Application of grey system theory and ARIMA model to forecast factors of tourism: A case of Binh Thuan Province in Vietnam

Nhu-Ty Nguyen 1, Bao-Phuong-Uyen Nguyen 1, Thanh-Tuyen Tran 2, *

1School of Business, International University - VNU-HCMC, Quarter 6, Linh Trung Ward, Thu Duc District, Ho Chi Minh City, Vietnam
2Scientific Research Center, Lac Hong University, No.10 Huynh Van Nghe Street, Dong Nai Province, Vietnam

Abstract
Tourism is becoming more and more popular, and this industry continues to develop strongly around the world. Thus, forecasting tourism demand plays an important role in development. In this study, the purpose is to provide some appropriate models for predicting the demand for tourism in Binh Thuan Province in Vietnam. There are five models applied in this study, namely GM (1, 1), DGM (1, 1), DGM (2, 1), Verhulst and ARIMA; the authors try to test these models to find which concise and accurate forecasting models being able to predict the best result about the tourism demand. So as to ensure the precision, the authors collected data of total revenue, domestic visitor, international tourists and top six countries having the biggest numbers of visitors (Russia, Germany, France, Korea, China and USA) in ten years (between 2008 to 2017) from Binh Thuan Department of Culture, Sports and Tourism. We apply MAPE, MSE, RMSE, and MAD to compare the forecasting model results. As a result, GM (1, 1), DGM (1, 1), Verhulst and ARIMA augment excellent results and minimum forecasted errors. In terms of total revenue, ARIMA is the best choice for prediction. About the domestic visitors and international tourists, GM (1, 1), DGM (1, 1) and Verhulst give better calculation than the other models. Besides, the performance of GM (1, 1), DGM (1, 1), Verhulst and ARIMA to forecast the number of visitors of the top six markets (Russia, Germany, France, Korea, China, and the USA) sending the largest number of tourists describes good results. For all the factors, DGM (2, 1) is rejected to predict due to the poor results. Moreover, recently, tourism industry has developed rapidly in Binh Thuan. Hence, the government has to propose suitable policies to develop local tourism industry.

Keywords:
Tourism
Forecasting
Vietnam
Grey models
ARIMA

1. Introduction
Since the late 1980s, thanks to the policy of reform and opening up of the state, tourism in Vietnam in general and Binh Thuan, in particular, has developed strongly and gained much success. Located in the South Central and Southern tourism area, Binh Thuan province owns strength in tourism potential. In recent years, the number of tourists traveling to Binh Thuan has increased rapidly, so that this "industry without a chimney" more and more contributes to the growth of the local economy.

* Corresponding Author:
Email Address: thanhtuyentran@hu.edu.vn (T. T. Tran)
https://doi.org/10.21833/ijaas.2020.01.009

© 2019 The Authors. Published by IASE. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

According to the annual report of BINH THUAN DEPARTMENT OF CULTURE, SPORTS, AND TOURISM, in the first six months of 2017, Binh Thuan province received about 2,300,000 travelers, reaching 45.87% of the plan, up 9.6% over the same period in 2016. Meanwhile, international arrivals reached about 295,000 people, up 17% over the same period in 2016; for example, The Korean market rocketed to 57.3%, China climbed to 27.8%, Russia increased by 18%, Thailand went up 4.1%, etc. On the other hand, some markets were having a downward trend significantly; for instance, Australia declined 17.5%, the USA fell down 15%, Germany reduced 11.5%; France dropped about 5%, etc. in the first six months in 2017. Additionally, both the number of domestic visitors and foreign sightseers has a growth yearly between 2008 and 2017 (described as Figs. 1 and 2). During, the total revenue from Binh Thuan tourism reached 10,812 billion VND, up approximately 20% in 2017 (Fig. 3).
Binh Thuan had stably maintained a constant innovation and improvement for the province's tourism over a ten-year period (2008-2017) which has the increasing figures of tourism indicators. Furthermore, the top six countries having the most outstanding visitors to Vietnam are indicated in Fig. 4. It can be seen that Russia is always the first top nation providing travelers to Binh Thuan province, but this proportion was equal in 2016 and overtaken in 2017 by the Chinese market; the others following are Germany, Korea, France, and the USA respectively.

In some journals, for instance, Song and Li (2008) stated that tourism demand forecasting scientists accumulate data from governments or other agencies. Besides, in a study of two Vietnamese researchers, Nguyen and Tran (2019) had to collect data from the Vietnamese Ministry of Tourism. It can be seen that conducting the research is compulsory to have all necessary figures, such as numbers of domestic visitors or also foreign arrivals in a nation and location, also tourist expenditure. In this study, the writer collected data from Binh Thuan Department Culture, Sports and Tourism.

Researchers apply different methods to analyze the forecasting tourism demand; there are some usual models, namely time-series model (such as GARCH), econometric model (such as ECM and VAR),
SES model, logistic growth model, neural network, etc. Also, combination methods are considered. According to Nguyen and Tran (2019), the correct approaches are dependent on determinants and separates into a month, quarter or annual demand.

