ISSN (Print) : 0974-6846
Indian Journal of Science and Technology, Vol 11(47), DOI: 10.17485/ijst/2018/v11i47/121923, December 2018 ISSN (Online) : 0974-5645

# Forecasting the Monthly Cases of Human Immunodeficiency Virus (HIV) of the Philippines

**Renato E. Apa-ap* and Hero L. Tolosa**

Polytechnic University of the Philippines, College of Science, Department of Mathematics and Statistics; r_apaap@yahoo.com, or3h.arym@gmail.com

## Abstract

**Objective**: To build a best model in forecasting the monthly cases of Human Immunodeficiency Virus (HIV) of the Philippines. **Methods/Statistical Analysis**: The researchers gathered time series data which is monthly HIV new cases from January 2011 to December 2015 from the monthly records of HIV/AIDS new cases in the Philippines from the HIV/AIDS and ART Registry of the Philippines (HARP) under the umbrella of Department of Health. The researchers utilized advanced statistical tool in developing the model using Univariate Box-Jenkins method in forecasting the HIV cases per month. **Findings**: The result shows that the monthly cases of HIV in the Philippines has an upward trend, it was observed that the highest peak was on June 2015 with 772 new cases. From the given time, December 2015, It shows that it increases by 82.39% number of HIV cases at the end of December 2016. Satisfying the assumptions of the said method and its model diagnostic checking, the researchers came up with the best model based on AIC is SARIMA (2,1,0) (0,0,1)[12] with drift. The model is Number of HIV Cases$_{(Predicted)}$ = (24.19) Inflation (All Items)– (61.39) Inflation(Health). The developed model of the researchers could help the Department of Health in allocating line item budget for the implementation of prevention projects for HIV cases and intensify measures in monitoring such cases. **Applications/Improvements**: Results from the study can be a basis/evidence to raise awareness about the epidemic cases of HIV in the country. Researchers suggest that a time series analysis of HIV cases by gender be explored.

**Keywords:** Cases, Forecasting, HIV, Monthly, Philippines

## 1. Introduction

HIV which stands for "Human Immunodeficiency Virus" is a serious disease that is cause by a virus that spread through the body fluids which attacks the body immune system just like cancer and can lead to death.

In the Philippines, HIV is also an alarming disease in which the government is doing enough effort to minimize if not eradicate it. Philippines undeniably is belonging to the third word and majority are Christians where safe sexual intercourse are still in argument of which is better and acceptable to the culture and beliefs of Filipinos.

Using the earliest known sample of HIV, scientists have been able to create a 'family-tree' ancestry of HIV transmission which allowing them to discover where HIV started. According to the study[1], that the first transmission of SIV to HIV in humans took place around 1920 in Kinshasa in the Democratic Republic of Congo. It's been known for having the most genetic diversity in HIV strains than anywhere else that reflects the number of different SIV was transferred to humans. It is also found that many of the first cases of AIDS were recorded on the said location.

The area around Kinshasa is full of transport ways such as roads, railways and rivers. Through this, most of the foreigner looks for a pleasure and had sex with others which is a growing trade around this region. The high population of migrants and sex trade will explain how HIV spread along these infrastructure routes, firstly to Brazzaville by[2].

*Author for correspondence

Based on the University of Oxford News (2014)[3], UniHIV-1 subtype B is now the most geographically spread subtype of HIV internationally, with 75 million infections to date. Some people says that the HIV in the United States of America started in 1980 but this was the first time that the people conscious in HIV and it is officially recognized as a new health condition. In 1981, a few cases of rare diseases were being reported among gay men in New York and California, such as Kaposi's sarcoma (a rare cancer) and a lung infection called PCP[4]. No one knew why these cancers and unscrupulous infections were spreading, and that there must be an infectious 'disease' causing them. Based on the study[5], the first disease was called relating to the word 'gay'. Based on the report of the Centers for Disease Control[6], it was realized that the disease was also spreading among other individual such as haemophiliacs and heroin users. By September that year, the 'disease' was finally named AIDS.

When cases of AIDS started to emerge in the USA, the majority of cases were among men who have sex with men (homosexuals), haemophiliacs and heroin users. Adding Haitians to above list, it was name as the 'Four-H Club' group which are at high risk of AIDS as revealed from the study of Farmer[7].

