



# House Price Forecast Model: Case of Vietnam Housing Market

*International Conference on Emerging Challenges:*

*Smart Business and Digital Economy*

NGUYEN Tai Quang Dinh<sup>1</sup>

DANG Bich Ngoc<sup>2</sup>

NGO Thu Giang<sup>3\*</sup>

<sup>1</sup> School of Applied Mathematics and Informatics, Hanoi University of Science and Technology, Hanoi, Vietnam

<sup>2</sup> School of Business, University of Lincoln, Lincoln, United Kingdom

<sup>3</sup> School of Economics and Management, Hanoi University of Science and Technology, Hanoi, Vietnam

\*Corresponding author: [giang.ngothu@hust.edu.vn](mailto:giang.ngothu@hust.edu.vn)

## Abstract

**Purpose** – *The purpose of this study is to build up a comprehensive model for house price forecast in Vietnam Housing market.*

**Design/methodology/approach** – *The data related to house price; macro economic indicators and housing industry were gathered from [www.euromonitor.com](http://www.euromonitor.com) which is source of data for market and industrial researches. The study reviewed house price forecasting models with different forecast techniques from pass research papers. In results, a new house price forecasting model with a suitable forecasting technique is proposed.*

**Findings** – *The results are obtained from the application of different forecasting techniques. The final model is confirmed based on the models' error consideration. The forecast value is also determined basing on the selected model and forecasting techniques. Two proposed models are ARIMA and VAR. With the ARIMA model, the study comes up the conclusion that the housing price index has a dependent relationship depending on the value of that index over the past 2 years. With VAR model, the research found out that Urban population Ratio has an impact on housing price value for 3 consecutive years; and Housing Completions Index has a strong impact on the housing price index. Remarkably, the use of VAR models can bring positive results as well as more accurate forecasts in forecasting housing prices.*

**Practical implications** – *The project of building a housing price forecast model in the real estate market in Vietnam. Therefore, the research results are important for investors in the housing market, suppliers, and policy makers in all fields from economics, finance and specifically in the real estate market in Vietnam.*

**Keywords** – Housing market, House Price Forecast, forecast models, forecast techniques

## 1. INTRODUCTION

The housing market is one of the important components of the real estate market. Products include apartments, villas, and townhouses. As a developing country in Asia, Vietnamese perceive housing as both a place to live and an important investment asset. Therefore, their home buying behavior is evaluated according to the investor's approach. Fluctuations in home prices will impact not only buying behavior but also decisions to keep or sell homes in the future.

The housing market has increased steadily over the long period. The year 2021 is an exceptional case of increasing period due to impact of Covid 19 on the economy. Vietnam's housing market has favorable conditions for development as people always expect long-term housing, and therefore they pay great attention on purchase transactions (GSO, 2019). However,

the market is facing problems of absorption reduction, and the emergence of tenancy trend (Savil,2023). This significantly affects the supply and demand of the housing market, and thereby can disturb the house pricing in the Vietnamese market.

Housing price forecasting is therefore an important and diverse field of study, with a mix of researchers, governments, financial institutions, and real estate companies. Housing price forecasts will help provide important information to understand and assess real estate market trends, support investment decisions, develop policies, and forecast economic risks.

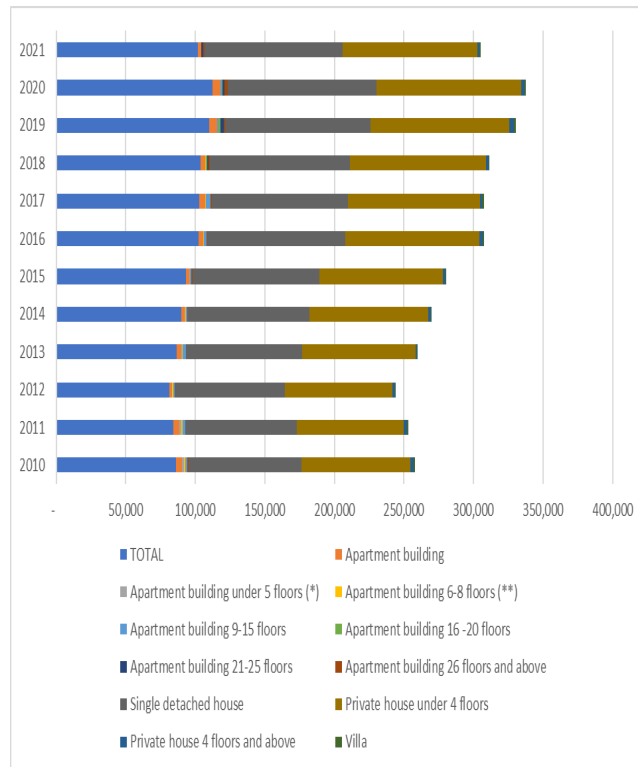


Fig.1 Residential construction by types of houses (Unit: '000. sqm)

Source: Vietnam General Statistics Office, 2023

The work of forecasting housing prices is done through a variety of methods and tools. Selection of appropriate and accurate methods and tools is a key issue of forecasting. This paper aims are reviewing methodologies applied in housing price forecast and proposing a proper methodology applied for house price forecasting in Vietnam housing market.

