**SEATTLE UNIVERSITY**

# Modelling and Forecasting of U.S. GDP

by

Haodi Tu

Huibo Jia

Nancy Jain

ECON 5305

Professor Karo Solat

June 6th, 2019

# CONTENTS

## Table of Contents

# Abstract

The GDP refers to the market value of all final goods and services produced within a country in a given period. According to the CBO, the United States GDP will increase steadily over the next decade from 20.23 trillion U.S. dollars in 2018 to 31 trillion U.S. dollars in 2029. GDP refers to the nation's overall economic activity. The purpose of this paper is to predict the GDP by computing short-term forecasts from the model which are compared with several standard forecasting alternatives. We are using annual and quarterly GDP data of US gathered from Fredcast.com. We did exploratory and other analysis on data to check for trend, seasonality and cyclic presence in data. We also checked if there is any autocorrelation in the data and if it follows any pattern or not. We have performed average, naïve, snaive and drift forecasting to perform and predict the GDP of US. Once a model has been selected and its parameters estimated, the model is used to make forecasts. The performance of the model can only be properly evaluated after the data for the forecast period have become available. Several methods have been developed to help in assessing the accuracy of forecasts. We are going to use those methods to evaluate the performance of the forecasting method we used.

# Introduction

A large part of the work in applied economic analysis for large businesses and governments of today is to forecast a future course for key macroeconomic variables such as Gross Domestic Product (GDP), inflation and unemployment to more effective adjust policies. A time series model can provide a reasonable benchmark to evaluate the values of variables in forecasting economic theory based on the past behavior of the variable. The growth rate of GDP is the most popular and common indicator of the nation's overall economic health.

The data we use for this analytical paper is from Bureau of Economic Analysis. It contains quarterly recorded GDP in billions of dollars from January 1st, 1947 to October 1st, 2018, along with other variables that calculated from GDP like GDP change, percent change of GDP from year ago, compounded annual rate of change, natural log of GDP in billions of dollars and etc… We primarily only use historical GDP records from data for this paper.

In this paper, we conduct a detailed analysis of the forecasting performance of multiple time series models for GDP growth focusing on the US from 1947 to 2018. Our main goal is to establish a time series forecast on US GDP growth and find out which forecasting method fit the real GDP growth the best. Doing so, we have divided our data from January 1947 to October 2009 as our training data set and a testing data set from January 2010 to October 2018.

We consider a large variety of models, including simple time series forecasting methods like average method, naïve method, drift method; more advanced methods like Seasonal and Trend decomposition using Loess (STL method), exponential smoothing methods, and Auto-Regressive Integrated Moving Averages (ARIMA) model. STL decomposition method handles well with any type of seasonality (our data has little seasonality) and is robust to outliers. Exponential smoothing methods allows us to control forecasting trend and seasonality in the data as well as to forecast with various build-in methods like additive, multiplicative and damped-controlling. ARIMA method has high performance in time series forecasting in general and is one of the most popular method in statistical forecasting. After implementing all of these forecasting methods, we conduct the testing data set to test models' accuracy by comparing root mean square error

(RMSE). RMSE is a measure of how spread out these residuals are and represents how concentrated the data is around the line of best fit. Therefore, we've decided to use RMSE as our model criteria to choose the best fitted forecasting model for our GDP data among all.

In next section relevant scholarly papers and issues related to the theory are discussed. Section three and four present the detailed forecasting methods are used in our research along with the data. The following section five presents the result from the forecasts and analysis. And the final section six summarizes the thesis with conclusions drawn.

# Literature Review

Gross Domestic Product (GDP) can reflect the economic development and people's living standards of a country and region, and its structure can reflect the proportional relationship between social production and use, investment and consumption. Although people's understanding of economic development has deepened and realized many of its limitations, at present, it still has a very important guiding significance for economic research and management (Henderson et al., 2012).

In general, GDP has four distinct components, including consumption, private investment, government spending, and net exports. Expressed as: GDP = CA + I + CB + X. Where: CA is consumption, I is private investment, CB is government expenditure, and X is net export value (Čižmešija et al., 2010).

Whether a country or region's economy is in a period of growth or recession can be observed from the change in GDP. In general, the form of GDP publication is no more than two, in terms of total and percentage ratios. When the GDP growth figure is positive, it shows that the region's economy is in an expansion phase; conversely, if it is in a negative number, it means that the region's economy has entered a recession. To be more specific, the sharp increase in a country's GDP reflects the country's booming economy, increased national income, and improved consumption level. In this case, the country's central bank will have the potential to raise interest rates and tighten the money supply, while good national economic performance and rising interest rates will increase the attractiveness of the country's currency. Conversely, if a country’s GDP shows negative growth, indicating that the country’s economy is in a recession. When consumption power is reduced, the country’s central bank will likely cut interest rates to stimulate economic growth again. At this time, interest rates fall and economic performance is sluggish. The attractiveness of the country’s currency has thus also been reduced (Ahmed et al., 2011). Therefore, in general, a high economic growth rate will promote the rise of the national currency exchange rate, while a low economic growth rate will cause the country's currency exchange rate to fall.