Last August 27, 1976 from the website of Uhavax Hartford statement[8], it cited that one-third of all new cases were coming from service men returning from the Philippines. Even the occurrence of the HIV and AIDS in the Philippines is still low, as reported in the article of UNICEF on their website, according to UNICEF[9] that the Philippines is one of the seven countries globally where the number of new HIV cases has increased by over 25 percent from 2001 to 2009. There are great focused among key population reported with specific risk behaviors, such as unprotected male-to-male sex, transactional sex and intravenous drug use. The primary prevention of HIV infection for the huge population concentrated in the Philippines which focus on the age bracket 20-29 years old. However, the commencement to sex and drug use is between 14 to 19 years old. Furthermore, there were five percent (5%) of HIV-positive pregnant women have received antiretroviral medicines to avoid mother-to-child transmission.

Bestowing to the article of Trivedi[10] reported by Anjani Trivedi that there were 24,000 cases reported last 2012 and the transmission is now growing rapidly. It was reported also that there were a new cases of HIV every three days. On the same article of these days, it is every three hours which is a terrifying statistics that makes the Southeast Asian nation one of only nine in the world where transmission is on the rise. Based on the report from the Department of Health in the Philippines 2013, it shows 415 new HIV cases in May, a 52% spike over the same period last year and the highest since 1984. According to the department, 9 out of 10 new cases are men, mostly under 30.

From January 1984 to December 2014, there were already 1,118 accumulated reported deaths among people with HIV disease infection in the Philippines. As of now, no vaccine or cure on this infection, only Anti-retroviral medication (ARVs) which only slows down the replication of the virus which still causes or lead to death of a patient. Based on the website of Index mundi it is estimated that there were 3,120 cases accumulated from year 1999 to year 2012. According also in the website of United Nation AIDS (UNAIDS)[11] that an estimated of less than 1000 deaths due to AIDS.

It was reported on the Philippine Daily Inquirer last 2014[12] by Tina G. Santos on April 1, 2014 according to Santos, there were sixteen (16) died of AIDS from January to February 2014. The highest number of deaths occurred in the 25-29 age group, followed by the 20-24 and the 30-34 age groups based on the Department of Health (DOH) data. For the month of February 2014 alone, eight deaths were reported. Moreover, the Sun Star Manila reported last April 2016[13] that there were 55 Acquired Immune Deficiency Syndrome (AIDS) deaths and 715 new cases of human immunodeficiency virus (HIV) reported in February, according to the Department of Health (DOH). "This was 16 percent higher compared to the same period last year, which is at 646," said the latest HIV/AIDS Registry of the Philippines.

In lieu from the said reports, researchers plan to develop a model in estimating the monthly cases of HIV cases in the Philippines using Box Jenkins statistical tool in predicting or forecasting the monthly HIV cases in the Philippines. The model may be used for budget allocation for the implementation of possible prevention policy. Moreover, it can be used to monitor on what category the Department of Health should focus on their campaign or project in fighting for the epidemic.

### 1.1 Objective of the Study

Using time series analysis, this study aims to develop a best model that can predict the monthly HIV cases in the Philippines by means of the Seasonal Autoregressive Integrated Moving Average (SARIMA) with Box-Jenkins method.

### 1.2 Statement of the Problem

Specifically, this research seeks to answer the following questions:

1. What is the behavior of the HIV monthly cases in the Philippines?
2. What is the best model to predict the HIV monthly cases in the Philippines?
3. What is the interpretation of the forecast produced by the time series analysis?

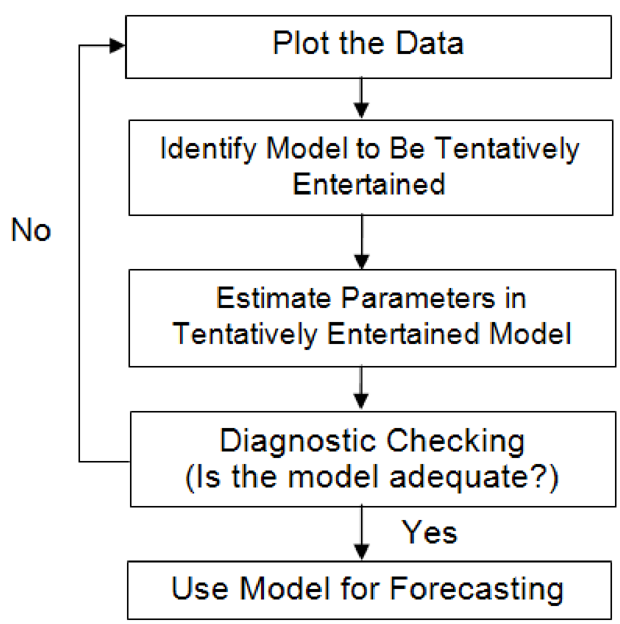
### 1.3 Scope and Limitation

The researcher's data used in this study came from the monthly records of HIV/AIDS new cases in the Philippines from January 2011 to December 2015 of the HIV/AIDS and ART Registry of the Philippines (HARP) under the umbrella of the Department of Health. This study will build a model that will forecast monthly HIV cases for the next 5 years.

