact in complex ways.

Research gaps

Despite considerable advances in tourism forecasting, several key limitations remain. Much of the existing literature assumes relatively stable demand conditions and does not adequately capture the structural disruptions introduced by external shocks such as the COVID-19 pandemic. At the same time, forecasting approaches are often applied in isolation, with limited efforts to systematically evaluate their performance across different data conditions. In addition, current studies rarely integrate diverse data sources to reflect the heterogeneous recovery patterns observed across sectors and regions. As a result, the ability of existing models to provide reliable and context-specific insights remains constrained. These gaps underscore the need for a more integrated forecasting framework that can account for structural changes while capturing the multi-dimensional dynamics of tourism systems.

Research contribution

To address these gaps, this study proposes a comprehensive multi-model forecasting framework that integrates descriptive statistical analysis, feature selection techniques, and multiple time series models. By systematically comparing six forecasting methods and evaluating their performance across different datasets, this research aims to provide a more accurate and reliable approach to tourism revenue forecasting. In addition, the study contributes to the literature by applying this framework to Vietnam’s tourism sector, offering empirical insights into its recovery and long-term development trends. This approach not only enhances forecasting accuracy but also provides a practical foundation for strategic decision-making in tourism management.

METHOD

Based on the methodology flowchart presented in Figure 1, the study adopts a structured three-stage framework for tourism revenue forecasting. In the first stage, data from multiple sources, including tourism business performance, economic sectors, and regional indicators, is collected, cleaned, and integrated into a unified database to ensure consistency and reliability for subsequent analysis.

The second stage focuses on deep feature selection to enhance model performance and reduce noise. This process involves a combination of statistical and analytical techniques, including Principal Component Analysis (PCA) for dimensionality reduction, Granger causality testing to identify causal relationships, and autocorrelation as well as partial autocorrelation analysis to examine temporal dependencies within the data. These techniques collectively enable the identification of the most relevant features that significantly influence tourism revenue.

In the final stage, a multi-model forecasting structure is implemented. Several time series models including decomposition, moving average, single and double exponential smoothing, Holt–Winters, and ARIMA are applied to generate first-layer forecasts. These individual model outputs are then aggregated using an ensemble approach to produce the final forecast. Model performance is systematically evaluated using multiple error metrics, including MAE, MAPE, MSE, and RMSE, to ensure robustness and reliability of the forecasting results.

*name of corresponding author

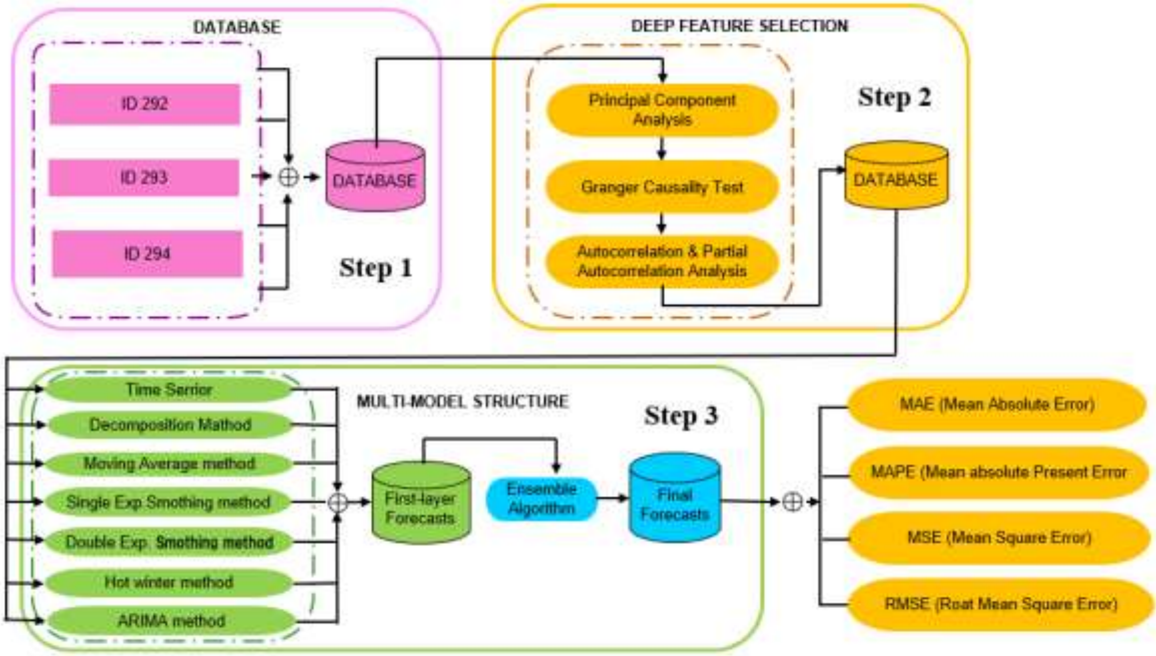


Fig 1. Research methodology

Step 1: This step describes the database collection and construction phase, where input data such as product codes, sales figures, time periods, regions, and other information are aggregated and stored. This is a foundational step that ensures the data is complete, consistent, and ready for analysis. This is crucial because the quality of the input data determines the accuracy and reliability of subsequent analysis and forecasting steps.

Step 2: Present the in-depth feature selection phase to identify and retain the variables (features) that have the greatest impact on revenue. This phase includes four main groups of analysis:

Correlation and multicollinearity analysis to eliminate duplicate variables (formula 1).

$$r_{xy} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \quad (1)$$

Where: x_i: the value of the independent variable (e.g., advertising cost, selling price...). y_i: the value of the dependent variable (e.g., revenue). r_{xy}: the correlation coefficient, showing the degree and direction of the relationship between the two variables ($-1 \leq r \leq 1$).

Component analysis and dimensionality reduction (PCA): extracting principal elements (formula 2).

$$\lambda = \frac{SS_{between}}{SS_{total}} \quad (2)$$

Where: $SS_{between}$: the sum of squared variances between the principal components. SS_{total}: the initial total variance of the data. λ: the proportion of variance explained by the principal component — the larger the λ, the more important the component.

Analysis of variance (ANOVA) and hypothesis testing: determining which variables significantly affect revenue (formula 3).

$$F = \frac{MS_{between}}{MS_{within}} \quad (3)$$

With: $MS_{between}$: the mean square between groups, representing the difference between data groups. MS_{within}: the mean square within a group, reflecting the variation within the group. F: the F statistic, used to test the hypothesis whether the groups are statistically significantly different.

Advanced feature selection: using models like Lasso Regression to remove weak variables (formula 4).

$$\min_{\beta}\left(\frac{1}{2n}\sum(y_i - X_i\beta)^2 + \lambda\sum|\beta_j|\right) \quad (4)$$

Where: y_i: the actual value of the dependent variable (revenue). X_i: vector of independent variables (input characteristics). β_j: regression coefficient of each variable (the level of influence of that variable). n: number of

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v10i3.16045 p-ISSN : 2541-044X

observed samples. λ: regularization parameter, controlling the level of penalty for β_j, and if λ is large → many variables are eliminated to zero.

These parameters help quantify the relationship, the level of influence, and the statistical significance of each variable in the model, thereby selecting the most important characteristics for revenue forecasting. Selecting and optimizing input variables plays a crucial role in improving the accuracy of the revenue forecasting model. Through statistical analyses such as correlation, PCA, ANOVA, and Lasso Regression, analysts can eliminate redundant variables, reduce noise, and retain only the factors that truly influence the results. The result is a more concise, stable model that accurately reflects the nature of the relationship between business factors and revenue.

Step 3: Present the phase of building a multi-model structure to forecast revenue more accurately. In this step, various forecasting methods such as ARIMA, trend, moving average, and exponential smoothing are combined to create a first-layer forecast. These results are then aggregated using the Ensemble algorithm to produce the final forecast, and accuracy is assessed using error metrics such as MAE, MAPE, MSE, and RMSE.

Trend Analysis: Modeling trends using regression (formula 5). Forecasting the factory's annual production output and analyzing population trends over time.

$$Y_t = a + bt \tag{5}$$

Where: a: constant (intersection point at t=0), b: trend coefficient (slope, rate of change over time).

Decomposition: Break down the time series into its components: trend, seasonality, and irregularity (formula 6). Forecast tourist numbers by month (with clear seasonality).

$$Y_t = T_t + S_t + e_t \tag{6}$$

With: T_t: trend component. S_t: seasonal component.e_t: irregular component.