The research implements literature review to find out current models, then, considering, and testing feasibility of models basing on data of housing market in Vietnam. The testing results are bases for proposing appropriate, applicable models for Vietnam house price forecasting.

## 2. LITERATURE REVIEW

### 2.1 House price and its impact factors

Housing is part of the real estate market, which is important for creating a sustainable economy. In their research paper Uyar et al. (2016) showed that house prices have a direct and indirect effect on financial stability, so monitoring the housing market and house prices has become a significant issue in the house market management. After the 2008 global financial crisis, which according to author Nuri Hacıevliyagil (2022), a home credit crisis has started in the United States and spread worldwide, making the whole world pay more attention to the real estate bubble, especially the house price bubble.

Therefore, monitoring house prices, and paying attention to monitoring the development of the housing market, thereby finding relationships with economic variables can help government authorities in monitoring and making appropriate and timely policies for the housing market. In addition, the rapid and long-term increase in housing prices combined with overinvestment can lead to speculative price fluctuations, resulting in extremely dangerous real estate bubbles. When the

real estate bubble burst, then the house prices are out of control, negative economic consequences occur in the entire economy. That is why house price analysis is one of the tools to help detect and prevent housing price bubbles (Beltratti and Morana 2010).

In addition, the authors found that housing prices can help forecast other economic indicators. For example, housing construction is an important part of the total activities of the economy, expressed through GDP. Therefore, housing price fluctuations can be seen as an indicator of GDP change and development (Case et al, 2005). On the other hand, as an economic indicator, house prices can help predict fluctuations in inflation rates (Gupta and Kabundi, 2010). Therefore, accurate forecasting of house prices can be a useful tool for both policymakers and housing market participants to have an overview of the economy to make appropriate choices or policies.

Assessing the fluctuation of housing prices, studies using the House price index (Quang Truong et al, 2020; Vasilios, 2015; John M. Clapp, 2002). However, many authors argue that housing price forecasts can be based on the variation of housing prices in previous forecasts with actual house prices (Carlos, 2023); or use the growth coefficient of housing prices as the forecast variable (Lasse Bork, 2014).

The determinants of housing prices are clarified into three groups:

- Group 1. Macroeconomic indicators: per capita income, long-term interest rates, population, stock market fluctuations, construction prices, unemployment rate, inflation, budget deficit rate. (Vasilios Plakandaras et al., 2015; Carlos Cañizares Martínez et al, 2023; Joël Vonlanthen, 2023).
- Group 2: sectoral indicators such as the degree of dispersion of the housing market (John M. Clapp and Carmelo Giaccotto, 2002); growth of demand (Lasse Bork and Stig Vinther Møller, 2014).
- Group 3. User indicators (John M. Clapp and Carmelo Giaccotto, 2002; Et Square, 2020; Sang Won Lee, 2022; Adeboye A. Akinwunmi, 2009).

In addition, there is one group of factors considered that have an impact on housing prices is the characteristics of the product (Square et al, 2020, Margot Geerts et al, 2023). However, because Vietnam is a developing country with very diverse housing needs and diversified house types which are difficult to be standardized. Therefore, this paper is not expected to include the product characteristics in the study.

## 2.2 Forecast techniques

Regarding research methods, when conducting research on housing prices, the research has reviewed the application of many methods applicable to give the most accurate housing price results.

Particularly, home pricing in the United State which is one of the most active real estate markets, Rapach and Strauss (2007) used an autoregressive distributed lag (ARDL) model with input data of 25 determinants to forecast housing prices in counties. And the results they obtained have shown that the ARDL model tends to perform better than the AR scoring model.

Vasilios Plakandaras (2015) proposes to combine two forecasting methods, Ensemble Empirical Mode Decomposition (EEMD), in progress analysis with Support Vector Regression (SVR) methods derived from machine learning models when forecasting actual house prices in the U.S. The results of the model have improved the error rate to only half compared with result obtained from a random movement model. The model also does not generate bias when forecasting other markets excluded in the research sample.

