Identifying rainfall patterns using Fourier series: A case of daily rainfall data in Sarawak, Malaysia

Aszila Asmat\textsuperscript{1,a}, Sharifah Norhuda Syed Wahid\textsuperscript{2,b}, Sayang Mohd Deni\textsuperscript{3,c}

\textsuperscript{1} Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Cawangan Pahang, Kampus Raub, 27600 Raub, Pahang Darul Makmur
\textsuperscript{2} Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Cawangan Pahang, Kampus Jengka, 26400 Bandar Tun Abdul Razak Jengka, Pahang Darul Makmur
\textsuperscript{3} Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor

\textsuperscript{a}aszila@uitm.edu.my, \textsuperscript{b}sha_norhuda@uitm.edu.my, \textsuperscript{c}sayan929@salam.uitm.edu.my

Abstract. The aim for this paper is to fit Fourier series on the daily rainfall data to describe the rainfall patterns and to identify the time frame of the wettest and the driest period for the studied stations. The estimated parameter of the Fourier series in modelling the mean daily rainfall amount received were obtained using the method of Generalized Linear Model and the performance was evaluated using the analysis of deviance. Following the objective, the daily rainfall data ranges from 41 to 58 years over four stations in Sarawak namely Kuching, Miri, Bintulu and Sibu are studied. Results showed three harmonics is appropriate in modelling the mean daily rainfall received for Kuching, Sibu and Miri stations. However, only two harmonics was found to be fitted the mean daily rainfall amount received significantly for Bintulu station. The wettest month of the studied stations was estimated to occur in late December to early January, whereas the driest month is in March for Miri station, April for Bintulu, July for Kuching and June for Sibu station. The information gained from this study can be used to monitor the occurrence of flood and drought events in the area so that future action can be taken to reduce the impacts.

1. Introduction
Malaysia is located in a stable geographical region and outside from the “Pacific Ring of Fire” which is free from dangerous natural disasters such as earthquakes, volcanic action and tropical typhoon. These natural disasters sporadically influence some of neighbor’s countries such as Indonesia and Philippine. According to [1], Malaysia experiences natural calamity such as regularly hit by haze, landslides and commonly by floods.

Flood is a natural disaster that occurred all over the world. According to the article published by the Department of Irrigation and Drainage (DID), flood can be defined as “body of rising, water, swelling and overflowing land not usually closed thus and also, overflowing of the bank of a stream, drainage system of water onto near land area as result of storm, tidal action, ice melt and channel obstruction” [2]. This definition is supported by [3] that stated flood as “the high of water flow that control the natural...
or artificial banks in any part of the river system. At the point when a riverbank is exceeding the high level of water, the water extends over the flood plain and become hazard to the society”.

Nowadays flood in Malaysia have become a common incidence due to serious climate change and global warming problems that always happened after heavy rainfall. [4] and supported by [5] in their study indicated that the climate changes, use of land, human activities and urban growth has changed the flood trends. No matter how hard or tough the government tried to minimize or stop it, this disaster still happened repeatedly every year. Flood can be divided into two different categories in Malaysia which is flash flood and monsoon flood. [6] in his study stated that monsoon flood usually occurred in seasonal pattern; southwest monsoon occurs late of May to September and northeast monsoon occurs in November to March, but the duration has changed slowly due to climate changes. Meanwhile, flash flood may happen immediately within less than half an hour because of heavy rainfall and the water has no place to drain properly.

Contrast to flood is drought phenomenon that also occurred at certain place in Malaysia. Drought refers as prolonged shortages in the water supply. A drought may happen after certain months or years or may be as early after 15 days. According to [7], flood and drought bring negative impact to society and also to the country including loss of lives and destruction of properties. Furthermore, [8] also agreed that flood is one of the natural disasters that prevents the country development. Therefore, the analysis of the rainfall pattern is very useful so that the findings may contribute to various sectors, such as hydrological, water resource management, tourism and agriculture. In addition, this analysis also can be used for the prediction of the natural disasters, floods and droughts.