Nguyen and Tran (2019) found that tourism demand forecasting supports the nation to catch the number of domestic visitors, also international arrivals, total revenue tourism; thus, that is the data that help to propose appropriate policies. The quantitative method is a common technique being applied to forecasting tourism demand.

Almost the previous papers, Time-series models namely ARIMA and GARCH (Condratov and Stanciu, 2012; Hadavandi et al., 2011; Radha and Thenmozhi, 2006) and econometric models viz. error correction model (ECM) and the vector autoregressive (VAR) models (Song and Witt, 2006) have been popular models using tourism demand forecasting techniques. Besides, Chang and Liao (2010) used a SARIMA model to forecast monthly outbound Taiwanese tourists traveling to Hong Kong, Japan, and the USA. Furthermore, Lin and Lee (2013) indicated econometric models adopting Multivariate Adaptive Regression Splines (MARS), Artificial Neural Network (ANN) and Support Vector Regression (SVR) to forecast monthly total arrivals visiting Taiwan.

Huang (2012) researched to find out the appropriate model improving the ability to forecast the demand for health tourism in Asian nations using a GM (1, 1). Nhu Ty Nguyen used Grey System Theory to test the concise models being able to predict the number of visitors in Vietnam. Otherwise, ARIMA illustrated better forecasting performance to predict the international tourism demand from four European nations to Seychelles (Du Preez and Witt, 2003).

The researchers have to apply the most appropriate model to obtain the best forecasting achievement because forecasting is one of the important factors affecting directly policy and decision-making in the future. In this study, the authors put models GM (1, 1), Verhulst, DGM (1, 1), DGM (2, 1) and ARIMA into practice. The goal of using these models is to check which models supervise the best appropriate forecasting the situation of Binh Thuan province’s tourism demand.

2. Data collection and description

The research analyzes four determinants to do the forecasting – a total number of domestic visitors, international arrivals, total revenue and six countries providing the most tourists to travel to Binh Thuan (Russia, China, Germany, Korea, France, and the USA).

We collect data between 2008 and 2017 that are gotten from Binh Thuan Department Culture, Sports and Tourism and Statistics Office of Binh Thuan.

The data composes of Total Revenue Index, Domestic Arrivals, International Tourists and Top Six Countries giving Visitors, etc. (Figs. 1, 2, 3 and 4).

In terms of the number of arrivals, we also obtain 4 variables datasets. They consist of reference sources for a decision, purposes of visiting, and forms of trip and means of transportation. In the context of Binh Thuan, the group reference sources for decision (Fig. 5) answers the question “why visitors decide to arrive in Binh Thuan province”, they are recommended by others who have ever gone to Binh Thuan. About the purposes of visiting (Fig. 6), this group wonders the free time, economic and social conditions, etc. Moreover, visitors also consider forms of the trip (Fig. 7) which makes them save much more money for their tours. Besides, the variable-means of transportation indicate that tourists choose transportation which is the most convenient choice for them (Fig. 8).

![Fig. 5: Reference sources for decision](image1)

![Fig. 6: Purposes of visiting](image2)

Table 1 is shown the descriptive statistics of the number of visitors arriving in Binh Thuan. The mean
of total revenue index, the number of domestic arrivals and the number of international visitors are 227.74, 3.007E6 and 366380, respectively. The top six countries include Russia, Germany, France, Korea, China, and the USA which are presented 104629.5, 31443.5, 15377.5, 25375.1, 50352.8 and 15166.3, respectively. It can be seen that Russia is the biggest market giving tourists to Binh Thuan.

Table 1: Descriptive statistics

|                          | Mean   | Minimum | Maximum | Std. Deviation (n-1) | Variance (n-1) |
|--------------------------|--------|---------|---------|----------------------|----------------|
| Total Revenue            | 227.74 | 61.15   | 464.24  | 135.81               | 1844E10        |
| Domestic Visitors        | 3.007E6| 1.806E6 | 4.542E6 | 90.025               | 8.137E11       |
| International Visitors   | 366380 | 195156  | 590636  | 1392659.93           | 1.697E10       |
| Russia                   | 104629.5| 29760   | 152855  | 40144.28             | 1.612E9        |
| Germany                  | 31443.5| 26743   | 34846   | 2591.27              | 6.715E6        |
| France                   | 15377.5| 13012   | 17835   | 1615.88              | 2.611E6        |
| Korea                    | 25375.1| 12466   | 66506   | 17138.65             | 2.937E8        |
| China                    | 50352.8| 4453    | 154274  | 51785.29             | 2.682E9        |
| USA                      | 15166.3| 13230   | 18215   | 1663.51              | 2.767E6        |

Note: Total Revenue by Million USD

![Fig. 7: Forms of trip](image_url)

3. Data analysis and result

The exact information and data sets influence significantly the accuracy of the forecasting process. In this paper, the data were collected from Binh Thuan Department Culture, Sports and Tourism and Statistics Office of Binh Thuan over a period of ten years (2008-2017) and absolutely, these data sets were never revised. It is easy to see that the tourism demand in Binh Thuan had an upward trend during the surveyed years.