In the United States, GDP is estimate by the Bureau of Economic Analysis, or BEA, which is part of the Department of Commerce, and the practice is estimated and counted quarterly. Each time after the preliminary estimates, there will be the first revision and the final revision, and the main publication time is in the third week of each month. GDP is usually used to compare with the same period of last year. If it increases, it means that the economy is developing faster, and it is conducive to the appreciation of its currency. If it is reduced, it means that economic development is slowing down, and its currency is depreciating. In the United States, the ideal level would be growth by 3%, which indicating that economic development is healthy. Above this level, there may existing inflation pressure. However, if below 1.5% growth, it shows signs of economic slowdown and recession.

There are usually two traditional methods for quarterly GDP forecasting (Lee et al., 2008):

1. Qualitative prediction method. This mainly refers to the method by which the forecaster judges the nature and direction of the future development of quarterly GDP based on practical. Due to its strong subjectivity, it is rarely used today.

2. Quantitative prediction method. It mainly refers to the method by which forecasters use traditional statistical data such as government statistics, survey data, and simulation data to predict quarterly GDP trends by using statistical methods or mathematical models, including cross-sectional data regression model, time series analysis model, panel data model and input-output model ( Jiang et al., 2017; Jansen et al.,2016).

This paper will explore and analyze the development of US GDP through time series analysis. Time series analysis is the most effective way to study the dynamic changes of things, involving almost every field of the economy. The basic idea is to establish a mathematical model that can accurately reflect the dynamic dependencies contained in the sequence based on the finite-length data. The model form is finally determined by testing the model, and then the model is used for prediction. The time series contains the autoregressive model (AR), moving-average model (MA), autoregressive—moving-average model (ARMA), autoregressive integrated moving average (ARIMA). Modern econometric studies have shown that most economic time series is non-stationary and the ARIMA model has a good effect in dealing with non-stationary economic time series, so it is widely used. Andrei and Bugudui (2011) used the ARIMA (1,1,1) model in the study of US GDP. The results show that the US economy is on a rising trend.

# Forecasting Model and Estimation Methods

## Model Identification

A preliminary analysis with a plot of the initial data should be run as the starting point in determining an appropriate model. The input data must also be adjusted to form a stationary series and identify seasonality in the dependent series (seasonally differencing it if necessary), and using plots of the autocorrelation and partial autocorrelation functions of the dependent time series to decide which (if any) autoregressive (AR) or moving average (MA) component should be used in the model. We can also visualize the relationship between different variables by plotting one series against the other. Using correlation values, we can determine the extent of linear relationship between two variables and using autocorrelation we can determine the linear relationship between lagged values of a time series.

## Model Estimation

The parameters of the selected ARIMA (p, d, q) model can be estimated consistently by least-squares or by maximum likelihood. Both estimation procedures are based on the computation of the innovations from the values of the stationary variable. The least-squares methods minimize the sum of squares. For the regression models, MLE gives exactly the same parameter estimates as least squares estimation. ARIMA models are much more complicated to estimate than regression models, and different software will give slightly different answers as they use different methods of estimation, and different optimization algorithms. R will report the value of the log likelihood of the data; that is, the logarithm of the probability of the observed data coming from the estimated model. For given values of p, d and q, R will try to maximize the log likelihood when finding parameter estimates.

## Diagnostic Checking

Before using the model for forecasting, it must be checked for adequacy and accuracy. The model is considered adequate if the residuals left over after fitting the model is simply white noise, that is, they are uncorrelated, and they have zero mean. Also, the pattern of ACF of the residuals may suggest how the model can be improved. It is also important to evaluate forecast accuracy. The size of the residuals is not a reliable indication of how large true forecast errors are likely to be. The accuracy of forecasts can be determined by considering how well a model performs on new data that were not used when fitting the model. We have determined forecast accuracy using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE) and Mean Absolute Scaled Error (MASE). We have also used

Akaike information criterion (AIC) and Bayesian information criterion (BIC) to determine the quality of each model and r-square to determine the goodness of fit of the model.

## Model Forecasting

Model forecasting states the difference between in-sample forecasting and out-of sample forecasting. In-sample forecasting for instance, explains how the chosen model fits the data in a given sample while Out-of-sample forecasting on the other hand, is concerned with determining how a fitted model forecasts future values of the regressand, given the values of the regressors. To build a reliable model, the following factors are highly considered in forecasting;

a) The level of accuracy required – forecasts should be prepared as accurately as possible to facilitate the decision-making process especially made based on the GDP forecasts.

b) Availability of data and information – a wealth of reliable and up-to-date GDP data results to a reliable model.

