Marginal Analysis of Gross Domestic Product using Best-Fitted Least Square Model

Rahela Abdul Rahim[#1], Fatinah Zainon[#2], Noraziah Man[#3]

[#]*School of Quantitative Sciences,Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia*

[1]rahela@uum.edu.my

[2]fatinah@uum.edu.my

[3]noraziah@uum.edu.my

***Abstract*— Gross Domestic Product (GDP) represents the total monetary value of all goods and services produced over a specific time period. This study models the GDP data of the state of Kedah, Malaysia for year 2005-2016. A mathematical model is developed for the state GDP data using Least Square Methods. The future GDP values are predicted using five empirical models: Linear model, Quadratic model, Cubic model, Exponential model and Polynomial model. The result has shown that Exponential model is the best fitted model for Kedah GDP data. The result of the derivative of the obtained model provide marginal analysis that shows the performance of GDP per year for the state.**

s document gives formatting instructions for authors preparing papers for publication in the said journal. The authors must follow the instructions given in the document for the papers to be published. You can use this document as both an instruction set and as a template into which you can type your own text.

***Keywords*—** *Gross Domestic Product, Least Square Methods, Empirical models, Marginal analysis.*

1. Introduction

The Gross Domestic Product (GDP) is one of the primary macroeconomic indicators used to assess the condition of a country's economy. It represents the total monetary value of all goods and services produced over a specific time period. Usually, GDP is expressed as a comparison to the previous quarter or year. For example, if the year-to-year GDP is up till 3%, that means the economy has grown by 3% over the last year. Economic production and growth, what GDP represents, has a large impact on nearly everyone within that economy. When the economy is healthy, typically observe low unemployment and higher wages as businesses demand labour to meet the growing economy. A significant change in GDP, whether up or down, usually has a significant effect on market. Investors always worry about negative GDP growth, because it is one of the factors economists use to determine, whether an economy is in a recession.

The graph in Figure 1 shows the Gross Domestic Product (GDP) of Kedah Darul Aman from year 2005 until 2016. As we can see, the horizontal axis presents the years and the vertical axis shows the GDP in (RM Million). The graph indicates that the GDP increased sharply between 2005-2007. Then for year 2007 to 2008 it increased a little before it decreased in year 2009. The graph increased dramatically in year 2010 and slightly became consistent until year 2016.

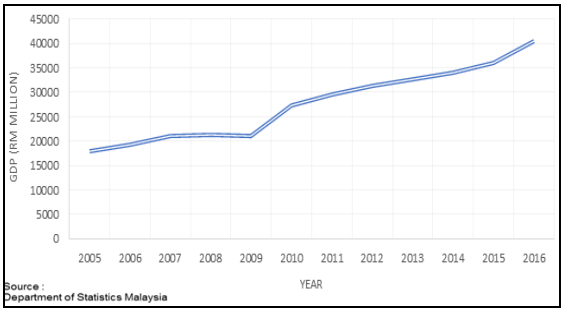

Figure 1: Graph of Gross Domestic Product of Kedah Darul Aman from 2005-2016.

2. Related Works

Probability distribution has been regularly used to model and quantify the GDP. By assuming the distribution of data, analyst can utilize the characteristics of the distribution to make predictions on outcomes. Probability distributions are regularly used to model financial data [13]. In many studies, assuming data normality is common, however the assumption can be resisted and best fitting distribution [3] and multivariate distribution [5] can be substituted. Forecasting of stock prices are regular example of financial data analysis such

International Journal of Supply Chain Management
IJSCM, ISSN: 2050-7399 (Online), 2051-3771 (Print)

as in [11] and [14] use forecasting with a two-state Markov-switching autoregressive model to forecast the stock prices. Differential equation is used to obtain the constant dividend and derived the installment option to owner [3]. Another approach to predict the new stock price is by combining the Hilbert-Huang transform [15]. Many studies also used statistical approach to model financial data [9],[10]. The method is used to analyze daily closed price stock market index. For continuous type of data, integral equation is used to value the installment derivatives where the asset price and time are assumed to be a continuous function [6]. Data mining and predictive technologies require the development of an automated computer program to analyze financial data such as work in [4], develop an automated computer program using data mining and predictive technologies to do a fair amount of trades in the markets. [2] uses a suite of statistical models to predict Latvian GDP. Sample forecasting step had been run out to find out GDP projections and to appraise forecasting accuracy of all individual statistical models.