In this portion, we use the data gathered from 2008 to 2017 to apply GM (1, 1), DGM (1, 1), DGM (2, 1), Verhulst and ARIMA to test the accuracy level of forecasting the demand of tourism in Binh Thuan:

- **GM (1, 1):**
  - a=0.1925; b=81611824.1276 and \((1-e^a)(x^{(0)}(1)-\frac{b}{a}) = 84943934.5560\) are calculated the Total Revenue.
  - The results of parameters connecting to the Domestic Visitors are \(a=-0.0983; b=1775695.8471(1-e^a)(x^{(0)}(1)-\frac{b}{a}) = 1860228.9725\)
  - The results of parameters connecting to the International Visitors are \(a=-0.1163; b=195467.1875(1-e^a)(x^{(0)}(1)-\frac{b}{a}) = 205954.3320\) are transmitted to the calculation of international visitors.
  - The results of parameters analyzing German visitors are \(a=-0.0732; b=78101.4692(1-e^a)(x^{(0)}(1)-\frac{b}{a}) = 77412.2129\) are analyzed the calculation of Russian visitors.
  - The results of parameters analyzing German visitors are \(a=-0.0132; b=14048.0298(1-e^a)(x^{(0)}(1)-\frac{b}{a}) = 14183.2281\) are calculated French visitors.
  - The results of parameters analyzing Chinese visitors are \(a=-0.2606; b=1277.9339(1-e^a)(x^{(0)}(1)-\frac{b}{a}) = 4645.7411\) are analyzed the calculation of Korean visitors.
  - The results of parameters analyzing Chinese visitors are \(a=-0.3625; b=6317.4331(1-e^a)(x^{(0)}(1)-\frac{b}{a}) = 6653.0704\) are related to the calculation of Chinese tourists.

- **DGM (1, 1) and DGM (2, 1):**
  - Total Revenue’s calculator is: \(\beta_1 = 1.2127; \beta_2 = 90475443.0508\) and \(x^{(0)}(1)(\beta_1 - 1) + \beta_2 = 103482866.9976\)
  - Calculation of Domestic Visitors: \(\beta_1 = 1.1033; \beta_2 = 1867795.7312\), so the equation \(x^{(0)}(1)(\beta_1 - 1) + \beta_2 = 2054349.8646\)
  - Calculation of International Visitors: \(\beta_1 = 1.1234; \beta_2 = 207602.7531\), so the equation \(x^{(0)}(1)(\beta_1 - 1) + \beta_2 = 231688.4552\)
  - With the same section, Russian visitors is calculated: \(\beta_1 = 1.0736; \beta_2 = 82069.6932\), so the equation \(x^{(0)}(1)(\beta_1 - 1) + \beta_2 = 84259.3222\)
  - We analyzed factor-German visitors: \(\beta_1 = 1.0047; \beta_2 = 31241.5299\), so the equation \(x^{(0)}(1)(\beta_1 - 1) + \beta_2 = 31367.6837\)
French visitors factor’s calculator is: $\beta_1 = 1.0127; \beta_2 = 14185.9725\text{and }x^{(1)}(1)(\beta_1 - 1) + \beta_2 = 14406.7166$

Calculation of Korean visitors is with the following parameters: $\beta_1 = 1.2955; \beta_2 = 1807.2379$, so the equation $x^{(0)}(1)(\beta_1 - 1) + \beta_2 = 1807.2379$

Similarly, we calculate Chinese Visitors: $\beta_1 = 1.4407; \beta_2 = 7937.3837$, so the equation $x^{(0)}(1)(\beta_1 - 1) + \beta_2 = 9899.6954$

Lastly, USA visitors factor’s calculator: $\beta_1 = 1.0124; \beta_2 = 14059.8029$ and $x^{(0)}(1)(\beta_1 - 1) + \beta_2 = 14267.5227$

- **Verhulst:**

Verhulst calculator of Total Revenues: $a=-0.2656$; $b=0$

$$\tilde{x}(k+1) = \frac{ax^{(1)}(0)}{bx^{(1)}(0)+[a-bx^{(1)}(0)]e^{ak}} \quad \text{(Ver. 7 mentioned in section 2)}$$

Verhulst calculator of Domestic Visitors: $a=-0.0961; b=0$ and Eq. Ver. 7 with $ax^{(1)}(0) = -173540.4603; a - bx^{(1)}(0) = -0.0990$; and $bx^{(1)}(0) = 0.0029$

International Visitors-factor’s calculation: $a=0.0927; b=0$ and Eq. Ver. 7 with $ax^{(1)}(0) = 18098.4270; a - bx^{(1)}(0) = -0.0147$; and $bx^{(1)}(0) = 0.1074$

Russian visitors: $a=-0.6932; b=0$ and $\tilde{x}(k+1) = \frac{ax^{(1)}(0)}{bx^{(1)}(0)+[a-bx^{(1)}(0)]e^{ak}} \quad \text{(Ver. 7 mentioned in section 2)}$

German visitors: $a=-0.4705; b=0$ and Eq. Ver. 7 with $ax^{(1)}(0) = -12581.4170; a - bx^{(1)}(0) = -0.0849$; and $bx^{(1)}(0) = -0.3856$

