CHAPTER 3

Rainfall Forecast Model

3.1 Introduction

3.1.1 Historical of rainfall forecasting

Namkhan 2 and 3 hydropower plants are two hydropower plants of Electricite Du Laos, which operated and managed by Electricite Du Lao-Generation public company (EDL-Gen). The electricity production previous of each hydropower plants have produced from order of the National Control Center (NCC) and according from water storage capacity in the reservoirs. Generally, if dry year, it will affect to electricity production, which produce not enough with energy demand of the socity. Differently, if wet year, hydropower plants must release water from spillway to control water level in reservoirs. More water released will impact to assets and living of peoples, whom live along the river on downstream. One cause is because not has the runoff forecaste and good planning for supporting operation of those reservoirs. Therefore, this section in research is studied the rainfall forecasting, which has important to plan and cope in several cases. Main objective of this lesson is to find the rainfall for supporting methodology of the reservoir operation and to avoid the damage from the released water and water shortage conditions.

3.1.2 Rationality of forecaste rainfall

Why do authors forecast the rainfall ?. This is a question that generated to find the reason and the answer of prediction in the future. Because, the forecasted rainfall is to analyze the possible happens in the future, which enables us to do something about it. Due, the forecasted in future is difficultly for finding the accuracy of prediction and control of any operation. Therefore, good forecasted will helps us to control and manage any operation in the future. It is a guideline for planning of the reservoirs operation and enables us to do something about it such as: electricity production, reservoir management,

water regulation to supply water on downstream river, and enables us to know the various constraints that relate with the reservoir operation in the next time.

3.2 Forecasted Methodology

Presently, the forecated rainfall has many of theories and methods as well as programs for forecasting have several softwares. This research studied the reservoir operational management for optimizing electricity production. So, the forecasted rainfall is an important section for planning and the analyzing operation of the electricity production and water regulation of these hydropower plants. The forecasted in this paper has four methods: (1) function furcating in the Microsoft excel (FFME), (2) Minitab software (MNT), (3) statistical package for the social sciences (SPSS) software, and (4) fast fourier transform (FFT) calculation online. These methods can be predicted the rainfall volume by using historical data of both hydropower plants. The forecasted detail will be described in the below of each methodologys. Related parameters for forecasted rainfall consisted by secular trend, seasonal variation, cyclical variation and irregular variation. That has effecting to results accuracy or error.

3.3 Forecast rainfall by FFME Method

The forecasted in Microsoft Excel will be used the statistic data for rainfall prediction. Parameter x must have a numeric value as: known_y and known_x must be arrays or cell ranges that contain equal numbers of numeric data values. The most common usage of forecasted includes a specific x value plus 2 ranges of cells that contain the data, such as forecaste (110, A1:A100, B1:B100). These steps illustrated the processing in forecast function as follow:

a) Create a worksheet in Microsoft Excel, After add data into colume or table.

b) Time loop definition of rainfall repeats for evaluating the historical data.

c) Choosed the inflow data in colume, Excel worksheet by using formulation as FORECAST (X, know_y, know_x). Then, entried data x in colume, which must same cell with both data y and x.

d) After that, selected the box sheet that added formulation, and then click until to go the target of forecasted.

3.3.1 Forecasted formulation

Forecast Error = ([Actual-Forecast]/Actual) * 100% (3.1)

Forecast Accuracy = Maximum (100% - Forecast Error) (3.2)

The equation for forecast is a + bx that:

$a = \bar{Y} + b\,\bar{X}$ and $b = [\Sigma (x - \bar{X})(y - \bar{Y})] / (\Sigma (x - \bar{X})^2)$ (3.3)

Source of forecast formulation as:

$F_t = F_{t-1} + \alpha (A_{t-1} - F_{t-1})$ / $\alpha = 0 - 1$ (3.4)

Where: F_t = new forecast,

F_{t-1} = previous period forecast,

A_{t-1} = previous period actual demand, and

α = smoothing (weighing) constant,

3.3.2 Forecast methodology

$F_i = ax_i + (1\text{-a})\, F_{i-1}$

$= ax_i + (1\text{-a}) [\, ax_{i-1} + (1\text{-a})\, F_{i-2}]$

$= ax_i + (1\text{-a}) [\, ax_{i-1} + (1\text{-a}) [ax_{i-2} + (1\text{-a})\, F_{i-3}]]$

...