As a conclusion, factor and bridge models are among the best individually performing models in the suite. Forecasting accuracy that was obtained by using disaggregated models of factor and bridge models can be considered as a good alternative to aggregated ones. Furthermore, weighted combination of the forecasts of the statistical models allows obtaining robust and accurate forecasts which leads to a reduction of forecasted errors. [12] uses Artificial Neural Network (ANN) to predict Palestine Gross Domestic Product (GDP). Time series of quarterly observations on (GDP) in Palestine is collected. Forecasting results of Artificial Neural Network (ANN) are compared with those of the Autoregressive Integrated Moving Average ARIMA and regression as benchmark methods. By using Root Mean Square Error (RMSE), the empirical results show that ANN performs better than the traditional methods in forecasting GDP. [1] uses the comparison between non-parametric models to find out the prediction of GDP in Turkey. Two alternative situations had been considered due to seasonal effects. In the first case, a semi-parametric model had been discussed where parametric component is dummy variable for the seasonality. Smoothing spline and regression spline methods had been used for prediction of the semi-parametric models. While in the second case, it had been considered the seasonal component to be a smooth function of time. Therefore, the model falls within the class of additive models. The results obtained by semi-parametric regression models are compared to those obtained by additive non-parametric.

ARIMA modelling techniques also is used by [8] to predict Nigeria GDP between the year 1980 until 2007. For statistical analysis, graphical methods had been used to display data distributions, Autocorrelation functions (ACF), Partial autocorrelation functions (PACF), Residuals and Forecasts, and differencing to check for stationarity. The results are summarized as the ARIMA (1, 2, 1) model was proposed for the data from the second differences which shows stability and invertibility. Then, forecasts were made for future observations up to thirteen years which shows an increasing trend over time. The Akaike Information Criterion (AIC) and the adjusted multiple correlation coefficient (Adjusted R-Square) provided a good summary of the total variability explained by the chosen fitted model. [7] use Kalman filter technique to estimate deseasonalized monthly series for Swiss gross domestic product at constant prices of 1990 for the period 1980-1998. The results are consistent with the quarterly figures estimated by the Federal Office for Economic Development and Labour. The work presented a general approach using the Kalman filter technique, nesting a great variety of interpolation setups. The work also evaluated competing models and provide a time series that can be used by other researchers.

3. Motivation

Data that is not monitored and analyzed indeed will affect the economic growth in the long term. Any significant change in the GDP, either up or down, can have a big effect on investing sentiment. For example, if investors believe the economy is improving and corporate earnings along with it, they are likely willing to pay more for any given stock. If there is a decline in GDP or the investors expect a decline, they would only be willing to buy a given stock for less, leading to a decline in the stock market. Apart from that, the economic data releases are also essential for each Forex trader, especially the important reports like the GDP of a country/state which reflects the overall state of the respective economy. Such data create volatility and plenty of speculation is always preceding them. Market players monitor this critical piece of economic data and consider either to enter a new position or add to a current one. Obviously economic activity and growth will be affected if data is not monitored and analyzed. So, to answer this issue, a mathematical model will be employed so that the data shall be monitored and analysed analytically and logically.

4. Motivation

This study uses the modeling methodology as shown in Figure 2. Five phases are involved begin with data collection, model development, data forecasting and data analysis.

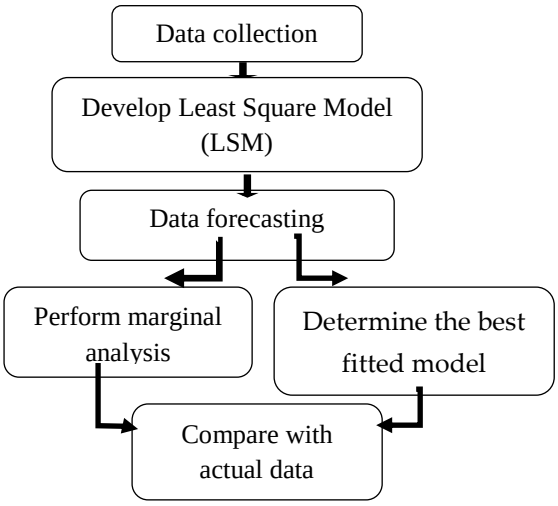

Figure 2: Framework of the proposed method

Phase 1 : **Data Collection**

Data of Gross Domestic Product (GDP) for Kedah state from year 2005 until 2016 were collected from Jabatan Perangkaan Malaysia website.

