International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 5 Volume No.7 September - October – 2023
DOI:10.46647/ijetms.2023.v07i05.027 ISSN: 2581-4621

Foreign Direct Investment in Vietnam: A Comparison of Forecasting Performance

Minh Ly Duc[1], Dong Phan Tan[2], Thu Nguyen Anh[2],
Ngoc Hoang Thi Mai[2], My Pham Ngoc Tra[2], Nhi Nguyen Thi Tuyet[2]
[1]*University of Van Lang, Faculty of Commerce, Ho Chi Minh City, Vietnam*
[2]*HCMC University of Technology and Education, Faculty for High-Quality Training, Department of Industrial Management, Ho Chi Minh City, Vietnam.*

***Abstract*—** In recent years, FDI has made a significant contribution to the economic development of Vietnam, especially in the manufacturing and processing sectors. To forecast the FDI inflow into Vietnam in the future, a research team used indicators such as the number of investment projects, the total registered capital, and the total implemented capital. These indicators were collected from the General Statistics Office to forecast the level of FDI attraction in Vietnam in the coming years. This study includes annual data on the number of FDI investment projects in Vietnam collected from 2000 to 2022. The forecasting methods used in this study include the Autoregressive Integrated Moving Average (ARIMA) model, additive time series forecasting, Holt-Winters forecasting, and linear regression forecasting using Mini-tab 18.0 software. After analyzing the data on the number of FDI investment projects in Vietnam from 2000 to 2023, the research team collected results and compared the accuracy of the forecasting models based on three metrics calculated from the forecasting models, which are MAPE, MAD, and MSE. The comparison of the metrics showed that the Time Series method is the most suitable forecasting method with the lowest MAPE and MAD values compared to the other three models, while the MSE value of the Time Series model is only higher than the ARIMA model and still lower than the other two models. With the comparison results of the forecasting models, the study found a suitable forecasting model for each time cycle.
***Keywords*— Autoregressive Integrated Moving Average, ARIMA, Time Series, Holt-Winters, Linear Regression.**

INTRODUCTION

An important and challenging task is to use time series data and apply models to forecast FDI in Vietnam. Statistical websites provide inconsistent data on the number of investment projects, total registered capital, and total implemented capital. Specifically, data providers often report the combined data of registered and adjusted capital. Evaluating the accuracy of forecasting models for FDI in Vietnam from 2023 to 2045 is a challenge due to the inconsistency of data from statistical websites and the unstable trend of FDI over the years. Each forecasting model has its own drawbacks and errors, and the accuracy of forecasting varies. It is necessary to use the same dataset and prediction analysis software to compare the performance and accuracy of each forecasting model.

Forecasting is a scientific prediction with a probabilistic nature and is an essential tool for management. There are various types of forecasting methods available to forecast the FDI situation, including: (1) Quantitative forecasting, this method uses mathematical models to predict the amount of FDI in the future. Quantitative methods use regression models such as linear regression, multiple regression, and time series models such as ARIMA, VAR, and their derivatives. (2) Qualitative forecasting, this method works based on non-numerical sources of information, such as economic trends, investment promotion policies, or political, security, geographical, and cultural factors. Qualitative methods usually use the Delphi method, expert opinion, or SWOT analysis. (3) Combination forecasting, this method combines both quantitative and qualitative methods to provide more accurate predictions. Combination methods can use quantitative models to predict the quantity of FDI and use qualitative factors to adjust the forecast results.

The forecasting methods for FDI can be applied to specific cases and need to be carefully evaluated before applying them to ensure the accuracy and reliability of the forecast results. Research objectives:

- Select the most effective forecasting model.
- Forecast the number of FDI projects from 2023 to 2026.
- Determine the variability of forecasting models.

The general structure of the research, title is Forecasting FDI into Vietnam based on the number of investment projects, total registered capital, and total implemented capital from 2023 to 2045. In this study, we analyze time series data from 2000 to 2022 and apply forecasting models to predict the trend of FDI capital inflows into Vietnam from 2023 to 2045. We will compare the performance and accuracy of forecasting models, including the ARIMA model, time series forecasting model, Holt-Winters forecasting model, and linear regression forecasting model. The ARIMA model (AutoRegressive Integrated Moving Average) is formed by combining two separate models, including the autoregressive model AR (p) and the moving average model MA (q). The time series forecasting model uses an additive model, including trend, seasonality, cyclical, and random components. The Holt-Winters forecasting model is an extension of the Holt model and is suitable for data with trends. The linear regression forecasting model is used to forecast a series of time numbers.

