Contents lists available at Science-Gate

International Journal of Advanced and Applied Sciences

Journal homepage: http://www.science-gate.com/IJAAS.html

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, *]

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

ARTICLE INFO

Article history:
Received 25 November 2018
Received in revised form
28 October 2019
Accepted 12 November 2019

Keywords:
Tourism
Forecasting
Vietnam
Grey models
ARIMA

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.

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.

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).

* Corresponding Author.
Email Address: thanhtuyentran@lhu.edu.vn (T. T. Tran)
https://doi.org/10.21833/ijaas.2020.01.009
Corresponding author's ORCID profile:
https://orcid.org/0000-0002-9900-8592

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 equaled in 2016 and overtaken in 2017 by the Chinese market; the others following are Germany, Korea, France, and the USA respectively.

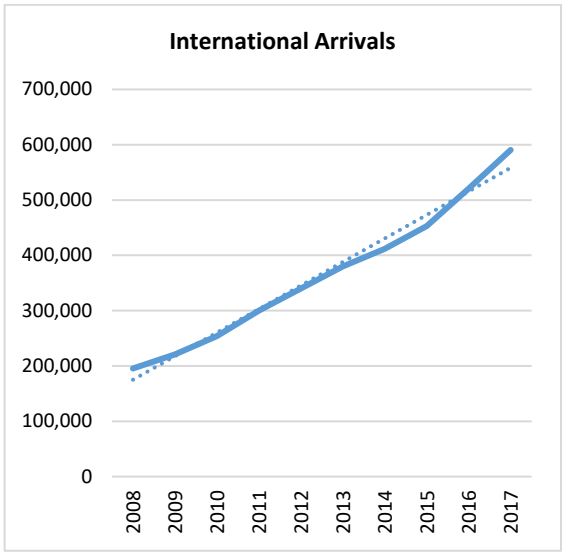

Fig. 1: Foreign visitor arrivals to Binh Thuan by year

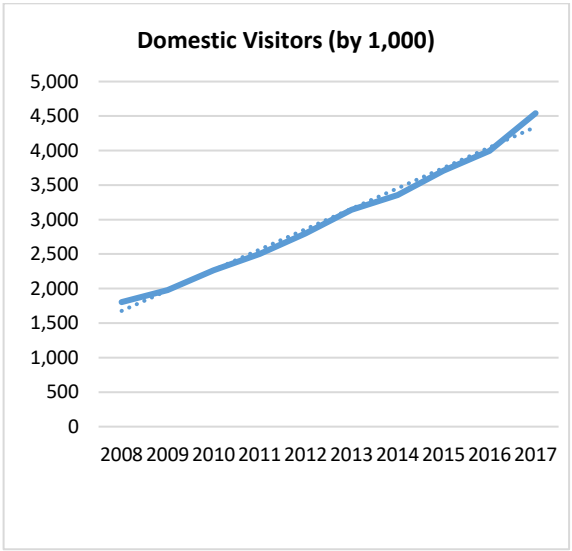

Fig. 2: Domestic visitor arrivals to Vietnam by year

Binh Thuan province has to give policies to promote tourism in the most absolute way to attract tourists and occupy a position on the map of Vietnamese tourism in particular and the world in general. To obtain a good strategic vision, Binh Thuan should forecast accurately tourism demand in the future. Tourism experts acknowledge that the improvement and accuracy of forecasting tourism are very necessary to research (Chandra and Menezes, 2001). Hence, the models of GM (1, 1), DGM (1, 1), DGM (2, 1), Verhulst model are demonstrated to find which models forecast exactly the situation.