French visitors: $a=-0.4705; b=0$ and Eq. Ver. 7 with $ax^{(1)}(0) = -12581.4170; a - bx^{(1)}(0) = -0.0849$; and $bx^{(1)}(0) = -0.3856$

Korean visitors- factor’s calculation: $a=0.0285; b=0$ and Eq. Ver. 7 with $ax^{(1)}(0) = 437.1976; a - bx^{(1)}(0) = -0.1098$; and $bx^{(1)}(0) = 0.1382$

Verhulst’ calculator of Chinese visitors: $a= -0.5224; b=0$ and Eq. Ver. 7 with $ax^{(1)}(0) = -2326.0560; a - bx^{(1)}(0) = -0.5143$; and $bx^{(1)}(0) = -0.0080$

Finally, USA visitors: $a=0.1239; b=0$ and Eq. Ver. 7 with $ax^{(1)}(0) = 2067.3654; a - bx^{(1)}(0) = 0.0164$; and $bx^{(1)}(0) = 0.1075$

- **ARIMA**

The model parameters of Total Revenue: $p=0$; $q=1; P=0; D=0; Q=0$ and $s=0$ with the confidence intervals being 95%

The model parameters of International Arrivals: $p=0$; $d=1; q=1; P=0; D=0; Q=0$ and $s=0$ with the confidence intervals being 95%

The model parameters of Russian Visitors: $p=0$; $d=1; q=1; P=0; D=0; Q=0$ and $s=0$ with the confidence intervals being 95%

The model parameters of German Visitors: $p=0$; $d=1; q=1; P=0; D=0; Q=0$ and $s=0$ with the confidence intervals being 95%

The parameters of French Visitors: $p=0$; $d=1; q=1; P=0; D=0; Q=0$ and $s=0$ with the confidence intervals being 95%

The parameters of Korean Visitors: $p=0$; $d=1; q=1; P=0; D=0; Q=0$ and $s=0$ with the confidence intervals being 95%

The parameters of Chinese Visitors: $p=0$; $d=1; q=1; P=0; D=0; Q=0$ and $s=0$ with the confidence intervals being 95%

The parameters of USA Visitors: $p=0$; $d=1; q=1; P=0; D=0; Q=0$ and $s=0$ with the confidence intervals being 95%

**Table 2** gives data on actual values, GM (1,1), DGM (1,1), DGM (2,1), Verhulst and ARIMA over a ten-year period (from 2008 to 2017) and the results of forecasting are used by five above models (GM (1,1), DGM (1,1), DGM (2,1), Verhulst and ARIMA) over a period of 5 constant years (2018-2022) for total tourism revenue.

**Table 3** describes the realistic numbers of domestic tourists and international visitors with the numbers of GM (1,1), DGM (1,1), DGM (2,1), Verhulst and ARIMA in ten years between 2008 and 2017 and forecasting results in the next five years (2018-2022).
Table 2: The true values and forecasting result for tourism revenue

| STAGES | Models | Actual (Million USD) | DGM(1,1) | DGM(2,1) | Verhulst | ARIMA |
|--------|--------|----------------------|-----------|-----------|----------|--------|
| 2009   | 81187724 | 61145942             | 61145942  | 61145942  | 61145942 | 61145942 |
| 2010   | 109016180| 126827266            | 126827266 | 126827266 | 126827266 | 126827266 |
| 2011   | 145529624| 151230811            | 151230811 | 151230811 | 151230811 | 151230811 |
| 2012   | 18771497 | 184336820            | 184336820 | 184336820 | 184336820 | 184336820 |
| 2013   | 235861358| 222369341            | 222369341 | 222369341 | 222369341 | 222369341 |
| 2014   | 276729984| 264960707            | 264960707 | 264960707 | 264960707 | 264960707 |
| 2015   | 328138939| 32777518             | 32777518  | 32777518  | 32777518  | 32777518  |
| 2016   | 386901550| 396012716            | 396012716 | 396012716 | 396012716 | 396012716 |
| 2017   | 464240111| 480207848            | 480207848 | 480207848 | 480207848 | 480207848 |

Table 3: The true values and forecasting result for domestic visitors and international Tourists

| STAGES | Models | Actual (Million USD) | DGM(1,1) | DGM(2,1) | Verhulst | ARIMA |
|--------|--------|----------------------|-----------|-----------|----------|--------|
| 2009   | 1808535 | 1805535              | 1805535   | 1805535   | 1805535 | 1805535 |
| 2010   | 1978463 | 2052323              | 2054350   | 2054350   | 2054350 | 2054350 |
| 2011   | 2229881 | 2264253              | 2266612   | 2266612   | 2266612 | 2266612 |
| 2012   | 2502338 | 2498069              | 2508087   | 2508087   | 2508087 | 2508087 |
| 2013   | 2880088 | 2756028              | 2759199   | 2868966   | 2868966 | 2868966 |
| 2014   | 3144785 | 3040626              | 3044429   | 3044429   | 3044429 | 3044429 |
| 2015   | 3534629 | 3570137              | 3596320   | 3596320   | 3596320 | 3596320 |
| 2016   | 3901804 | 3820103              | 3840286   | 3840286   | 3840286 | 3840286 |
| 2017   | 4515452 | 4504849              | 4511254   | 4511254   | 4511254 | 4511254 |
| 2018   | 4993517 | 4975723              | 4991155   | 4991155   | 4991155 | 4991155 |
| 2019   | 5482560 | 5491651              | 5498073   | 5498073   | 5498073 | 5498073 |