### 1.4 Significance of the Study

This study could be used to forecast HIV cases monthly and allow people to prepare for the figures resulted from the forecast thus, reducing the decision risk. Furthermore, this study can be also used by the concerned government agencies in making decisions in implementing their precautionary measures in reducing the wide spread of HIV cases in the Philippines. For government officials, this may help them form a new policy in Health and Education of the country.

### 1.5 Research Paradigm




## 2. Methodology

The HIV new cases data obtained from the department of Health through HARP from year 2011 to 2015 were exploited in building a mathematical model done by advanced statistical tool SARIMA through the use of Box-Jenkins method with the help of Statistical software for accuracy of computations.

Time series analysis accounts for the fact that data points taken over time may have an internal structure (such as autocorrelation, trend or seasonal variation) that should be accounted for.

### 2.1 Box-Jenkins Models

The method of model building is a combination of the Autoregressive (AR) and Moving Average (MA) models and a self- forecasting method.

$$Xt = \delta + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} \quad [14]$$
$$+A_t - \theta_1 A_{t-1} - \theta_2 A_{t-2} - \cdots - \theta_q At - q,$$

where, the terms in the equation have the same meaning as given for the AR and MA model.

The Box-Jenkins model assumes that the time series is stationary. Box and Jenkins recommend differencing non-stationary series one or more times to achieve stationarity. Doing so produces an ARIMA model, with the "I" standing for "Integrated".

Some formulations transform the series by subtracting the mean of the series from each data point. This yields a series with a mean of zero. Whether you need to do this or not is dependent on the software you use to estimate the model.

Box-Jenkins models can be extended to include seasonal autoregressive and seasonal moving average terms. Although this complicates the notation and mathematics of the model, the underlying concepts for seasonal autoregressive and seasonal moving average terms are similar to the non-seasonal autoregressive and moving average terms.

The most general Box-Jenkins model includes difference operators, autoregressive terms, moving average terms, seasonal difference operators, seasonal autoregressive terms, and seasonal moving average terms. As with modeling in general, however, only necessary terms should be included in the model. Those interested in the mathematical details can consult Box[15–17].

There are three primary stages in building a Box-Jenkins time series model.

- Model Identification
- Model Estimation
- Model Validation

The Model Identification comprises several methods of diagnostics to ensure that the assumptions are meeting, one is Stationarity and Seasonality.

The first step in developing a Box-Jenkins model is to determine if the series is stationary and if there is any significant seasonality that needs to be modeled. Detecting stationarity can be accessed from a run sequence plot. The run sequence plot should show constant location and scale. It can also be detected from an autocorrelation plot. Specifically, non-stationarity is often indicated by an autocorrelation plot with very slow decay. Detecting seasonality (or periodicity) can usually be assessed from an autocorrelation plot, a seasonal subseries plot, or a spectral plot. Differencing to achieve stationarity in Box and Jenkins recommend the differencing approach to achieve stationarity. However, fitting a curve and subtracting the fitted values from the original data can also be used in the context of Box-Jenkins models. Seasonal differencing at the model identification stage, our goal is to detect seasonality, if it exists, and to identify the order for the seasonal autoregressive and seasonal moving average terms. For many series, the period is known and a single seasonality term is sufficient.

For example, for monthly data we would typically include either a seasonal AR 12 term or a seasonal MA 12 term. For Box-Jenkins models, we do not explicitly remove seasonality before fitting the model. Instead, we include the order of the seasonal terms in the model specification to the ARIMA estimation software. However, it may be helpful to apply a seasonal difference to the data and regenerate the autocorrelation and partial autocorrelation plots. This may help in the model identification of the non-seasonal component of the model. In some cases, the seasonal differencing may remove most or all of the seasonality effect.

Once stationarity and seasonality have been addressed, the next step is to identify the order (i.e., the p and q) of the autoregressive and moving average terms.

The primary tools for doing this are the autocorrelation plot and the partial autocorrelation plot. The sample autocorrelation plot and the sample partial autocorrelation plot are compared to the theoretical behavior of these plots when the order is known. Specifically, for an AR(1) process, the sample autocorrelation function should have an exponentially decreasing appearance. However, higher-order AR processes are often a mixture of exponentially decreasing and damped sinusoidal components. For higher-order autoregressive processes, the sample autocorrelation needs to be supplemented with a partial autocorrelation plot. The partial autocorrelation of an AR(p) process becomes zero at lag  and greater, so we examine the sample partial autocorrelation function to see if there is evidence of a departure from zero. This is usually determined by placing a 95 % confidence interval on the sample partial autocorrelation plot (most software programs that generate sample autocorrelation plots will also plot this confidence interval). If the software program does not generate the confidence band, it is approximately with N denoting the sample size.