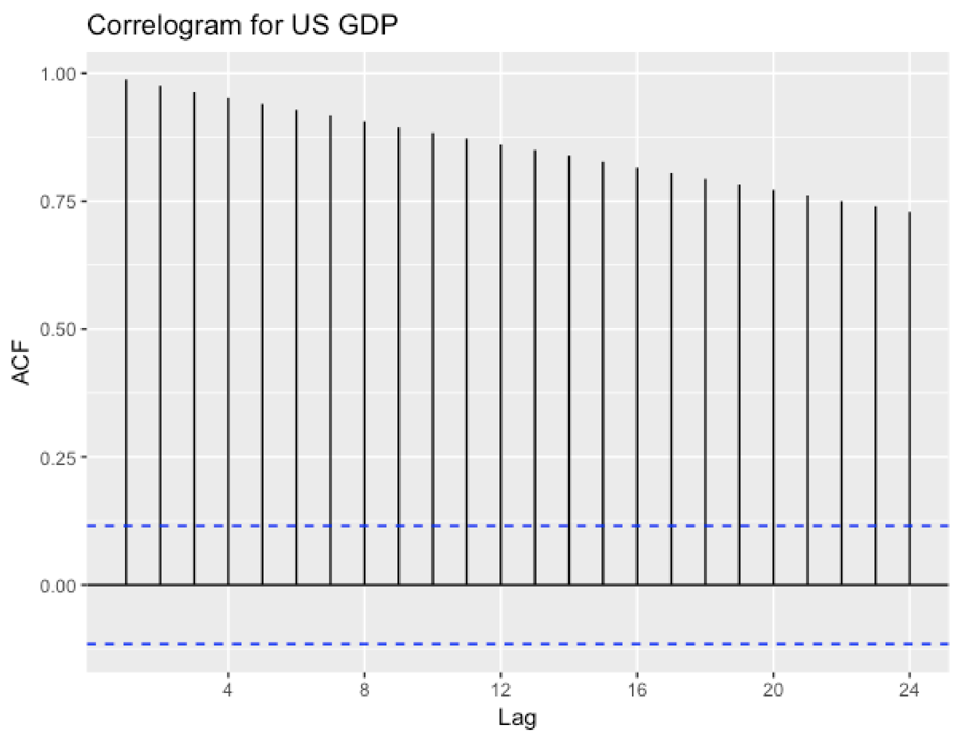
c) The time horizon that the GDP forecast is intended to cover.

# Data

An extensive time series data is required for univariate time series forecasting. Chatfield (1996) recommends more than 50 observations to build a reliable ARIMA model. In this study, we obtained our data from FRED website. The original source of our data is the Bureau of Economic Analysis. Our quarterly dataset consists of 288 records from 1947 to 2018 with variable of US GDP in billions of dollars. The summary statistics is shown in the table:

| **Descriptive Statistics** | | | | | | |
|---|---|---|---|---|---|---|
| **Statistic** | N | Mean | St. Dev. | Min | Pctl(25) | Pctl(75) |
| **GDP** | 288 | 5,983.445 | 6,069.967 | 243.164 | 712.672 | 10,447.490 |

The empirical characteristics of the univariate time series data were checked by obtaining time plots for the data. To gain an insight into univariate processes, autocorrelation functions ACF was considered. The ACF measures the ratio of the covariance between observations k lags apart and the geometric average of the variance of observations.




# Results and Discussion

## Basic Analysis

This study used a single set of data for Modeling that comprised of quarterly levels of GDP for US. The preliminary analysis of the data was done by use of time plots for the series as shown by Figure 1.

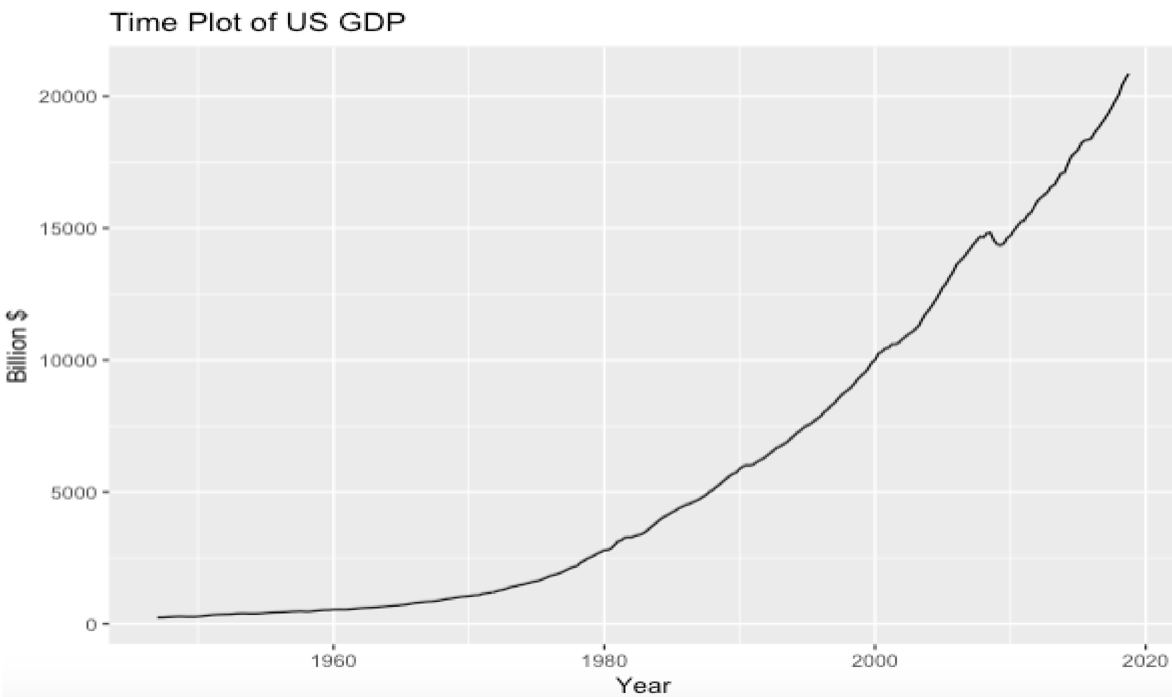



Figure 1

From Figure 1 above, a visual inspection of the time plots indicates that US GDP has shown the trend of exponential growth and there is not seasonal pattern. This implies that both the mean and the variance are not constant. Therefore, we regard it as a non-stationary time series. Also, there is a slight dip in GDP in 2008, that is probably due to Financial Crisis which happened between 2007 and 2008 in US.