Table 4: presents information about both the figures of tourists from Russia and Germany markets with actual, GM(1,1), DGM(1,1), DGM(2,1), Verhulst and ARIMA starting from 2008 till 2017 and the forecasted numbers in five years from 2018 to 2022. However, the model DGM(2,1) can not run to forecast the proportion of the German Market because the number sequence from 2011 to 2017 is negative totally; thus, they are errors.

Table 5: demonstrates the quantities of French and Korean travelers beginning from 2008 till 2017 by actual, GM(1,1), DGM(1,1), DGM(2,1), Verhulst and ARIMA; it also shows the forecasting consequences in next five years. Meanwhile, the proportion of the Russian market is not applied Verhulst model to forecast in 2011 so it is an error.
Table 6 shows data about China and USA markets in the period (2008-2017) with real numbers and figures of GM (1,1), DGM (1,1), DGM (2,1), Verhulst and ARIMA and the consequences of forecasting in constant five years later (2018-2022).

As can be seen in Fig. 9 that all factors namely actual total revenue, GM (1,1), DGM (1,1), DGM (2,1), Verhulst and ARIMA increased in ten years regularly from about 61,145,941.62 million USD to approximately 500,000,000 million USD.

![Forecasting result of tourism revenue](image)

Table 6: The results of China and USA markets

| STAGES | Models         | Actual | GM(1,1) | DGM(1,1) | DGM(2,1) | Verhulst | ARIMA | Actual | GM(1,1) | DGM(1,1) | DGM(2,1) | Verhulst | ARIMA |
|--------|----------------|--------|---------|----------|----------|----------|-------|--------|---------|----------|----------|----------|-------|
| 2008   | 4458           | 4453   | 4453    | 4453     | 4453     | 4453     | 16688 | 16688  | 16688   | 16688    | 16688    | 16688    | 16688 |
| 2009   | 4547           | 9560   | 9900    | 5915     | 7429     | 19874    | 13565 | 14225  | 14268   | 17313    | 16402    | 16325    |
| 2010   | 18846          | 13736  | 14262   | 9441     | 12310    | 17732    | 14935 | 14412  | 14445   | 19497    | 16098    | 13334    |
| 2011   | 13203          | 19737  | 20547   | 14078    | 20168    | 24485    | 13593 | 14602  | 14625   | 24290    | 15758    | 14501    |
| 2012   | 27607          | 28361  | 29602   | 20179    | 32455    | 24914    | 14492 | 14794  | 14807   | 34813    | 15383    | 13276    |
| 2013   | 45074          | 40751  | 42646   | 28204    | 50821    | 45510    | 14404 | 14909  | 14991   | 57910    | 14988    | 14076    |
| 2014   | 42013          | 58556  | 61439   | 38761    | 76496    | 61532    | 16970 | 15187  | 15178   | 108611   | 14564    | 14029    |
| 2015   | 78750          | 84138  | 88514   | 52648    | 109222   | 50092    | 15821 | 15387  | 15367   | 219904   | 14112    | 16473    |
| 2016   | 120711         | 120988 | 127520  | 70916    | 146364   | 107984   | 15571 | 15589  | 15558   | 464294   | 13634    | 17774    |
| 2017   | 154274         | 173718 | 183714  | 94947    | 183344   | 142949   | 13230 | 15794  | 15752   | 1008648  | 13130    | 15314    |
| 2018   | 249614         | 264672 | 168142  | 240883   | 159247   | 16002    | 15948 | 4761589| 12055   | 13327    |
| 2019   | 358670         | 381305 | 222884  | 258763   | 16213    | 16146   | 10433662| 11498  |
| 2020   | 513572         | 549335 | 294802  | 278710   | 16427    | 16347   | 22884834| 10911  |
| 2021   | 740536         | 791412 | 389459  | 278332   | 16643    | 16551   | 50215137| 10322  |
| 2022   | 1064073        | 1140165| 513977  | 283859   | 16862    | 16757   | 110208790| 9728   |

Fig. 12 mentions the actual and ARIMA model quantities of Russian citizens had a growth inconstantly; in detail, the actual number is declined from 152,855 to 116,086 in 2016 and ARIMA is similar to the actual which reduce 38,183 people (from 165,657 to 127,474). The others kept going their increases in ten years.

Fig. 13 summarizes that the DGM (2,1) model gives negative numbers so they are errors; the actual number of Germany market fluctuated over the entire period shown and the other models climbed slowly during the surveyed period.

France market which is described in Fig. 14 provides that only DGM (2,1) had an upward tendency and it details that the number increased from 17,323 to 2,554,016. Besides, all the lines of the others wavered in different years of the period.