$= a \sum_{j=0}^{i} (1-a)^2 (x_{i-j})$ (3.5)

3.3.3 Conclusions of forecast function

The process for finding data is summarized the squared deviations of average data (actual data). It has procedures in the forecast function, which should be performed as bellowing:

- Should be find the average values of actual data

- Calculated each squared deviation
- Summarize the squared deviations

The method is rather accuracy than the general prediction. This procedure is the easy performent for the calculator formula. Due, the method was suitable for using to calculate the data points. The methodology process performs as following.

- To find out the data summation of squares of all data, the data size, and to summation of all atual data.

- Forecated results equal to previous period forecast plus result of previous period actual demand minus previous period forecast, then multiply the smoothing (weighing) constant.

This methodology is one of many other functions for predicting data to estimate the event in the future. However, before used should be understood about behavior of historical data. It help to decide for definiting the time loop of data repeat.

3.3.4 Data analysis of forecast function in Microsoft Excel model

Data analysis to find the rainfall in the future used the statistical data 50 years (since 1960 to 2009) of Namkhan river, which has will be analyze to find the rainfall in 2010-2016. The forecasted rainfall used the forecaste function in Microsoft excel, which consists the previous period actual, smoothing constraint (weighing), and previous period forecast. that formulation follow the algorithm above (2.1.2).

After analyzed and formulated will be received the forecastes rainfall results, which shown in figure 5.1. Note, before the forecasted by using forecaste function in Microsoft Excel should be known the behavior of previous rainfall. Because, the smoothing constant is important parameter that affect to change the forecasted results. This method may be not best correct but it will help for comparing with others methods in the decision time. [3]

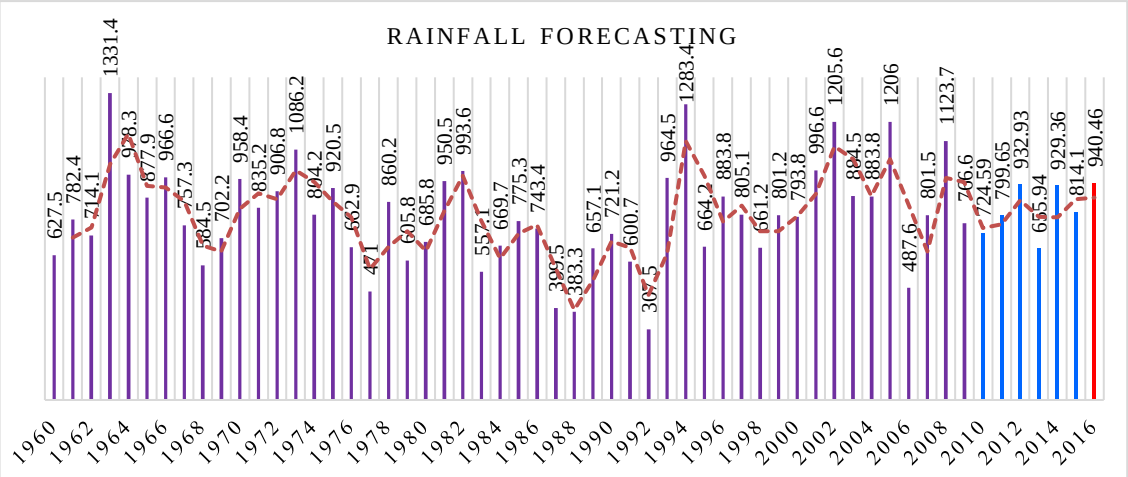

Figure 3.1 The rainfall result of function forecast in Microsoft Excel.