## In-Sample and Out-of-Sample forecasts

For the purpose of forecasting, we have split the data into two parts: training data - which is from January 1947 to October 2009 and test data – which is from January 2010 to October 2018. We have used the following forecasting methods as benchmarks and applied it to the training data: average method, naïve method, drift method. Figure 2 shows the three methods applied to the quarterly beer production data and how does it vary from the test data.

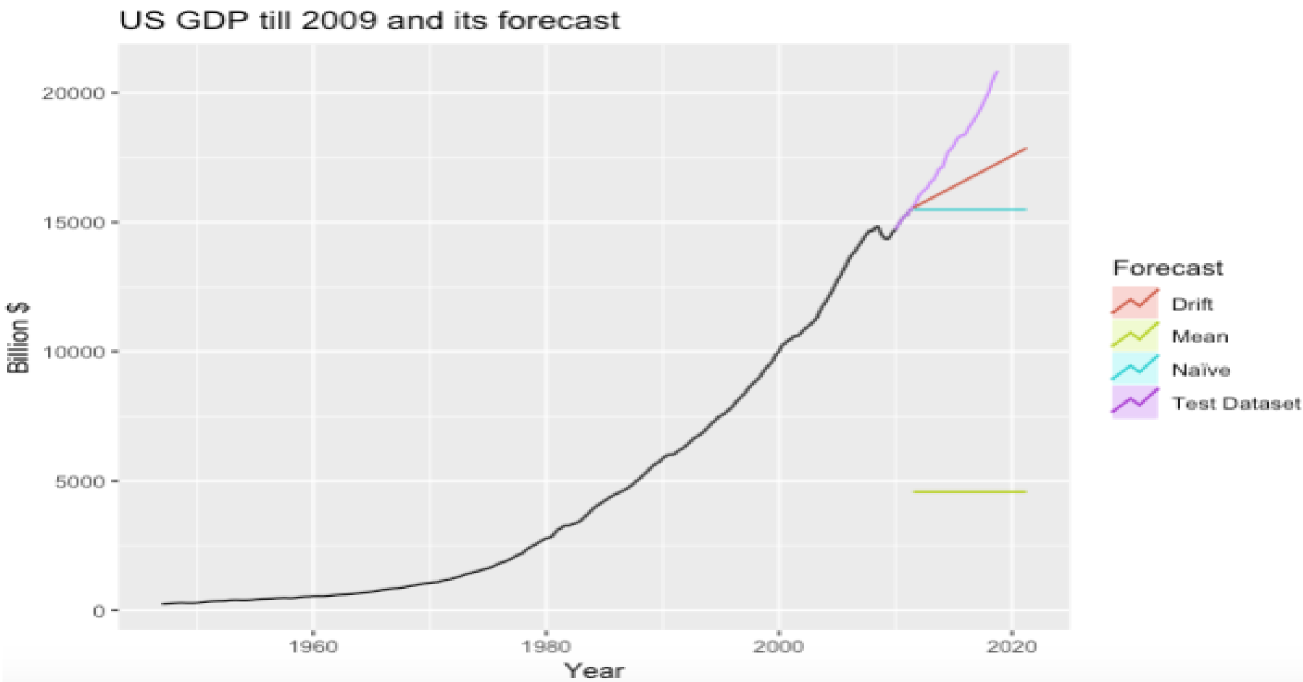



Figure 2

```
                        ME       RMSE        MAE         MPE       MAPE      MASE      ACF1 Theil's U
Training set 7.921312e-14  4676.911  3964.684 -335.44294 371.15933 16.07327 0.9875972        NA
Test set     1.174938e+04 11757.686 11749.382   71.92218  71.92218 47.63329 0.6436701  67.12593
> accuracy(naiveforecast, testing)
                   ME       RMSE        MAE      MPE     MAPE      MASE      ACF1 Theil's U
Training set  59.35029  89.19854  63.76922 1.597384 1.702792 0.2585275 0.7135934        NA
Test set     835.38510 945.02424 835.38510 5.045628 5.045628 3.3867436 0.6436701  5.558683
> accuracy(driftforecast, testing)
                       ME      RMSE       MAE       MPE     MAPE      MASE      ACF1 Theil's U
Training set -1.264754e-13  66.5877  52.04726 -3.969612 4.597716 0.2110053 0.7135934        NA
Test set      5.089585e+02 577.5361 508.95849  3.073527 3.073527 2.0633740 0.6028009   3.39837
```

Figure 3

From Figure 2 and 3 it is evident that Drift method of Forecasting give results closest to test data set.