The autocorrelation function of a MA (q) process becomes zero at lag q+1 and greater, so we examine the sample autocorrelation function to see where it essentially becomes zero. We do this by placing the 95 % confidence interval for the sample autocorrelation function on the sample autocorrelation plot. Most software that can generate the autocorrelation plot can also generate this confidence interval. The sample partial autocorrelation function is generally not helpful for identifying the order of the moving average process. The following table summarizes how we use the sample autocorrelation function for model identification.

| SHAPE | INDICATED MODEL |
|---|---|
| Exponential, decaying to zero | Autoregressive model. Use the partial autocorrelation plot to identify the order of the autoregressive model. |
| Alternating positive and negative, decaying to zero | Autoregressive model. Use the partial autocorrelation plot to help identify the order. |
| One or more spikes, rest are essentially zero | Moving average model, order identified by where plot becomes zero |
| Decay, starting after a few lags | Mixed autoregressive and moving average model. |
| All zero or close to zero | Data is essentially random. |
| High values at fixed intervals | Include seasonal autoregressive term. |
| No decay to zero | Series is not stationary. |

In practice, the sample autocorrelation and partial autocorrelation functions are random variables and will not give the same picture as the theoretical functions. This makes the model identification more difficult. In particular, mixed models can be particularly difficult to identify. Although experience is helpful, developing good models using these sample plots can involve much trial and error. For this reason, in recent years information-based criteria such as FPE (Final Prediction Error) and AIC (Akaike Information Criterion)[18] and others have been preferred and used. These techniques can help automate the model identification process. These techniques require computer software to use. Fortunately, these techniques are available in many commercial statistical software programs that provide ARIMA modeling capabilities.

For additional information on these techniques[16]. In Box-Jenkins Model Estimation the use of Statistical Software is recommended.

Estimating the parameters for the Box-Jenkins models is a quite complicated non-linear estimation problem. For this reason, the parameter estimation should be left to a high quality software program that fits Box-Jenkins models. Fortunately, many commercial statistical software programs now fit Box-Jenkins models.

The main approaches to fitting Box-Jenkins models are non-linear least squares and maximum likelihood estimation.

Maximum likelihood estimation is generally the preferred technique. The likelihood equations for the full Box-Jenkins model are complicated and are not included here. In[16] the mathematical details.

Model diagnostics for Box-Jenkins models is similar to model validation for non-linear least squares fitting.

That is, the error term is assumed to follow the assumptions for a stationary univariate process. The residuals should be white noise (or independent when their distributions are normal) drawings from a fixed distribution with a constant mean and variance. If the Box-Jenkins model is a good model for the data, the residuals should satisfy these assumptions.

If these assumptions are not satisfied, we need to fit a more appropriate model. That is, we go back to the model identification step and try to develop a better model. Hopefully the analysis of the residuals can provide some clues as to a more appropriate model.

As discussed in the EDA chapter, one way to assess if the residuals from the Box-Jenkins model follow the assumptions is to generate a 4-plot of the residuals and an autocorrelation plot of the residuals. One could also look at the value of the Box-Ljung (1978) statistic.

## 3. Results and Discussion

The monthly cases of HIV in the Philippines time series was normal on its level but stationary at first differenced. The rages of monthly cases occurred from year 2011-2015 is from 150-780 cases and the highest peak happened on June 2015 with 772 cases, furthermore the seasonal peaks happens during month of September. The result shows that there is a serial correlation within the series and seasonality is present at one percent level.

Since all time series data for this study are identified as seasonal by the Kruskal -Wallis H Statistic, seasonal patterns were identified by the researchers using visualized graphs. In general, seasonal fluctuations were found mostly months of November and April.

Moreover, the analysis of residuals proved that the "best" models created by the specifications of X-12 ARIMA obtained no violation of assumptions of a time series model. Models passed the normality, independence, and constant variance tests.