3.4 Forecaste Rainfall by MNT Method

Minitab software is one of all methods to analyze standard deviations (σ) of the repeat or replicate measurements. This method is based on a linear model with a log link function: $\ln(\sigma) = A\gamma$, where A is the design matrix and γ is a vector of parameters to be estimated. One advantage of using the log link function is the fitted values, which are always positive. First method uses least squares estimation and second method uses maximum likelihood estimation (MLE). The methods produce equivalent results in the saturated model, when the number of parameters equals the number of data points. For the least squares estimation, Minitab uses weighted least squares regression. If there are an equal number of repeats or replicates, the weights are equal. For MLE, Minitab assumes that the original data come from a normal distribution. The distribution of the sample variance is related to the χ^2 distribution.

3.4.1 Forecaste algorithm of Minitab software

The ARIMA abbreviated from autoregressive integrated moving average, which written by Professor William Q. Meeker, Jr., of Iowa State University. The works of professor Meeker can help in the development become to Minitab software. Forecasts are calculated repeat, depend on the model and the parameters. For example, if an ARIMA model is fit with 1 autoregressive term (AR_1) and one seasonal differencing term with a seasonal period of 12 months. To estimate Y_{k+1}, the first forecast, where k is the origin. Forecast' method will uses the volume, trend, and seasonal components to predict the

data. The forecast formulation of (m) periods ahead from a point at time (t) shows algorithm as:

$$Y_{k+1} = L_t + mT_t \quad (3.6)$$

Where L_t is the data volume. T_t is the trend during (t) time. m is period of data set. The seasonal component is one of the same period from the previous time. forecast method uses historical data to forecasts origin time and the forecasts result. The fitted value at time t is the unscented moving average at time t-1. The forecasts are the fitted values at the forecast origin. If you forecast 10 time units ahead, the forecasted value for each time will be the fitted value at the origin. Data up to the origin are used for calculating the moving averages.

3.4.2 Forecaste methodology in Minitab software

The forecast method can use the linear method for performing the consecutive volume averages. It is analyzed a trend of the data set. First process, calculated and stored the original data in the series model. next, analyzed the previously stored data in column to obtain a behavoir data. The predict model for time (t) is the data value during time t-1. Average process of length one will be helped in forecasting of data set. Decomposition function will calculates the forecast volume in the linear regression or has condition that relates the seasonal. The function forecast used is the decomposition module in the Minitab software, which shown in figure 3.2 as below for using the time series decomposition.

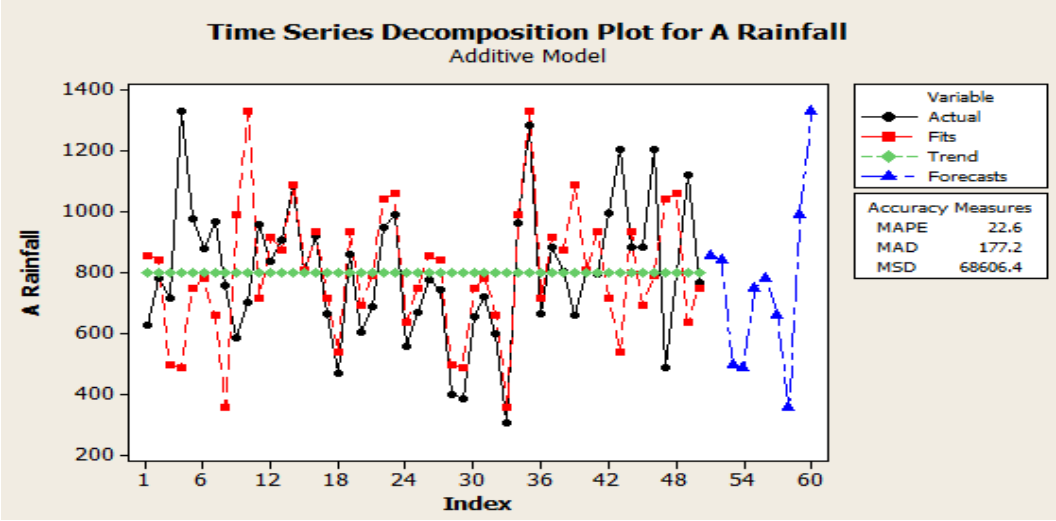

Figure 3.2 Time series decomposition plot of the Minitab software.