By comparing the forecasting results of these models on the same dataset and using the same prediction analysis software, we will provide significant insights into the accuracy and performance of each model. This study has the following important contributions:

- Conducting research and experiments on time series datasets to examine the performance of time series forecasting models in predicting FDI capital inflows into Vietnam.
- Comparing the accuracy and performance of forecasting models such as ARIMA, time series forecasting model, Holt-Winters forecasting model, and linear regression forecasting model. As a result, we determine which model provides the best results in forecasting FDI capital inflows into Vietnam.
- Providing detailed experimental results of time series forecasting models applied on the same dataset and using the same prediction analysis software. In this case, we use Minitab 18.0 software to perform analysis and compare results

Through this study, we hope to provide accurate evaluations of FDI capital inflows into Vietnam in the future based on factors such as the number of investment projects, total registered capital, and total implemented capital. The results from this study will provide important information for economic policy decisions and support for attracting foreign investment into Vietnam.

Theoretical Foundation

Overview of FDI in Vietnam.

FDI has made significant contributions to the economic development of Vietnam in recent years. Since opening up to foreign investment in the early 1990s, Vietnam has attracted a large amount of FDI, particularly showing a notable growth trend in the volume and value of FDI projects invested in Vietnam. According to reports from the Ministry of Planning and Investment of Vietnam, the amount of FDI invested in Vietnam from 2000 to 2020 increased 28 times, from \$2.7 billion in 2000 to \$62.1 billion in 2020. In addition, in 2020, Vietnam attracted over 2,000 new FDI projects with a total registered capital of \$31 billion. The main sectors attracting FDI in Vietnam include manufacturing and industrial processing, construction, real estate, and retail business. Foreign direct investment has contributed to the creation of thousands of jobs domestically, while also improving the quality and production capacity, introducing new products and services, and enhancing Vietnam's economic infrastructure. The major countries investing directly in Vietnam are Japan,

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 5 Volume No.7 September - October – 2023
DOI:10.46647/ijetms.2023.v07i05.027 ISSN: 2581-4621

South Korea, Singapore, Taiwan, and China. Vietnam has also signed various free trade agreements with different countries and regions, particularly the CPTPP and EVFTA, which have helped attract more FDI and expand export markets for Vietnamese businesses.

Forecasting models

1) Autoregressive Integrated Moving Average (ARIMA) Model: The Autoregressive Integrated Moving Average (ARIMA) model is formed by combining the autoregressive model (AR) of order p and the moving average model (MA) of order q. Here, p represents the order of the autoregressive component, and q represents the lag of the moving average component. The ARIMA(p, q) model is expressed as follows:

$$\hat{Y}_t = \emptyset_0 + \emptyset_1 Y_{t-1} + \emptyset_2 Y_{t-2} + \cdots + \emptyset_p Y_{t-p} + \varepsilon_t - \omega_1 \varepsilon_{t-1} - \omega_2 \varepsilon_{t-2} - \cdots - \omega_q \varepsilon_{1-q} \quad (1)$$

The simplified formula is as follows:

$$\widehat{Y}_t = \underbrace{\theta_0 + \sum_{i=1}^{p} \theta_i Y_{t-i}}_{\text{AR Model}} + \underbrace{\varepsilon_t - \sum_{j=1}^{q} \omega_j \varepsilon_{t-j}}_{\text{MA Model}}$$

With:

$\hat{Y}_t$: Forecast value at time t.

$\emptyset_0$:Initial constant.

$\emptyset_i$:Auto-regression coefficient.

ω_i:Estimated coefficient; note the convention of placing the minus sign.

ε_t:Forecast error at time t.

$\varepsilon_{t-1}, \varepsilon_{t-2}$:Past forecast error.

Based on the autocorrelation function (ACF) plot and the partial autocorrelation function (PACF) plot, users can choose one of the three models: autoregressive AR(p), moving average MA(q), or autoregressive integrated moving average ARIMA(p,q) for forecasting. If the ACF plot of the time series decreases and abruptly stops at lag q, and the PACF plot decreases and abruptly stops at lag p, then the combined ARIMA(p, q) model should be applied. According to Hanke [6], the values of p and q are usually smaller than 2. Refer to Table 1 for criteria for selecting the forecasting model based on ACF and PACF.