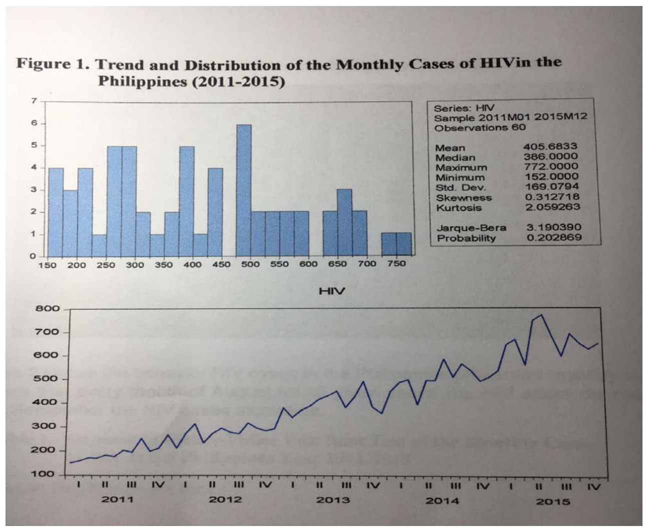



**Figure 1.** Trend and distribution of the monthly cases of HIV in the Philippines (2011-2015).

The fitted number of HIV infections was calculated by optimum ARIMA (2,1,0) model from 2011-2015. The fitted number or the inbound forecast was similar to the observed number of HIV cases, with a MAPE of 5.05%. The forecasted number of new HIV cases for January 2016 was 672.

Figure 1 shows the descriptive statistics of the monthly HIV cases from 2011 to 2015. It shows that the behavior of the monthly HIV cases has an increasing trend, somewhat has a cycle and seasonality as well. It found out that the highest increase of the HIV cases is on April to May 2015 with 33.57% and the highest decrease of the HIV cases in on October 2013 December 2013 with 21.79%. Base on the results, it shows that the distribution of the monthly HIV cases is positively skewed and it also shows that the data is found to be platykurtic. Moreover, the value of the computed probability of Jarqu-bera test is 0.2029 compare to alpha 5% which means that the data follows normal distribution.

Figure 2 shows the behavior HIV cases in the Philippines presented monthly every year. It shows that every month of August for all given years, the HIV cases decreases while every September the HIV cases increases.

Table 1 shows the Augmented Dickey Fuller test to identify whether the data is stationary or not. Base on the computed p-value of ADF test which is equal to 0.8474, there is no sufficient evidence to Reject the Ho. This means that the given data has a unit root or the data is not stationary.

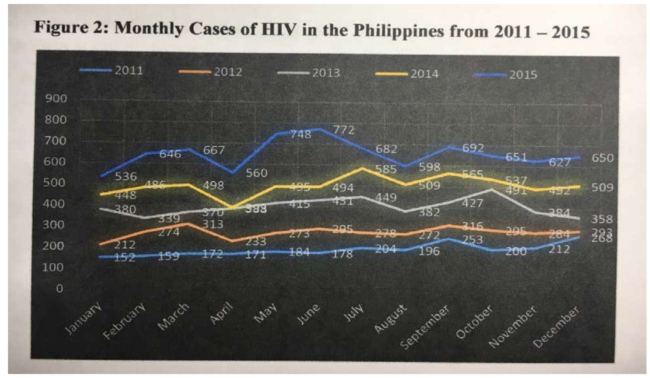



**Figure 2.** Monthly cases of HIV in the Philippines from 2011-2015.

**Table 1.** Augmented Dickey-Fuller Unit Root Test of the Monthly Cases of HIV in the Philippines Year2011-2015

| | | |
|---|---|---|
| **Null Hypothesis: HIV has a unit root** | | |
| **Exogenous: Constant** | | |
| **Lag Length: 2 (Automatic - based on SIC, maxlag=10)** | | |
| t-Statistic<br>Augmented Dickey-Fuller test statistic<br>-.662818543380155<br>Test critical values:<br>1% level<br>-.550395660058178<br>5% level<br>-.913549248241188<br>10% level<br>-2.594521200061805 | | Prob.*<br>0.8474315970945149 |

**Table 2.** Augmented Dickey-Fuller Unit Root Test of the Transformed Monthly Cases of HIV in the Philippines Year 2011-2015

| | | |
|---|---|---|
| **Null Hypothesis: DLOG(HIV) has a unit root** | | |
| **Exogenous: Constant** | | |
| **Lag Length: 1 (Automatic - based on SIC, maxlag=10)** | | |
| t-Statistic<br>Augmented Dickey-Fuller test statistic<br>-10.34194872203377<br>Test critical values:<br>1% level<br>-3.550395660058178<br>5% level<br>-2.913549248241188<br>10% level<br>-2.594521200061805 | | Prob.*<br>1.529715467077728e-11 |

Table 2 shows the transformed data to test unit root test. Base on the computed p-value of ADF test which is equal to 0.0000, there is a sufficient evidence to Reject the Ho. This means that the given data does not have a unit root or the data is stationary.