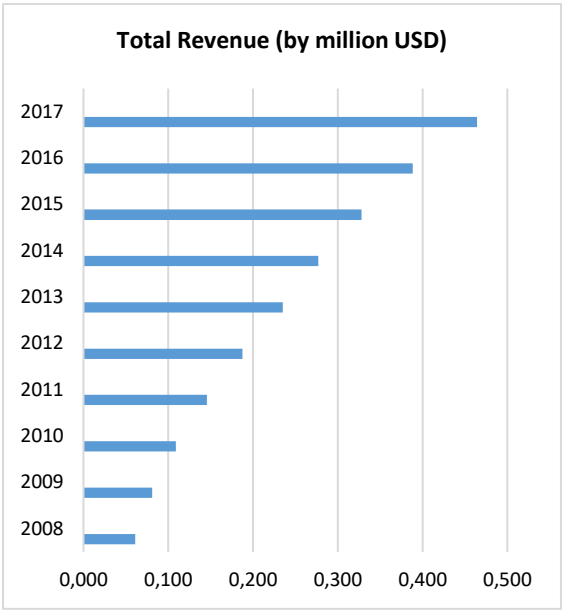

Fig. 3: Total revenue tourism in 2008-2017

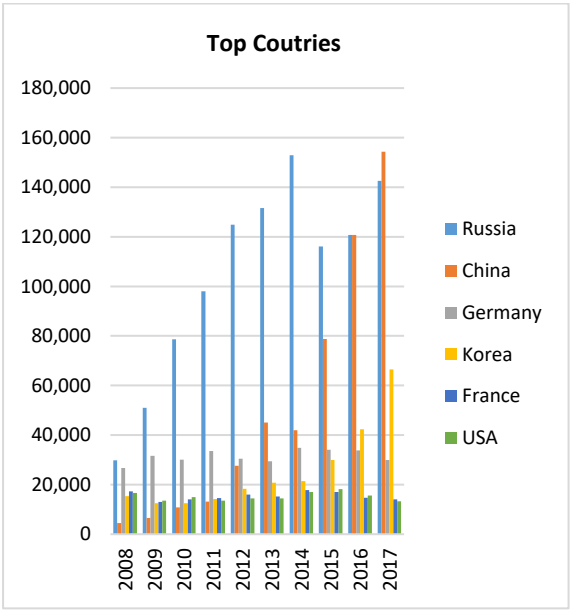

Fig. 4: The top 6 countries providing most travelers to Binh Thuan

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).

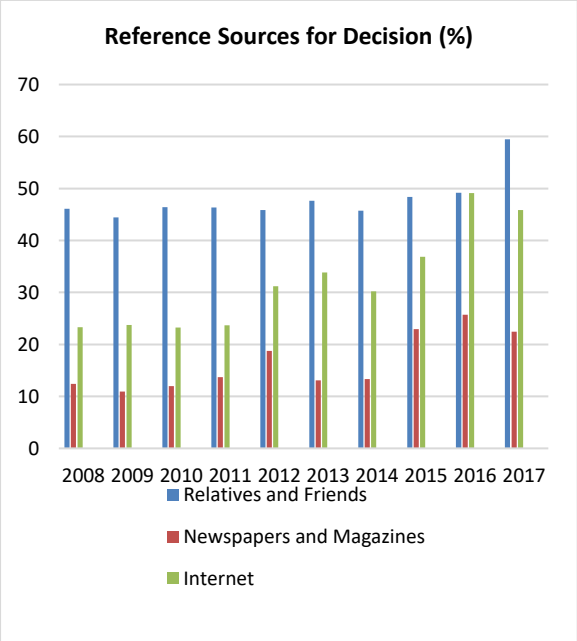

Fig. 5: Reference sources for decision

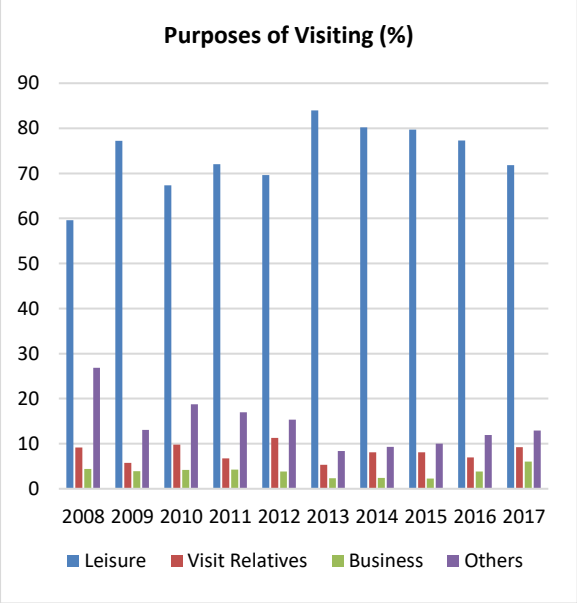

Fig. 6: Purposes of visiting

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	1.844E10
Domestic Visitors	3.007E6	1.806E6	4.542E6	9.02E5	8.137E11
International Visitors	366380	195156	590636	130265.93	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