TABLE I
CRITERIA FOR SELECTING FORECASTING MODELS BASED ON ACF AND PACF

Model	**Autocorrelation Function (ACF)**	**Partial Autocorrelation Function (PACF)**
MA (q)	Abrupt stop at lag q	Decreasing
AR (p)	Decreasing	Abrupt stop at lag p
ARIMA (p, q)	Decreasing and abrupt stop at lag q	Decreasing and abrupt stop at lag p

The AR(p), MA(q), or ARIMA (p, q) forecasting models can be used to forecast a time series, provided that the time series is stationary. In the case of a non-stationary time series, differencing must be applied to make the time series stationary. Specifically, the initial series Y_t is non-stationary and the first-order difference of Y_t is: which is a stationary series. The combined regression - differenced moving average model is:

$$\hat{Y}_t = Y_{t-1} + \emptyset_0 + \emptyset_1 (Y_{t-1} - Y_{t-2}) + \varepsilon_t - \omega_1 \varepsilon_{t-1} \quad (3)$$

In principle, it is necessary to take the differences of the series until it becomes stationary before applying the AR(p), MA(q), or ARIMA(p,q) forecasting models. Here, d is the symbol for the total number of differencing required to achieve stationarity. The combined differencing with AR(p) and

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 5 Volume No.7 September - October – 2023
DOI:10.46647/ijetms.2023.v07i05.027 ISSN: 2581-4621

MA(q) models is called the ARIMA(p, d, q) model. If the initial series is already stationary, d = 0, and the ARIMA model becomes the ARIMA(p, q) model.

2) Time Series Forecasting using Additive Model: The forecasting formula for a time series following an additive model is as follows:

$$\hat{Y}_t = T_t + S_t + I_t + C_t \ \mathbf{(4)}$$

With:

$\hat{Y}_t$: Forecast value for at time t

T_t: Trend component

S_t: Seasonal ingredients

I_t: Unusual ingredients

C_t: Cycle component

The time series forecasting based on the additive model is performed (Table.II).

TABLE II
STEP BY STEP FOR TIME SERIES FORECASTING METHOD

Step	Explain detail for step by step
Step 1	Determine the seasonal coefficients. The centered moving average at time (t + 0.5) is calculated as: $$Y^*_{t+0.5} = \sum_{j=-\frac{s}{2}+1}^{s/2} \frac{Y_{t+j}}{S} \ \mathbf{(5)}$$ With: s=12 (data follow months)
Step 2	The centered moving average at time t is calculated as: $$Y^*_t = \frac{Y^*_{t-0.5} - Y^*_{t+0.5}}{2} \ \mathbf{(6)}$$
Step 3	The average seasonal factor for period i is the average of the seasonal factors for period i across all periods (months), and is calculated using the following formula: $$\bar{S}_i = \frac{\sum_{j=0}^{r-1} S_{jp+i}}{r} \ \mathbf{(7)}$$ With: r:Season cycle (month), p:Number of seasons in a cycle (in months), S_i:Seasonal factors in period i
Step 4	Seasonal adjustment factor: $$k = \frac{\sum \bar{S}_i}{p} \ \mathbf{(8)}$$ With, $\bar{S}_i$:Average seasonal factor. It is necessary to adjust the seasonal indices so that the sum of the seasonal factors is always zero. Adjusting the seasonal values to zero will result in discrepancies in the forecasting results.
Step 5	Determine the trend component T_t by constructing a linear regression equation to estimate the trend of the time series. The dependent variable Y in the regression equation is assigned the deseasonalized time series data. The independent variable t is assigned the time with values increasing from 1 to n time points. The parameters α and β are calculated using the following formulas: $\beta_0 = \bar{y} - b_1\bar{t} \ \mathbf{(9)}$

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 5 Volume No.7 September - October – 2023
DOI:10.46647/ijetms.2023.v07i05.027 ISSN: 2581-4621