Fig. 15 outlines that there was an upward trend in DGM (2,1) which shows that the number of Korean visitors upsurged from 15,349 to 76,121 (2008-2017). Verhulst gives the result in 2017 being an error. The others such as actual, GM (1,1), DGM (1,1) and ARIMA had the oscillations in ten years.

In terms of China market (Fig. 16), it can be seen that actual, GM (1,1), DGM (1,1), DGM (2,1) and Verhulst are the factors that had the rocketed tendency. Notwithstanding, the number of ARIMA dropped between 2009 and 2010 (reducing from 19,874 to 17,732).

According to Fig. 17, USA market is observed that there is only DGM (2,1) having growth gradually during the examined years. Otherwise, all of the others (actual, GM (1,1), DGM (1,1), Verhulst and ARIMA) palpitated from 2008 to 2017.

Domestic tourism is the factor which had an upward trend year by year (Fig. 10). It is clear that these numbers went up from 1,805,535 in 2008 to more than 4,500,000 in 2017 for all models.

Similarly, Fig. 11 represents the proportions of international visitors of all models rose constantly during the examined years from 195,156 to more than 590,000.
Fig. 10: Forecasting result of domestic visitors

Fig. 11: Forecasting result of international arrivals

Fig. 12: Forecasting result of Russia visitors
Fig. 13: Forecasting result of Germany visitors

Fig. 14: Forecasting result of France visitors

Fig. 15: Forecasting result of Korea visitors
3.1. Analyzing the ability of forecasting models by MAPE, MSE, RMSE and MAD methods

It is well-known that a variety of methods is used to evaluate the accuracy of forecasting models. First, MAPE (Mean Absolute Percentage Error) is applied as a proportion of merit to recognize whether a data mining method is showing well or not. The MAPE is lower, the data mining method is better performance:

\[ \text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{\text{Actual} - \text{Forecast}}{\text{Actual}} \right) \times 100; \]

n: forecasting number of step.

Meanwhile, the evaluation follows to these results:

- MAPE < 10% => Excellent
- 10% < MAPE < 20% => Good
- 20% < MAPE < 50% => Reasonable
- MAPE > 50% => Poor

Next, the Mean Squared Error (MSE) summarizes the way a regression line is next to a set of points. The distances from the points to the regression line are the errors and then square them. It is estimated by squaring the MAD:

\[ \text{MSE} = \frac{1}{h+1} \sum_{t=2}^{h} \hat{y}_{t-1} - y_t \right)^2 \]

Root Mean Square Error (RMSE) is the standard deviation of the residuals (prediction errors). RMSE is usually used in forecasting. The smaller errors, the more exact the ability to forecast.

\[ \text{RMSE} = \sqrt{\frac{1}{h+1} \sum_{t=2}^{h} \left( \hat{y}_{t-1} - y_t \right)^2} \]

The last is Mean Absolute Deviation (MAD) is the average distance between actual data sets and forecasted data sets. The forecasting model is more accurate when the MAD’s value is lower.

\[ \text{MAD} = \sum_{i=1}^{n} |e_i| \]

Table 7 indicates the efficiency of five models GM (1, 1), DGM (1, 1), DGM (2, 1), Verhulst and ARIMA to
forecast tourism revenue. It is clear that GM (1, 1), DGM (1, 1) and ARIMA are good to forecast total revenue with MAPES being lower than 10% and MSE, RMSE, and MAD also being low. Verhulst is only reasonable in the process. According to the results, the evaluation of DGM (2, 1) is poor, so it is chosen.

Table 8 presents a similar method because the parameter of MAPE, MSE, RMSE, and MAD are lower than 10%, the performance of GM (1, 1), DGM (1, 1) Verhulst and ARIMA are good to do the forecasting; therefore, they are efficient models for this process. DGM (2, 1) shows a poor calculation, so it is not chosen to forecast this factor.

Table 9 illustrates the same method, GM (1, 1), DGM (1, 1), Verhulst and ARIMA are also the most appropriate models since the parameter of MAPE, MSE, RMSE, and MAD are lower than 10%. Also, DGM (2, 1) is rejected to forecast international visitors.

Table 10 also applies the same method, by contrast, Table 9, Verhulst has an excellent evaluation with low MAPE, MSE, RMSE, and MAD (lower than 10%) and it is chosen for forecasting. GM (1, 1), DGM (1, 1), and ARIMA are also useful in this section with low MAPE, MSE, RMSE, and MAD. DGM (2, 1) is not accepted for forecasting.

Table 11 compares the above five models, there are four good models in this situation, viz. GM (1, 1), DGM (1, 1), Verhulst and ARIMA; all of them are accepted to forecast Germany Visitors with MAPE, MSE, RMSE, and MAD are low. Only DGM (2, 1) is rejected with poor results.

Table 12 describes the same method, it is obvious that GM (1, 1), DGM (1, 1), Verhulst and ARIMA have low MAPE, MSE, RMSE and MAD (lower 10%), so they are allowed because they give the most accurate results. With the poor calculation, DGM (2, 1) is not accepted for the prediction.