Moving Average: Smooth data using a moving average (formula 7). Smooth sales data to eliminate short-term fluctuations. Track the 7-day average stock price.

$$MA_t = \frac{\boldsymbol{Y_t + Y_{t-1} + \cdots + Y_{t-n+1}}}{\boldsymbol{n}} \tag{7}$$

Where: n: number of moving average periods. $\boldsymbol{Y_t}$: observed value at time t.

Single Exponential Smoothing: Short-term forecast for a chain with no clear trend (formula 8). Forecast of daily demand for convenience stores.

$$\boldsymbol{F_{t+1} = \alpha Y_t + (1 - \alpha) F_t} \tag{8}$$

Where: α: smoothing coefficient (0 < α < 1). $\boldsymbol{Y_t}$ actual value at time t. $\boldsymbol{F_t}$: forecast value at time t.

Double Exponential Smoothing (Holt's Method): Applied to trending chains (formula 9). Forecasts quarterly revenue to show a steady upward trend over time.

$$\begin{cases} \boldsymbol{L_t = \alpha Y_t + (1 - \alpha)(L_{t-1} + T_{t-1})} \\ \boldsymbol{T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1}} \\ \boldsymbol{F_{t+m} = L_t + mT_t} \end{cases} \tag{9}$$

Where: α: level smoothing coefficient. β: trend smoothing coefficient. $\boldsymbol{L_t}$: level value at time t. $\boldsymbol{T_t}$: trend slope. m: number of forecast periods.

Winters' Method (Holt-Winters): Forecasts product lines that incorporate both trends and seasonality (Formula 10). Forecasts monthly sales with seasonal factors (such as beer, air conditioners).

$$\begin{cases} \boldsymbol{L_t = \alpha \frac{Y_t}{S_{t-s}} + (1 - \alpha)(L_{t-1} + T_{t-1})} \\ \boldsymbol{T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1}} \\ \boldsymbol{S_t = \gamma \frac{Y_t}{L_t} + (1 - \gamma)S_{t-s}} \\ \boldsymbol{F_{t+m} = (L_t + mT_t)S_{t-s+m}} \end{cases} \tag{10}$$

Where: α: smooths the level. β: smooths the trend. γ: smooths the season. s: length of the season cycle. $\boldsymbol{L_t}$, $\boldsymbol{T_t}$, $\boldsymbol{S_t}$: are the level, trend, and season, respectively.

RESULT

Present the entire process of data processing, analysis, and experimental results of the study. This includes a description of the dataset, data collection methods, and data architecture, along with visual and statistical analysis steps to clarify data characteristics. Following this are the analysis results at each model step (from data collection and feature selection to forecasting), helping to evaluate the accuracy and effectiveness of the proposed method. Present the origin, structure, and characteristics of the data used in the study. This includes a flowchart of the data collection method, a description of the architecture and key data variables, along with visual and descriptive

*name of corresponding author

DOI : https://doi.org/10.33395/sinkron.v10i3.16045
e-ISSN : 2541-2019
p-ISSN : 2541-044X

statistical analysis to understand trends, distributions, and relationships between variables. This is a crucial foundation for ensuring the reliability and suitability of the data for subsequent analysis steps.

Data information

The data is structured into three main groups: ID 292 is tourism business performance data categorized by group (such as group (Revenue of accommodation establishments), group (Revenue of travel agencies), group (Number of international visitors including visitors served by accommodation establishments + visitors served by travel agencies)). ID 293 is tourism revenue data at current prices, categorized by economic type into two groups: State-owned and Non-State-owned. ID 294 is tourism revenue data at current prices categorized by locality into regional groups (such as group (Red River Delta), group (Northern Midlands and Mountains), group (North Central and Central Coastal Region), group (Central Highlands), group (Southeast Region), group (Mekong Delta)). Information for each data set is shown in Table 1.

Table 1
Data Information

No.	Items
ID 292	Tourism Business Results: -ID 292-1: Revenue of Accommodation Establishments (Billion VND) -ID 292-2: Revenue of Travel Agencies (Billion VND) -ID 292-3: Number of Domestic Tourists (Guests served by Accommodation Establishments + Tourists served by Travel Agencies) -ID 292-4: Number of International Tourists (Guests served by Accommodation Establishments + Tourists served by Travel Agencies)
ID 293	Tourism revenue at current prices, categorized by economic sector: - ID 293-1: State-owned economy - ID 293-2: Non-state-owned economy
ID 294	Tourism revenue at current prices, broken down by region: - ID 294-1: Red River Delta - ID 294-2: Northern Midlands and Mountains - ID 294-3: North Central and Coastal Regions - ID 294-4: Central Highlands - ID 294-5: Southeast Region - ID 294-6: Mekong Delta

Data collection flowchart

The flowchart describes a tourism information system consisting of two main stages. The interface layer collects request from users through channels such as mobile applications, Google Maps, and websites. This data goes to the processing layer, where the system performs filtering and grouping of data by type/region. Then, it processes the data from the tourism database and returns the approved results to the user (Figure 2).

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v10i3.16045 p-ISSN : 2541-044X

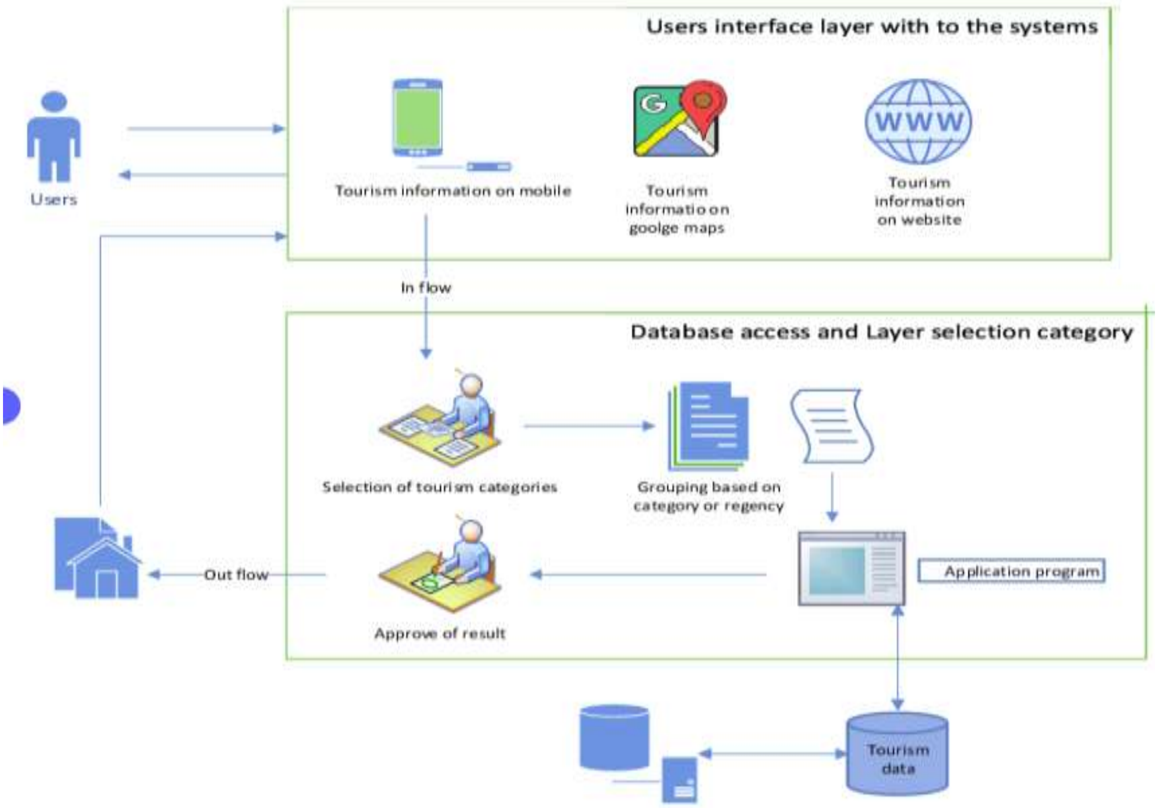


Fig. 2 Data collection method flowchart

Figure 2 presents the data collection and processing workflow of the tourism information system. Users access the system through multiple platforms such as mobile applications, Google Maps, and websites, generating input data. This data is then processed through a filtering and categorization stage, where it is grouped based on predefined criteria. The system interacts with the central database to retrieve and refine relevant information, which is subsequently validated and delivered back to users. This workflow ensures efficient data organization and supports reliable analysis for forecasting purposes.