Table 3 shows the model identification from the monthly working data from 2011-2013 of HIV cases in the Philippines. The following data use was transformed into logarithmic for remedial and satisfaction of the assumption of the model. It is found that the best model to forecast the monthly is ARIMA (2,1,0) (0,0,1)[12] with drift.

**Table 3.** Estimated Model based on the Akaike Information Criterion (AIC)

| MODELS | | |
|---|---|---|
| SARIMA(2,1,2) (1,0,1)[12] | with drift | Inf |
| ARIMA(0,1,0) | with drift | -33.3685 |
| SARIMA(1,1,0) (1,0,0)[12] | with drift | -37.3569 |
| SARIMA(0,1,1) (0,0,1)[12] | with drift | Inf |
| ARIMA(0,1,0) | | -34.3421 |
| ARIMA(1,1,0) | with drift | -39.2717 |
| SARIMA(1,1,0) (0,0,1)[12] | with drift | -37.388 |
| SARIMA(1,1,0) (1,0,1)[12] | with drift | Inf |
| ARIMA(2,1,0) | with drift | -51.9119 |
| ARIMA(2,1,1) | with drift | -50.3113 |
| ARIMA(3,1,1) | with drift | Inf |
| ARIMA(2,1,0) | | -43.7188 |
| SARIMA(2,1,0) (1,0,0)[12] | with drift | -51.7386 |
| SARIMA(2,1,0) (0,0,1)[12] | with drift | -52.3523 |
| SARIMA(3,1,0) (0,0,1)[12] | with drift | -50.7413 |
| SARIMA(2,1,1) (0,0,1)[12] | with drift | -50.7137 |
| SARIMA(3,1,1) (0,0,1)[12] | with drift | Inf |
| SARIMA(2,1,0) (0,0,1)[12] | | -45.2257 |
| SARIMA(2,1,0) (1,0,1)[12] | with drift | Inf |
| SARIMA(2,1,0) (0,0,1)[12] | with drift | (BEST MODEL) |

Table 4 shows the model estimation of the best model to identify the significant of its parameter. Based on the computed value of the coefficient for each parameter and its standard error, the absolute quotient value of the AR1 and AR2 is 6.0932 and 5.3102 respectively, is greater than 3, it means that there is statistical sufficient evidence to say that the parameters are significant.

**Table 4.** Estimated value of the parameter of the best model

| **Coefficients:** | | | | |
|---|---|---|---|---|
| | ar1 | ar2 | sma1 | drift |
| | -0.8171 | -0.6659 | 0.4233 | 0.0293 |
| s.e. | 0.1341 | 0.1254 | 0.3023 | 0.0086 |

Figure 3 shows the residual plot of the best model created as part of residual diagnostics of the model. This shows that the variance of the error term are seems to be constant. It also shows that the average of the residual is equal to -0.001444353 which is approximately equal to zero.

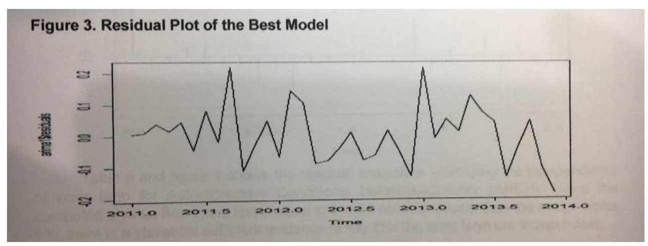



**Figure 3.** Residual plot of the best model.

Table 5 and Figure 4 show the residual analysis in identifying the independency of error term for Autoregressive Conditional Heteroskedasticity (ARCH). Since the computed p-value Box-Ljung test is equal to 0.9355 which is greater than the assign alpha 5%, there is a statistical sufficient evidence to say that the error term is independent.

Table 6 as the residual analysis in identifying the independency of error term for Generalized Autoregressive Conditional Heteroskedasticity (GARCH). Based on the computed p-value of the Box-Ljung test is equal to 0.2675 which is also greater than the assign alpha 5%, it means that there is a statistical sufficient evidence to say that the error term is independent. Table 7 shows the residual analysis to identify the normality of error terms. Since the computed p-value of JarqueBera test is equal to 0.6162, there is statistical evidence to do not reject or fail to reject the null hypothesis of the normality of error term. This means that the error term is normally distributed.