3.4.3 Data analysis in Minitab software

Data analysis used the statistical data for during 50 years (since 1960 to 2009) of Namkhan river. The objective is analysis to find the forecasted rainfall of 2010-2016 year. The forecasted rainfall in MNT method can set for showing results such as: graph, frequency value, season analysis and others. Analysis is used by Time Series Decomposition for predicting rainfall. The software shows the accuracy measure, which relate with forecasted such as: mean absolute percentage error (MAPE), mean absolute deviation (MAD) and mean squared deviation. The forecasted results in 2016 is average normal year, which shows annual water inflow of 65.1 m^3/s, 89.4 m^3/s for Namkhan 2 and 3 hydropower plants respectively. Accuracy measured in Minitab software is shown MAPE = 22.6, MAD = 177.2, which are in the acceptable ranges. The result from Minitab software shown in figure 3.2 for forecasting yearly. Furthermore, the Minitab software can be shown the result in the main page of program, which easy to link with Excel. The time loop determination is important for using this method of repeat origin within the software. Then, authors should be checked the accuracy measure that has minimum value or acceptanable ranges, which showing that optimal result of simulation program.[3]

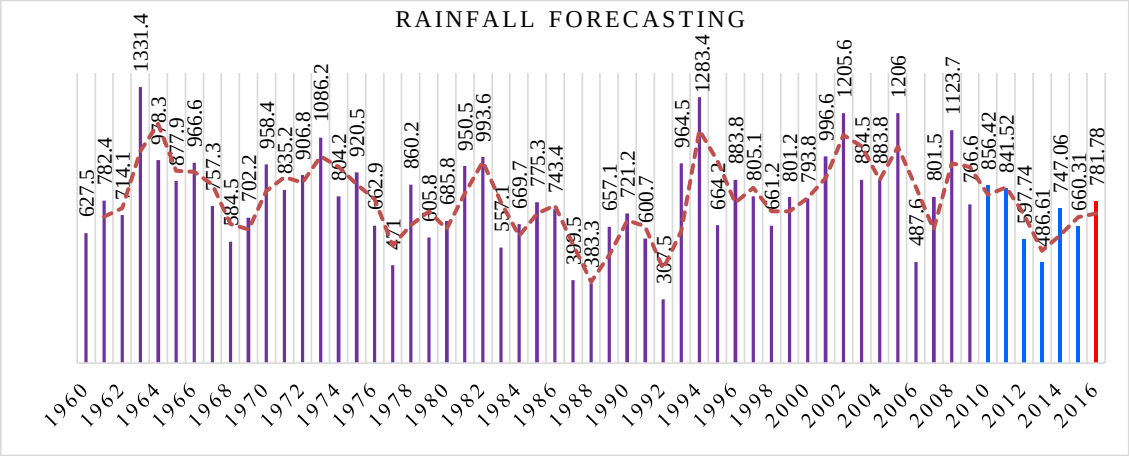

Figure 3.3 The rainfall result of the Minitab software.

3.5 Forecast Rainfall by SPSS Method

Full name of SPSS program is statistical package for the social sciences, which is a computer software to analyze the statistic. SPSS was bought by IBM on June 2552 and it was changed the new name “SPSS: An IBM company”. The SPSS program is developed

by Norman H. Nie and C. Hadlai Hull Noman which presently, they are professor in Stamford University.