Statistical data analysis

Provide the Max, Min, After, and Sigma values for each ID to assess the accuracy of the data, as shown in Table 2.

Table 2
Statistical data analysis

Data name	Max	Min	After	Sigma	Evaluate
ID 292-1	94101	23690	58198	28349	Pass
ID 292-2	78991	8999	42405	29979	Pass
ID 292-3	215242	63770	142703	66715	Pass
ID 292-4	70280	3396	29514	28625	Pass
ID 293-1	2716	279	1524	867	Pass
ID 293-2	68938	8262	37340	26447	Pass
ID 294-1	29636	3938	16088	11084	Pass
ID 294-2	1070	133	574	414	Pass
ID 294-3	12907	1061	6111	5304	Pass
ID 294-4	273.7	51	146.7	92.5	Pass
ID 294-5	32412	3502	17948	12105	Pass
ID 294-6	2692	314	1537	1089	Pass

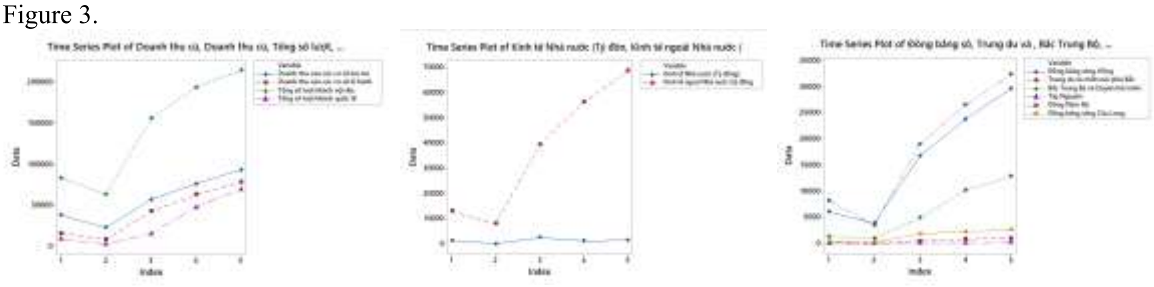
*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v10i3.16045 p-ISSN : 2541-044X

Table 2 presents descriptive statistical indicators for each dataset, including maximum (Max), minimum (Min), mean (After), and standard deviation (Sigma), to assess data distribution and consistency. The results show that all variables exhibit reasonable value ranges and acceptable variability, indicating no abnormal fluctuations or extreme outliers. The “Pass” evaluation across all IDs confirms that the data is reliable and suitable for subsequent analysis and forecasting. In particular, the relatively high Sigma values in some variables reflect natural variability in tourism-related indicators, which is expected in real-world data and does not compromise overall data quality.

Data visualization

Here, we will describe the three IDs using a Time Series model to provide the most intuitive view, as shown in Figure 3.



ID 292 ID 293 ID 294

Fig. 3 Data visualization of ID292_ ID293_ID294

Figure 3 presents the time series visualization of three main data groups (ID 292, ID 293, and ID 294), highlighting the trends and dynamics of Vietnam’s tourism indicators over time. For ID 292, all variables including accommodation revenue, travel service revenue, and tourist volumes show a clear upward trend after an initial decline, reflecting a strong recovery phase in tourism activities. Notably, domestic tourist numbers increase significantly, indicating their dominant role in driving overall tourism demand. In ID 293, the non state economic sector exhibits a rapid and consistent growth trend, while the state sector remains relatively stable at a lower level. This suggests that the non-state sector plays a key role in the expansion of tourism revenue. For ID 294, regional disparities are clearly observed. The Southeast region and the Red River Delta demonstrate strong growth patterns, whereas other regions show more moderate increases. This indicates an uneven spatial distribution of tourism development across regions.

Evaluating the performance of the models

MAPE (Mean Absolute Percentage Error) (Eq. 11) measures the mean absolute error as a percentage, helping to assess the relative accuracy of the model. While easy to interpret, this metric is less suitable when the actual value is too small or zero.

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right| \tag{11}$$

With: n: number of observations (data points), y_t: actual value at time t, $\hat{y}_t$: predicted (forecasted) value at time t.

MSE (Mean Squared Error) (Eq. 12) reflects the degree of deviation through the square of the error, strongly highlighting large deviations in forecasts. This is an effective measure for model optimization, but its square unit makes it difficult to visualize.

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2 \tag{12}$$

With: n: number of observations,

MSD (Mean Squared Deviation) (Eq. 13) measures the sign error, indicating whether the forecasting model is biased up or down. However, because the error can be canceled out, MSD is not used to assess overall accuracy but mainly to determine bias.

$$MSD = \frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2 \tag{13}$$

Experimental results

The decomposition and Winter’s Method tend to provide better forecasting performance across most IDs, as they often yield lower MAPE and MSD values. For example, in ID 292-1, Decomposition has the lowest MAPE

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v10i3.16045 p-ISSN : 2541-044X

(22) and MSD (100,342,297), while Winter's Method performs well with MSD of 163,718,766. In ID 292-3, Decomposition again gives the best accuracy with an MAPE of 18 compared to 51 in Double Smoothing. However, in some cases, like ID 294-3, the Moving Average performs best with the lowest MAPE (28) and MSD (4,424,067). Notably, Double Smoothing frequently shows poor performance, such as in ID 292-1 with a very high MSD (9,551,063,736). Therefore, Decomposition is generally the most consistent method, while Moving Average or Winter's Method may be preferable depending on specific datasets, as shown in Table 4

Table 4
Experimental results

ID.	Eva.	**Trend Analysis**	**Decomposition**	**Moving Average**	**Single Smoothing**	**Double Smoothing**	**Winter's Method**
ID 292-1	MAPE	24	**22**	41	33	66	26
	MAD	7445	9534	29856	16172	24707	11539
	MSD	103520651	100342297	911806902	384742039	9551063736	163718766
ID 292-2	MAPE	48	44	54	**39**	116	58
	MAD	6161	10488	31255	13119	21447	14546
	MSD	70767014	122646970	122646970	313853878	671804780	237841905
ID 292-3	MAPE	21	**18**	38	23	51	24
	MAD	18289	20874	68901	31272	56667	28036
	MSD	467022772	600674867	911806902	2059757747	4524261523	1035115137
ID 292-4	MAPE	104	103	**64**	80	277	137
	MAD	9286	9356	28730	10800	19853	18833
	MSD	104122359	104115368	1015217406	162705925	474374718	431026324
ID 293-1	MAPE	97	**52**	161	104	105	71
	MAD	593	550	1307	594	675	696
	MSD	559608	462314	2308263	733119	708523	683299
ID 293-2	MAPE	46	**41**	47	120	52	56
	MAD	5377	8911	16418	24081	6199	12724
	MSD	49931144	93560326	362737637	628334299	63379316	190049143
ID 294-1	MAPE	41	37	45	**34**	46	49
	MAD	2182	3570	6972	4718	2537	4905
	MSD	8667100	14594360	63577194	43253230	11005451	27611290
ID 29	MAPE	44.7	43.5	47.4	**36.6**	49.5	59.4
	MAD	88.9	135.2	254	142	96.7	211

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v10i3.16045 p-ISSN : 2541-044X

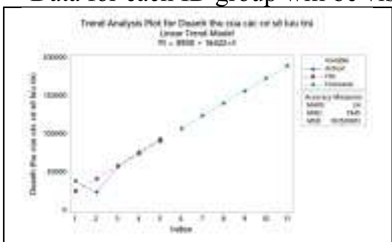
4-2	MSD	12682.6	19970.3	83645.6	628334299	16134.1	47837.0
ID 294-3	MAPE	66	64	45	**28**	75	147
	MAD	1183	1425	3038	1505	1342	7631
	MSD	1693912	2284243	12562658	4424067	2162238	102399379
ID 294-4	MAPE	28.466	**20.560**	34.53	26.99	31.741	21.990
	MAD	21.536	20.065	58.6	44.03	24.221	21.372
	MSD	580.538	440.293	4925.86	3788.58	736.835	648.736
ID 294-5	MAPE	55	**48**	66	52	63	21.990
	MAD	2917	4847	8418	6264	3399	21.372
	MSD	15556417	26653325	88766723	67044197	19752154	648.736
ID 294-6	MAPE	52	40	40	**32**	57	61
	MAD	286	393	630	457	322	572
	MSD	108782	241871	707977	544087	137915	506123

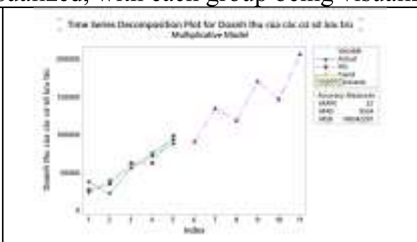
Table 4 presents the comparative performance of six forecasting models across different datasets using multiple error metrics, including MAPE, MAD, and MSD. Overall, the results indicate that no single model consistently outperforms others across all datasets, highlighting the importance of model selection based on data characteristics.