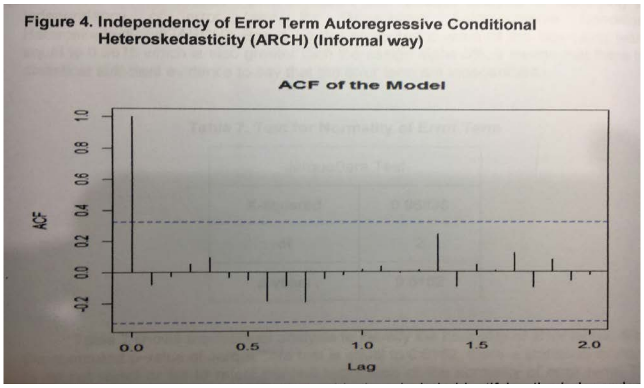



**Figure 4.** Independency of error term autoregressive conditional heteroskedasticity (ARCH) (informal way).

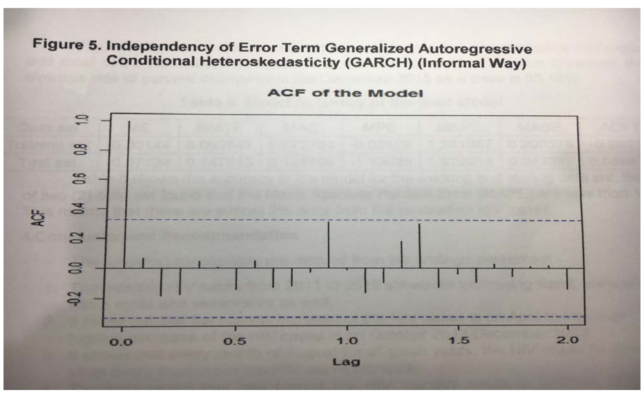



**Figure 5.** Independency of error term generalized autoregressive conditional heteroskedasticity (GARCH) (informal way).

**Table 5.** Independency of error term for Autoregressive Conditional Heteroskedasticity (ARCH) (Formal way)

| Box-Ljung test | |
|---|---|
| X-squared | 11.388 |
| df | 20 |
| p-value | 0.9355 |

**Table 6.** Independency of Error Term for Generalized Autoregressive Conditional Heteroskedasticity (GARCH) (Formal way)

| Box-Ljung test | |
|---|---|
| X-squared | 23.444 |
| df | 20 |
| p-value | 0.2675 |

**Table 7.** Test for Normality of Error Term

| JarqueBera Test | |
|---|---|
| X-squared | 0.96838 |
| df | 2 |
| p-value | 0.6162 |

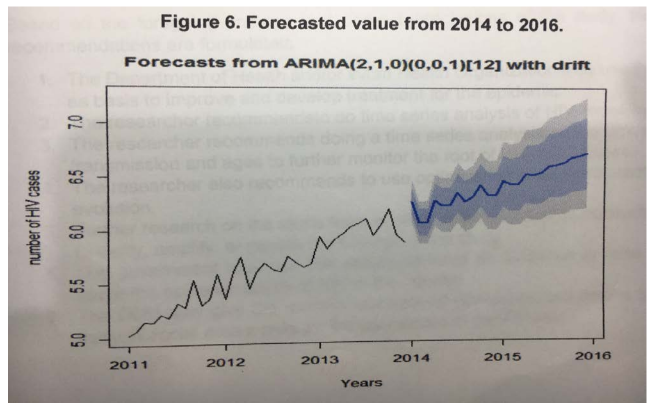



**Figure 6.** Forecasted value from 2014 to 2016.

Figure 5 shows the point forecast (blue line), 80% high and low forecast (dark gray) and 95% high and low forecast (light gray). It shows that the forecasted value from the created model has an increasing trend from 2014 to 2016 while Figure 6 shows the forecasted value with realization of HIV cases from 2014 to 2016. Figure 6 shows that some forecasted value lie in the realization HIV cases and most of the realization HIV cases lie in the 80% forecasted variation Moreover, the average rate of percent change from the December 2015 as a base is 56.16%.

Table 8 shows the accuracy of the model for the working and testing data set. Both of two (2) data set found that the Mean Absolute Percent Error (MAPE) are less

**Table 8.** Model accuracy of the best model

| Data set | ME | RMSE | MAE | MPE | MAPE | MASE | ACF1 | Theil's U |
|---|---|---|---|---|---|---|---|---|
| Training set | -0.00144 | 0.092547 | 0.072784 | -0.03129 | 1.281067 | 0.200375 | -0.08111 | NA |
| Test set | -0.07124 | 0.147513 | 0.124706 | -1.13686 | 1.972014 | 0.343317 | 0.089877 | 1.09631 |

than 10. This means that there is almost 2% error from the realization HIV cases.