A forecasting method is two-variable forecasting method. In conducting an analysis of house price growth in 16 economies in the Organization for Economic Cooperation and Development, N. Kundan Kishor and Hardik A. Marfatia (2017) used information from national macroeconomic indicators as well as global indicators of housing markets applying in the two-variable forecasting method. The result was much better than the conventional univariate autoregression models. The obtained results show that national income, industrial production, and the stock market are also considered as key impact factors of the changes in real house prices in the future. When conducting research on price forecasts for retail properties in 10 major Chinese cities with monthly data from 2005-2021 with Gauss regression Gauss regression method. Along with the use of 10 cores, 4 basic functions, 2 standard options through Bayesian optimization and cross-validation for each variable. Results from the RMSE model range from 0.0113% to 0.4835%, providing a tool to help investors or policymakers assess the retail real estate market.

In Vietnam, understanding factors affecting pricing of the real estate market; particularly the housing is an important research direction. Vietnam's real estate market began to change markedly after the economic reform in 1986. As foreign

investment inflows began to flow into Vietnam, there was a remarkable development in the housing market, especially high-end condominiums in the country's two largest cities, Hanoi and Ho Chi Minh City. Through the analysis of Geographic Information System (GIS) and Hedonic models for apartments in Vietnam's two largest cities, Hanoi and Ho Chi Minh City, You Seok Chung et al. (2018) found out the relationship between the geographical factors of the apartment and the price of the apartment. Thereby, the author has explained the impact of government policies along with other factors affecting house prices. The result is considered as a guideline for real estate manufacturers, and investors in making related decisions.

Another method of analysis is time series data regression which is also applied in housing price forecasts in Vietnam. Specifically, Nguyen Ngoc Vinh (2019) used the pseudo-variable regression method with 1078 observations collected in 2015 and 2016, to build the house price index in Ho Chi Minh City; Nguyen Quynh Nga (2014) researched to build a land price appraisal model for Go Vap district, Ho Chi Minh City. The results obtained from these studies show that housing prices are greatly influenced by the geographical values and properties of the house such as length, area, orientation of the house, etc.

In summary, there are many forecasting methods that have been used in forecasting housing prices in the world as AR scoring, Autoregressive distributed lag (ARDL), Ensemble Empirical Mode Decomposition (EEMD); Support Vector Regression (SVR); Two-variable forecasting; Root Mean Squared Error; Gauss regression and OLS Linear Regression.

Each method has its own conditions of application and application advantages. However, it is possible to classify forecasting methods into 2 main groups: (1) univariate forecasting and (2) multivariate forecasting.

In conclusion, the current research related to the housing price forecasting model are divided into two directions: forecasting the house price index of the economy and forecasting the house price index by each specific region with specific segments/groups of customers. According to each research direction, the composition of the forecasting model is also different. At the same time, forecasting methods are quite differentiated in diversified research contexts. Therefore, the purpose of this research project is to propose a housing forecasting model with an appropriate forecasting method for the Vietnamese whole housing market.

### **3. RESEARCH METHODOLOGY**

#### **3.1 Sample and Data Collection**

The research period is from 1990 to 2022, the beginning of the transition from a centralized economy to a state-regulated market economy. The housing market shifted from a distribution mechanism from the government to a purchase and sale mechanism according to the needs and economic conditions of households. The State encourages private ownership, financial autonomy and allows participating financial institutions to grant credit for housing purchases.

With this transformation, Vietnam's housing market officially prospered and entered a hot growth period with the increase in the emergence of real estate companies and investment projects. Thus, the market already has supply and demand operating.

Therefore, the study chose to collect data for the period from 1990 to 2022. The data collected includes data on housing prices, macroeconomic indicators, the industry's housing supply growth index and the population index from [www.euromonitor.com](http://www.euromonitor.com).

#### **3.2 Variables and forecast models**

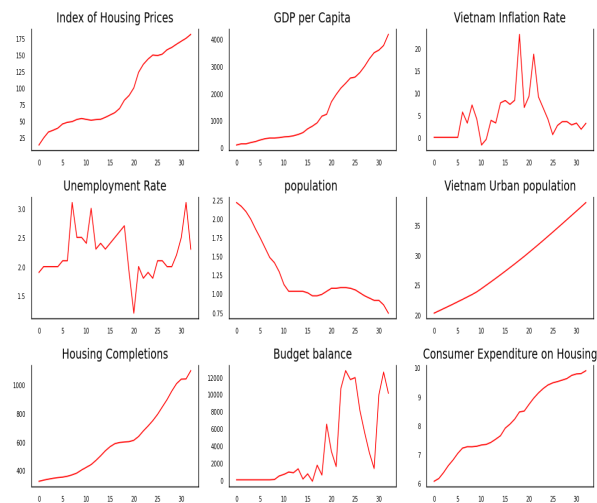
As outlined in the literature review, the study conducted a review of research related to housing price forecasts. In terms of scope of application, the valuation models are classified into two groups of models: the group of models that forecast housing prices on an economy-wide scale and the group of models that forecast housing prices of a specific type of house and on a specific market segment. As confirmed, the study focuses on building up a house price forecast model for the whole housing market. Hence, the variables considered in the study are summarized in the following table.

