

Forecasting Korean Arrivals in the Philippines

Enrique Raphael Versoza¹

Sofia Elaine Romarate²

Jacque Bon-Isaac Aboy³

School of Management, University of the Philippines Cebu, Cebu City, Philippines^{1,2}

College of Science, University of the Philippines Cebu, Cebu City, Philippines³

Abstract:- This paper investigates the rise of South Korean tourism in the Philippines from 2014 to 2018 and explain its behavior year-to-year, and the other part is to forecast it's growth or decline in the next following years; all of this is done through a Seasonal ARIMA (SARIMA) modelling framework. Results reveal that Korean arrivals were best modelled through a ARIMA(1,0,0)(2,1,0)₁₂ model, with residuals that are randomly distributed and contain no autocorrelations and an AICc value of -36.18, the lowest among the tested variations of the model, the model is the most appropriate to forecast the data for a 3-year period.

Keywords:- South Korean arrivals; SARIMA; Philippine tourism; prediction

➤ Subject classification codes: C53, F22, F47

I. INTRODUCTION

Since the successful invention of the first airplane by the Wright brothers in the 1900s, air travel has become more frequent and favorable to the point wherein it is more affordable and efficient to travel across countries by air than it is by boat. With this rise in air transportation has also come the rise of international travel. Every year more and more people travel abroad, either for vacation or on business. From a measly 25 million international visits in 1950 to a record 1.4 billion international visits in 2018 (Blackall, 2019). Hand in hand with this rise in international visits is the rise of global tourism and its contribution to the world economy. In 2018, tourism was worth about \$1.7 trillion, or roughly 2% of global GDP (Blackall, 2019).

The Asia Pacific region comprised close to a quarter of the world's tourist arrivals in 2017, with Southeast Asia having 120 million tourist arrivals. An interesting statistic here is that 80% of the traffic has been intraregional, with China being the largest contributor to international arrivals and South Korea as the second largest contributor (Horwath HTL Market Report, 2018).

Speaking of intra-Asia Pacific tourism, tourism in the Philippines has become an essential pillar of the Philippine economy. In 2018, tourism contributed Php 2.2 trillion to the economy 14.3% higher compared to 2017's 1.9 trillion, calculated by comparing the total Tourism Gross Value Added to GDP (Mapa, 2019).

The countries with the largest foreign arrivals coming to the Philippines are South Korea, China, United States of America, Japan, and Australia, but what separates South Korea from the others is that it has consistently been top 1 when it comes to this for more than a decade. In 2018 alone, there were 1.6 million foreign arrivals coming from South Korea. One reason as to the cause of this large and consistent wave of South Korean arrivals is South Korea's increasing prosperity which has caused South Korean investment to increase in the Philippines and South Korean children to study in the Philippines (Miralao, 2007).

This study can further analyze South Korean arrivals into the Philippines, contributing to information on the frequency and magnitude of arrivals in a year. It will also predict South Korean arrivals for the next year. Helping local businesses prepare for influxes or shortages of tourists, and helping the government prepare to accommodate increasing tourist arrivals.

II. REVIEW OF RELATED LITERATURE

South Korea has not always been the industrialized economic powerhouse it is today. There was a time in its history wherein Korea was a mess due to warring lords and the Japanese invasion which led to poverty and starvation. Lee (2005) called the 1st wave of migration involuntary emigration because farmers and peasants were forced to migrate to China, Russia, or Japan to escape the Japanese exploitation of Korea's land. Lee (2005) called the 2nd wave of migration voluntary emigration because the Korean migration at that time was sponsored by the government to alleviate population pressures, earn foreign currency for Korea, and aid other countries.

According to (Gomez, 2013) what strengthened South Korean-Philippines diplomatic relationship were 2 things: 1) Philippines had sent the Philippines Expeditionary Force to Korea (PEFTOK) as part of the UN during the Korean war. This was not lost on older Koreans. 2) There is always the possibility of attack on South Korea by North Korea, and South Korea has hinted the Philippines might be an alternative refuge in that event together with Japan and Taiwan. Miralao (2007) also emphasizes that intensified economic cooperation between the Philippines and South Korea contributes to the recent wave of Korean migrations to the Philippines, with some Korean company personnel and businessmen moving to the Philippines and bringing their families along with them. It is also significant to note that (Gomez, 2013) mentioned that the Philippines has also been considered a fall-back for middle-class parents who

want to keep their children competitive by letting them learn English in the Philippines. Lastly (Gomez, 2013) noted the Department of Tourism's program or enticing foreigners to retire in the Philippines.

Rufino (2013) supports Miralao's statement of Korea's economic rise fueling the recent wave of Korean arrivals in the Philippines. In his paper, Rufino (2013) found out that tourism demand from the Korean market is highly income elastic providing evidence that a prospering Korean economy bodes well for Philippine tourism.