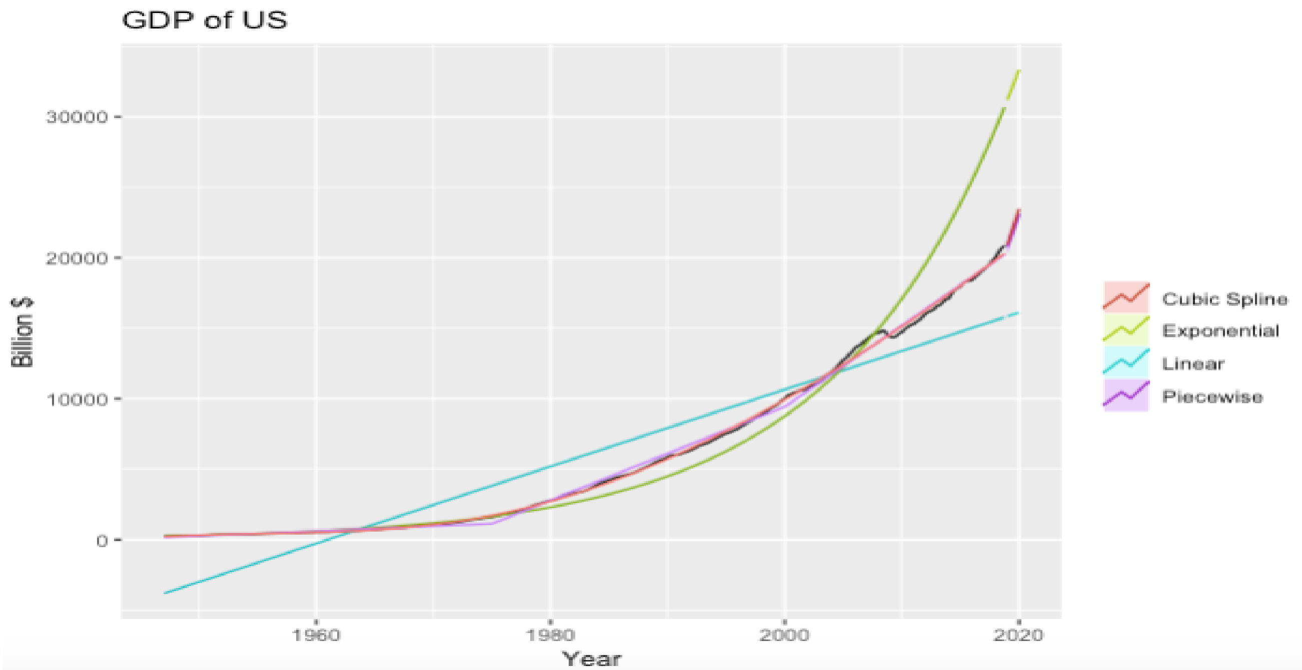



Figure 4

Figure 4 above shows the fitted lines and forecasts from linear, exponential, piecewise linear, and cubic spline trends. The best forecasts appear to come from the piecewise linear trend, while the cubic spline gives the best fit to the historical data but poor forecasts.

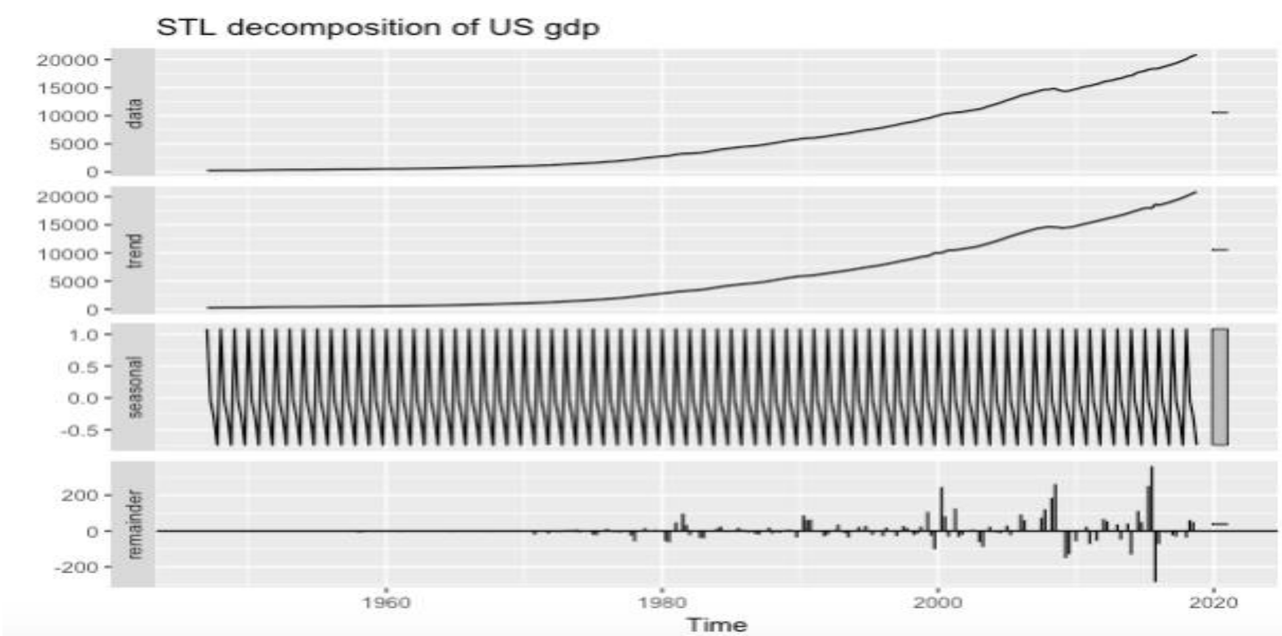



Figure 5

We have used STL method for time series decomposition. Figure 5 above shows the US GDP and its three additive components obtained from a robust STL decomposition with flexible trend-cycle.