**Table 1.** Variables and measurements

No	Variables	Measurements
	House price	House price index
Macro variables		
	Urbanization	Urbanization population
	Population	Population growth rate
	Unemployment	Unemployment rate
	Inflation	Consumer price index
	Budget balance	Budget deficit rate
Industrial variables		
	Housing Completions	No of houses completed.
User characteristics		
	Average personal income	GDP per capita
	Housing Expenditure	Consumer expenditure on housing

Source: Authors, 2023

In this study, the data used for forecasting is time series data and to facilitate observation and analysis, the data is displayed as Fig. 2. Through that we can observe properties of the data. Data such as trends, seasons, fluctuations... From there, you can choose the appropriate pre-processing method.

**Fig. 2** Data Visualization

Source: Authors, 2023

The above factors are considered as determinants of house pricing. Hence, the study proposes a house price forecast model as mentioned in Fig. 3.

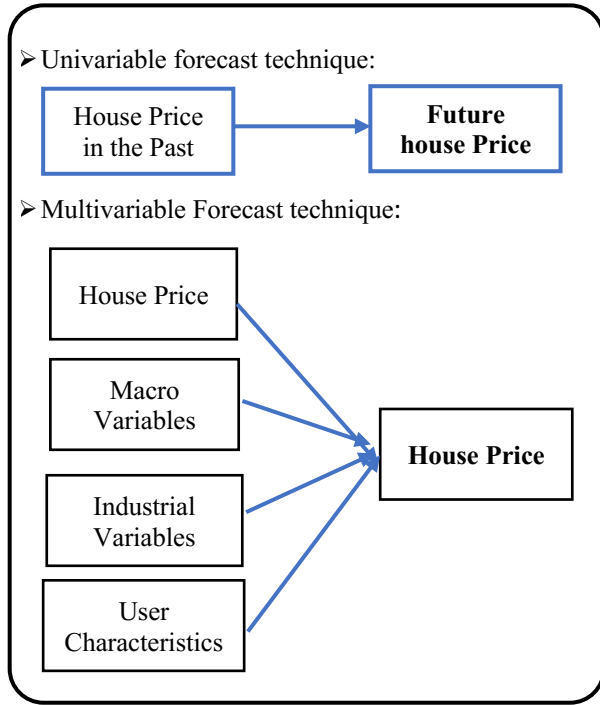


Fig. 3 Proposed house price forecasting model

Source: Authors, 2023

### 3.3 Forecast techniques.

According to the literature review, forecasting methods are classified into 2 main groups: (1) univariate forecasting and (2) multivariate forecasting.

The advantage of the univariate forecasting group is that it focuses on the trend and cyclicity of the data to forecast the forecast value in the next period. The principle of this forecasting method is that the characteristics of the forecast data are already reflected in the trend and cycle of the data, so when the right trend and cycle is identified, the forecast value will be determined. However, the disadvantage of this method is assuming other factors do not have sudden fluctuations. This is not feasible in practice.

The multivariate forecasting method has the advantage of considering the impact of other factors. However, under the strong influence of other factors, trends, and distinctive characteristics such as the periodicity of forecast variable data will be distorted.

Thus, to consider the advantages of forecasting methods, the study proposes two forecasting models with variables as in Table 1. and apply trend forecasting in combination with impact assessment of other variables, namely ARIMA and VAR.

Algorithm  $\hat{A}$  (Autoregressive Integrated Moving Average) is one of the algorithms frequently used in analyzing and forecasting time series for economic variables. ARIMA is based on the hypothesis that values in a time series correlate past values to values in the present time. ARIMA is built as a combination of three smaller models as Autoregression (AR), moving average (MA) and intergration (I) to model patterns and trends in time series data. The AR component of a model is the set of delays of the current variable. AR(p) is the model that will use the past value p of the current variable. Specifically, the AR(p) is represented as follows:

$$AR(p) = \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p}$$

Moving Average (MA) is the process of simulating the process of shifting or changing the average value of a series over time. The MA process will find a linear connection between stochastic terms. And to ensure the above property, the noise must be white noise, or the chain must be a stop chain. However, many time data series do not guarantee stationary, so the Intergrated process will bring the required chain to the form of a stop chain through proper copper taking or wrong analysis. And when combining the above 3 components, ARIMA model is formed as following:

$$Y(t) = c + \phi_1 * Y(t - 1) + \dots \phi_p * Y_{t-p} + \theta_1 * \epsilon_{t-1} + \dots + \theta_q * \epsilon_{t-q} + \epsilon_t$$

And the ARIMA model will be written as ARIMA(p,d,q), in which where:

- p is the number of elements in the autoregressive model.
- d is the number of differentiating times needed to transform the data into stationary data.
- q is the number of elements in the moving average model.