2. Time Series Analysis and Fourier Series
Floods in Malaysia are commonly caused by the natural and human activities. Previous study stated that many factors will cause the flood to happen such as temperature, rainfall distribution, wind movements and the natural terrain [9]. This statement is supported by other study on the urban flooding which revealed that pollution, improper drainage system, urbanization management, weather, environment factor and dam break will cause flood to happen [10]. The same environmental issue repeated yearly in Malaysia is flood and drought phenomenon that much related to rainfall distribution. Looking to this phenomenon, the researcher interested to use time series analysis to model flood and drought phenomenon in Sarawak, which is the largest state in Malaysia.

Time series analysis has become an important tool in many applications or sectors including environmental management fields such as rainfall distribution analysis. Currently, time series analysis of rainfall distribution become very important tool in rainfall prediction. The need for accurate rainfall prediction is very useful in providing useful information for reservoir operation, would provide for river control, flood mitigation and also forestry interests [11]. Many methods have been introduced for rainfall modelling and prediction including AR (auto-regressive), ARMA (auto-regressive moving average), ARIMA (auto-regressive integrated moving average), ANNs (artificial neural networks) and that SARIMA (seasonal autoregressive integrated moving average).

Fourier series become famous tool in rainfall prediction [12, 13] whereby the application of Fourier series analysis to seasonal Fourier time series model that can lead to reliable forecast [14]. Rainfall data becomes the key input in a variety of hydrological applications but must be handled properly due to the rainfall process varies in time and space. The generated rainfall models must be able to absorb and represent the variations of time and space which employ different sets of parameters at different time of the year. Therefore, harmonic analysis of Fourier series is suitable technique to capture all the variations while keeping the models parsimonious. As found by [15], the Fourier time series model best fit the results for modelling and forecasting rainfall data than the SARIMA model. In contrast, another study conducted in Nigeria revealed that SARIMA (seasonal autoregressive integrated moving average) model performs better in modelling the rainfall data in Port Harcourt Rivers State from 2000-2014 compared to Fourier series models [11].

According to [16] in his study revealed that Fourier series become very powerful method in connection with ordinary and partial differential equations. In addition, he also suggested that to
examine annual variation of precipitation that is the difference of observed and computed precipitation is analyzed as time series after harmonic simulation of rain gauge data. Furthermore, [17] found that the Fourier series analysis indicates that the data selected belongs to a uniform statistical series which are random numbers. A study conducted by [12] mentioned that periodic functions occur frequently in terms of simple periodic functions such as cosine and sins. For this reason, the occurrences of a series of wet and dry years usually tend to return to the mean pattern even the series is random and abnormal.

3. Data Description
In total 189 river basins (89 of the river basins are in peninsular Malaysia, 78 in Sabah and 22 in Sarawak) are located in Malaysia. Sarawak with an area of 124,450km² stretches over some 700 km along the northwestern coast of the Borneo island is a tropical rainforest with annual rainfall ranging between 3300 mm near the coastland to 4600 mm in further inland [18]. Sarawak experiences a similar climate as other states in Malaysia. However, the time frame for the minimum and maximum amount of rainfall received is almost similar to Sabah. Kuching station located in Sarawak recorded the highest annual rainfall of 5,423.0 mm with an increase of 877.5 mm in 2016 as compared to 2015 (4,545.5 mm). In addition, there are three other stations located in Sarawak namely Miri station, Sibu station and Bintulu station as shown in the following figure 1.

![Figure 1. Location of Study Area](image)

Following the objectives of the study, daily rainfall data were collected from the Malaysian Meteorological Department and analyzed. Four stations involved for the analysis and the length of the data is shown in table 1.

| Station Name | Data               |
|--------------|--------------------|
| Kuching      | 58 years (1951 - 2009) |
| Miri         | 41 years (1968 - 2009) |
| Bintulu      | 58 years (1951 - 2009) |
| Sibu         | 56 years (1953 - 2009) |

4. Methodology
The rainfall characteristics of mean rainfall amount received per rainy day is used in the study and it will be fitted using Fourier series to identify the daily rainfall pattern in Sarawak. A mathematical function, $g(t)$ of Fourier series is
\[ g(t) = A_0 + \sum_{j=1}^{m} (A_j \sin(jt') + B_j \cos(jt')) \]  

where \( j \) is the harmonic, \( m \) is the maximum harmonic required for the series, \( A_j \) and \( B_j \) are the parameter coefficients and \( t' = \pi(t - 183)/183 \).