3.5.1 Time series modeling and forecasting

The forecasting add-on module provides two procedures for accomplishing the tasks of creating models and producing forecasts. The Time Series Modeler procedure estimates exponential smoothing, unilabiate autoregressive integrated moving average (ARIMA), and multivariate ARIMA (or transfer function models) models for time series, and produces forecasts. The procedure includes an expert odeler that automatically identifies and estimates the best-fitting ARIMA or exponential smoothing model for one or more dependent variable series. Thus eliminating the needs to identify an appropriate model through trial and error. Alternatively, you can specify a custom ARIMA or exponential smoothing model.

- Create models, use the expert modeler to find the best fitting ARIMA or Exponential Smoothing model for your time-series data.
- Apply model, apply your saved time-series models to create new forecasts.
- Seasonal decomposition, estimate multiplicative or additive seasonal factors for time series.
- Spectral plots, analyze the variation of the series as a whole into periodic components of different frequency.

3.5.2 SPSS methodology

a) Linear models.

Algorithm used in SPSS method is linear models to predict a values. the predication for one or more values is depend on the relationship data that affect to forecast results. Linear models are easily explained in the mathematical model for formulating values. The features are easily understood and can be built model that very quickly compared with other model. example as the neural networks or decision trees model of the same data set. This feature model is based on available of the statistics data. furthermore, can see the content detail of linear models in more information.

b) Generalized linear mixed model.

Linear model combination or mixed models is to extend the dataset model. the calculation target is to get the forecast results that nearly correct with actual data. The related factors and variable cause to forecast function. The results can have a non-normal distribution and the observations correlated. Generalized linear mixed models can cover many models from the simple linear regression to complex multilevel models for non-normal longitudinal data. This feature can be used in the advanced statistics that complex multilevel models .

3.5.3 Data analysis in SPSS software

Data analysis used is the statistical data for a period of 50 years (since 1960 to 2009) of Namkhan river, which has objective to find the forecasted rainfall of 2010-2016. The forecasted function used in SPSS software is the analyze forecasting function for creating the forecaste model. The forecasted results indicated the rainfall prediction in graphic and data model. Accuracy measure of SPSS softeware consist of mean absolute percentage erorr (MAPE), mean asolute erorr (MAE), and R-squared (R^2). These accuracy measures are index for checking acceptable ranges of software accuracy.

The forecasting by SPSS software is shown in figure 3.4 for describing the forecasted results. The forecasted results consist the predicted rainfall, lower confidence limits, and upper confidence remits. The forecasted results in 2016 is average normal year, which shows annual water inflow of 64.25 m^3/s, 88.19 m^3/s of Namkhan 2 and 3 hydropower plants respectively. Accuracy measured in SPSS software is shown MAPE = 19.03, MAD = 13.55, R^2 = 0.8, which are in the acceptable ranges.

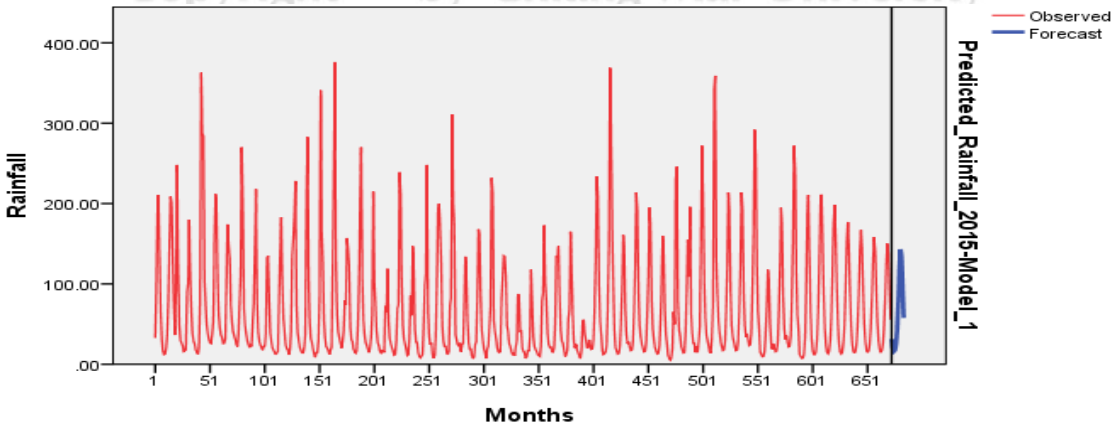

Figure 3.4 The predicted result of the SPSS program.