According to Miralao (2007), low cost is not the only reason why Korean parents choose to send their kids to the Philippines to study English. Koreans also perceive English education in the Philippines to be "the best place in the world to learn English as a second language". Miralao (2007) also notes that older Koreans retain the memory of the Philippines being more developed than Korea in the 1960 and that their schools provide quality education. However precisely because of this, Korean migration to the Philippines can be considered temporary as Korean's view the Philippines as a 'stepping-stone' to being competitive job-wise in Korea or in other foreign nations.

Despite this "temporary migration", there is some form of transculturation that goes on here. Gomes (2013) notes Koreans' substantial impact on Filipino culture, especially recently, with Filipinos picking up a taste for Koreans' dishes such as *japchae* or *samgyeupsal* or Filipinos' love for Korean dramas. On the other way around, Gomez (2013) theorizes that this transculturation can be viewed as the Philippines performing its role as a

bridge between the east and the west; acclimating Korean youths to what the west might be like in a place that is much cheaper where they can live comfortably.

III. RESEARCH METHODOLOGY

This study uses time series analysis to create a SARIMA model to predict future South Korean arrivals in the Philippines based on prior data. An ARIMA model describes the autocorrelations in the data and by regressing the data with its lagged terms is able to forecast future data. An ARIMA model can segregate the autocorrelations in the data from the noise or errors.

A total of 60 observations were gathered from the Tourism Demand Statistics of the Department of Tourism of the Philippines ranging from the years 2014-2018 with the observations taken monthly. The Tourism Demand Statistics were taken from the Philippines' Department of Tourism website http://www.tourism.gov.ph/tourism_dem_sup_pub.aspx. (Table of raw data can be found in Appendix B).

The study utilized Microsoft Excel to collate the data, which was then passed on to R Studio as a dataset. R studio is an integrated development environment (IDE) for R, a programming language for statistical computing and graphics. The data was then analyzed and modeled using R studio to come up with the best fitted Seasonal Auto Regressive Integrated Moving Average (SARIMA) model to forecast the data. After the data was passed to R studio as a dataset, the individual observations were transformed to their time series forms and plotted:

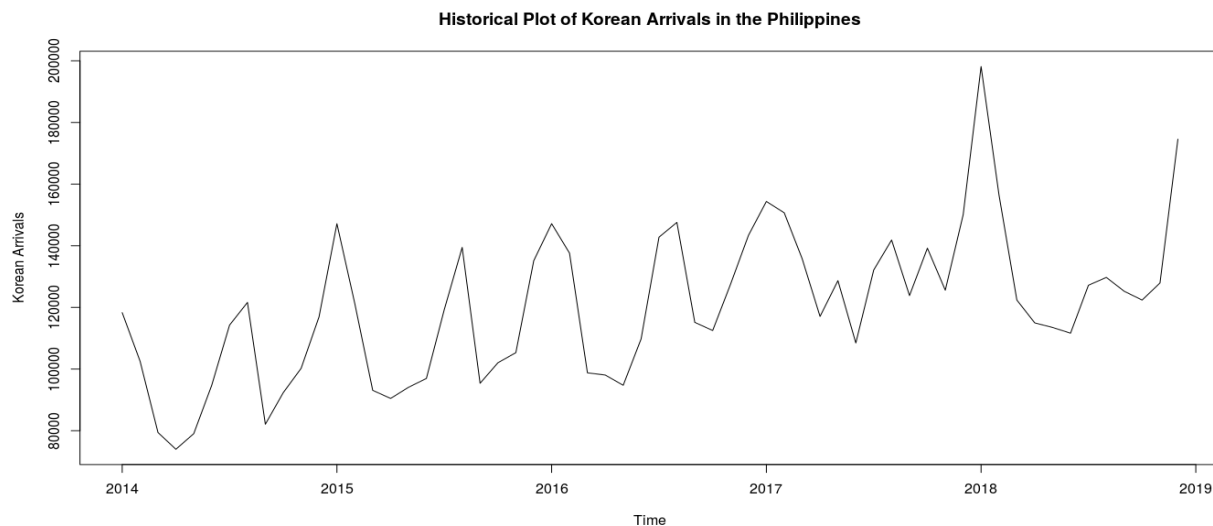


Fig 1:- Historical plot of Number of South Korean Arrival to the Philippines over time

The data shows an upward trend with dips in the years 2017 and 2018 where there was some volatility. The data also shows seasonality with 3 peaks each year. The first and thirds peaks are a rise in South Korean arrivals every January and February of each year which were spillovers from December of the last year. The second peak is a rise in South Korean arrivals during the months. of July and August. The volatility in the data affected the seasonality of the data during the years 2017 and 2018 with the second peak decreasing and the first and third peaks being more pronounced.