Table 13 outlines a similar method, DGM (1, 1) and ARIMA are accepted to forecast this situation thanks to good calculation MAPE, MSE, RMSE, and MAD. GM (1, 1) and DGM (2, 1) obtain reasonable level. With a high parameter of MAPE, MSE, RMSE, and MAD, Verhulst is not chosen for forecasting.
Similarly, Table 14 represents only GM (1, 1) is a good calculation with MAPE, MSE, RMSE, and MAD accepted. DGM (1, 1) belongs to a reasonable level. Besides, there are three models evaluated that they are poor, so they are rejected in this section.

Finally, Table 15 gives information on the ability to forecast USA Visitor. It can be seen that GM (1, 1) and DGM (1, 1) are chosen as excellent results and accurate calculation with low MAPE, MSE, RMSE, and MAD (lower 10%). The models summarizing the good results are Verhulst and ARIMA, so they are accepted. Notwithstanding, DGM (2, 1) is rejected with a poor calculation for forecasting.

| Table 13: Evaluating models with Korea visitors forecasting errors |
|-------------------|-------------------|-------------------|-------------------|-------------------|
| Models | GM (1, 1) | DGM (1, 1) | DGM (2, 1) | VERHULST | ARIMA |
| MAPE | 20.99% | 18.8% | 22.18% | 349.21% | 17.9% |
| MSE | 47155485.1 | 3736671.5 | 4197123.9 | 117201E+11 | 37127144.09 |
| RMSE | 68669.985 | 6112.788 | 64785.528 | 342345.863 | 6093.205 |
| MAD | 4895.1 | 4281.9 | 5431.5 | 14051.7 | 4289.02 |
| Evaluation | Reasonable | Good | Reasonable | Poor | Good |

| Table 14: Evaluating models with China visitors forecasting errors |
|-------------------|-------------------|-------------------|-------------------|-------------------|
| Models | GM (1, 1) | DGM (1, 1) | DGM (2, 1) | VERHULST | ARIMA |
| MAPE | 19.33% | 23.41% | 214.3% | 27.07% | 46.41% |
| MSE | 76011286.8 | 147230546.3 | 703478132 | 37287367.5 | 185252000 |
| RMSE | 87184.45 | 1213.860 | 26523.162 | 19308479 | 13610.73 |
| MAD | 5902.6 | 897.25 | 16573.6 | 13953.3 | 10690.2 |
| Evaluation | Good | Reasonable | Poor | Poor | Poor |

| Table 15: Evaluating models with USA visitors forecasting errors |
|-------------------|-------------------|-------------------|-------------------|-------------------|
| Models | GM (1, 1) | DGM (1, 1) | DGM (2, 1) | VERHULST | ARIMA |
| MAPE | 6.75% | 6.76% | 1256.97% | 10.46% | 10.52% |
| MSE | 19911732.2 | 1992614.8 | 122744E+11 | 4155559.4 | 334749.94 |
| RMSE | 14118.01 | 1411.60 | 350348.678 | 20385.19 | 1829.60 |
| MAD | 1027.2 | 1030.2 | 181203.5 | 1617 | 1578.4 |
| Evaluation | Excellent | Excellent | Poor | Good | Good |

4. Conclusion and discussion

Tourism is defined as an important integrated economic sector with the content of deep culture, interdisciplinary fields, and socialization. Developing tourism means that we respond to the needs of domestic citizens and international tourists for sightseeing, recreation, and relaxation which contribute to improving the intellectual standards of the people, job creation and socio-economic development. Moreover, this topic supports to study the current trend of tourism and proposes the best solutions for the long-term period of the local tourism industry. Tourism is the strongest developing industry all over the world and it also plays a significant role in economic growth (Akama and Kieti, 2007; Cortez, 2008). Vietnam is one of the nations in top of Asian area having developed tourism market, so Binh Thuan—the one of the provinces in Vietnam consider that tourism is a key economic sector in province; recently, Binh Thuan has attracted a large number of both domestic visitors and international tourists and these numbers are predicted that they more and more rocker considerably.

Therefore, this study is focused on finding the best method describing the most accurate result easily to forecast tourism demand. In this research, we applied five models, namely GM (1, 1), DGM (1, 1), DGM (2, 1), Verhulst and ARIMA to test and look for the models which augment best results and minimum the forecasting errors. As can be seen from the above tables (Tables 7–15), GM (1, 1), DGM (1, 1), Verhulst and ARIMA are better to predict all the factors, viz. the tourism revenue, the proportion of tourists (both domestic visitors and international arrivals) because the parameter of MAPE, MSE, RMSE, and MAD are accepted for the process. Nevertheless, DGM (2, 1) is a poor model to forecast the demand for tourism in Binh Thuan Province (cf. Chia-Nan and Ty, 2013; Nguyen et al., 2015; Nguyen and Tran, 2018).