The simulation results from the SPSS program is shown in figure 3.5. Especially, this SPSS program is not appropriately for predicting a long time, which may be more error. Therefore, the SPSS program is appropriately used for predicting in year to year. Furthermore, SPSS program can be indicated several values that relate with the rainfall forecasting.[3]

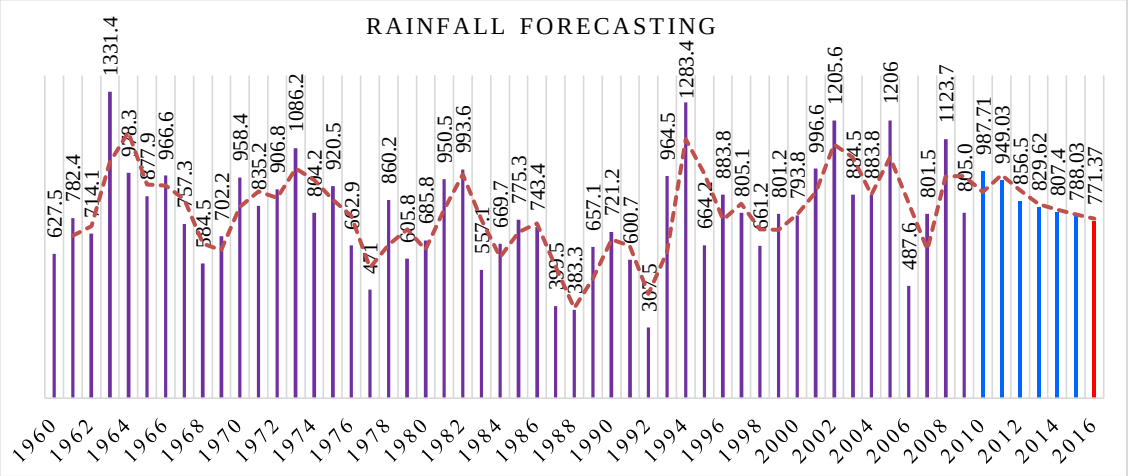

Figure 3.5 The forecasted result of the SPSS program.

3.6 Forecast Rainfall by FFT Software Calculation Online

The forecasted rainfall by using FFT software is one of all methods, which studied the forecaste methodology for finding rainfall in future. The FFT software is online calculation, which is formulation to find the time loop of repeat possibilities for next time. The FFT methodology cannot directly forecasted the rainfall but it will be helped to find out the possible for repeating the rainfall loops. This research is used water criterion of Namkhan 2 hydropower plants for describing water criterions of each years. Namkhan 2 hydropower plants has determined the water criterion as follow. The wet year is annual average of water inflow more than 100 cubic meter per second (m^3/s). The average wet year is annual average of water inflow during 80 to 90 m^3/s. The normal year is annual average of water inflow during 60 to 79 m^3/s. The average dry year is annual average of water inflow during 40 to 59 m^3/s. The dry year is annual average of water inflow less than 40 m^3/s.

3.6.1 Fast fourier transform algorithm

This is formulation of the DFT for computing the efficiently.

$$F(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) e^{-2\pi i x u / N} \quad (3.7)$$

The DFT calculation will evaluates from N values of u. Generally, if maked in the obvious way, which clearly takes N^2 multiplications. The DFT calculation is possible that more efficiently than this, if used the fast fourier transform (FFT) algorithm. those results will be computed the sum of results, is a power of 2 (actually and complex).