These observations can be further seen when we decomposed the data:

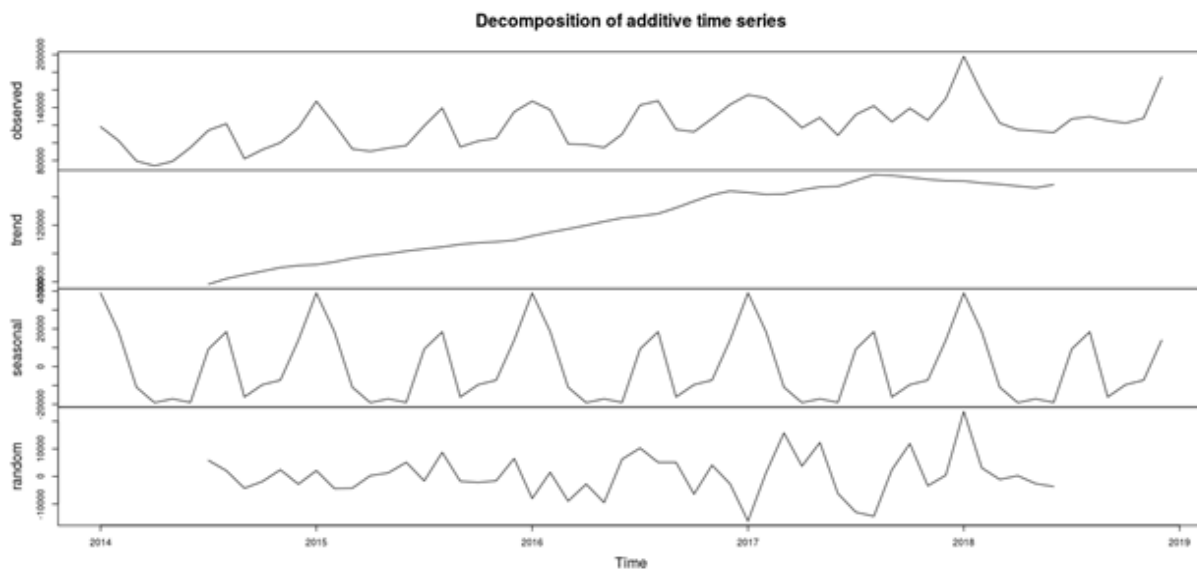


Fig 2:- Decomposed Plot of South Korean Arrivals to the Philippines

Here we see a steady upward trend from the years 2014 up until 2017, where the data dipped during 2017 and slowly rose back the same year only to dip again during the end of 2017 up to 2018. Here our initial descriptive analysis of the seasonality of the series is confirmed with each season containing 3 peaks with the first and third peak being larger than the second peak. Our analysis of the volatility of the data is correct but incomplete with the data already starting to show signs of volatility during mid 2016 becoming more volatile during the start of 2017 continuing until 2018.

➤ *Differencing and Testing for Stationarity*

Before we difference the data and test it for stationarity, we first log transformed the data to avoid any exponential trends. Once log-transformed, we can then difference the data to achieve stationarity. We differenced the data straight away because upon visual analysis of the data we saw a trend and seasonality therefore we can conclude that the initial data is non-stationary.

We initially seasonally differenced the data due to its strong seasonal pattern and plotted the data to see whether there were still obvious trends or seasons in the differenced data:

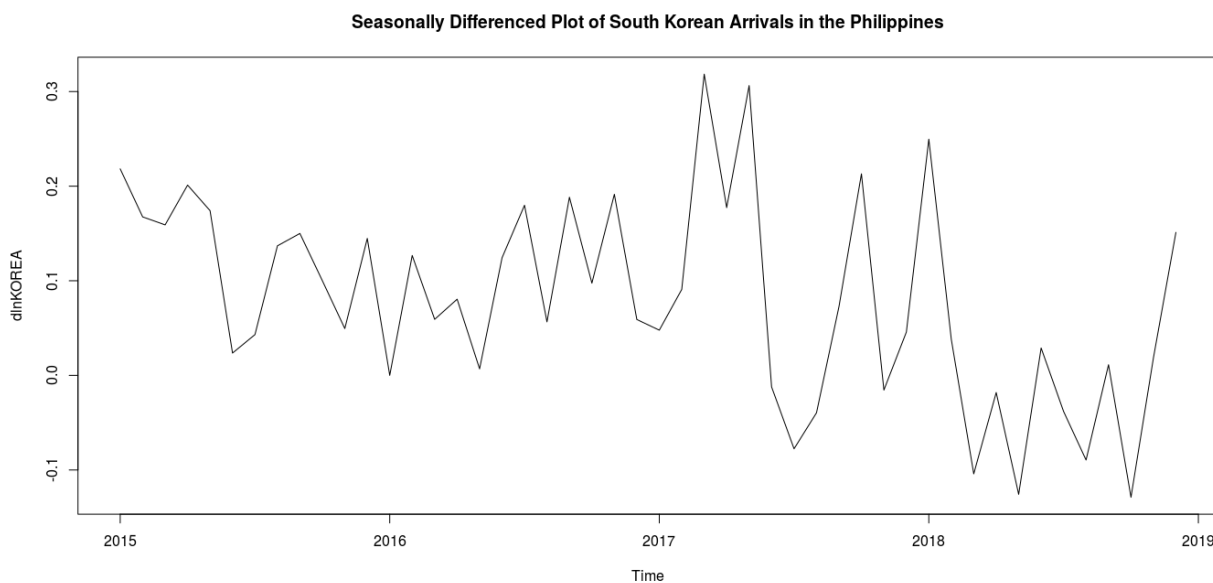


Fig 3:- Seasonally Differenced Plot of South Korean Arrivals in the Philippines

We can see that the upward trend of the data has disappeared and that its seasonality is gone. This can be confirmed by conducting the Weibel and Ollech (WO) test for seasonality on the differenced data which produced the following outcome:

Weibel and Ollech test for seasonality	
p-value	0.8711424

Table 1:- Results of WO test on the Seasonally Differenced Data

Since the p-value of the WO is greater than 0.05 we can accept the alternative hypothesis and say that there is no longer seasonality in the differenced data. Therefore based from the visual analysis and result of the WO test we can say that the seasonally differenced data is now stationary. To further confirm this we conducted 2 unit root tests on the seasonally differenced data, we conducted the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test and the Augmented Dickey-Fuller (ADF) test and got these results:

Tests	p-value
KPSS test	0.02564
ADF test	0.07693

Table 2:- Results for the KPSS and ADF test on the Seasonally Differenced Data

Since the p-value of the KPSS test is less than 0.05 we accept the null hypotheses and can conclude that the differenced data is stationary. For the ADF test, since the p-value is greater than 0.05 we accept the alternative hypothesis that the data is stationary.

➤ Using the Autocorrelation Function (ACF) Plot and Partial Autocorrelation Function (PACF) Plot to Model the SARIMA $(p,d,q)(P,D,Q)_{12}$