	$\beta_1 = \frac{\sum ty - nt\bar{y}}{\sum t^2 - nt^{-2}}$ **(10)** With, t:Independent variable at time t of the time series. $\bar{t}$ ::Average of t. y:Dependent variable Y_t -S_i . $\bar{y}$ ::Average of the variable y. n:Number of periods (sample size).
Step 6	Evaluate the model's errors and calculate the forecasted value. The sum of squared residuals R^2 is used to evaluate the model's errors. It is necessary to identify the outlier component I of the time series before calculating the sum of squared residuals R^2. The trend component T and seasonal factor S are known from the precalculated series. The irregular component I is calculated using the following formula: $I_t = Y_t - (T_t - S_i)$ **(11)** With, I_t:Unusual ingredients. T_t:Trend component. Y_t:Time data t. S_i:Adjusted Seasonal Factor.
Step 7	The estimation of forecast model error is calculated using the sum of squared errors R^2, according to the following formula: $R^2 = 1 - \frac{\sum I_t^2}{\sum (Y_t - \bar{Y})^2}$ **(12)** With, $\bar{Y}$: Average of data t period. The R^2 value is always less than 1,the R^2 value is to 1, the model is qualified.
Step 8	To test the autocorrelation of the model, the Durbin-Watson test is applied. Autocorrelation is the correlation of I_t by the time. If autocorrelation exists, it indicates that the model is not suitable due to non-random residuals. The statistical value of the Durbin-Watson test is calculated using the following formula: $DW = \frac{\sum_{t=2}^{n} (I_t - I_{t-1})^2}{\sum_{t=1}^{n} (I_t)^2}$ **(13)** With, I_{t-1}:Abnormal component slips 1 unit of time.
Step 9	The forecast value for period t is calculated using the formula: $F_t = T_t + S_i$ **(14)**

3) Step by Step for Holt-Winters Forecasting: Winters' exponential smoothing is an extension of Holt's exponential smoothing, applied to data with both trend and seasonality. The forecast value for Winters' exponential smoothing for period (t + 1) is calculated using the following formula 15 and step by step for Holt-Winters Forecasting (Table III).

$$\hat{Y}_{t+1} = (F_t + T_t) \times S_{t+1} \textbf{ (15)}$$

TABLE III
STEP BY STEP FOR HOLT-WINTERS FORECASTING

Step	**Explaint detail for step by step**
Step 1	The forecast for n future periods after time t is calculated using the formula: $\hat{Y}_{t+n} = (F_t + nT_t) \times S_{t+n}$ **(16)**
Step 2	The forecast value for the seasonal exponential component is calculated using the formula:

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 5 Volume No.7 September - October – 2023
DOI:10.46647/ijetms.2023.v07i05.027 ISSN: 2581-4621

	$F_t = \alpha\left(\frac{Y_t}{S_t}\right) + (1-\alpha)(F_{t-1} + T_{t-1})$ **(17)**
Step 3	The smoothing parameter for the trend is calculated using the formula: $T_t = \beta(F_t - F_{t-1}) + (1-\beta)T_{t-1}$ **(18)**
Step 4	The seasonal index is calculated using the formula: $S_{t+p} = \gamma\left(\frac{Y_t}{F_t}\right) + (1-\gamma)S_t$ **(19)** With, α :Exponential factor. β :Exponential of the navigation factor. γ :Exponential of the seasonal index. Y_t :Time data t. $\hat{Y}_{t+1}$:Winters exponential predictive value for time t+1. F_t :Exponential Forecasting (Winters). T_t :Navigation factor. S_t :Season coefficient. n:Number of forecast periods for the future after time t. p:Number of seasons in the cycle (months).

4) *Linear Regression Forecasting: Two variables, X and Y, where Y is linearly dependent on X, with the formula (20) and step by step for Linear Regression Forecasting (Table. IV).*