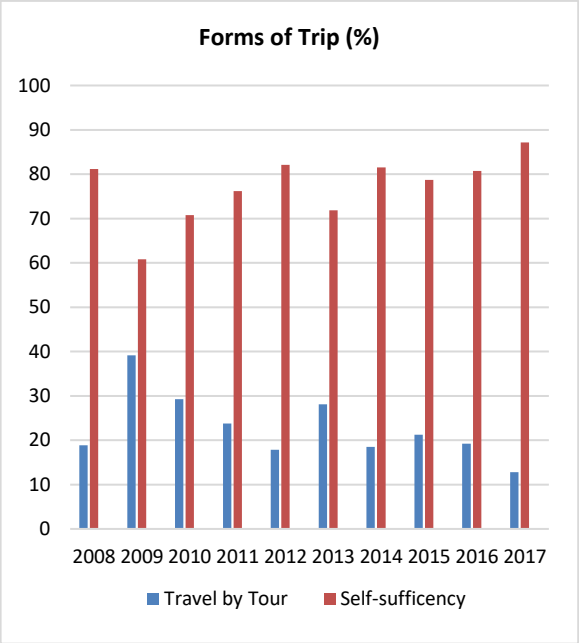

Fig. 7: Forms of trip

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 Thuanover 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=81611824.1276and $(1-e^{a})\left(x^{(0)}(1)-\frac{b}{a}\right)=84943934.5560$ are calculated the Total Revenue.
- ✓ The results of parameters connecting to the Domestic Visitors are a=-0.0983; b=1775695.8471so$(1-e^{a})\left(x^{(0)}(1)-\frac{b}{a}\right)=1860228.9725$
- ✓ a=-0.1163; b=195467.1875and $(1-e^{a})\left(x^{(0)}(1)-\frac{b}{a}\right)=205954.3320$ are transmitted to the calculation of international visitors.
- ✓ a=-0.0732: b=78101.4692 and $(1-e^{a})\left(x^{(0)}(1)-\frac{b}{a}\right)=77412.2129$ are analyzed the calculation of Russian visitors.
- ✓ The results of parameters analyzing German visitors are a=-0.0050; b=31120.9859so$(1-e^{a})\left(x^{(0)}(1)-\frac{b}{a}\right)=31176.5122$
- ✓ a=-0.0132; b=14048.0298and $(1-e^{a})\left(x^{(0)}(1)-\frac{b}{a}\right)=14183.2281$ are calculated French visitors.
- ✓ a=-0.2606; b=1277.9339and $(1-e^{a})\left(x^{(0)}(1)-\frac{b}{a}\right)=4645.7411$ are analyzed the calculation of Korean visitors.
- ✓ a=-0.3625; b=6317.4331and $(1-e^{a})\left(x^{(0)}(1)-\frac{b}{a}\right)=6653.0704$ are related to the calculation of Chinese tourists.
- ✓ The results of parameters analyzing USA visitors are a=-0.0131; b=13913.9127and $(1-e^{a})\left(x^{(0)}(1)-\frac{b}{a}\right)=14040.1869$

• 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: $\beta_1 = 1.0127$; $\beta_2 = 14185.9725$and$x^{(0)}(1)(\beta_1 - 1) + \beta_2 = 14406.7166$
- ✓ 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=0and $\hat{x}(k+1) = \frac{ax^{(1)}(0)}{bx^{(1)}(0)+\left(a-bx^{(1)}(0)\right)e^{ak}}$ (Ver. 7 – mentioned in section 2) in which $ax^{(1)}(0) = -16238333.8319$; $a - bx^{(1)}(0)$=-0.2506; and $bx^{(1)}(0) = -0.0150$
- ✓ 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=0and $\hat{x}(k+1) = \frac{ax^{(1)}(0)}{bx^{(1)}(0)+\left(a-bx^{(1)}(0)\right)e^{ak}}$ (Ver. 7 – mentioned in section 2) in with $ax^{(1)}(0) = -20630.4220$; $a - bx^{(1)}(0)$= -0.5462; and $bx^{(1)}(0) = -0.1470$
- ✓ 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; d=1; q=1; P=0; D=0; Q=0 and s=0 with the confidence intervals being 95%
- ✓ The model parameters of Domestic Tourists: 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 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.