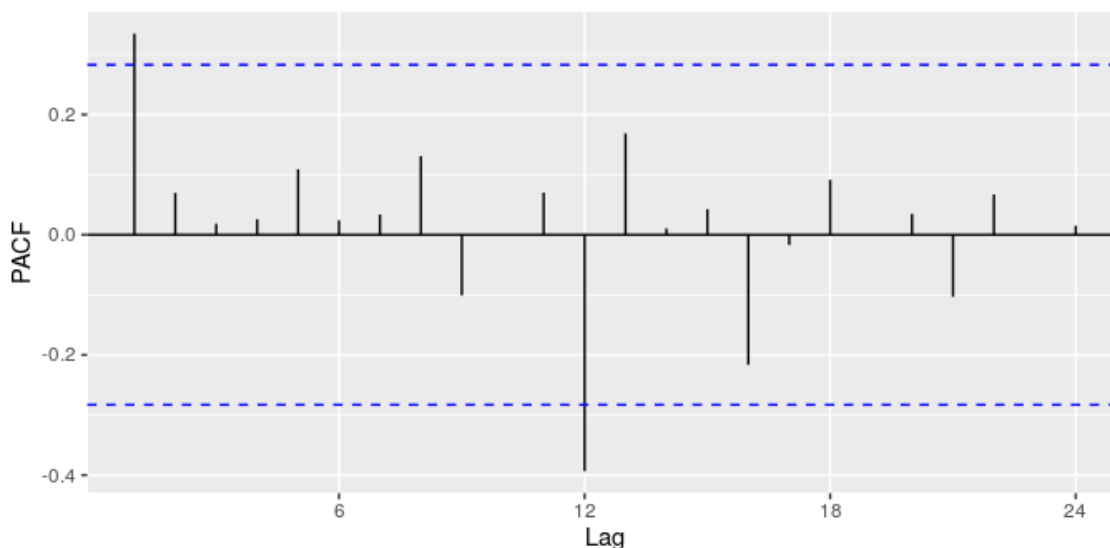


Fig 4:- Plot of Partial Autocorrelation Function of Series dlnKOREA

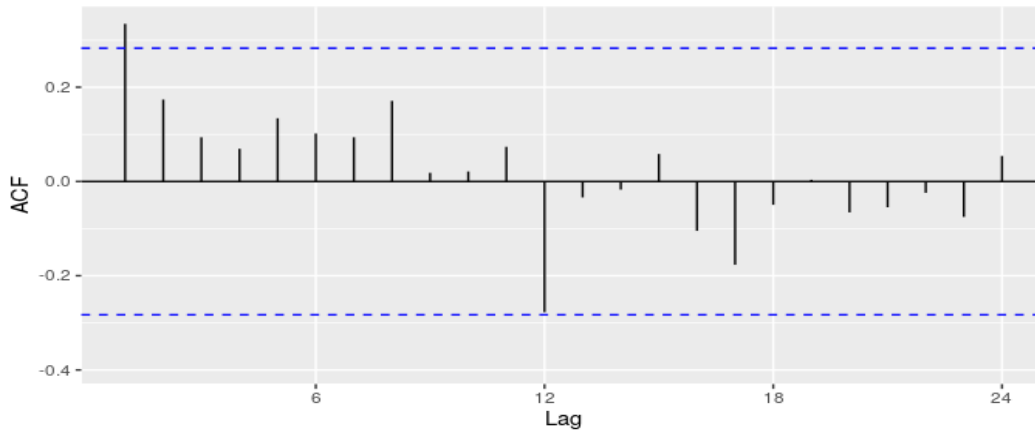


Fig 5:- Plot of Autocorrelation Function of Series dlnKOREA

➤ *Picking the Model Order*

There is a significant spike at lag 1 in the ACF plot suggesting a non-seasonal MA(1) component. There is also a significant spike at lag 1 in the PACF plot suggesting a non-seasonal AR(1) component. There is a significant spike at lag 12 in the PACF plot suggesting a seasonal AR(1) component. With that we got the order of the model $ARIMA(1,0,1)(1,1,0)_{12}$ indicating a non-seasonal MA and AR component, a seasonal AR component, and a seasonal difference, however, we also tried orders without non-seasonal AR or MA terms as well such as $ARIMA(1,0,0)(1,1,0)_{12}$, and $ARIMA(0,0,1)(1,1,0)_{12}$. We then compared their AICc values to determine the best order of our ARIMA model (the one with the lowest AICc value) moving forward.

Model order	AICc values
$ARIMA(1,0,1)(1,1,0)_{12}$	-32.68
$ARIMA(1,0,0)(1,1,0)_{12}$	-34.46
$ARIMA(0,0,1)(1,1,0)_{12}$	-34.08

Table 3:- Different initial ARIMA model orders and their respective AICc values

Since the model $ARIMA(1,0,0)(1,1,0)_{12}$ has the lowest AICc value among the three, we will be using this model moving forward.

➤ *Analyzing the Residuals to Determine Lacking Terms*

After choosing the ARIMA model we then plotted the ACF and PACF of the residuals of the $ARIMA(1,0,0)(1,1,0)_{12}$ model to determine any terms we might have missed.