ARIMA is considered a highly flexible model because of its ability to model many types of temporal data patterns. The model is highly reliable because the ARIMA model is an econometric model that has been evaluated and tested using a variety of statistical methods. Application for economic forecasting is proved already. However, The ARIMA model is a single input multiple output model, so it is usually only suitable for forecasting variables with short fluctuations and in relatively stable conditions. But the model cannot analyse the impact of others factors on analyzed factor which leads to limitation of future scenarios' proposal.

And to overcome the above disadvantage, the study proposes an improved version of the ARIMA model, which is the Vector AutoRegression (VAR) model. VAR is a multivariate statistical model used to model the self-regression relationship between past variables and forecast their value in the future.

The VAR model is described using p latency of past dependent variables. The formula of the VAR(p) model after being modified is as follow:

$$Y(t) = c + \Phi(1) * Y(t-1) + \Phi(2) * Y(t-2) + \dots \\ + \Phi(p) * Y(t-p) + \varepsilon(t)$$

In which:

- Y(t) is the dependent variable vector in the time series at time t.
- c is the constant vector.
- $\Phi(1), \Phi(2), \dots, \Phi(p)$  are autoregressive coefficient matrices of size (k x k), where k is the number of dependent variables
- $\varepsilon(t)$  is a vector of random components with a Gaussian distribution with a mean of 0 and a constant covariance matrix.

## 4. RESULT AND DISCUSSION

### 4.1 Forecast Model 1 – ARIMA

According to the theory of the ARIMA model, to request the input data of the model must be a stationary series. To test for stationary, the Dickey Fuller Argument test, also known as the unit test, is applied. And the results returned when checking the ADF with each different differential error value are presented in the table below. The sequence is the stationary when we mis determine 3<sup>rd</sup> order differential.

**Table 2.** ARIMA and differential analysis

differential	ADF Statistic	p-Value	Critical Values (5%)
0	0.2799	0.9764	-2.98648896
1	-2.3996	0.1418	-2.9812
2	-2.1316	0.2319	-2.9922
<b>3</b>	<b>-2.9206</b>	<b>0.0430</b>	<b>-2.6252</b>

*Source: Authors, 2023*

With the formula of differential is determined by:

$$1^{\text{st}} \text{ order differential: } I(1) = \Delta(x_t) = x_t - x_{t-1}$$

$$d \text{ th order differential: } I(d) = \Delta^d(x_t) = \Delta(\dots \Delta(x_t))$$

The ARIMA model automatically helps identify pairs of p and q values with predetermined parameter d, thereby finding the ARIMA model that best fits the given data. With a data set of 33 observations from 1990 -2022. The dataset is

separated into a training set to train the model of 28 observations from 1990-2017 and the dataset from 2018-2022 which are used as test sets to evaluate the effectiveness of the model.

**Table 3.** Result Auto ARIMA with  $d=3$

Model	AIC	BIC
ARIMA (0,3,0)	164.636	165.814
ARIMA (0,3,1)	150.150	152.506
ARIMA (0,3,2)	151.679	155.214
ARIMA (1,3,0)	161.343	163.699
ARIMA (1,3,1)	151.856	155.391
ARIMA (1,3,2)	149.757	154.466
<b>ARIMA (2,3,0)</b>	<b>147.89</b>	<b>151.424</b>
ARIMA (2,3,1)	149.828	154.54
ARIMA (2,3,2)	146.872	152.762

Source: Authors, 2023

As a result, when running the Auto ARIMA model with the value  $d=3$  and the  $p,q$  values in paragraphs 0 to 3, we determine that the best model is the **ARIMA(2,3,0) model**. The result when running the model with the value ARIMA (2,3,0) returns indicators Akaike Information Criterion (AIC) =147.89 and Bayesian Information Criterion (BIC) =151.424. The two indicators AIC and BIC are measures to assess the suitability of the model, the lower this value proves the better the model. And the two AIC and BIC values are the 2 lowest values of all the models taken into consideration by ARIMA.

And when using the ARIMA model (2,3,0) to determine the test set, the predictive results are obtained and shown in the **Table 4** below.