Furthermore, the DFT algorithm can use with a little modification that perform an inverse DFT too. Then, it back to the DFT definitions.

$$F(x) = \frac{1}{N} \sum_{x=0}^{N-1} f(u) e^{-2\pi i x u / N} \quad (3.8)$$

This DFT formulation that apart from a factor of 1/*N*. it is a forward DFT, after results will inverse from DFT. Thus results is to calculate an inverse DFT:

- using data in the fourier space data,
- Calculation by using the DFT algorithm,
- The results used from calculated, at the same time, be multiplying by N value.

3.6.2 Forecasted results by using FFT method

The forecasted result from the FFT methology is described by graphical, which shows the time loops for repeating possibility. The data used is historical data of Namkhan River (rainfall 50 years) of water inflow. The forecasted results from FFT online calculation will be compact model. Example, $A \pm Bj$; where A and B is atual amount, j is complex amount. So, the time loop will get from magtitude calculation, which equal to $\sqrt[2]{A^2 + B^2}$ of all values that got from FFT results. The time loop graphical is shown in figure 3.6. Detail results of FFT calculation is described conclusion as below.

- The time loop of rainfall 16,000 m^3/y will be happen in every 41 year.
- The time loop of rainfall 10,000 m^3/y will happen in every 32 year.
- The time loop of rainfall 4,000 m^3/y will happen in every 24 year.
- The time loop of rainfall 2,000-1000 m^3/y will happen in every 12 year.
- The time loop of rainfall 999-500 m^3/y will happen in every year.

- The time loop of rainfall less than 500 m^3/y will happen in every 12 year

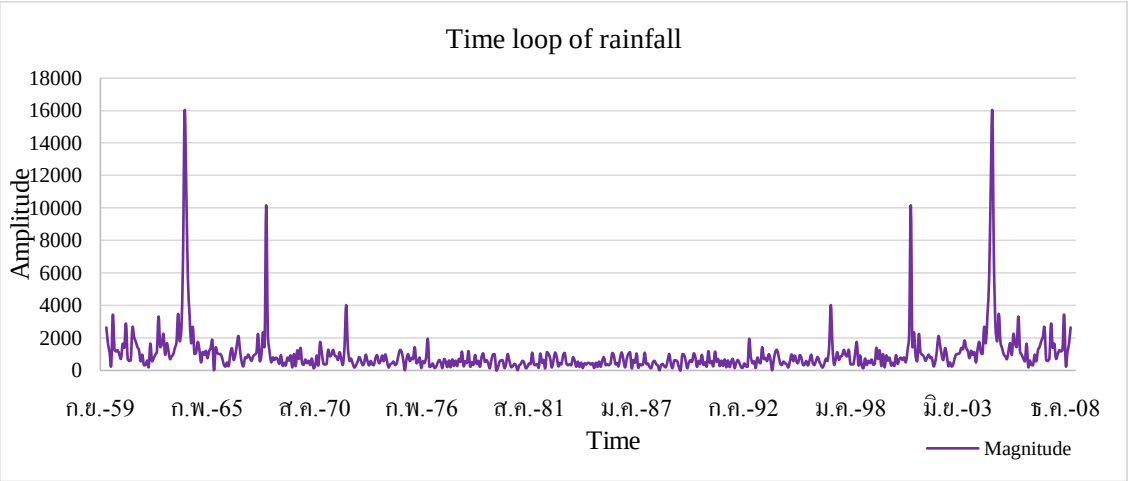

Figure 3.6 The time loop of rainfall result of yhe FFT program.

3.7 Results Conclusion of Four Methodology

3.7.1 Simulation result of rainfall forecasting by four methods

After analyzed and simulated from the rainfall prediction found different values of the simulation result. Due, a features of each softwares have different, which affects to the simulation results for forecasting rainfall since 2010-2016. The simulation results is shown in the table 3.1 and 3.2 of Namkhan 2 and 3 hydropower plants respectively. The forecasted rainfall is considered by accoding the catchment area of Namkhan 2 hydropower plant. The catchment area of Namkhan 2 HPP has 5167 square kilometer (km^2) of 7620 km^2 of total catchment area of Namkhan river. The forecasted results of Namkhan 3 hydropower plant is getted from comparison the catchment area. The catchment area of Namkhan 3 HPP has 7049 square kilometer.