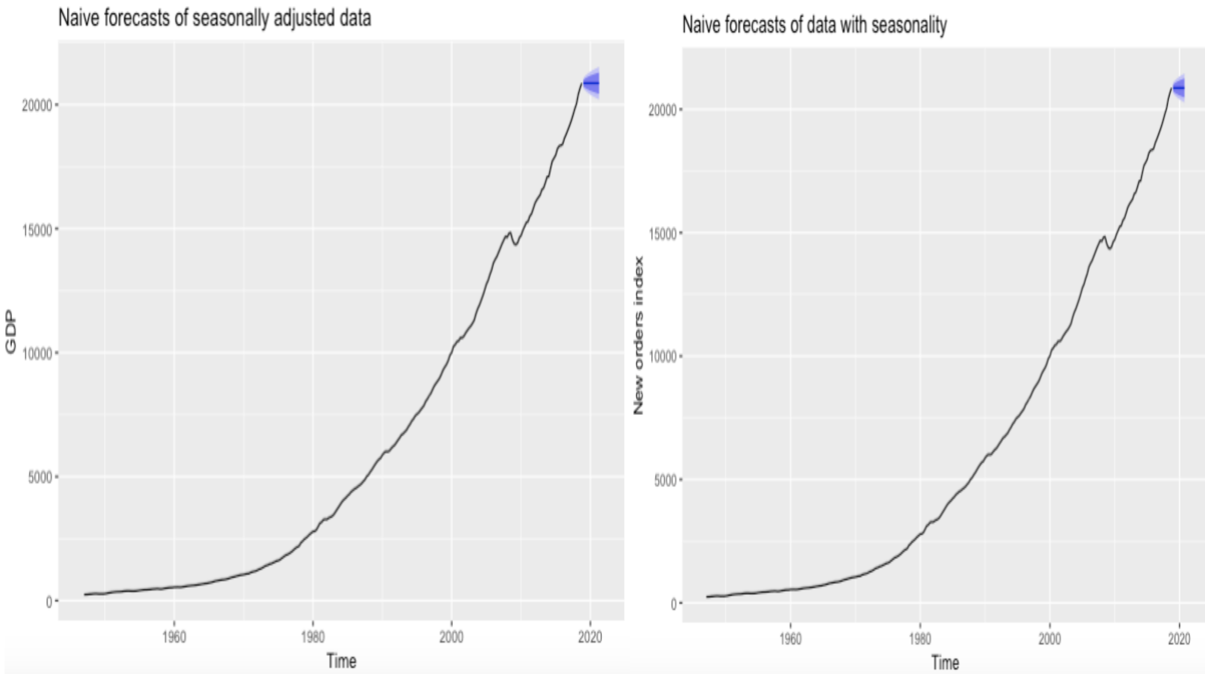



Figure 6

Figure 6 shows Naïve forecasts of the seasonally adjusted data and seasonal data obtained from an STL decomposition of the US GDP data.

```
> accuracy(fc_linear, testing)
                     ME     RMSE      MAE         MPE      MAPE       MASE       ACF1 Theil's U
Training set   1.186922 44.77991 24.41030  0.05383081 0.7508890 0.09896205 0.03848949        NA
Test set     -17.663559 51.24444 43.64119 -0.10805088 0.2670524 0.17692621 0.18720850 0.2831526
> accuracy(fc_damped, testing)
                     ME      RMSE       MAE       MPE      MAPE       MASE        ACF1 Theil's U
Training set   4.902057  44.48517  24.65645 0.1472486 0.7517423 0.09995998 0.009361005        NA
Test set     103.411280 142.11510 113.77948 0.6191269 0.6856244 0.46127460 0.361034951 0.8253753
> accuracy(fc_damped_add, testing)
                    ME     RMSE      MAE       MPE      MAPE      MASE          ACF1 Theil's U
Training set  5.112978 44.20560 25.41573 0.1528393 0.9170015 0.1030382 -0.0007383263        NA
Test set     75.362102 98.44304 88.58951 0.4591002 0.5439356 0.3591517  0.3073697358 0.6344833
> accuracy(fc_add, testing)
                     ME     RMSE      MAE         MPE      MAPE      MASE      ACF1 Theil's U
Training set   1.145218 44.46074 25.17949  0.04988808 0.9376359 0.1020805 0.0319895        NA
Test set     -17.057326 53.02296 43.02836 -0.10411573 0.2657467 0.1744417 0.1547240  0.319293
> accuracy(fc_damped_mul, testing)
                    ME      RMSE       MAE       MPE      MAPE       MASE        ACF1 Theil's U
Training set  6.485445  43.16447  24.54954 0.1863622 0.7717929 0.09952653 -0.01977141        NA
Test set     93.410123 124.81273 108.22813 0.5681843 0.6632212 0.43876882  0.35031372 0.8056608
> accuracy(fc_mul, testing)
                     ME     RMSE      MAE         MPE      MAPE      MASE         ACF1 Theil's U
Training set  0.7338059 44.46487 25.10790  0.02443038 0.7952332 0.1017902 -0.007586001        NA
Test set     -4.3608424 57.33389 47.93285 -0.02594917 0.2958277 0.1943251  0.035696634 0.3572936
```

Figure 7

We have used exponential smoothing technique for smoothing time series US GDP data. The above figure 7 shows that Holt's Linear Method provides best forecast with lowest RMSE value.