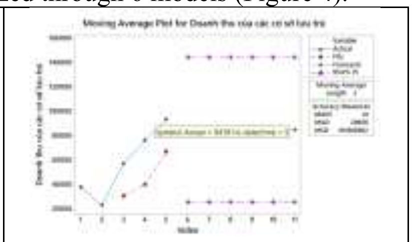
Among the evaluated methods, the decomposition approach generally achieves the lowest MAPE values in several cases ID 292-1, ID 292-3, indicating strong performance in capturing both trend and seasonal patterns. Similarly, Holt–Winters (Winter's Method) also demonstrates relatively stable and competitive results across multiple datasets, particularly in cases with clear seasonality.

In contrast, the double exponential smoothing method frequently produces significantly higher error values, especially in terms of MSD, suggesting its limitation in handling complex seasonal structures. The moving average method performs well in specific cases ID 294-3, indicating its suitability for smoothing short-term fluctuations in certain datasets. Furthermore, the results reveal variations in model performance across different groups. For example, datasets in ID 292 (tourism demand indicators) tend to favor decomposition methods, while some regional datasets (ID 294) show better performance with simpler smoothing techniques.

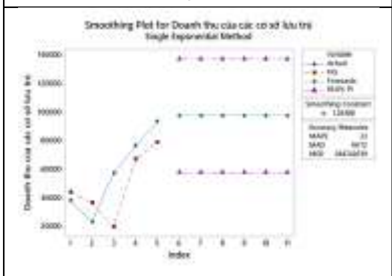
Data for each ID group will be visualized, with each group being visualized through 6 models (Figure 4).

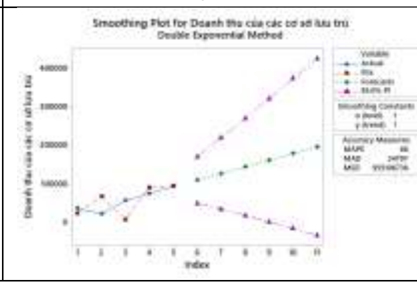


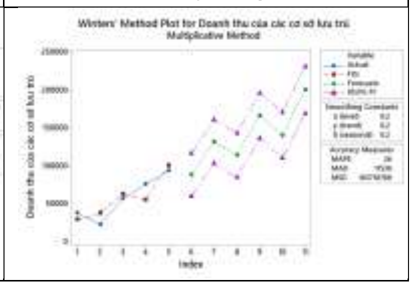




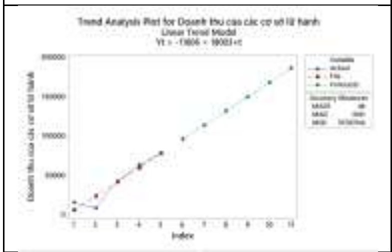
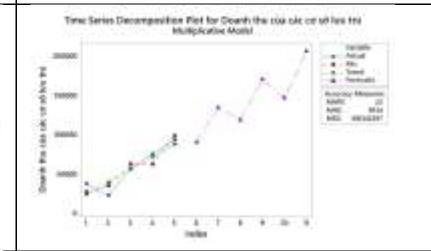
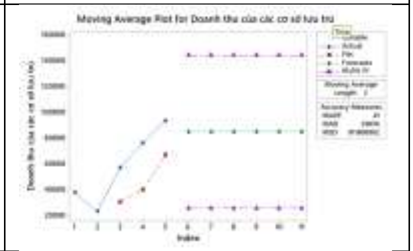
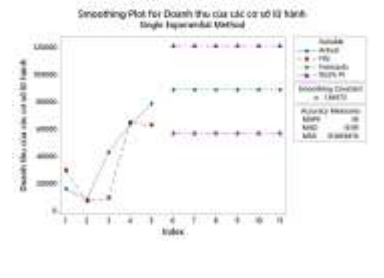
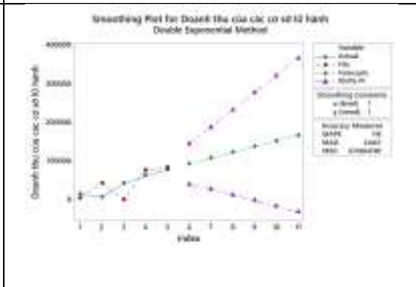
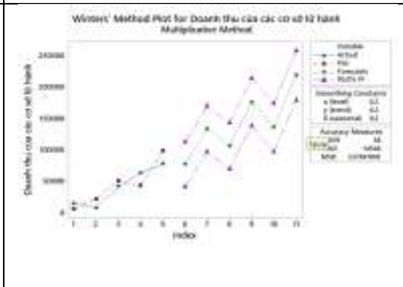
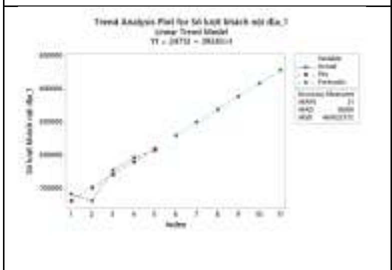
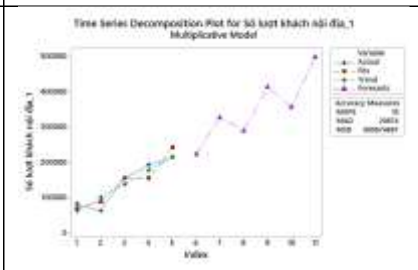
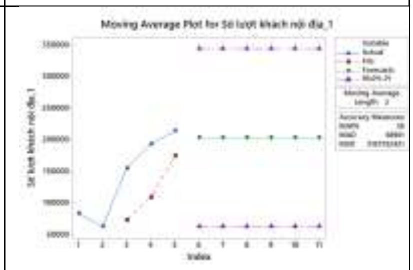
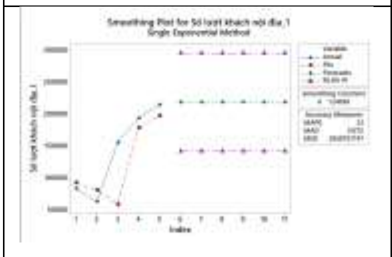
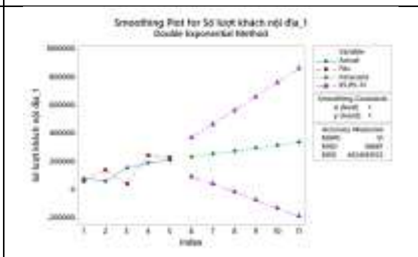
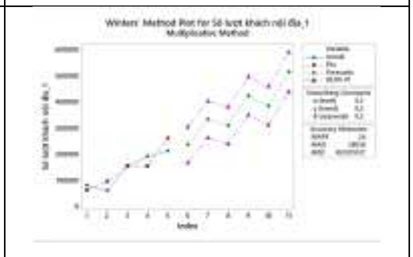
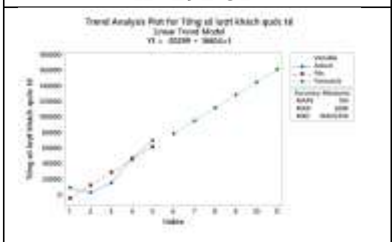
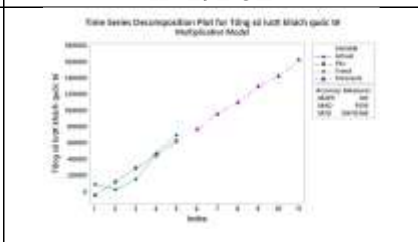
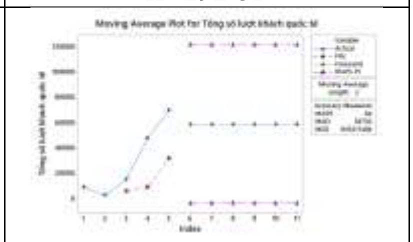
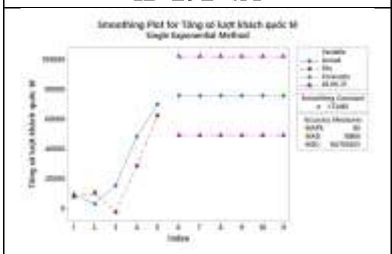
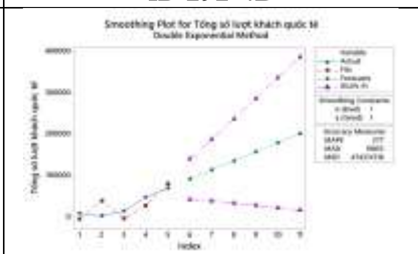
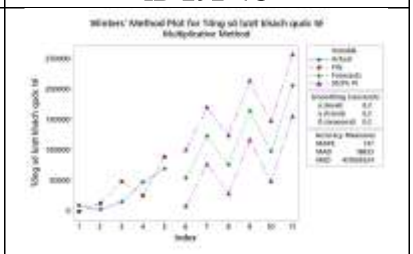
ID 292-1A ID 292-1B ID 292-1C