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \ \mathbf{(20)}$$

TABLE IV
STEP BY STEP FOR LINEAR REGRESSION FORECASTING

Step	**Explain detail for step by step**
Step 1	To predict a time series using a linear regression model, the time order t needs to be established. The variable t usually takes the value 1 for the first observation and increases by 1 for each subsequent observation. The regression model is calculated using the following formula: $\hat{Y} = b_0 + b_1 t$ **(21)** With, $\hat{Y}$ ::Forecast value. t :Time sequence (t = 1, 2, ..., n). b_0 *and* b_1 :Regression coefficient.
Step 2	The least squares method (LSM) will draw a straight line through the observations with the samples $(Y_1, t_1), (Y_2, t_2), \ldots (Y_n, t_n)\varepsilon_i b_0,\ b_1 \beta_0,\ \beta_1$, the line is drawn in such a way that the residual error is minimized with the values to estimate the parameters. The sum of squared residuals (SSE) is used to avoid the case where the total residual difference is zero. The total sum of squares (SST) is calculated as the squared deviation of the observed values Yi from their mean value. The formulas for SSE and SST are as follows: $\sum_{i-1}^{n} \varepsilon_i^2 = \sum_{i-1}^{n} (Y_i - \hat{Y}_i)^2 = \sum_{i-1}^{n} (Y_i - b_0 - b_1 t_i)^2$ **(22)**
Step 3	By taking partial derivatives of the equation for the squared residual difference with respect to b_0 and b_1, and then setting the expressions to zero, we obtain a system of equations:

DOI:10.46647/ijetms.2023.v07i05.027 ISSN: 2581-4621

	$\begin{cases} \sum_{i=1}^{n} y_i = nb_0 + b_1 \sum_{i=1}^{n} t_i (1) \\ \sum_{i=1}^{n} t_i y_i = b_0 \sum_{i=1}^{n} t_i + b_1 \sum_{i=1}^{n} t_i^2 (2) \end{cases}$ **(23)**
Step 4	From formula (23), we can calculate b0 and b1 using the following formulas. The base coefficient b_1 is calculated as: $b_1 = \frac{\sum tY - \frac{\sum t \sum Y}{n}}{\sum t^2 - \frac{(\sum t)^2}{n}}$ **(24)**
Step 5	The origin coordinates b_0 is calculated as: $b_0 = \bar{y} - b_1\bar{t}$ **(25)** With, $\bar{\boldsymbol{t}} = \frac{\sum_{i=1}^{n} t_i}{n}$:Average of the independent variable. $\bar{\boldsymbol{Y}} = \frac{\sum_{i=1}^{n} Y_i}{n}$:Average of the dependent variable.
Step 6	To evaluate the adequacy of a regression prediction model, we can rely on the coefficient of determination R^2 and the variance of the regression. The coefficient of determination R^2 is calculated using the following formula: $R^2 = 1 - \frac{SSE}{SST}$ **(26)** With, value of coefficient R^2, $0 \leq R^2 \leq 1$
Step 7	SSE: Sum of Squared Errors, the squared difference between $\boldsymbol{Y_i}$ and $\hat{\boldsymbol{y}}_i$, is calculated using the formula: $SSE = \sum (Y_i - \hat{y}_i)^2$ **(27)**
Step 8	SST: Total Sum of Squares, the squared deviation of Y_i around the mean value of $\bar{Y}$, calculated using the following formula: $SST = \sum (y_i - \bar{y})^2$ **(28)** With, $\bar{Y}$:Mean value of dependent variable. Y_i :Observed value. $\hat{Y}_i$:Predictive value corresponding to a value of X_i
Step 9	The value of the coefficient of determination R^2 is applied to evaluate the errors of the regression model and measure the usefulness of the regression model in predicting time series. The estimated variance is calculated when the dependent variable Y follows a normal distribution, meaning that the estimated parameters b_0, b1 are also typically distributed. It is calculated using the following formula: $S_e^2 = \frac{SSE}{n-2}$ **(29)** With, n: Sample size
Step 10	The estimated standard error Se is calculated using the formula (30). The larger the

DOI:10.46647/ijetms.2023.v07i05.027 ISSN: 2581-4621

	estimated standard error of S_e, the larger the average error of the regression model. $S_e = \sqrt{S_e^2}$ **(30)**

III. THE COMPARATIVE STUDY

Data source

The database on Foreign Direct Investment (FDI) in Vietnam covers various aspects including the total number of investment projects, the total registered capital, and the total implemented capital from 2000 to 2022 (Table. 5). The data is extracted from the book "Statistical Yearbook 2021" published by the General Statistics Office and the website https://www.mpi.gov.vn/. The dataset is organized on an annual basis, where the unit of measurement for the "Total number of investment projects" aspect is the number of projects, while the "Total registered capital" and "Total implemented capital" aspects are measured in million US dollars. This dataset is compiled annually by the General Statistics Office and the Ministry of Planning and Investment, covering all localities within the territory of Vietnam, and it is published in the Statistical Yearbook 2021 by the General Statistics Office and the Electronic Information Portal of the Ministry of Planning and Investment. The data used in this research is collected over time (annually) from the websites https://www.gso.gov.vn and https://www.mpi.gov.vn.