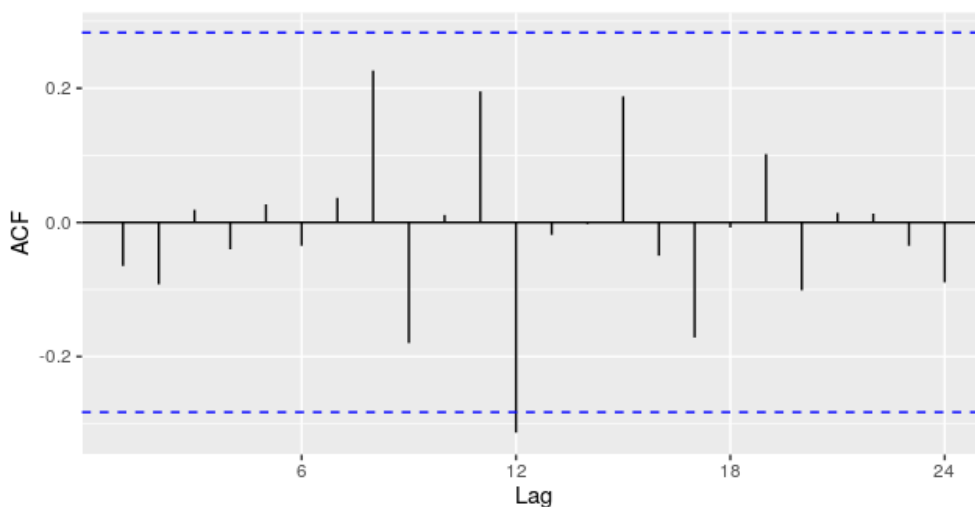


Fig 6:- ACF of the Residuals of $ARIMA(1,0,0)(1,1,0)_{12}$

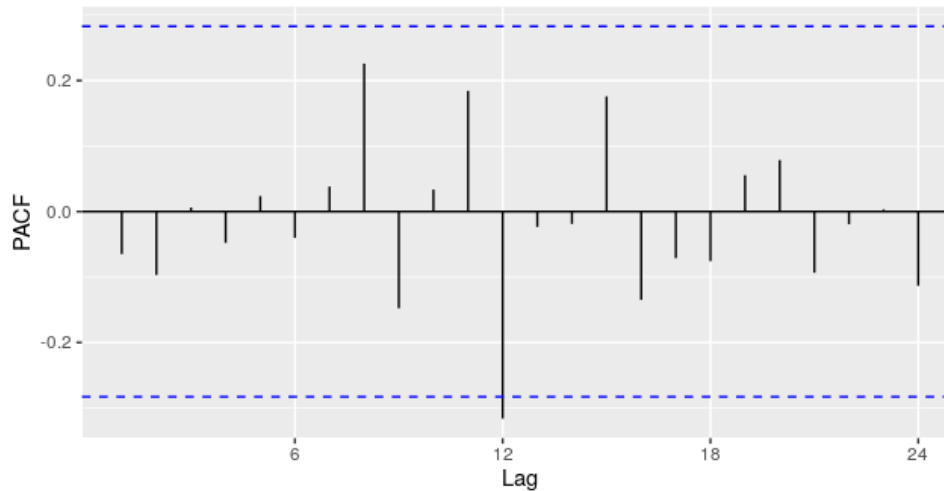


Fig 7:- PACF of the Residuals of ARIMA(1,0,0)(1,1,0)₁₂

In both the ACF and PACF plots of the residuals there was a significant spike in lag 12 and an almost significant spike at lag 8. This meant that there was an additional seasonal term that should be included in the model. We also tried to include another non-seasonal term in the model and compared the different variations based on their AICc values to determine the best fitted model. The different variations to the model we compared were an ARIMA(1,0,0)(1,1,1)₁₂, an ARIMA(1,0,0)(2,1,0)₁₂, and an ARIMA(2,0,0)(0,1,2)₁₂

Model order	AICc values
ARIMA(1,0,0)(1,1,1) ₁₂	-35.52
ARIMA(1,0,0)(2,1,0) ₁₂	-36.18
ARIMA(2,0,0)(0,1,2) ₁₂	-33.96

Table 4:- Different improved ARIMA model orders and their respective AICc values

Since the model ARIMA(1,0,0)(2,1,0)₁₂ has the lowest AICc value among the three, we used this model moving forward.

➤ *Reanalyzing the Residuals of the Improved ARIMA Model*

We once again plotted the residuals of the ARIMA(1,0,0)(2,1,0)₁₂ model to determine if there were any terms we might have missed.

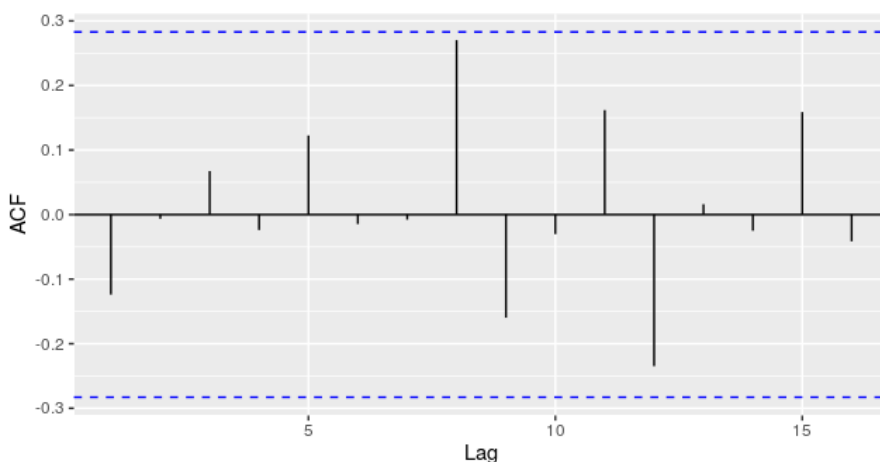


Fig 8:- ACF of the Residuals of ARIMA(1,0,0)(2,1,0)₁₂

Now all the spikes appeared to be within the threshold suggesting that the residuals are now acting like white noise. Furthermore, a Ljung-Box test conducted on the residuals of the data confirmed that there are no more autocorrelations remaining.

Ljung-Box Test

p-value	0.3395
---------	--------

Table 5:- P-value of the Ljung-Box test of the Residuals of ARIMA(1,0,0)(2,1,0)₁₂

This suggests our model is ready to be used to forecast the data.

➤ *Forecasting the Data (Point Forecast Equation)*

First we multiplied the seasonal and non-seasonal AR components which are represented by $(1 - \phi_1 B^{12} - \phi_2 B^{24})$ and $(1 - \phi_1 B)$ respectively. We then forecasted the data, with the point forecast equation of the data.

$$(1 - \phi_1 B^{12} - \phi_2 B^{24})(1 - \phi_1 B)(12 - B)y_t = \varepsilon_t \quad (3.6.1)$$

with ϕ_1 and ϕ_2 representing the seasonal AR(1) and AR(2) terms, ϕ_1 representing the non-seasonal AR(1) term, and $(12 - B)y_t$ representing the seasonal differencing.

$$[1 - \phi_1 B - \phi_1 B^{12} + (\phi_1)(\phi_1)B^{13} - \phi_2 B^{24} + (\phi_2)(\phi_1)B^{25}](12 - B)y_t = \varepsilon_t \quad (3.6.2)$$