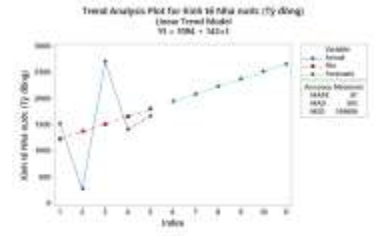
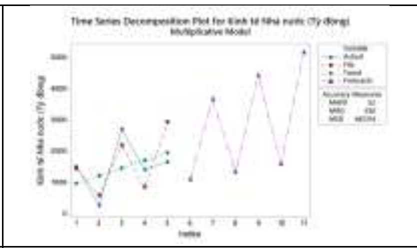
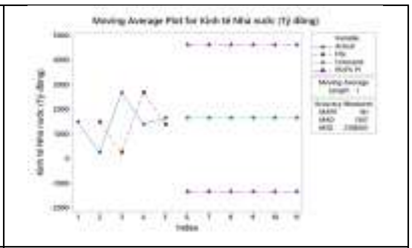
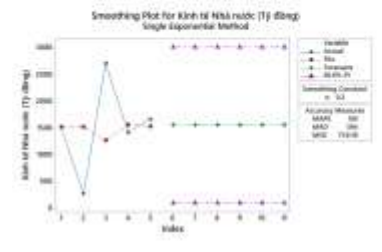
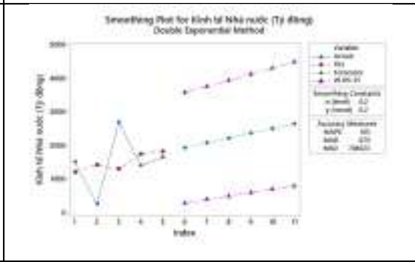
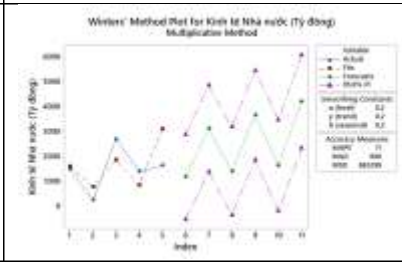
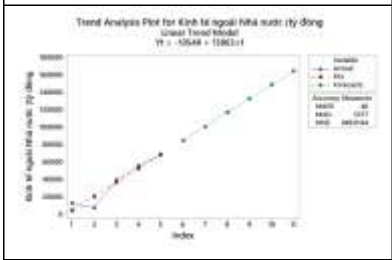
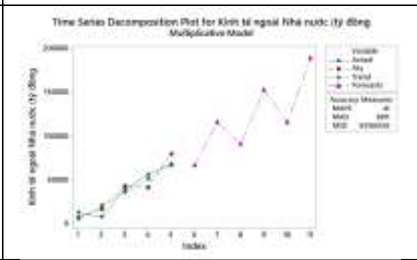
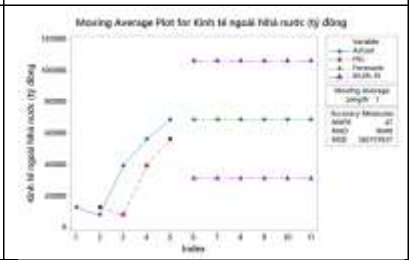
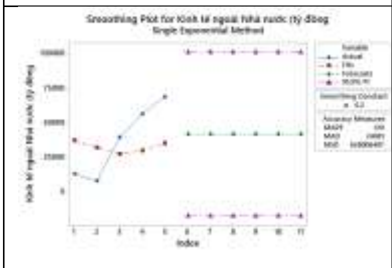
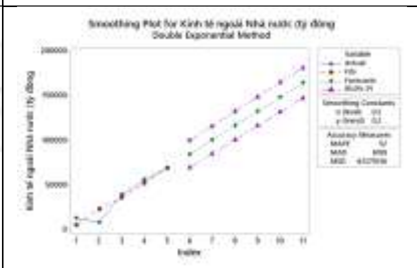
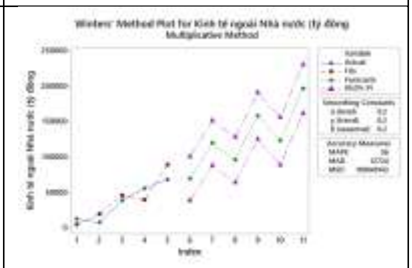
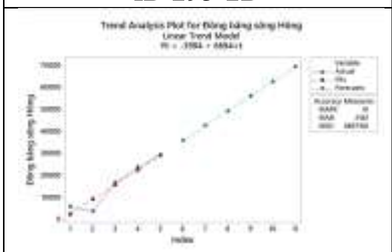
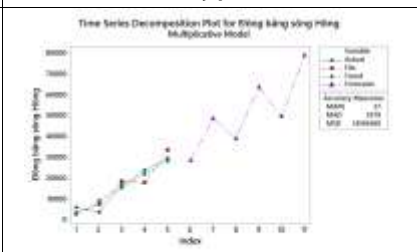
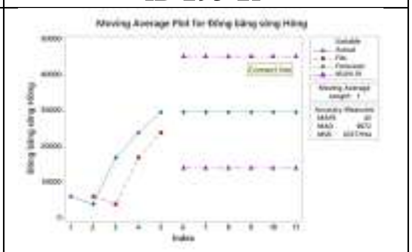
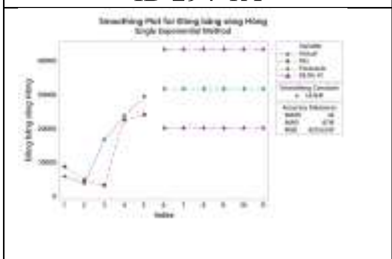
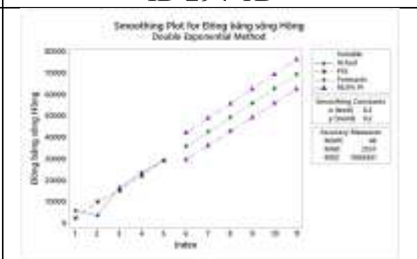
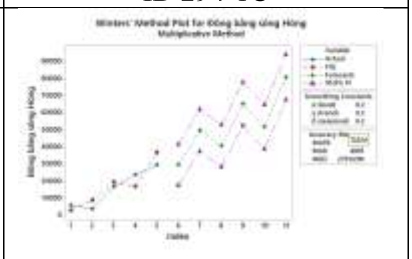