TABLE V

DATA OF FDI PROJECTS INVESTED IN VIETNAM FROM 2000 TO 2022

Year	**Number of projects**	**Year**	**Number of projects**
2000	391	2012	1287
2001	555	2013	1530
2002	808	2014	1843
2003	791	2015	2120
2004	811	2016	2613
2005	970	2017	2741
2006	987	2018	3147
2007	1544	2019	4028
2008	1171	2020	2610
2009	1208	2021	1818
2010	1237	2022	2036
2011	1186	-	-

Accuracy Evaluation Index

Forecast error is the difference between the forecasted value and the actual data at a specific time point.

$\varepsilon_t = Y_t - \hat{Y}_t$ **(31)**

With

Y_t:Actual value at time t.

$\hat{Y}_t$:Forecast value at time t.

ε_t:Forecast error at time t.

The smaller the prediction error, the better the predictive model (Table. 6). However, the randomness of the parameters is also an important factor in assessing the accuracy of the forecast. Assuming that the original data is random, the calculations, evaluations, and experiments are also

DOI:10.46647/ijetms.2023.v07i05.027 ISSN: 2581-4621

based on assumptions of randomness and often follow a certain distribution. If the model is correct, the errors should not exhibit any directional bias.

TABLE VI

Step	Explain detail of step by step
Step 1	Mean Absolute Error (MAE) eliminates the positive and negative errors by taking their absolute values to reduce the error value for absolute values. MAE is calculated using the following formula: $MAE = \frac{\sum_{t=1}^{n}\lvert\varepsilon_t\rvert}{n} = \frac{\sum_{t=1}^{n}\lvert Y_t - \hat{Y}_t\rvert}{n}$ **(32)**
Step 2	To determine the percentage ratio for normalization and avoid ambiguity in the data, mean absolute percentage error (MAPE) is calculated. For instance, a 1 mistake is low compared to 1000 but high compared to 10. This formula is used to calculate MAPE. $MAPE = \frac{\sum_{t=1}^{n}\frac{\lvert\varepsilon_t\rvert}{Y_t}}{n} = \frac{\sum_{t=1}^{n}\frac{\lvert Y_t - \hat{Y}_t\rvert}{Y_t}}{n}$ **(33)**
Step 3	The formula below is used to compute mean square error (MSE). $MSE = \frac{\sum_{t=1}^{n}\varepsilon_t^2}{n} = \frac{\sum_{t=1}^{n}\left(Y_t - \hat{Y}_t\right)^2}{n}$ **(34)**
Step 4	The following formula is used to determine Root Mean Square Error (RMSE). $RMSE = \sqrt{\frac{\sum_{t=1}^{n}\varepsilon_t^2}{n}} = \sqrt{\frac{\sum_{t=1}^{n}\left(Y_t - \hat{Y}_t\right)^2}{n}}$ **(35)**

These two criteria are also seen as being equivalent because the RMSE criterion is the square root of the MSE, nevertheless, the RMSE will have a lower value than the MSE. The properties, functions, and output of the MAE, MSE, and RMSE criteria are the same. The MSE or RMSE criteria are chosen as the evaluation criteria if the _t values are different, and vice versa, if the _t values are too different, the MAE criteria are chosen for evaluation.The relative error can be estimated using the MAPE criterion. Therefore, the MAPE criterion should be applied while assessing the forecast error using various data sets. In contrast, the MAPE criterion should not be used due to its complex calculation when using the same data set but employing different forecasting techniques.

result of DESIGN for THE SYSTEM Data

In evaluating the performance of time series forecasting models, we often use indicators such as MAPE (Mean Absolute Percentage Error), MAD (Mean Absolute Deviation), and MSE (Mean Squared Error). Based on the obtained results, we can compare and select the most appropriate evaluation method (Table VII).