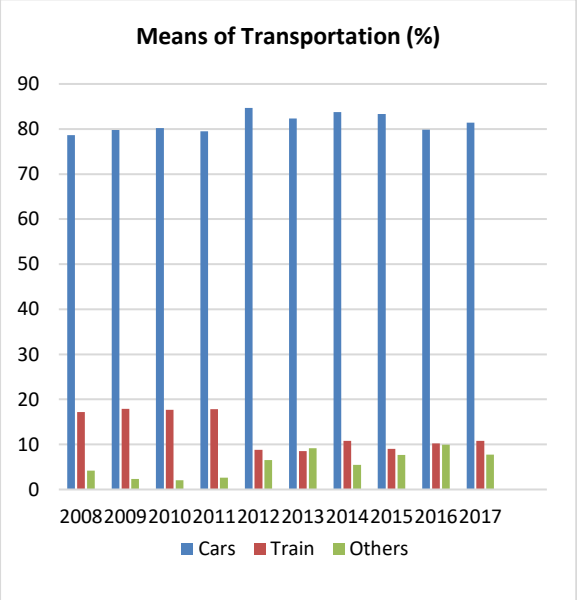

Fig. 8: Means of transportation

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

		Total Revenue (by Million USD)					
STAGES	Models	Actual	GM(1,1)	DGM(1,1)	DGM(2,1)	Verhulst	ARIMA
MODEL BUILDING	2008	61145942	61145942	61145942	61145942	61145942	61145942
	2009	81187204	102972419	103482867	72520576	78397414	101039711
	2010	109016100	124827266	125496520	97252960	100039290	113733369
	2011	145529624	151320581	152193082	125252969	126900244	151030240
	2012	187714897	183436832	184568737	156952318	159800068	187178369
	2013	235061354	222369431	223831585	192839761	199448690	234066886
	2014	276992396	269565077	271446721	233468626	246307941	281671919
	2015	328138939	326777518	329190907	279465346	300429838	319173617
	2016	388394155	396132716	399218869	331539118	361304028	381478114
	2017	464240011	480207848	484143705	390492837	427763999	440153522
FORECASTING	2018		582127071	587134390	532796107	569737920	484093298
	2019		705677610	712034027	618339749	640476955	
	2020		855450492	863503253	715185352	707867979	
	2021		1.037E+09	1.047E+09	824826124	769990599	
	2022		1.257E+09	1.27E+09	948952553	825543021	

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

		Domestic Tourists						International Tourists					
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
MODEL BUILDING	2008	1805535	1805535	1805535	1805535	1805535	1805535	195156	195156	195156	195156	195156	195156
	2009	1978463	2052323	2054350	1895965	1993699	2107844	221643	231352	231688	207545	220333	236443
	2010	2249881	2264253	2266612	2088949	2202167	2293274	250321	259882	260283	234140	248712	257142
	2011	2502338	2498069	2500807	2301497	2433283	2556770	300060	291930	292406	263688	280687	290864
	2012	2800008	2756028	2759199	2535590	2689696	2810251	340181	327930	328495	296515	316698	350948
	2013	3144785	3040626	3044289	2793413	2974407	3103822	380052	368370	369037	332986	357235	377319
	2014	3354029	3354612	3358835	3077371	3290828	3443850	411897	413797	414582	373506	402840	426646
	2015	3701375	3701021	3705882	3390115	3642844	3665224	453105	464825	465749	418523	454115	446073
	2016	3994084	4083203	4088786	3734561	4034898	4000886	520754	522147	523231	468538	511724	502888
	2017	4541582	4504849	4511254	4113923	4472093	4297579	590636	586537	587807	524104	576396	578310
FORECASTING	2018		4970036	4977372	4991915	5506335	4518951		658867	660352	654425	730209	599489
	2019		5483260	5491651	5498735	6118097			740117	741851	730625	821175	
	2020		6049482	6059068	6056932	6804842			831387	833409	815283	922856	
	2021		6674174	6685112	6671713	7577450			933912	936266	909339	1036355	
	2022		7363373	7375841	7348814	8448803			1049080	1051818	1013836	1162846	

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 and 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 Germany 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; additionally, it shows the forecasting consequences in next five years. Meanwhile, the proportion of the Korean market is not applied Verhulst model to forecast in 2011 so it is an error.

Table 4: The results of Russia and Germany markets

		Russia						Germany					
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
MODEL BUILDING	2008	29760	29760	29760	29760	29760	29760	26743	26743	26743	26743	26743	26743
	2009	50982	83290	84259	39225	49110	42353	31689	31332	31368	23170	28686	28476
	2010	78638	89614	90459	57141	72764	63781	30048	31489	31516	8322	30049	29861
	2011	98060	96418	97114	73627	95844	91588	33517	31646	31664	Error	30969	30729
	2012	124914	103739	104260	88799	113907	110807	30412	31804	31814	Error	31572	31389
	2013	131650	111616	111931	102760	125757	137846	29378	31963	31964	Error	31961	31487
	2014	152855	120091	120166	115608	132656	144090	34846	32123	32115	Error	32209	31938
	2015	116086	129209	129008	127431	136398	165657	34119	32283	32266	Error	32366	32657
	2016	120711	139019	138499	138310	138348	127474	33797	32445	32418	Error	32464	33190
	2017	142639	149575	148690	148322	139345	133137	29886	32607	32571	Error	32526	33435
FORECASTING	2018		160932	159630	166014	140102	142869		32770	32725	Error	32590	33080
	2019		173151	171375	173815	140228			32933	32879	Error	32605	
	2020		186298	183984	180995	140292			33098	33034	Error	32614	
	2021		200443	197521	187602	140324			33263	33190	Error	32620	
	2022		215663	212054	193682	140340			33429	33347	Error	32624	