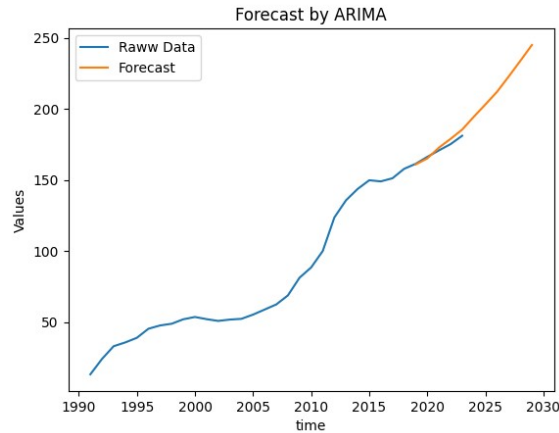
**Table 4.** The Result of ARIMA (2,3,0)

Year	Test Data	Predict Data
2018	161.3853	160.9390
2019	166.2712	165.1684
2020	170.8138	172.7794
2021	175.2215	178.9217
2022	181.1282	185.4631
2023		194.4686
2024		203.1626

Source: Authors, 2023

According to the results, house prices in Vietnam will tend to increase continuously in the coming years, with values increasing by 3-5% per year.

The forecast result is shown graphically in **Fig. 4**.



**Fig 4.** House price forecast with ARIMA model

Source: Authors, 2023

#### 4.2 Forecast Model 2 – VAR

VAR(p) is one of many Regression Models. Therefore, it is necessary to check the relationship and impact between variables before adding variables to the model. In this study, Principal Component Analysis is used to find the main variables that have a strong impact on the housing price index. PCA is a method to reduce the dimensionality of input data thereby finding the values that have the strongest impact on the predicted value. The results returned when conducting PCA analysis with the above data set are that the 4 variables that have the strongest impact on the housing price index are: GDP per Capita, Vietnam Urban Population, Housing Completions, Housing Expenditure.

Similarly, when applying the ARIMA model, the input data of the VAR model also requires stationary sequences. The results of running the ADF test of the selected variables to find appropriate differential as follows:

**Table 4.** VAR and Differential test.

Diff	0	1
GDP per Capita	ADF Stat: 3.3880 p-value: 1.0	<b>ADF Stat: -3.323</b> <b>p-value: 0.0138</b>
Vietnam Urban Population	ADF Stat: 1.5195 p-value: 0.9976	<b>ADF Stat: -3.677</b> <b>p-value: 0.0044</b>
Housing Completions	ADF Stati 1.6520 p-value: 0.9980	<b>ADF Stat: -7.768</b> <b>p-value: 9.04e-12</b>
Housing Expenditure	ADF Stat: -0.528 p-value: 0.8864	<b>ADF Stat: -3.004</b> <b>p-value: 0.0345</b>

Source: Authors, 2023

From the results we get with the House Price Index variable, it is necessary to analyze 3 order differentials. The rest of the variables will be stationary at taking the 1 order differential.

To determine the lag that best suits the model, the study determines through the VAR Order Selection algorithm with a maximum lags value of 4. The result is shown in following Table 5.

**Table 5.** VAR Order Selection

Lag	AIC	BIC	FPE	HQIC
<b>0</b>	6.556	6.798	703.6	6.226

<b>1</b>	1.958	3.411	7.401	2.377
<b>2</b>	2.388	5.050	14.47	3.158
<b>3</b>	-0.256	3.615	2.157	0.8590
<b>4</b>	-9.685*	-4.604*	0.0015*	-8.222*

Source: Authors, 2023 (\* highlights the minimums)

In short, with a delay of 4 the model will obtain the most optimal results. Therefore, the forecasting model with input vector includes 5 variables: Housing Price Index, GDP per capita, Vietnam Urban Population, Housing Completion, Housing Expenditure.

With the lag value  $p=4$ , the prediction results of the model **VAR (4)** application are shown in Table 6 below.

**Table 6.** The Result of VAR (4)

<b>Variable</b>	<b>Coeff</b>	<b>Std err</b>	<b>P &gt; z </b>
<b>intercept</b>	90.2227**	0.001	0.000
L1. Housing Prices Index	0.2653**	0.064	0.000
L1.GDP per Capita	0.0062	0.022	0.564
L1. Vietnam Urban population	-0.4779**	0.007	0.000
L1. Housing Completions	-0.1585**	0.047	0.001
L1. Housing Expenditure	-3.6980**	0.005	0.000
L2. Housing Prices Index	0.6819**	0.086	0.000
L2.GDP per Capita	-0.0147	0.017	0.397
L2. Vietnam Urban population	-1.2897**	0.008	0.000
L2. Housing Completions	0.0707	0.084	0.401
L2. Housing Expenditure	-3.5942**	0.005	0.000
L3. Housing Prices Index	0.7020**	0.031	0.000
L3.GDP per Capita	- 0.0192	0.021	0.367
L3. Vietnam Urban population	-1.2260**	0.007	0.000
L3. Housing Completions	0.1529*	0.067	0.023
L3. Housing Expenditure	-1.9525**	0.005	0.000

L4. Housing Prices Index	0.4136 **	0.084	0.000
L4.GDP per Capita	-0.0119	0.018	0.501
L4. Vietnam Urban population	-0.9631**	0.007	0.000
L4. Housing Completions	0.0790*	0.026	0.002
L4. Housing Expenditure	-1.2840**	0.004	0.000

Where (\*), (\*\*) correspond to the significance level of 0.05 and 0.01, respectively.