TABLE VII

Result / Method	**MAPE**	**MAD**	**MSD**	**Dự báo**			
				2023	**2024**	**2025**	**2026**
ARIMA	20	759	211112	2044.49	1956.55	1883.64	1822.5

DOI:10.46647/ijetms.2023.v07i05.027 ISSN: 2581-4621

Time series	18	332	241517	2966.94	3078.56	3190.18	3301.8
Holt-Winter	24	429	349604	2904.36	2758.59	3227.03	3135.76
Linear Regression	19	334	290024	3504.21	3782.45	4082.78	4406.96

Firstly, we compare the Time series and ARIMA methods which have been used to forecast and have different results:

- With the Time series method, we obtained a MAPE of 18, MAD of 332, and MSE of 241,517. This shows that the average percentage deviation of the forecast from the actual value is about 18%. The value of MAD is 332, indicating that the average absolute deviation between the forecast and the actual value is 332 units. MSE is 241,517 square units, indicating that the average deviation of the squared error between the forecast and the actual value is 241,517.
- With the ARIMA method, we obtained a MAPE of 20, MAD of 759, and MSE of 211,112. This shows that the average percentage deviation of the forecast from the actual value is about 20%. The value of MAD is 759, indicating that the average absolute deviation between the forecast and the actual value is 759 units. MSE is 211,112 square units, indicating that the average deviation of the squared error between the forecast and the actual value is 211,112.
- From the above results, we can see that the Time series method has lower MAPE and MSE than the ARIMA method, indicating that the Time series method has higher accuracy in forecasting. However, the ARIMA method has a lower value of MAD, indicating that the average absolute deviation between the forecast and the actual value is less than that of the Time series method.

When comparing the indicators of Time series with Holt-Winters, we obtained the following results.

- With the Time series method, we obtained a MAPE of 18, MAD of 332, and MSE of 241,517. The low MAPE indicates that the average percentage error between the forecast and the actual value is small (about 18%), indicating that this method has good forecasting ability. The low MAD (332) also indicates that the average absolute deviation between the forecast and the actual value is small, which is another advantage of the Time series method. However, the MSE is quite high (241,517), indicating that the average deviation of the squared error between the forecast and the actual value is large.
- With the Holt-Winters method, we obtained MAPE of 24, MAD of 429, and MSE of 349,604. The higher MAPE compared to the Time series method (about 24%) indicates that the Holt-Winters method has a larger average percentage error. However, the MAD is 429, indicating that the average absolute deviation between the forecast and the actual value is still small. The MSE is 349,604, higher than the Time series method, indicating that the average deviation of the squared error between the forecast and the actual value is large.
- From the above results, we can see that the Time series method shows better results than the Holt-Winters method in forecasting this specific time series. The Time series method has lower MAPE, MAD, and MSE, indicating that this method has higher accuracy in forecasting.

Finally, we compare the indicators obtained from the Time series forecasting method with the Linear Regression method, and obtained the following results.

- With the Time series method, we obtained MAPE of 18, MAD of 332, and MSE of 241,517. The low MAPE indicates that the average percentage error between the forecast and the actual value is small (about 18%), indicating that this method has good forecasting ability. The low MAD (332) also indicates that the average absolute deviation between the forecast and the actual value is small, which is another advantage of the Time series method. The MSE is also low (241,517), indicating that the average deviation of the squared error between the forecast and the actual value is small.
- With the Linear Regression method, we obtained MAPE of 19, MAD of 334, and MSE of 290,024. The higher MAPE compared to the Time series method (about 19%) indicates that the

DOI:10.46647/ijetms.2023.v07i05.027 ISSN: 2581-4621

Linear Regression method has a larger average percentage error. However, the MAD is 334, indicating that the average absolute deviation between the forecast and the actual value is still small. The MSE is higher (290,024) than the Time series method, indicating that the average deviation of the squared error between the forecast and the actual value is larger.

- From the above results, we can see that the Time series method has better MAPE, MAD, and MSE compared to the Linear Regression method in forecasting this specific time series. The Time series method shows higher accuracy, and the average error between the forecast and the actual value is smaller.

In conclusion, based on the obtained results, we can see that the Time series method is the most suitable method for FDI forecasting in the future because it satisfies all the indicators (MAPE, MAD, MSE) at a better level compared to the ARIMA, Holt-Winters, and Linear Regression methods.