## Estimation Results

Modeling results of an ARIMA (1, 2, 2) process have been estimated by use of the Gaussian MLE Criterion and are presented below (Figure 8).

```
Series: gdp_quarter_col
ARIMA(1,2,2)

Coefficients:
         ar1      ma1     ma2
      0.7383  -1.3527  0.3817
s.e.  0.0931   0.1152  0.1053

sigma^2 estimated as 2556:  log likelihood=-1526.9
AIC=3061.81   AICc=3061.95   BIC=3076.43

Training set error measures:
                   ME     RMSE      MAE       MPE      MAPE       MASE        ACF1
Training set 5.51826 50.11815 28.66665 0.1573635 0.6848674 0.09747061 -0.01803509
```

Figure 8

```
 Ljung-Box test

data:  Residuals from ARIMA(1,2,2)
Q* = 8.328, df = 5, p-value = 0.1391

Model df: 3.   Total lags used: 8
```

Figure 9

The fitted ARIMA models were diagnosed using AIC, SBC and the log likelihood ratio test. Parameter estimation for the ARIMA models was done using the Gaussian MLE criterion. The ARIMA models fitted were adequate since the standardized residuals were not significantly correlated as shown by the Ljung-Box Q statistics. The ACF plot of the residuals from the ARIMA (1,2,2) model shows that all autocorrelations are within the threshold limits, indicating that the residuals are behaving like white noise. A portmanteau test returns a large p-value, also suggesting that the residuals are white noise.