Source: Authors, 2023.

According to the returned results, we see that with 20 input variables, 16 variables are statistically significant at the 5% significance level. In addition, we can see that GDP per Capita with all four lags is not statistically significant in predicting the housing price index. And the result of house price forecast with VAR (4) is shown in the Table. 7.

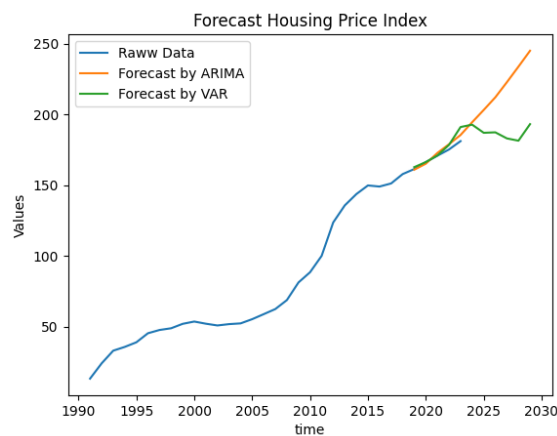
**Table 7.** VAR (4) and house price forecast result

Year	Test Data	Predict Data
2018	161.3853	162.7250
2019	166.2712	166.3270
2020	170.8138	170.9991
2021	175.2215	178.4419
2022	181.1282	191.0826
2023		192.7676
2024		187.0086

Source: Authors, 2023.

### 4.3. Proposed model for the house price forecasting

Based on forecast results obtained when applying two forecasting techniques. The results will be displayed as shown. 6.



**Fig 6.** Forecast Housing Price by proposed models.

Source: Authors, 2023.

It can be seen that the forecast results of the ARIMA (2,3,0) model tend to increase while the results of the VAR (4) model also increase but are more volatile. And to check the model better, we use to compare 2 models using 4 evaluation

indexes including: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Criterion Akaike information (AIC), Bayesian information criterion (BIC). In which the RMSE and MAE indexes help determine the level of forecast error compared to test data, and the AIC and BIC indexes are indicators that evaluate the effectiveness of the model compared to the input data.

The results obtained when conducting the analysis are returned to table 8, showing that the error results when determined by the ARIMA (2,3,0) model are lower than the VAR(4) model. However, the VAR(4) model fits the data better than the ARIMA(2,3,0) model. Therefore, we propose that the ARIMA model can be used to forecast house prices in the short term. And if you want to forecast house prices in the long term, you must use the VAR model.

**Table 7:** Test 2 models

Model	ARIMA (2,3,0)	VAR (4)
RMSE	2.7481	3.6065
MAE	1.8784	2.9028
AIC	147.89	-9.685
BIC	151.43	-4.604

Source: Authors, 2023.

#### 4. CONCLUSION

By building an ARIMA model to help forecast the Housing Price Index, the study concludes that the housing price index has a relationship that depends on the value of the housing index in the past with a 2-year lag. In addition, through error checking with test data set for 5 years from 2018-2022. The results returned when measuring error using RMSE and MAE indexes show relatively low error values. Thereby, we can use the ARIMA model to forecast the house price index and macroeconomic indicators in the short term.

The results returned when evaluating the VAR (4) model using the AIC and BIC indexes help us conclude that the VAR (4) model is highly compatible with the data set, from which we can propose a VAR model. (4) in predicting long-term housing price index. In addition, we also draw the conclusion that macroeconomic variables have a direct impact on housing index forecasts. Especially in the model, housing spending with lag index = 1 has the strongest impact on fluctuations in the housing price index.