We then multiplied the product of the AR components with the seasonal difference which is $(12 - B)y_t$, and applied the backshift notation.

$$[12 - (1 - 12\phi_1)B + \phi_1 B^2 - 12\phi_1 B^{12} + (\phi_1 + 12\phi_1 \phi_1)B^{13} - \phi_1 \phi_1 B^{14} - 12\phi_2 B^{24} + (\phi_2 + 12\phi_2 \phi_1)B^{25} - \phi_2 \phi_1 B^{26}]y_t = \varepsilon_t \quad (3.6.3)$$

$$12y_t - (1 - 12\phi_1)y_{t-1} + (\phi_1)y_{t-2} - (12\phi_1)y_{t-12} + (\phi_1 + 12\phi_1 \phi_1)y_{t-13} - \phi_1 \phi_1 y_{t-14} - 12\phi_2 y_{t-24} + (\phi_2 + 12\phi_2 \phi_1)y_{t-25} - \phi_2 \phi_1 y_{t-26} = \varepsilon_{t-26} \quad (3.6.4)$$

We then substituted ϕ_1, ϕ_2, ϕ_1 with their respective AR terms:

$$12y_t - 6.9604y_{t-1} + 0.4967y_{t-2} + 14.1684y_{t-12} - 8.22174428y_{t-13} + 0.58645369y_{t-14} + 7.716y_{t-24} - 4.4755372y_{t-25} + 0.319378y_{t-26} = \varepsilon_t \quad (3.6.5)$$

Then we transpose all values from the left-hand side to the right-hand side except $12y_t$:

$$12y_t = 6.9604y_{t-1} - 0.4967y_{t-2} - 14.1684y_{t-12} + 8.22174428y_{t-13} - 0.58645369y_{t-14} - 7.716y_{t-24} + 4.4755372y_{t-25} - 0.319378y_{t-26} + \varepsilon_t \quad (3.6.6)$$

We then replaced t with $T + h$, with y_T representing our observations up to time T and h representing the additional period of our data we will be forecasting, and we replaced ε_{T+h} with 0:

$$12y_{T+h}|T = 6.9604y_T - 0.4967y_{T-1} - 14.1684y_{T-11} + 8.22174428y_{T-12} - 0.58645369y_{T-13} - 7.716y_{T-23} + 4.4755372y_{T-24} - 0.319378y_{T-25} + \varepsilon_{T+h} \quad (3.6.7)$$

For a forecast of y_{T+2} we replace t with $T + 2$ in (3.6.6).

$$12y_{T+2}|T = 6.9604y_{T+1} - 0.4967y_T - 14.1684y_{T-10} + 8.22174428y_{T-11} - 0.58645369y_{T-12} - 7.716y_{T-22} + 4.4755372y_{T-23} - 0.319378y_{T-24} + \varepsilon_{T+2} \quad (3.6.8)$$