Table 5: The results of France and Korea markets

		France						Korea					
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
MODEL BUILDING	2008	17323	17323	17323	17323	17323	17323	15349	15349	15349	15349	15349	15349
	2009	13012	14372	14407	18326	16859	16087	12466	6029	6344	15683	17272	20041
	2010	14021	14564	14590	22013	16421	15003	12522	7823	8218	16601	19828	14335
	2011	14553	14758	14776	30677	16005	14820	14133	10151	10647	18054	23388	17647
	2012	16057	14954	14965	51035	15611	14952	18336	13173	13794	20354	28689	18159
	2013	15202	15153	15155	98870	15237	15115	20827	17094	17870	23993	37415	24896
	2014	17835	15355	15348	211267	14880	15358	21377	22182	23152	29752	54454	24364
	2015	17037	15560	15544	475366	14541	15674	29950	28785	29994	38867	102472	25708
	2016	14737	15767	15742	1095916	14218	15739	42285	37353	38859	53292	1105121	39498
	2017	13998	15977	15943	2554016	13909	15631	66506	48471	50343	76121	Error	50790
FORECASTING	2018		16190	16146	14030336	13333	15468		62898	65222	169427	Error	77890
	2019		16406	16351	32945900	13063			81620	84498	259916	Error	
	2020		16624	16560	77391617	12804			105915	109470	403124	Error	
	2021		16846	16771	181825292	12556			137441	141824	629765	Error	
	2022		17070	16985	427212095	12318			178351	183738	988447	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.

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.

Table 6: The results of China and USA markets

		China						USA					
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
MODEL BUILDING	2008	4453	4453	4453	4453	4453	4453	16688	16688	16688	16688	16688	16688
	2009	6547	9560	9900	5915	7429	19874	13565	14225	14268	17313	16402	16325
	2010	10846	13736	14262	9441	12310	17732	14935	14412	14445	19497	16090	13334
	2011	13203	19737	20547	14078	20168	24485	13593	14602	14625	24290	15750	14501
	2012	27657	28361	29602	20179	32455	24914	14492	14794	14807	34813	15383	13276
	2013	45074	40751	42646	28204	50821	45510	14404	14989	14991	57910	14988	14076
	2014	42013	58556	61439	38761	76495	61532	16970	15187	15178	108611	14564	14029
	2015	78750	84138	88514	52648	109222	50092	18215	15387	15367	219904	14112	16473
	2016	120711	120898	127520	70916	146364	107984	15571	15589	15558	464204	13634	17774
	2017	154274	173718	183714	94947	183344	142949	13230	15794	15752	1000468	13130	15314
FORECASTING	2018		249614	264672	168142	240843	159247		16002	15948	4761589	12055	13327
	2019		358670	381305	222844	258763			16213	16146	10433662	11490	
	2020		515372	549335	294802	270710			16427	16347	22884434	10911	
	2021		740536	791412	389459	278332			16643	16551	50215137	10322	
	2022		1064073	1140165	513977	283059			16862	16757	110208790	9728	

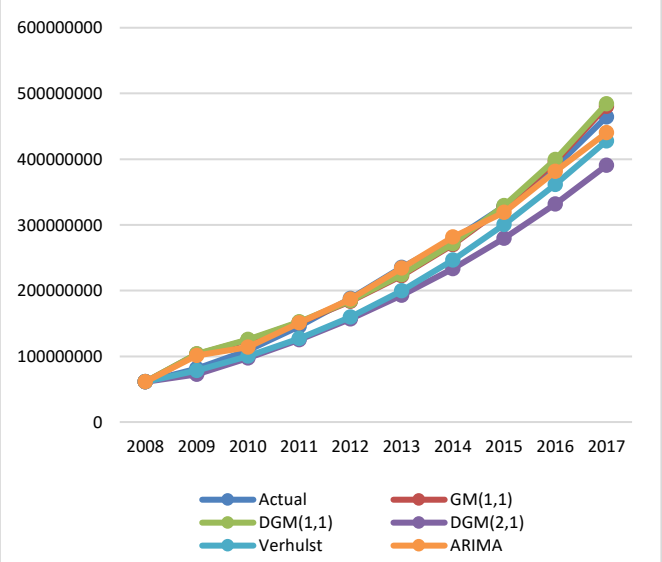

Fig. 9: Forecasting result of tourism revenue

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 waved 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.