Phase 2 : Develop Least Square Method (LSM) Model

Trendline models were created based on the graph by choosing types of line to estimate data points. Different types of Least Square models such as quadratic, cubic, order of 6 and exponential were developed. Type of line that has the lowest Sum of Square Error (SSE), will be chosen as the best fitted model for the data.

Phase 3 : Data Forecasting

The predicted Gross Domestic Product were calculated using the equation from the best fitted model.

Phase 4 : Marginal Analysis

Marginal analysis was performed using the first derivative of the best fitted equation of Least Square model.

Phase 5 : Determine The Best Fitted Model

The Sum of Square Error (SSE) was conducted for model comparison. In this study we used Discrete Dynamical System (DDS) model as the comparative model. The comparative results obtained from DDS model and the Least Square model which had the lower SSE will be choosen as the best fitted model for the data.

4.2 Data

The GDP for Kedah state from year 2005 to year 2016 are shown in Table 1 below.

Table 1: GDP of Kedah Darul Aman from 2005-2016

Year	Gross Domestic Product (RM Million)
2005	17829
2006	19255
2007	21033
2008	21209
2009	21092
2010	27356
2011	29585
2012	31241
2013	32740
2014	34105
2015	35999
2016	40596

5. Result and Finding

The least squares method is a form of mathematical regression analysis that finds the line of best fit for a dataset, providing a visual demonstration of the relationship between the data points. Sometimes they are called statistical method used to determine a line of best fit by minimizing the sum of square created by a mathematical function. By using Microsoft Excel, an initial graph was plotted based on the data given in Table 1, where *x* represents the number of year and *y* represent the GDP data. Linear trendline was created based on the graph by choosing types of line to estimate data points. The analysis of each type of trendline is given in Figure 3 – 6.

Table 2: GDP of Kedah Darul Aman from 2005-2016

Year	*x*	*y* (GDP)
2005	5	17829
2006	6	19255
2007	7	21033
2008	8	21209
2009	9	21092
2010	10	27356

2011	11	29585
2012	12	31241
2013	13	32740
2014	14	34105
2015	15	35999
2016	16	40596

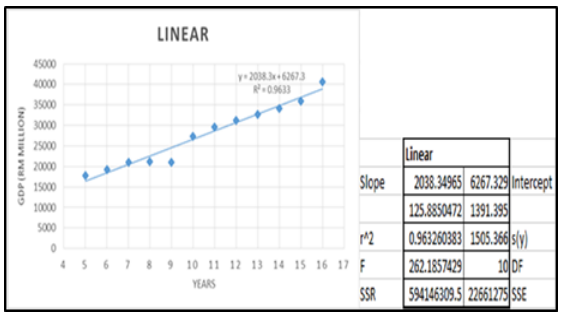

Figure 3: Graph and Result of Linear Least Square Method.

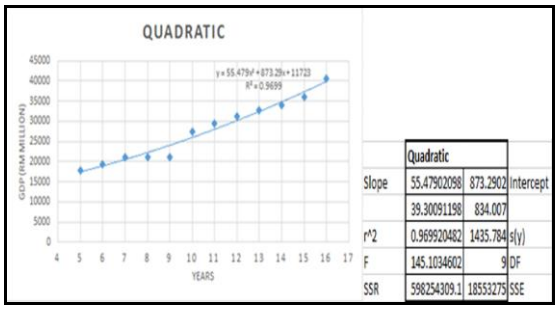

Figure 4: Graph and Result of Quadratic Least Square Method.