Table 3.1 The forecasted results of the four methods for Namkhan 2 HPP

No.	Forecast Methods	Forecasted Rainfall (m^3) / Year (12 Months)						
		2010	2011	2012	2013	2014	2015	2016
1	By Microsoft Excel	724.59	799.65	932.93	655.94	929.36	814.1	940.49
2	By Minitab software	856.42	841.52	597.74	486.61	747.06	660.31	781.78
3	By SPSS software	987.71	949.03	856.5	829.62	807.4	788.03	771.37
4	Total Average	856.2	863.4	795.7	657.4	827.9	754.1	831.2

Table 3.2 The forecasted results of the four methods for Namkhan 3 HPP

No.	Forecast Methods	Forecasted Rainfall (m^3) / Year (12 Months)						
		2010	2011	2012	2013	2014	2015	2016
1	By Microsoft Excel	988.5	1,090.9	1,272.7	894.9	1,267.9	1,110.6	1,283
2	By Minitab software	1,168.4	1,148	815.5	663.9	1019.2	900.8	1,066.5
3	By SPSS software	1,347.5	1294.7	1,168.5	1,131.8	1,101.5	1,075.1	1,052.3
4	Total Average	1,168.1	1,177.9	1,085.6	896.8	1,129.5	1,028.8	1,134

3.7.2 Conclusion of forecasred results

These forecasted results get from the four methods for analyzing the rainfall prediction, which concluded the content of forecasted rainfall as follow. The forecasted by using forecast function in Microsoft Excel (FFME) is studied the behavior of rainfall in each years to predicts in the future. The forecasted by using Minitab software (MNT) can predicts yearly of the forecasted rainfall. The forecasted by using SPSS software can predicts monthly, which appropreiates year to year of the forecasted. The SPSS is not appropriate for forcasting a long period because it has more deviation value. The FFT is unable directly the forecasted but it can be shown the frequency loop for repeating possibilities. These results are shown in figure 3.7 for showing the different values of each methodologies.

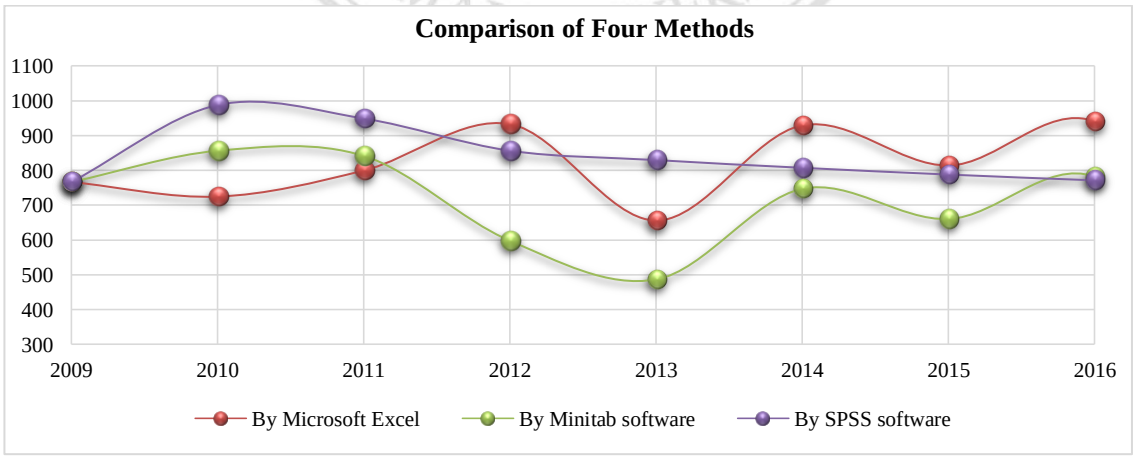

Figure 3.7 The result comparison of four methods.

Note: most importance of forecasted rainfall must studies and analyze to understand the model or behavior of rainfall historical for predicting in future.