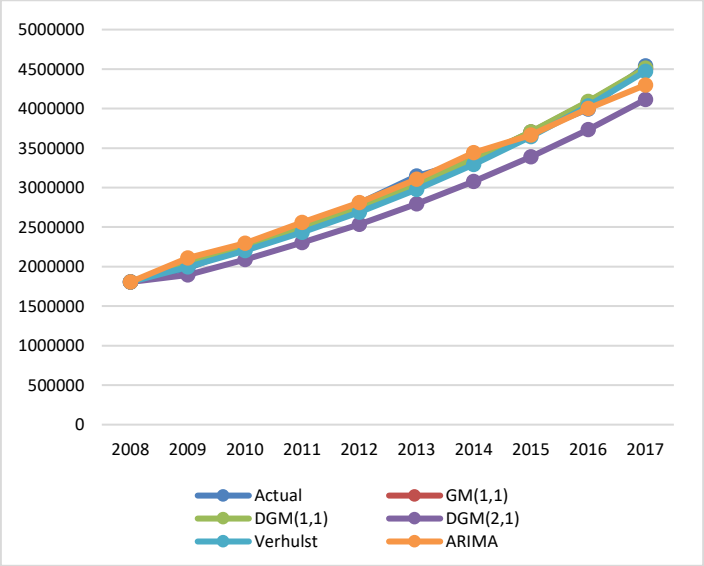

Fig. 10: Forecasting result of domestic visitors

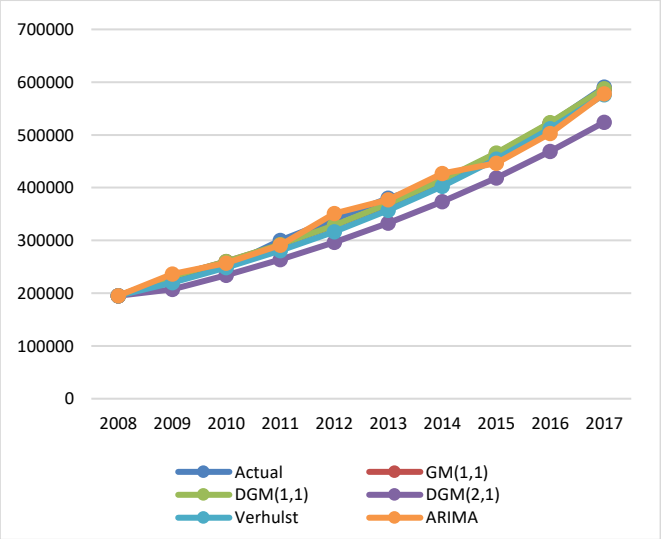

Fig. 11: Forecasting result of international arrivals

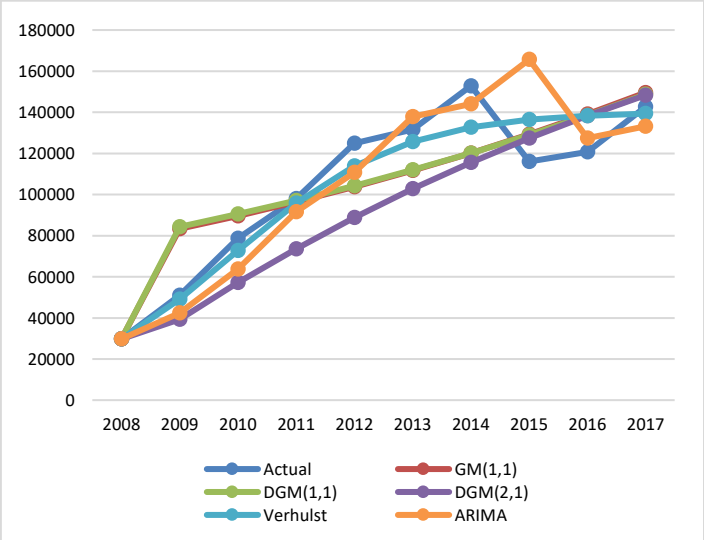

Fig. 12: Forecasting result of Russia visitors

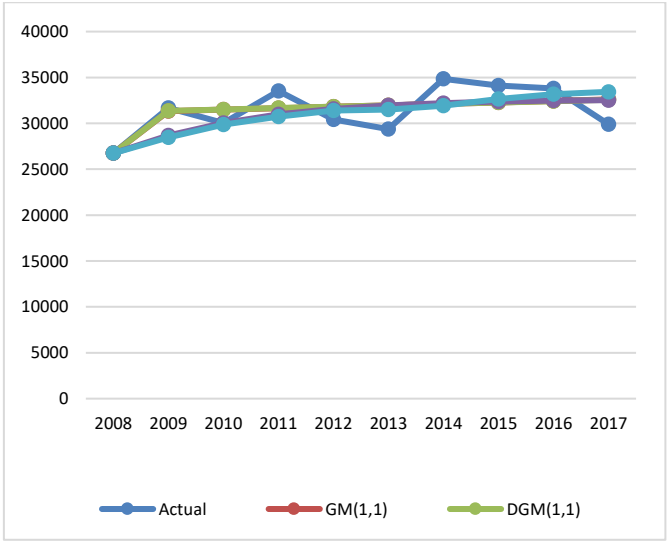

Fig. 13: Forecasting result of Germany visitors

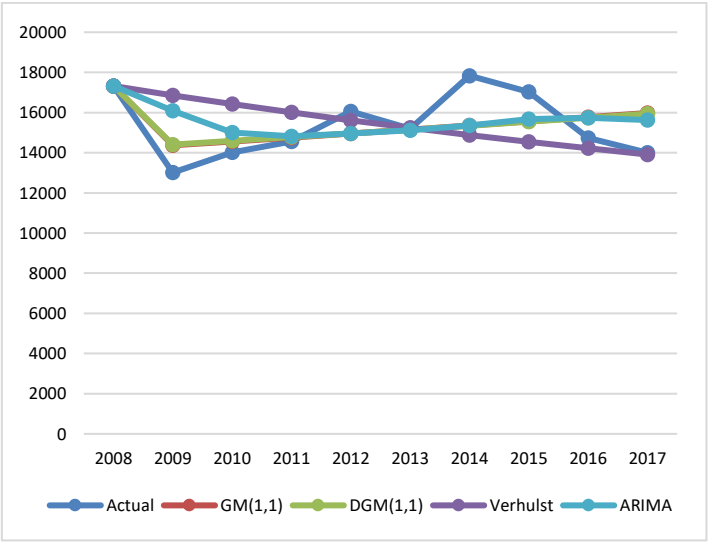

Fig. 14: Forecasting result of France visitors

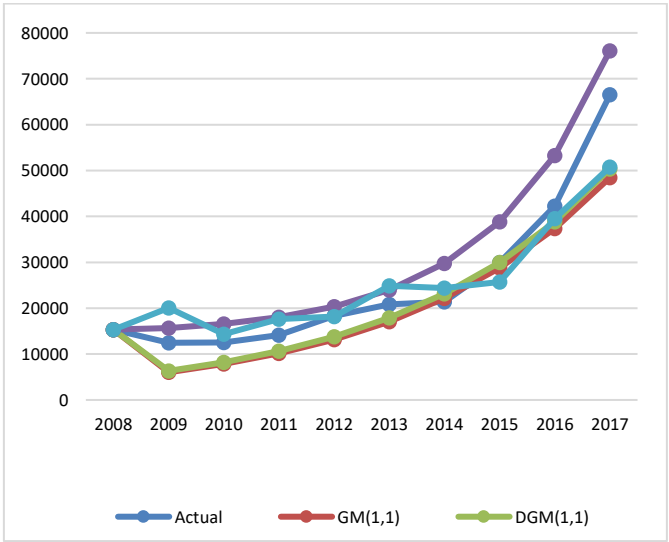