*name of corresponding author

ID 292-1D	ID 292-1E	ID 292-1F
		
ID 292-2A	ID 292-2B	ID 292-2C
		
ID 292-2D	ID 292-2E	ID 292-2F
		
ID 292-3A	ID 292-3B	ID 292-3C
		
ID 292-3D	ID 292-3E	ID 292-3F
		
ID 292-4A	ID 292-4B	ID 292-4C
		
ID 292-4D	ID 292-4E	ID 292-4F

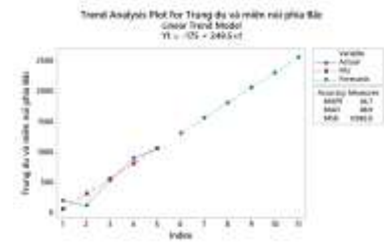
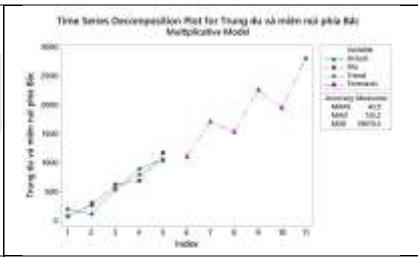
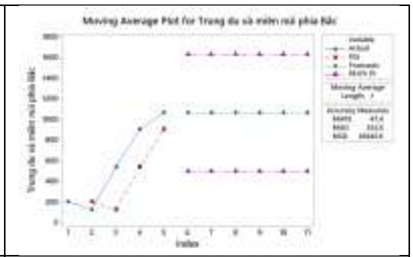
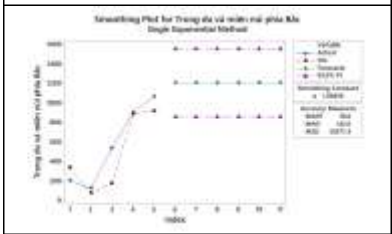
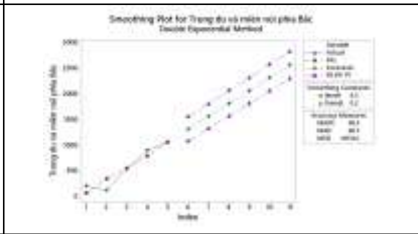
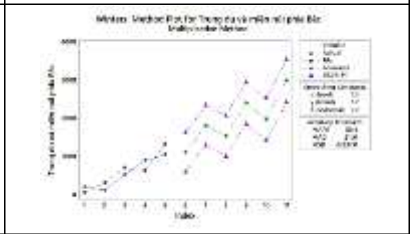
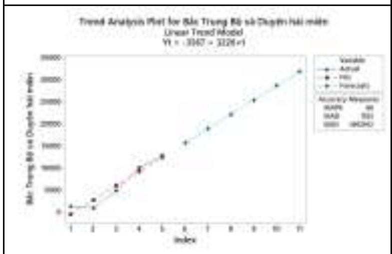
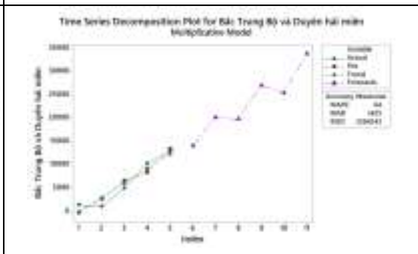
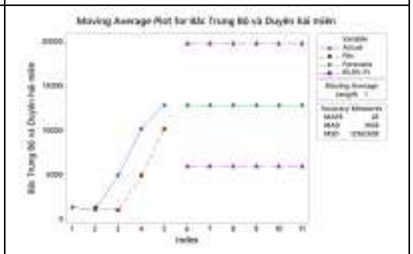
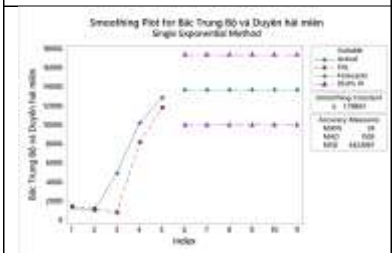
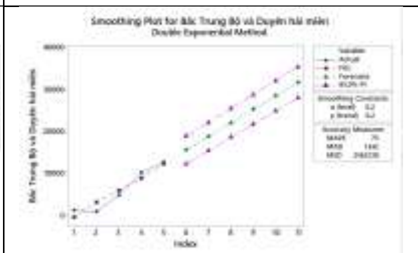
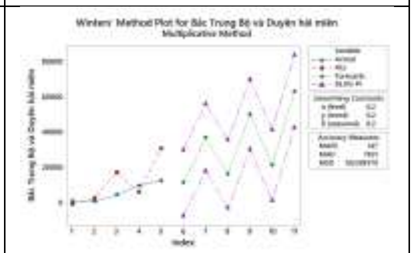
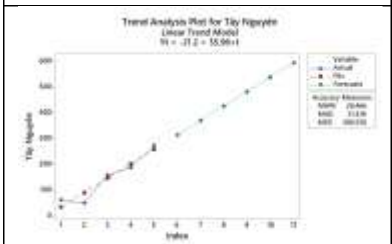
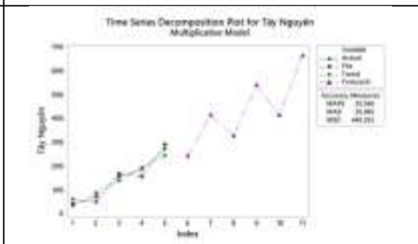
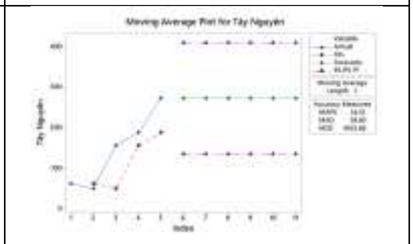
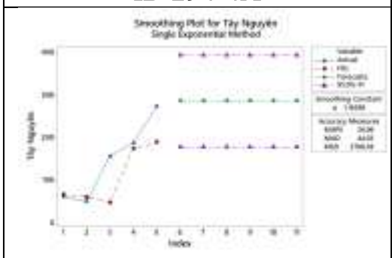
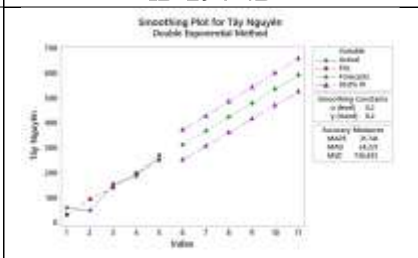
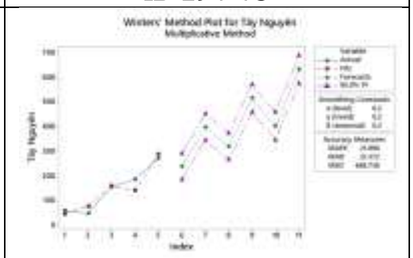
*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v10i3.16045 p-ISSN : 2541-044X

		
ID 293-1A	ID 293-1B	ID 293-1C
		
ID 293-1D	ID 293-1E	ID 293-1F
		
ID 293-2A	ID 293-2B	ID 293-2C
		
ID 293-2D	ID 293-2E	ID 293-2F
		
ID 294-1A	ID 294-1B	ID 294-1C
		
ID 294-1D	ID 294-1E	ID 294-1F

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v10i3.16045 p-ISSN : 2541-044X

		
ID 294-2A	ID 294-2B	ID 294-2C
		
ID 294-2D	ID 294-2E	ID 294-2F
		
ID 294-3A	ID 294-3B	ID 294-3C
		
ID 294-3D	ID 294-3E	ID 294-3F
		
ID 294-4A	ID 294-4B	ID 294-4C
		
ID 294-4D	ID 294-4E	ID 294-4F

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v10i3.16045 p-ISSN : 2541-044X