Conclution

With the results of the study, the research team has found the best forecasting method for a dataset similar to the one used in this study, which is the Time series forecasting method. In addition, by using the Time series method to analyze and forecast the dataset used in the research, the research team was able to forecast the number of FDI investment projects in Vietnam from 2023 to 2026. This result can help investment management agencies to plan their FDI development strategy for the next 4 years. The study also identified the variability of the Integrated Autoregressive Moving Average (ARIMA), Time series forecasting, Holt-Winters forecasting, and Linear Regression forecasting models, which can provide experience for future research and the selection of appropriate forecasting models. However, due to the limited scope of the research, there are still some limitations in the results, such as:

- The dataset and topic used in the study were synthesized from previous domestic research, which may not ensure objectivity and transparency in the analysis and forecasting of the number of FDI investment projects in Vietnam from 2000 to 2023.
- The study only used and analyzed the results of four commonly used forecasting models, namely the Integrated Autoregressive Moving Average (ARIMA), Time series forecasting, Holt-Winters forecasting, and Linear Regression forecasting. Therefore, the study cannot accurately determine the most effective model for forecasting economic data in the future.
- The research team only provided a forecast for the next 4 years from 2023 to 2026. Thus, investment management agencies in Vietnam cannot use the forecast results of this study to plan their long-term investment strategy for FDI projects beyond 2026.

With the benefits and limitations of the study mentioned above, the research team has also planned for future research to serve investment management and research on the situation of foreign direct investment in Vietnam, including:

- Conducting research on the registration and disbursement of FDI capital by foreign enterprises in Vietnam, thereby identifying the causes and finding solutions to the inconsistent FDI registered capital and actual capital disbursement in each year.
- Conducting research to determine the main types of FDI investment in Vietnam, thereby providing proposals for managing and planning the economy by sector to maintain balance for Vietnam's market economy in the future.
- Conducting research to analyze and optimize the adjustment parameters of forecasting models such as the Integrated Autoregressive Moving Average (ARIMA), Holt-Winters forecasting, and Linear Regression forecasting. From there, the research can provide the optimal parameters for each case for the above forecasting models and support future forecasts to provide the best results when using these forecasting models on commonly used statistical software.

DOI:10.46647/ijetms.2023.v07i05.027 ISSN: 2581-4621

ACKNOWLEDGMENTS

We acknowledge Van Lang University for supporting this study.

References

[1] Anh ,D, B, H. (2016) Business Forecasting, Ho Chi Minh City Economic Publishing House, 2016, ISBN: 978-604-922-348-8.

[2] Pragya, V., Sai, V, R., Lata, R. and Debabrata, D. (2021) Comparison of Time-Series Forecasting Models. *2021 International Conference on Intelligent Technologies (CONIT),* No. 21174630, pp. 1-7, DOI: 10.1109/CONIT51480.2021.9498451.

[3] Qiang, D., Chuancheng, Zh., Jingyang, Zh., Sai, D., Dan, X., and Zhiqiang, L. et al. (2017)Day-Ahead Probabilistic Load Flow Analysis Considering Wind Power Forecast Error Correlation.Energy and Power Engineering. Vol. 9, pp. 292-299, https://doi.org/10.4236/epe.2017.94B034

[4] Rita, G., Francesco, L., Bianca, R., and Fabio, S. (2010) Forecasting of Sporadic Demand Patterns with Seasonality and Trend Components: An Empirical Comparison between Holt-Winters and (S)ARIMA Methods.*Mathematical Problems in Engineering*. Vol. 2010, pp. 1-14, Doi:10.1155/2010/579010.

[5] John, E, H., and Dean, W, W. (2005) Business Forecasting, 8th Edition, Chapter 9.

[6] Yu, X., and Zhezhi, J. (2021) The Forecast Research of LinearRegression Forecast Model in National Economy. *Open Access Library Journal*, Vol. 8, pp. 1-17.

[7] Jun, F, L., and Qun, Z. (2012) The Forecasting of the Elevator Traffic Flow Time Series Based on ARIMA and GP. *Advanced Materials Research*, Vol.588-589,pp.1466-1471.

[8] Yin, P, C., Ai, P, W., Cui, L, W., Hai, Y, Zh., and Shu, X, F. (2013) Time Series Analysis of Pulmonary Tuberculosis Incidence: Forecasting by Applying the Time Series Model. *Advanced Materials Research*, Vol. 709, pp. 819-822.

[9] Zhuo, M, W., and Hong, C, S. (2013) Prediction Scheme of Railway Passenger Flow Based on Multiplicative Holt-Winters Model. *Applied Mechanics and Materials*, Vol. 416-417, pp. 1949-1953.

[10] General Statistics Office. (2022). Statistical year work of Viet Nam 2022. Statistical Publishing House.