Fig. 15: Forecasting result of Korea visitors

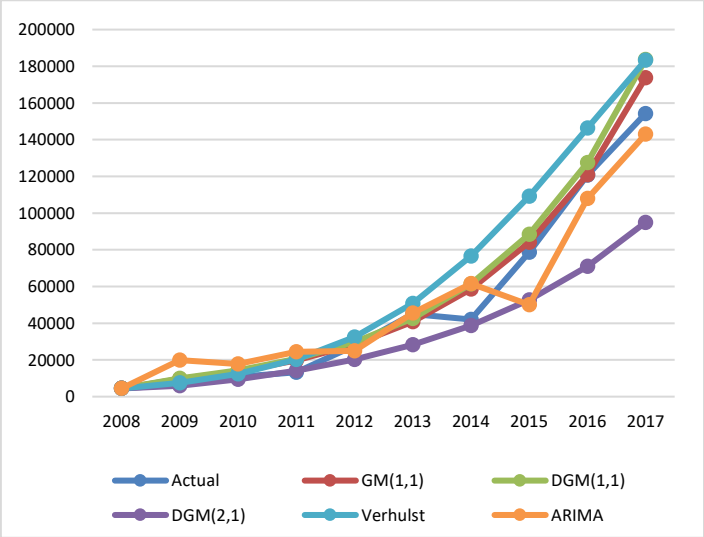

Fig. 16: Forecasting result of China visitors

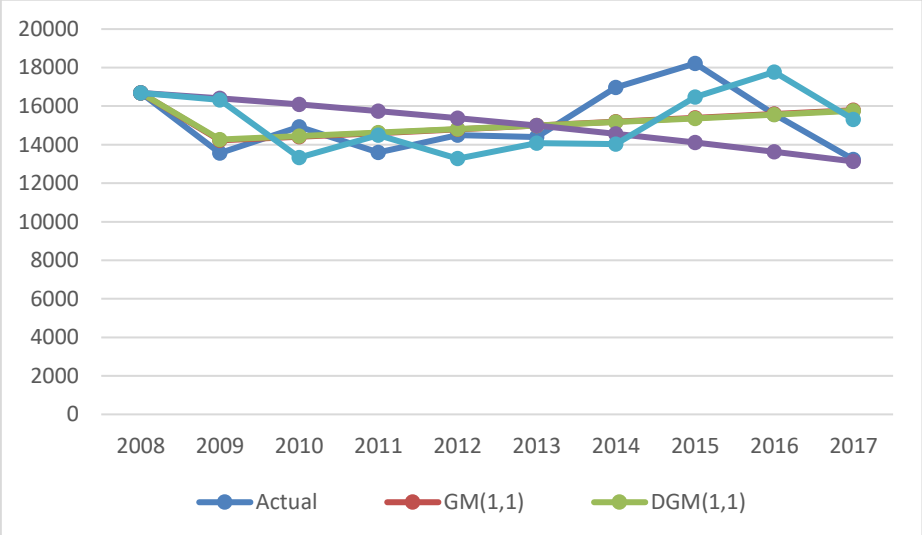

Fig. 17: Forecasting result of USA visitors

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:

$$MAPE=\frac{1}{n}\sum\frac{|Actual-Forecast|}{Actual}\times 100;$$

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:

$$MSE=\frac{1}{h+1}\sum_{t=s}^{s+h}\hat{x}_{t-1}(1)-x_t)^2$$

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.