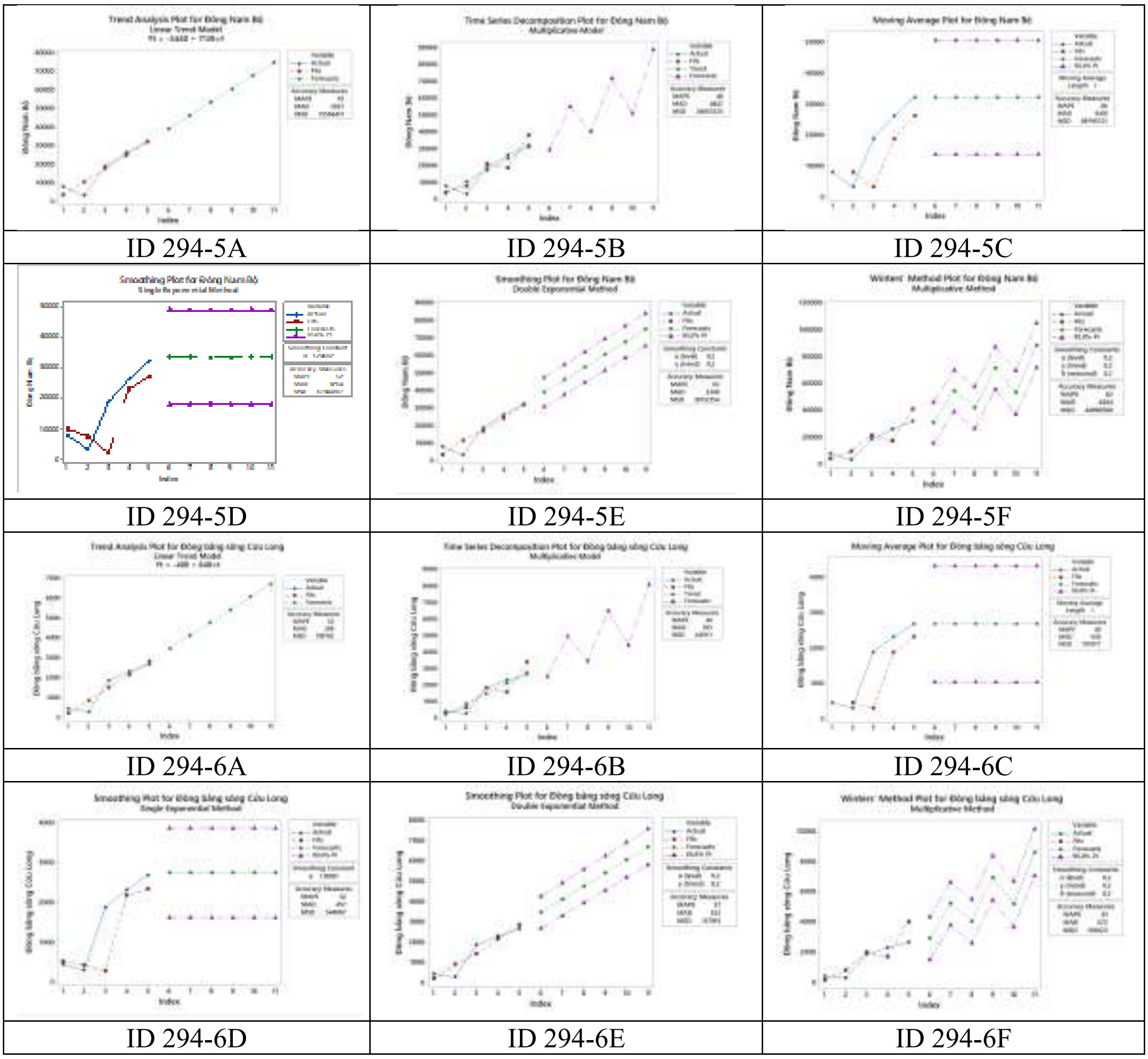


Fig. 4 Visual for 6 model

Figure 4 provides a comprehensive visualization of forecasting results across six-time series models applied to multiple datasets (ID 292, ID 293, and ID 294), covering tourism demand, economic structure, and regional indicators. The figure enables a direct comparison between actual data, fitted values, and forecasted outputs for each method, thereby revealing the strengths and limitations of different modeling approaches under varying data conditions. Overall, models that incorporate both trend and seasonality, particularly the decomposition method and Holt–Winters approach, consistently produce forecasts that closely follow the underlying data patterns. These models demonstrate strong capability in capturing long-term growth trends and cyclical fluctuations, especially in datasets with clear seasonal behavior such as tourism demand and regional activity.

In contrast, simpler approaches such as moving average and single exponential smoothing tend to generate overly smoothed forecasts, often failing to respond adequately to rapid changes or structural shifts in the data. Double exponential smoothing shows unstable behavior in several cases, particularly when handling datasets with strong seasonal variations, leading to larger deviations from observed values. Furthermore, the results highlight significant variation across datasets. For ID 292 (tourism demand indicators), most models capture the upward recovery trend, but only seasonal models accurately reflect fluctuations. For ID 293 (economic sectors), the non-state sector exhibits strong growth patterns that are better captured by trend-based and decomposition methods. For ID 294 (regional data), clear spatial disparities are observed, with some regions showing stable growth while others display higher volatility, affecting model performance. Figure 4 confirms that forecasting performance is highly dependent on both model structure and data characteristics. No single model consistently dominates across all cases, reinforcing the necessity of a multi-model framework to improve forecasting robustness and adaptability, particularly in the context of post-COVID tourism recovery where structural changes and uncertainties remain significant.

*name of corresponding author

DISCUSSIONS

The empirical findings highlight several important patterns in Vietnam's tourism recovery and long-term growth trajectory. As illustrated in Figure 3 and Figure 4, the time series analysis confirms a clear upward trend across most indicators, particularly after the post-pandemic recovery phase. The results suggest that tourism demand in Vietnam is transitioning from short-term recovery to a more stable and sustainable growth stage. This is further supported by the forecast results for the 2025–2030 period, which indicate substantial expansion in both revenue and visitor volume. From a methodological perspective, the comparative analysis of six forecasting models demonstrates that no single model consistently outperforms others across all datasets. However, decomposition and Holt–Winters methods generally provide the most reliable performance, as evidenced by lower MAPE and MSD values in Table 4. These findings are consistent with previous studies emphasizing the effectiveness of decomposition techniques in capturing both trend and seasonality (Iftikhar et al., 2023), as well as the robustness of Holt–Winters models in handling seasonal tourism data (Lallawmzuali & Muniyan, 2025). In contrast, the relatively poor performance of double exponential smoothing in several cases reflects its limitation in modeling complex seasonal structures.

At the structural level, the analysis of dataset groups (ID 292, 293, and 294) reveals a strong imbalance in the drivers of tourism growth. As shown in Table 5, domestic tourism (ID 292-3) continues to dominate in terms of volume, reaching projected levels of over 500 million visitors by 2030, while international tourism (ID 292-4) exhibits slow recovery and stagnation. This imbalance raises concerns regarding the sustainability of revenue growth, as domestic tourists typically generate lower spending per capita. Similar observations have been reported in recent tourism studies, where post-pandemic recovery is often driven by domestic demand rather than international markets (Lin et al., 2023). Furthermore, the non-state economic sector (ID 293-2) emerges as the primary contributor to tourism revenue growth, significantly outperforming the state sector. This finding highlights the increasing role of private enterprises in driving innovation and service quality within the tourism industry. From a regional perspective, the Southeast region (ID 294-5) maintains its leading position, while the Central Highlands (ID 294-4) demonstrates exceptional growth potential, suggesting a shift toward more geographically diversified tourism development. This trend aligns with the concept of decentralized tourism expansion observed in emerging markets (Nguyen-Da et al., 2023).

Table 5
Analysis of research results

ID	2020–2024	2025–2030	Remarks
ID 292-1: Accommodation revenue	Recovered from the lowest level of 23,690.4 billion VND in 2021 to 94,101.4 billion VND in 2024	Continues stable growth, reaching 207,262 billion VND by 2030	Recovery and sustainable growth of services
ID 292-2: Travel agency revenue	Increased from 16,492 to 78,990.5 billion VND	Continues to grow, reaching 89,142.2 billion VND	Shows the important role of travel services in high-value tourism
ID 292-3: Domestic tourists	Reached 215,241.5 thousand visitors in 2024	Continues to grow, reaching 500,511 thousand visitors by 2030	Domestic demand is the main driver in terms of volume
ID 292-4: International tourists	Slow recovery, reaching 70,280.3 thousand visitors in 2024	Remains stable at around 59,313 thousand visitors throughout the period	More investment is needed to increase international demand
ID 293-1: State sector	Reached 1,673 billion VND in 2024	Slow growth trend, reaching 5,168.49 billion VND	Accounts for a small proportion

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16045
e-ISSN : 2541-2019
p-ISSN : 2541-044X

ID 293-2: Non-state sector	Strong growth from 8,261.5 (2021) to 68,937.9 billion VND (2024)	Reaches 189,449 billion VND by 2030	The main driver of the industry should be prioritized for development
ID 294-5: Southeast region	Leading in revenue: 32,412.3 billion VND (2024)	Continues strong growth, reaching 633,638 billion VND	The high-revenue region needs focus to maintain its position
ID 294-4: Central Highlands	Reached 273.7 billion VND in 2024	Breakthrough potential, reaching 665,679 billion VND	Expected rapid growth requires strategic positioning
ID 294-2: Northern midlands and mountainous region	Reached 1,070.4 billion VND in 2024	Steady growth, reaching 42,082.7 billion VND	Needs cultural preservation and high-quality tourism development
Other regions (294-1, 294-3, 294-6)	Reached 29,635.8; 12,906.7; 2,691.6 billion VND respectively in 2024	Reached 31,911.3; 13,737.8; 2,754.1 billion VND	Maintain stable growth; focus on quality to increase visitor spending and stay duration

Compared with previous studies, this research provides a more comprehensive and multidimensional perspective by integrating 12 variables and multiple forecasting models. While prior studies often rely on single-model approaches or a limited set of indicators (Li et al., 2024), the findings demonstrate that a multi-model framework enhances forecasting reliability and better captures the complexity of tourism dynamics. The results reveal a clear post-COVID recovery trajectory of Vietnam's tourism sector, accompanied by notable structural shifts. Tourism revenue is projected to reach nearly 300 trillion VND by 2030, reflecting not only overall growth but also the increasing contribution of the non-state sector, which emerges as the primary driver of expansion. In contrast, the state sector maintains a relatively modest role, indicating a shift toward a more market-driven tourism economy.