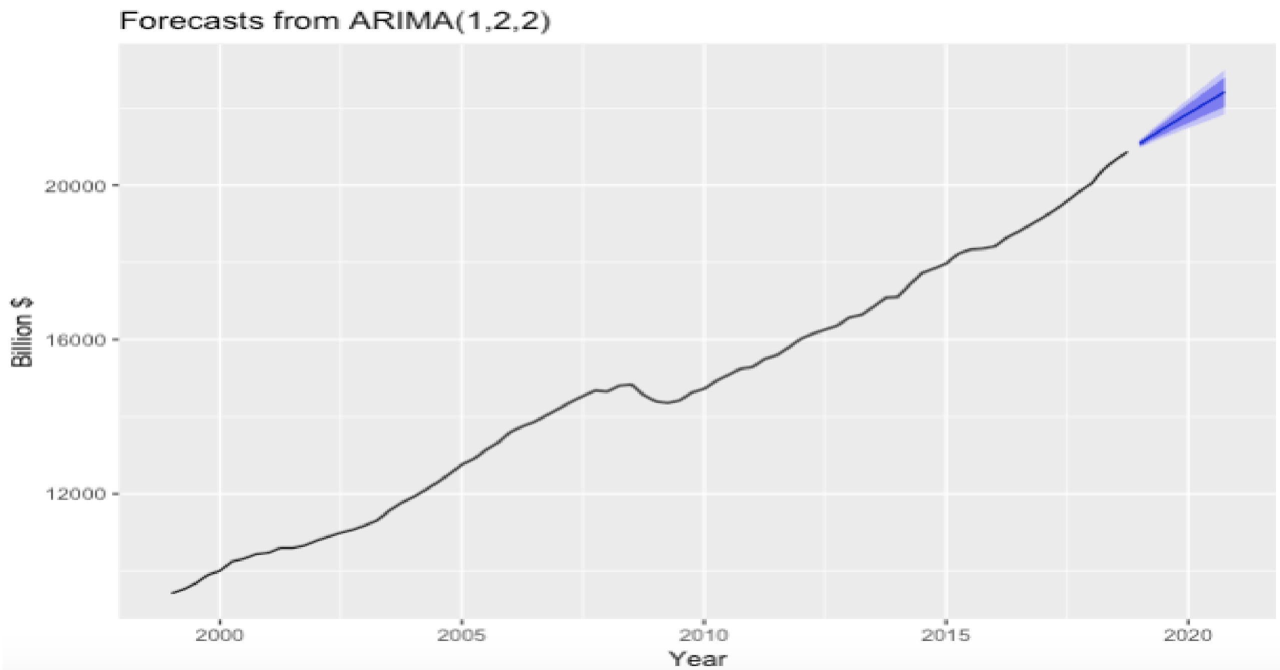



Figure 10

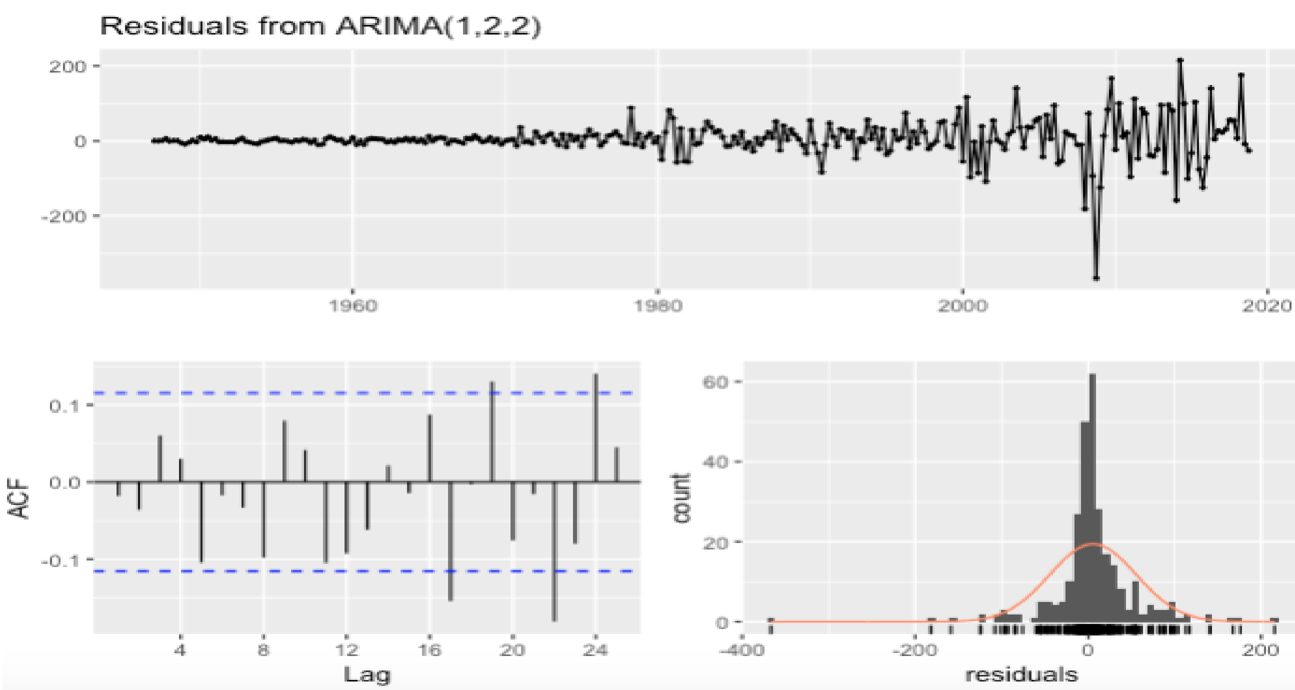



Figure 11

The study emphasized on forecast performance which suggests more focus on minimizing out-of-sample forecast errors than on maximizing in-sample goodness of fit. The approach adopted was therefore one of model mining with the objective of optimizing forecast performance. The model's efficiencies were evaluated using the Mean Squared Errors (MSE). The model that had the minimal MSE was considered the most efficient. However, other statistical properties especially the diagnostics and goodness of fit tests were considered in choosing the most efficient model. The chosen model, therefore, is justified by its relatively lower values of residual Kurtosis and MSE in addition to the other diagnostics considered.

# Conclusion

In this paper we have provided an extensive evaluation of the role of sophisticated nonlinear time series models for GDP growth. Our main conclusion is that ARIMA model fits the best for our GDP data with the minimal MSE and most fitted diagnostics. This finding is particularly evident when using real time data or considering only the period starting in the mid- ‘40s to 2018. We’ve come out with US GDP forecast for next 8 periods based on the ARIMA model forecasting quarterly result (from quarter 1, 2019 to quarter 4 2020). The predicted US GDP in first quarter of 2019 is approximately in the range of 21.010 trillion at low 80% interval level and 21.174 trillion at high 95 interval level; and the predicted US GDP in fourth quarter of 2020 is approximately in the range of 22.035 trillion at low 80% interval level and 22.988 trillion at high 95 interval level.

However, we have also identified some important cases where the adoption of a more complicated benchmark, like inflation rate, unemployment rate and etc. These variables can alter the conclusions of economic analyses about the driving forces of GDP growth. Therefore, comparing theoretical models also with more sophisticated time series benchmarks can guarantee more robust conclusions.

Last but not the least, from the findings of the study, the following areas are suggested for further research:

i. Analysis of GDP Dynamics in US using different models.

ii. Examination of individual components of the GDP.

# References


Henderson, J. V., Storeygard, A., & Weil, D. N. (2012). Measuring economic growth from outer space. *American economic review*, *102*(2), 994-1028.

Jiang, Y., Guo, Y., & Zhang, Y. (2017). Forecasting China's GDP growth using dynamic factors and mixed-frequency data. *Economic Modelling*, *66*, 132-138.

Jansen, W. J., Jin, X., & de Winter, J. M. (2016). Forecasting and nowcasting real GDP: Comparing statistical models and subjective forecasts. *International Journal of Forecasting*, *32*(2), 411-436.

Andrei, E. A., & Bugudui, E. (2011). Econometric modeling of GDP time series. *Theoretical and Applied Economics*, *10*(10), 91.

Chatfield, C. (2016). *The analysis of time series: an introduction*. Chapman and Hall/CRC.

Čižmešija, M., & Sorić, P. (2010). Assessing Croatian GDP components via economic sentiment indicator. *Economic research-Ekonomska istraživanja*, *23*(4), 1-10.

Ahmed, A. E. M., & Suliman, S. Z. (2011). The long-run relationship between money supply, real GDP, and price level: Empirical evidence from Sudan. *Journal of Business Studies Quarterly*, *2*(2), 68.

Amadeo, Kimberly. “What Is the Ideal GDP Growth Rate?” *Thebalance*, 22 Apr. 2019, www.thebalance.com/what-is-the-ideal-gdp-growth-rate-3306017.

Lee, C. K., Song, H. J., & Mjelde, J. W. (2008). The forecasting of International Expo tourism using quantitative and qualitative techniques. *Tourism Management*, *29*(6), 1084-1098.