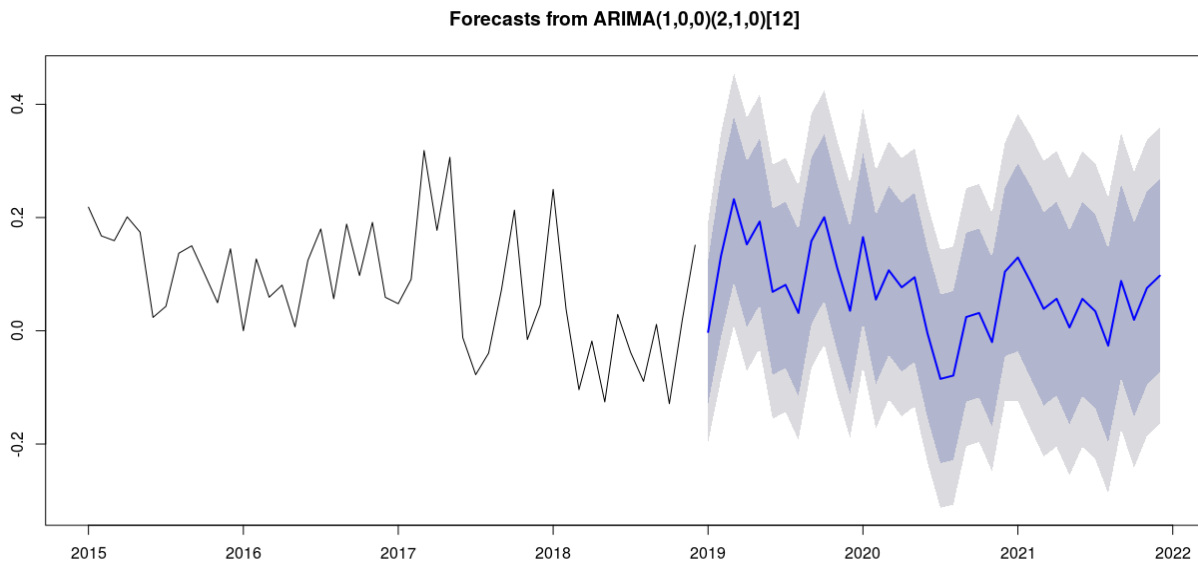


Fig 9:- Forecast of South Korean Arrivals into the Philippines for 2019-2021

IV. RESULTS AND DISCUSSION

Through our statistical analysis of South Korean arrivals into the country we discovered an upward trend of increasing South Korean arrivals over time. Similarly, according to data we got from <https://www.macrotrends.net/countries/KOR/south-korea/gdp-gross-domestic-product>, South Korea's economy has also been on an upward trend. These 2 data support Miralao's (2007) theory that South Korean arrivals in the Philippines is correlated to South Korea's increasing prosperity.

Through the statistical analysis we also discovered 3 peak seasons of South Korean arrivals every year. With the 1st peak being in January-February, the 2nd peak being in July -August, and the 3rd peak being in December. The 1st and 3rd peaks of every year can be explained by the Christmas to New Year winter holiday season, while the 2nd peak can be explained by Korea's summer season which is also on July-August.

We forecasted 3 years into the future to 2021. The plot of the forecast of the data showed 4 prediction intervals. 2 of the prediction intervals allowed for the data to increase but stabilized at a certain level. The other 2 prediction levels allowed for the data to trend downwards. The point forecast of the data trends slightly downwards. The forecast of the data allows South Korean arrivals to the Philippines to increase slightly or decrease, however it does not allow the data to continue its upward trend. The forecasted data also shows that the 3 peaked seasonality of the data is maintained although the 2nd peak has seen a significant decrease while the 1st and 3rd peaks maintain their usual levels.

V. CONCLUSION AND RECOMMENDATIONS

A. Conclusion

From the forecast of the model we can see that South Korean arrivals will soon either cease to increase or, worse, decrease. This has a large implication for a lot of sectors in Philippine society, mainly local tourist-inclined businesses. Niche-market businesses that cater to mostly Korean customers such as Korean KTVs and Korean supermarkets may lose a significant portion of their customer base. English-training schools may also be affected if these predictions were to come true because this would mean a decline of new Korean students coming to the Philippines for English training.

These predictions are also indicative of a change in either Korean society and culture or Philippine society and culture. One theory we have is that the rise in online platforms for education are now preferable to Koreans to learn English because of their accessibility and cost rather than coming to the Philippines for English training.

B. Recommendations

- **Sample Size** - The data used for South Korean arrivals into the Philippines ranged from January 2014-December 2018. A larger sample size may provide a clearer image of the trend of South Korean arrivals into the Philippines.
- **Methodology**- This study uses an ARIMA modelling approach to time series forecasting, however an exponential smoothing approach may yield better results due to the strong presence of trend and seasonality in the data.

- **Comparing the results with other studies of other nationalities' foreign arrivals to the Philippines-** The forecast of the data predicts a downward trend of South Korean arrivals into the Philippines. It would be interesting to know whether this forecast is isolated to South Korean arrivals or is this a decline of Philippine tourism as a whole.

REFERENCES

- [1]. *ASIA PACIFIC: Regional Tourism Trends* (Market Report). (2018, May). Retrieved May 24, 2020, from Horwath HTL website: https://corporate.cms-horwathhtl.com/wp-content/uploads/sites/2/2018/05/MR_AP_REGIONAL-TOURISM-TRENDS.pdf
- [2]. Blackall, M. (2019, July 1). Global tourism hits record highs - but who goes where on holiday? *The Guardian*. Retrieved May 24, 2020, from <https://www.theguardian.com/news/2019/jul/01/global-tourism-hits-record-highs-but-who-goes-where-on-holiday>
- [3]. Gomez, J. A., Jr. (2013). The Korean Diaspora in Philippine Cities: Amalgamation or Invasion? In J. Hou (Ed.), *Transcultural Cities: Border-Crossing and Placemaking* (pp. 77-90). New York, NY: Routledge.
- [4]. Lee, H.-K. (2005). The Korean Diaspora and its Impact on Korea's Development. *Asian and Pacific Migration Journal*, 14(1-2), 149-168. <https://doi.org/10.1177/011719680501400108>
- [5]. Mapa, C. S. (2019, May 6). *Press Release* (2019-093). Retrieved May 24, 2020, from Philippine Statistics Authority website: <https://psa.gov.ph/sites/default/files/2018%20PTSA%20Press%20Release.pdf>
- [6]. Rufino, Cesar & Professor, Distinguished. (2013). *KOREAN INBOUND TOURISM TO THE PHILIPPINES – A MACROECONOMETRIC EVALUATION*.
- [7]. Tourism Demand Statistics. Retrieved May 25, 2020, from http://www.tourism.gov.ph/tourism_dem_sup_pub.asp

x