## 4. Conclusions and Recommendations

The following conclusions are derived from the findings presented:

1. The monthly HIV cases from 2011 to 2015 shows an increasing trend, somewhat has a cycle and seasonality as well.
2. It found out that the highest increase of the HIV cases is on April to May and highest decrease of the HIV cases in on October 2013 December 2013
3. It shows that every month of August for all given years, the HIV cases decreases while every September the HIV cases increases.
4. The best model that can predict the HIV monthly cases is SARIMA (2,1,0)x (0,0,1)[12] with drift
5. The forecasted value of the created model has an increasing trend
6. The average forecasted value is more than the half of the actual value from December 2015

Based on the foregoing findings and derived conclusions of the study, the following recommendations are formulated:

1. The Department of Health and/or World Health Organization may use these results as basis to improve and develop treatment for the epidemic.
2. The researcher recommends to do time series analysis of HIV cases by gender.
3. The researcher recommends doing a time series analysis on the modes of transmission and ages to further monitor the root of HIV/AIDS cases.
4. The researcher also recommends using optimization such as differential evolution.
5. Further research on the same topic in other locales should be conducted in order to verify, amplify, or negate the findings of the study.
6. The government may use these results to have an evidence to raise awareness about the epidemic cases of HIV in the country.
7. The DOH may give the monthly statistics of HIV cases and deaths to the public through social media sites for the awareness of the Filipino
8. The government may also give seminars about the HIV/AIDS and encourage people to submit their cases in the HIV/AIDS and ART Registry of the Philippines (HARP) in the country.
9. Lastly, DepEd and CHED may also include social health subjects to their curriculum to raise awareness among the Filipino youth.

## 5. References


1. Faria NR. The early spread and epidemic ignition of HIV-1 in human populations. Science. 2014; 346(6205):56–61. https://doi.org/10.1126/science.1256739. PMid:25278604 PMCid:PMC4254776
2. Aids: Origin of pandemic ‘was 1920s Kinshasa [Internet]. [cited 2014 Oct 03]. Available from: https://www.bbc.com/news/health-29442642.
3. HIV pandemic’s origins located [Internet]. [cited 2014 Oct 03]. Available from: http://www.ox.ac.uk/news/2014-10-03-hiv-pandemics-origins-located.
4. Hymes KB. Kaposi’s sarcoma in homosexual men: A report of eight cases. Lancet. 1981; 2(8247):598–600. https://doi.org/10.1016/S0140-6736(81)92740-9
5. Brennan RO, Durack DT. Gay compromise syndrome. Lancet. 1981; 2(8259):1338–9. https://doi.org/10.1016/S0140-6736(81)91352-0
6. Centers for Disease Control. Opportunistic infections and Kaposi’s sarcoma among Haitians in the United States. 1982; 31(26):353–4.
7. Farmer P. AIDS and Accusation: Haiti and the Geography of Blame. University of California Press; 1992. p. 1–338. PMid:1431026
8. Uhavax Hartford website [Internet]. [cited 1992 Mar 19]. Available from: http://uhavax.hartford.edu/bugl/rise.htm.
9. Unicef [Internet]. [cited 2018 Nov 24]. Available from: https://en.wikipedia.org/wiki/UNICEF.
10. Trivedi [Internet]. [cited 2018 Nov 24]. Available from: https://en.wikipedia.org/wiki/Trivedi.
11. Philippine Daily Inquirer [Internet]. [cited 2014 Apr 01]. Available from: http://newsinfo.inquirer.net/590915/hiv-cases-rose-43-to-486-in-february-16-aids-deaths-reported-doh.
12. Sun Star Manila [Internet]. [cited 2016 Apr 05]. Available from: http://www.sunstar.com.ph/manila/local-news/2016/04/05/55-aids-deaths-700-hiv-cases-recorded-february-doh-466140.
13. Time Series Analysis [Internet]. [cited 2018]. Available from: http://th-www.if.uj.edu.pl/zfs/gora/timeseries18/.
14. Jenkins G, Reinsel G. Time Series Analysis. 2nd edition. John Wiley & Sons, Inc.; 1994. p. 1–8.

15. Chatfield C. Model Uncertainty and Forecast Accuracy. Journal of Forecasting. 1996; 15(7):495–508. https://doi.org/10.1002/(SICI)1099-131X(199612)15:7<495::AID-FOR640>3.0.CO;2-O
16. Brockwell P, Davis R. Introduction to Time Series and Forecasting. 2nd Edition. Springer; 2002. p. 1–449. https://doi.org/10.1007/0-387-21657-X_1. https://doi.org/10.1007/b97391
17. Time Series Analysis [Internet]. [cited 2018 Nov 17]. Available from: https://en.wikipedia.org/wiki/Time_series.