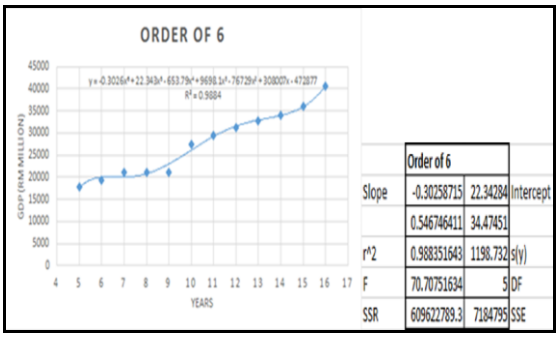

Figure 5: Graph and Result of Polynomial (Order of 6) Least Square Method.

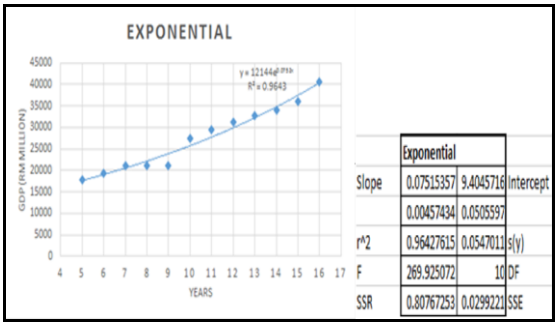

Figure 6: Graph and Result of Exponential Least Square Method.

Table 3: Comparison of Coefficient of Determination Between Models.

Empirical Model	r^2
Linear,	0.7283
Quadratic	0.4648
Cubic	0.4495
Polynomial order 6	0.6913
Exponential	0.7412

From the results, it is obvious that exponential line has the smallest Sum of Square Error (SSE) which is 0.0299221 compared to other models where SSE is 22661275 for linear, 18553275 for quadratic, 17597583 for cubic and 7184795 for polynomial of order 6. Therefore, exponential model is best representing the GDP data for state of Kedah for year 2005 – 2016.

5.1 Marginal Analysis

The marginal analysis is performed to see the change of the state GDP performance by year. The first derivative of the obtained exponential model was derived as the following;

$$f(x) = 1.2144e^{0.0752x}$$

$$f'(x) = 0.09132288e^{0.0752x} \quad \text{(i)}$$

The respective value of x is substituted into (i) to obtain the GDP marginal for that year. The results of the yearly GDP marginal from 2005 to 2016 are shown in Figure 7.

GDP for Kedah state from year 2005 to year 2016 are shown in Table 4 below.

Table 4: Marginal of Gross Domestic Product for Kedah in 2005-2016.

Year	x	Marginal Value, f '(x)
2005	5	0.133006947
2006	6	0.143394756
2007	7	0.154593851
2008	8	0.166667591
2009	9	0.179684289
2010	10	0.193717586
2011	11	0.208846881
2012	12	0.22515777
2013	13	0.242742536
2014	14	0.261700667
2015	15	0.282139424
2016	16	0.304174442

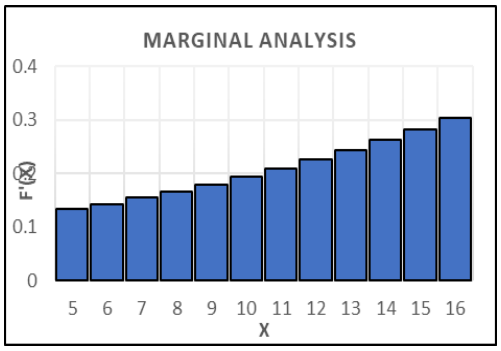

Figure 7: Marginal Analysis Graph of Gross Domestic Product of Kedah Darul Aman from

The average marginal for the Kedah state GDP data from 2005 to 2016 is 0.20798556067. It can be concluded that during that period the average increment in the GDP data for state of Kedah is nearly RM 21 million. The least square method states that the curve that best fits a given set of observations, is said to be a curve that have a minimum sum of the squared residuals (or deviations or errors) from the given data points. The least squares method is a very useful method of curve fitting. Despite of many advantages, it has few limitations too. One of the main limitations is, in the process of regression analysis which utilizes least squares method for curve fitting, it is implicitly assumed that the errors in the independent variable are negligible or zero. In the case, when independent variable errors are non-negligible, the models are subjected to the measurement errors. In this case, least square method may even lead to hypothesis testing, parameter estimates, and confidence intervals are taken into consideration due to the presence of errors occurring in the independent variables.