$$RMSE=\sqrt{\frac{1}{h+1}\sum_{t=s}^{s+h}\hat{x}_{t-1}(1)-x_t)^2}$$

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.

$$MAD=\sum_{i=1}^{n}|e_i|$$

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 7: Evaluating models with total revenue forecasting errors

Models	GM (1, 1)	DGM (1, 1)	DGM (2, 1)	VERHULST	ARIMA
MAPE	6.15%	6.3%	13.08%	8.89%	4.47%
MSE	1.3E+14	1.49E+14	1.63E+15	6.26E+14	1.18E+14
RMSE	11442841.42	12224161.03	40358644.72	25013244.34	10854371.43
MAD	9285246.4	9714151.8	33649017	21588316.95	7624876.2
Evaluation	Good	Good	Poor	Reasonable	Excellent

Table 8: Evaluating models with domestic tourists forecasting errors

Models	GM(1, 1)	DGM(1, 1)	DGM(2, 1)	VERHULST	ARIMA
MAPE	1.25%	1.26%	7.25%	2.1%	2.15%
MSE	2775539068	2773759466	69008814139	6238950282	9232644844
RMSE	52683.385	52666.490	262695.290	78987.026	96086.653
MAD	36742.9	36979.7	233516.1	64473	65518.9
Evaluation	Excellent	Excellent	Poor	Good	Good

Table 9: Evaluating models with international visitors forecasting errors

Models	GM(1, 1)	DGM(1, 1)	DGM(2, 1)	VERHULST	ARIMA
MAPE	2.16%	2.17%	8.84%	2.72%	2.7%
MSE	69804107.7	69783873.7	1572822995	181905273.7	121162032.6
RMSE	8354.885	8353.674	39658.832	13487.226	11007.363
MAD	7044.5	7099.7	34910.4	10192.9	9628.9
Evaluation	Excellent	Excellent	Poor	Good	Good

Table 10: Evaluating models with Russia visitors forecasting errors

Models	GM(1, 1)	DGM(1, 1)	DGM(2, 1)	VERHULST	ARIMA
MAPE	16.4%	16.43%	17.89%	7.43%	11.91%
MSE	364570117	365199271.3	519431419.6	134129156.4	324466984
RMSE	19093.719	19110.188	22791.038	11581.414	18012.967
MAD	15726.6	15586.7	19456.6	8830.4	12486.3
Evaluation	Good	Good	Poor	Excellent	Good

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, MRSE, 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 11: Evaluating models with Germany visitors forecasting errors

Models	GM(1, 1)	DGM(1, 1)	DGM(2, 1)	VERHULST	ARIMA
MAPE	5.118%	5.117%	14189.18%	5.52%	5.58%
MSE	3434183	3434751	9.15477E+13	4230107	4709137.9
RMSE	1853.155	1853.308	9568056.907	2056.722	2170.055
MAD	1627.8	1627.8	4424462.5	1765.8	1779.96
Evaluation	Excellent	Excellent	Poor	Good	Good

Table 12: Evaluating models with France visitors forecasting errors

Models	GM(1, 1)	DGM(1, 1)	DGM(2, 1)	VERHULST	ARIMA
MAPE	6.66%	6.68%	3023.37%	9.5%	8.03%
MSE	1671475.4	1672145.6	7.8767E+11	3810717.7	2338220.6
RMSE	1292.856	1293.115	887507.565	1952.106	1529.124
MAD	1022.6	1025.6	442103.4	1423.9	1199.07
Evaluation	Excellent	Excellent	Poor	Good	Good

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	37366171.5	41971323.9	1.17201E+11	37127144.09
RMSE	6866.985	6112.788	6478.528	342345.863	6093.205
MAD	4895.1	4281.9	5431.5	140512.7	4288.02
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%	21.43%	27.07%	46.41%
MSE	76011286.8	147230546.3	703478132	372817367.5	185252000
RMSE	8718.445	12133.860	26523.162	19308.479	13610.731
MAD	5902.6	8392.5	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	1991173.2	1992614.8	1.22744E+11	4155559.4	3347439.4
RMSE	1411.089	1411.600	350348.678	2038.519	1829.600
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 – 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(2): 215-221. https://doi.org/10.1080/10941661003630001

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 Abbasian-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 CJ 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): 1075-1089. https://doi.org/10.1007/s00521-017-3058-9

Nguyen NT, Tran TT, Wang CN, and Nguyen NT (2015). Optimization of strategic alliances by integrating DEA and grey model. Journal of Grey System, 27(1): 38-56.

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.876556

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