#### 5. REFERENCES

- Adeboye A. Akinwunmi, 2009, "An Investigation Into Factors Affecting Housing Finance Supply In Emerging Economies: A Case Study Of Nigeria", Doctor dissertation, the University of Wolverhampton.
- Beltratti, A., & Morana, C. (2010). International house prices and macroeconomic fluctuations. *Journal of Banking & Finance*, 34(3), 533-545.
- Carlos Cañizares Martínez et al, 2023, Forecasting housing investment, Working Paper Series, European Central Bank, Paper No. 2807.
- Case, K. E., Quigley, J. M., & Shiller, R. J. (2005). Comparing wealth effects: the stock market versus the housing market. *Topics in Macroeconomics*, 5(1), 20121001.
- Chung, Y. S., Seo, D., & Kim, J. (2018). Price determinants and GIS analysis of the housing market in Vietnam: the cases of Ho Chi Minh City and Hanoi. *Sustainability*, 10(12), 4720.
- GSO, 2019, "20 major indicators of the 2019 population and housing census", [www.gso.gov.vn](https://www.gso.gov.vn/en/data-and-statistics/2019/12/infographic-20-major-indicators-of-the-2019-population-and-housing-census/), access link: <https://www.gso.gov.vn/en/data-and-statistics/2019/12/infographic-20-major-indicators-of-the-2019-population-and-housing-census/>
- Gupta, R., & Kabundi, A. (2010). The effect of monetary policy on house price inflation: A factor augmented vector autoregression (FAVAR) approach. *Journal of Economic Studies*, 37(6), 616-626.

- Hacievliyagil, N., Drachal, K., & Eksi, I. H. (2022). Predicting house prices using DMA method: evidence from Turkey. *Economies*, 10(3), 64.
- Heping Liu, Jing Shi, (2013), "Applying ARMA–GARCH approaches to forecasting short-term electricity prices", *Energy Economics* 37 (2013) 152–166.
- Joël Vonlanthen, 2023, Interest rates and real estate prices: a panel study", *Swiss Journal of Economics and Statistics* (2023) 159:6 <https://doi.org/10.1186/s41937-023-00111-0>
- John M. Clapp and Carmelo Giaccotto, 2002, Evaluating House Price Forecasts, *The Journal of Real Estate Research*, Taylor & Francis, Ltd., Vol. 24, No. 1 (2002), pp. 1-26 (26 pages)
- Kishor, N. K., & Marfatia, H. A. (2017). The dynamic relationship between housing prices and the macroeconomy: Evidence from OECD countries. *The Journal of Real Estate Finance and Economics*, 54, 237-268.
- Lasse Bork and Stig Vinther Møller, 2014, Forecasting house prices in the 50 states using Dynamic Model Averaging and Dynamic Model Selection, Working Paper Series of "Macroeconomic Methodology, Theory and Economic Policy (MaMTEP)", ISBN 9788791646737
- Margot Geerts et al, 2023, "A Survey of Methods and Input Data Types for House, International Journal of Geo-Information, Price Prediction", 2023, 12, 200. <https://doi.org/10.3390/ijgi12050200>
- Mohan Munusamy et al, 2015, "An Overview of the Forecasting Methods Used in Real Estate Housing Price Modelling", *Journal of Technology*, Penerbit UTM Press.
- Nguyen Ngoc Vinh, Building a housing price index: Empirical study in Ho Chi Minh City, *Journal of Finance* Vol 2 November 2019.
- Plakandaras, V., Gupta, R., Gogas, P., & Papadimitriou, T. (2015). Forecasting the US real house price index. *Economic Modelling*, 45, 259-267.
- Rapach, D. E., & Strauss, J. K. (2007). Forecasting real housing price growth in the eighth district states. *Federal Reserve Bank of St. Louis. Regional Economic Development*, 3(2), 33-42.
- Sang Won Lee, 2022, "A comparative study on determinants of housing mortgage prepayment of individual borrowers", *Journal of Derivatives and Quantitative Studies*, Vol. 30 No. 4, 2022, pp. 278-295, Emerald Publishing Limited, e-ISSN: 2713-6647, p-ISSN: 1229-988X. DOI 10.1108/JDQS-05-2022-0013
- Savil, 2023, "Q1 2023 HCMC Residential Market Report", access link: "[HCMC-Residential-MRQ12023-EN.pdf](https://www.savills.com.vn/HCMC-Residential-MRQ12023-EN.pdf) ([savills.com.vn](https://www.savills.com.vn))
- Uyar, K. S. G., & Yayla, N. (2016). Konut fiyatlarının hedonik fiyatlama yaklaşımına göre mekânsal ekonometrik modeller ile tahmini: İstanbul konut piyasası örneği. *Social Sciences (NWSASOS)*, 3C0153, 11(4), 326-342.
- Vasilios Plakandaras et al (2014), "Forecasting the U.S. real house price index", *Economic Modelling* 45 (2015) 259–267.
- Vasilios Plakandaras et al, 2015, "Forecasting the U.S. Real House Price Index", *Economic Modelling* Volume 45, February 2015, Pages 259-267; <https://doi.org/10.1016/j.econmod.2014.10.050>
- Xu, X., & Zhang, Y. (2023). A Gaussian process regression machine learning model for forecasting retail property prices with Bayesian optimizations and cross-validation. *Decision Analytics Journal*, 100267.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