At the demand level, domestic tourism continues to dominate in terms of volume, while international tourism shows slower and more unstable recovery, highlighting the need for targeted strategies to improve visitor quality and spending capacity. From a spatial perspective, regional disparities remain evident, with the Southeast region maintaining its leading position, while emerging regions such as the Central Highlands exhibit significant growth potential. From a policy perspective, these findings suggest the need to enhance high-value international tourism, strengthen support for the non-state sector as a key growth engine, and promote balanced regional development to avoid over-concentration. Overall, Vietnam's tourism industry is not only recovering but also undergoing structural transformation, and the application of a multi-model framework provides valuable insights for both short-term forecasting and long-term strategic planning.

CONCLUSION

In this study develops a structured multi-model time series framework to forecast Vietnam's tourism revenue in the post-COVID-19 recovery period. By integrating 12 variables and multiple forecasting techniques, the proposed approach improves analytical depth and forecasting reliability compared to conventional single-model studies. The empirical results show that decomposition and Holt–Winters methods consistently achieve the highest forecasting accuracy, with MAPE values generally lower than other models across most datasets, while moving average models perform better only in specific cases with smoother data patterns.

From a quantitative perspective, Vietnam's tourism sector demonstrates a strong recovery trajectory, with total tourism revenue projected to approach 300 trillion VND by 2030. Domestic tourism demand is expected to remain the dominant contributor in terms of volume, while the non-state sector emerges as the primary driver of revenue growth. In contrast, international tourism shows a slower and less stable recovery, indicating structural imbalances in the post-COVID tourism system.

Compared with previous studies that focus on single-model forecasting or limited indicators, this research provides a more comprehensive and data-driven framework that captures both structural changes and multi-dimensional dynamics of tourism development. The findings highlight the importance of adopting multi-model approaches to improve forecasting robustness in volatile environments.

From a practical perspective, the results suggest that policymakers should prioritize strategies to enhance the quality and spending capacity of international tourists, strengthen support for the non-state sector, and promote more balanced regional development. From a theoretical perspective, this study contributes to the literature by demonstrating the effectiveness of integrating multiple time series models and multi-source data in capturing complex recovery patterns.

However, several limitations remain. The current framework relies primarily on historical time series data and does not incorporate external factors such as macroeconomic conditions, policy interventions, or global shocks. Future research should extend this approach by integrating econometric and machine learning models with exogenous variables such as GDP, exchange rates, and visa policies. This would enable more adaptive, scenario-based forecasting and improve the robustness of predictions, particularly for sensitive segments such as international tourism demand.

ACKNOWLEDGMENT

The authors are extremely grateful to Van Lang University, Vietnam, for supporting this research.

REFERENCES

Chen, S., Ning, Y., Wang, L., & Wang, S. (2023). Research on the factors influencing tourism revenue of Shandong Province in China based on uncertain regression analysis. *Mathematics, 11*(21), 4490. https://doi.org/10.3390/math11214490

Kumar, R. R., Stauvermann, P. J., & Dau, L. T. M. (2025). Exploring the tourism and economic growth relationship in Vietnam: A cointegration analysis with model-specific structural breaks. *Economies, 13*(2), 29. https://doi.org/10.3390/economies13020029

Tran, P. K. T., Nguyen, H. K. T., Nguyen, L. T., Nguyen, H. T., Truong, T. B., & Tran, V. T. (2023). Destination social responsibility drives destination brand loyalty: A case study of domestic tourists in Danang city, Vietnam. *International Journal of Tourism Cities, 9*(1), 302–322. https://doi.org/10.1108/IJTC-03-2022-0069

Lin, W.-C., Wu, C. K., Le, T. K. T., & Nguyen, N. A. (2023). Assessment of Vietnam tourism recovery strategies after COVID-19 using multi-criteria decision-making approach. *Sustainability, 15*(13), 10047. https://doi.org/10.3390/su151310047

Nguyen, T. A., & Nguyen, L. A. (2024). Attracting and retaining international students in emerging countries: A case of Vietnam. *Journal of Contemporary East Asia Studies, 13*(1), 118–142. https://doi.org/10.1080/24761028.2024.2445457

Li, X., Xu, Y., Law, R., & Wang, S. (2024). Enhancing tourism demand forecasting with a transformer-based framework. *Annals of Tourism Research, 107*, 103791. https://doi.org/10.1016/j.annals.2024.103791

Fatima, S. S. W., & Rahimi, A. (2024). A review of time-series forecasting algorithms for industrial manufacturing systems. *Machines, 12*(6), 380. https://doi.org/10.3390/machines12060380

Ma, Q. (2025, June). A novel prediction of tourism economic development trends based on machine learning. In *Proceedings of the 2025 International Conference on Intelligent Computing and Knowledge Extraction (ICICKE)* (pp. 1–7). IEEE. https://doi.org/10.1109/ICICKE65317.2025.11136217

Kucharavy, D., Damand, D., & Barth, M. (2023). Technological forecasting using mixed methods approach. *International Journal of Production Research, 61*(16), 5411–5435. https://doi.org/10.1080/00207543.2022.2102447

Grandón, T. G., Schwenzer, J., Steens, T., & Breuing, J. (2024). Electricity demand forecasting with hybrid classical statistical and machine learning algorithms: Case study of Ukraine. *Applied Energy, 355*, 122249. https://doi.org/10.1016/j.apenergy.2023.122249

Kaur, J., Parmar, K. S., & Singh, S. (2023). Autoregressive models in environmental forecasting time series: A theoretical and application review. *Environmental Science and Pollution Research, 30*(8), 19617–19641. https://doi.org/10.1007/s11356-023-25148-9

Iftikhar, H., Bibi, N., Canas Rodrigues, P., & López-Gonzales, J. L. (2023). Multiple novel decomposition techniques for time series forecasting: Application to monthly forecasting of electricity consumption in Pakistan. *Energies, 16*(6), 2579. https://doi.org/10.3390/en16062579

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16045
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Mulla, S., Pande, C. B., & Singh, S. K. (2024). Time series forecasting of monthly rainfall using seasonal autoregressive integrated moving average with exogenous variables (SARIMAX) model. *Water Resources Management, 38*(6), 1825–1846. https://doi.org/10.1007/s11269-024-03756-5

Kahraman, E., & Akay, O. (2023). Comparison of exponential smoothing methods in forecasting global prices of main metals. *Mineral Economics, 36*(3), 427–435. https://doi.org/10.1007/s13563-022-00354-y

Lallawmzuali, M., & Muniyan, S. (2025). Rainfall prediction in Mizoram using the modified Holt–Winters method. *Theoretical and Applied Climatology, 156*(2), 125. https://doi.org/10.1007/s00704-025-05354-w

Montero-Manso, P. (2023). How to leverage data for time series forecasting with artificial intelligence models: Illustrations and guidelines for cross-learning. In *Forecasting with artificial intelligence: Theory and applications* (pp. 123–162). Springer Nature. https://doi.org/10.1007/978-3-031-35879-1_6

Dhankhar, S., Dhankhar, N., Sandhu, V., & Mehla, S. (2024). Forecasting electric vehicle sales with ARIMA and exponential smoothing method: The case of India. *Transportation in Developing Economies, 10*(2), 32. https://doi.org/10.1007/s40890-024-00216-y

Fang, P., Gao, Z., & Tsay, R. S. (2023). Supervised kernel principal component analysis for forecasting. *Finance Research Letters, 58*, 104292. https://doi.org/10.1016/j.frl.2023.104292

Anh, H. H., Phuong, D. N. D., Ha, P. T., Tu, L. H., Hanh, T. M. D., & Loi, N. K. (2024). Sustainable agricultural development in Gia Lai Province, Central Highlands Vietnam: A holistic approach to addressing socio-economic and climate challenges. *Environment, Development and Sustainability*. https://doi.org/10.1007/s10668-024-05761-5

Nguyen-Da, T., Li, Y.-M., Peng, C.-L., Cho, M.-Y., & Nguyen-Thanh, P. (2023). Tourism demand prediction after COVID-19 with deep learning hybrid CNN–LSTM—Case study of Vietnam and provinces. *Sustainability, 15*(9), 7179. https://doi.org/10.3390/su15097179

*name of corresponding author