According to the results, it is easy to consider a realistic consequence. It is a fact that applying ARIMA for prediction of total revenue is the best choice. Otherwise, about the domestic visitors and international tourists, GM (1, 1), DGM (1, 1) and Verhulst give better calculation than the other models. Besides, the application of GM (1, 1), DGM (1, 1), Verhulst and ARIMA to forecast the number of visitors of top six markets (Russia, Germany, France, Korea, China and USA) sending the largest number of tourists describes good results and these numbers will go up in next 5 years. During the forecasting process, the number of Chinese tourists has the strongest upward trend, the number of Russian and Korean arrivals also increases and the numbers of others fluctuate by year. For all the factors, DGM (2, 1) is rejected to predict due to the poor results. In general, GM (1, 1), DGM (1, 1), Verhulst and ARIMA are concise and accurate models for forecasting tourism demand in Binh Thuan.

In conclusion, it is no doubt that the tourism industry has developed rapidly for recent years in Binh Thuan. Hence, the government has to propose suitable policies to develop the local tourism industry to serve a large number of tourists, also attract investors and invest in construction potential projects.
Compliance with ethical standards

Conflict of interest

The authors declare that they have no conflict of interest.

References

Akama JS and Kieti D (2007). Tourism and socio-economic development in developing countries: A case study of Mombasa Resort in Kenya. Journal of Sustainable Tourism, 15(6): 735-748. 
https://doi.org/10.2167/jost543.0

Chandra S and Menezes D (2001). Applications of multivariate analysis in international tourism research: The marketing strategy perspective of NTOs. Journal of Economic and Social Research, 3(1): 77-98.

Chang YW and Liao MY (2010). A seasonal ARIMA model of tourism forecasting: The case of Taiwan. Asia Pacific journal of Tourism Research, 15(6): 735-748. 
https://doi.org/10.116/jost543.0

Chia-Nan W and Ty NN (2013). Forecasting the manpower requirement in Vietnamese tertiary institutions. Asian Journal of Empirical Research, 3(5): 563-575.

Condratov I and Stanciu P (2012). The use of ARIMA models for forecasting the supply and demand indicators from tourism sector. The USV Annals of Economics and Public Administration, 12(2 (16)): 234-244.

Cortez N (2008). Patient without borders: The emerging global market for patients and the evolution of modern health care. Forthcoming Indiana Law Journal, 83(1): 71-132.

Du Preez J and Witt SF (2003). Univariate versus multivariate time series forecasting: An application to international tourism demand. International Journal of Forecasting, 19(3): 435-451. 
https://doi.org/10.1016/S0169-2070(02)00057-2

Hadavandi E, Ghanbari A, Shahanaghi K, and Abbasiyan-Naghneh S (2011). Tourist arrival forecasting by evolutionary fuzzy systems. Tourism Management, 32(5): 1196-1203. 
https://doi.org/10.1016/j/tourman.2010.09.015

Huang YL (2012). Forecasting the demand for health tourism in Asian countries using a GM (1, 1)-Alpha model. Tourism and Hospitality Management, 18(2): 171-181.

Lin Q and Lee TS (2013). Tourism demand forecasting: Econometric model based on multivariate adaptive regression splines, artificial neural network and support vector regression. Advances in Management and Applied Economics, 3(6): 1-18.

Nguyen NT and Tran TT (2018). A two-stage study of grey system theory and DEA in strategic alliance: An application in Vietnamese fertilizing industry. International Journal of Advanced and Applied Sciences, 5(9): 73-81. 
https://doi.org/10.21833/ijaas.2018.09.011

Nguyen NT and Tran TT (2019). Optimizing mathematical parameters of Grey system theory: An empirical forecasting case of Vietnamese tourism. Neural Computing and Applications, 31(2): 1-18.

Hadavandi E, Ghanbari A, Shahanaghi K, and Abbasiyan-Naghneh S (2011). Tourist arrival forecasting by evolutionary fuzzy systems. Tourism Management, 32(5): 1196-1203. 
https://doi.org/10.1016/j/tourman.2010.09.015

Huang YL (2012). Forecasting the demand for health tourism in Asian countries using a GM (1, 1)-Alpha model. Tourism and Hospitality Management, 18(2): 171-181.

Lin Q and Lee TS (2013). Tourism demand forecasting: Econometric model based on multivariate adaptive regression splines, artificial neural network and support vector regression. Advances in Management and Applied Economics, 3(6): 1-18.

Nguyen NT and Tran TT (2018). A two-stage study of grey system theory and DEA in strategic alliance: An application in Vietnamese fertilizing industry. International Journal of Advanced and Applied Sciences, 5(9): 73-81. 
https://doi.org/10.21833/ijaas.2018.09.011

Nguyen NT and Tran TT (2019). Optimizing mathematical parameters of Grey system theory: An empirical forecasting case of Vietnamese tourism. Neural Computing and Applications, 31(2): 1-18.

Radha S and Thenmozhi M (2006). Forecasting short term interest rates using ARMA, ARMA-Garch and ARMA-EGARCH models. Indian Institute of Capital Markets 9th Capital Markets Conference Paper. 
https://doi.org/10.2139/ssrn.875556

Song H and Li G (2008). Tourism demand modelling and forecasting: A review of recent research. Tourism Management, 29(2): 203-220. 
https://doi.org/10.1016/j/tourman.2007.07.016

Song H and Witt SF (2006). Forecasting international tourist flows to Macau. Tourism Management, 27(2): 214-224. 
https://doi.org/10.1016/j/tourman.2004.09.004