6. Conclusion

This study proposed a method of studying the trend of Gross Domestic Product data for any particular state wherein Least Square models are employed on the basis of the SSE from the actual data. The fitting of the linear, quadratic, cubic, polynomial of order 6 and exponential model also follows the determination of the function. The results exhibit that exponential model is significantly the best fitted line that can be used to represent the GDP data for that respected years. The marginal analysis should be interpreted from an economic perspective. An increasing trend in the GDP marginal for state of Kedah is observed between 2005 and 2016. However, there is no drastic increment is seen at any particular year. Any changes in the economic development during that year might not have affected the state GDP. By looking at the trend, it is expected that the GDP marginal for state of Kedah will continue to increase slowly at the near future.

Acknowledgments

This research is funded by Universiti Utara Malaysia under the university grant with S/O Code 13876.

References

[1] Aydin, D. "Comparison of regression models based on nonparametric estimation techniques: Prediction of GDP in Turkey". International Journal Of Mathematical Models And Methods In Applied Sciences, 1, 70-75, 2007.

[2] Bessonovs, A., "Suite of statistical models forecasting Latvian GDP." Procedia - Social and Behavioral Sciences, 110, 1094-1105, 2014.

[3] Beiranvand, A. and Ivaz, K.. Valuation of Installment Option by Penalty Method. *Journal of Computational Methods for Differential Equations*, 3 (4): 298-310, 2015.

[4] Cannan, K. S., Sekar, P. S., Sathik, M. M. and Arumugam, P. (2010). Financial Stock Market Forecast Using Data Mining Techniques. Proceeding of the International MultiConference of Engineers and Computer Science (1).

[5] Chandrasekara, N.V., Mammadov, M.A. and Tilakaratne, C.D.. Estimating Parameters of Multivariate scaled t Distribution of GSPC and its associated financial indices. *International Journal of Statstics and Economics*, 16 (3): 48-62, 2015

[6] Ciurlia, P. (2011). On a General Class of Free Boundary Problems for European-style Installment Options with Continuous Payment plan. *Communication on Pure and Applied Analysis*, 10 (4): 1205-1224.

[7] Cuche, N. A. & Hess, M. K, "Estimating Monthly GDP In A General Kalman Filter Framework: Evidence From Switzerland". Working Paper, Study Centre Gerzensee, *2*, 1999.

[8] Fatoki, O., Ugochukwu M. & Abbas, O. "An Application of Arima Model to the Nigeria Gross Domestic Product (GDP)". International Journal of Statistics and Systems, 5, 63-72, 2010.

[9] Jaber, M. A., Trend Extraction Using Empirical Model Decomposition and Statistical Empirical Mode Decomposition: Case Study: Kuala Lumpur Stock Market. *AIP Conference Proceedings*, 1635 (1): 776-779, 2014

[10] Jaber, M. A., Ismail, M. T. and Altaher, A. M. Application of Empirical Model Decompositon with Local Linear Quantile Regression in Finanncial Time Series Forecasting. The Scientific World Journal, 2014.

[11] Induruwage, D., Tilakaratne, C. D. and Rajapaksha, S.R.M.S.P.. Forecasting Black Tea Auction Prices by Capturing Common Seasonal Patterns. *Sri Lankan Journal of Applied Statistics*, 16 (3), 23-34. 2016.

[12] Safi, S.K., "A Comparison of Artificial Neural Network and Time Series Models for Forecasting GDP in Palestine". American Journal Of Theoretical And Applied Statistics, 5, 58-63, 2016.

[13] Vincent A.. Financial interaction analysis using best-fitted probability distribution in Vincent Ang (eds.), *IFC Bulletins Chapters*, 39. 2015.

[14] Wang, D., Qing N., Lei, M. and Chang X.. Dynamic Relation of Chinese Stock Price-Volume Pre and Post the Split Share Structure Reform: New Evidence from a Two-State Markov-Switching Approach. *China Finance Review International*, 5 (1): 386-401. 2015.

[15] Xuan L., Guozhong C. and Fulong, L. Stock Data Analysis Based on Hilbert-Huang Transform. *IEEE International Conference on Grey System and Intelligent Services (GSIS)*, DOI:10.1109/GSIS.2015.7301816. 2015.