The maximum likelihood estimates of the coefficients the Fourier series, \( \beta_j \) is expressed as:

\[ \hat{\beta}^{(l-1)} = \hat{\beta}^{(l-1)} + I(\hat{\beta}^{(l-1)})^{-1}U(\hat{\beta}^{(l-1)}) \]  

where;

\[ U(\beta) = \sum_{t=1}^{n} \frac{n(t)}{\mu^2(t)} (x(t) - \mu(t)) \frac{\partial \mu(t)}{\partial \beta_j} \]  

\[ I(\beta) = \sum_{t=1}^{n} \frac{n(t)}{\mu^2(t)} \frac{\partial \mu(t)}{\partial \beta_j} \frac{\partial \mu(t)}{\partial \mu_r} \]

The analysis of deviance will be used to assess the maximum number of harmonics required [19] in the model. According to [20], the analysis of deviance is useful in describing the performance of the Fourier series. This method will be used to determine the significant harmonics and also the number of parameters required for the model. The analysis of deviance uses statistical F-test and the procedures of the statistical test include several steps. Firstly, the deviance is calculated at each time the model is fitted. Then, the models with different number of harmonics will be compared by reduction of the deviance. The deviance, \( D^2 \) is expressed as:

\[ D^2 = 2 \sum_{i=1}^{n} n(t) [\ln(\hat{\mu}(t)) - \ln x(t)] \]  

where \( \hat{\mu}(t) \) is the fitted value of \( \mu(t) \) and

\[ \ln x(t) = \frac{\sum_{i=1}^{n(t)} \ln x_i(t)}{n(t)} \]

5. Results and Discussion

The mean amount of rainfall received per rainy day from four stations in Sarawak, Malaysia were analyzed with different number of harmonics to illustrate the fitting of the Fourier series. The deviance is calculated for each different number of harmonics that is fitted to the model. Then, the analysis of deviance is conducted and showed in table 2. The observed values and fitted curves for each station were also plotted and showed in figure 2 to justify the fitness of the selected model.

The models with different number of harmonics are then compared by considering reduction in deviance. The maximum numbers of harmonics required for the model can be determined when no further harmonics reduced the deviance significantly.

The analysis of deviance in table 2 showed that when one harmonic is fitted to the mean amount of rainfall received in Kuching, the deviance is significant since p-value is less than 0.05. When applying two harmonics and three harmonics into the model, the deviances are still reduced significantly. However, when four harmonics is fitted, the reduction of deviance is insignificant (p-value is greater than 0.05) suggests the three harmonics are adequate to represent the amount of rainfall received in Kuching. The model for Kuching station is expressed as follows:

\[ g(t) = 2.898 - 0.102 \sin(t) + 0.114 \sin(2t) - 0.124 \sin(3t) - 0.26 \cos(t) + 0.072 \cos(2t) - 0.024 \cos(3t) \]
Table 2. Analysis of deviance

| Source   | df | Kuching          | Sibu          | Bintulu         | Miri          |
|----------|----|------------------|---------------|-----------------|---------------|
|          |    | Deviance | P-value | Deviance | P-value | Deviance | P-value | Deviance | P-value |
| 1-harmonic | 2  | 603.64 | 0.000 | 61.67   | 0.000 | 59.86 | 0.000 | 57.86 | 0.000 |
| 2-harmonic | 2  | 146.93 | 0.000 | 10.25   | 0.000 | 31.22 | 0.000 | 28.19 | 0.000 |
| 3-harmonic | 2  | 103.09 | 0.000 | 12.54   | 0.000 | 3.29 | 0.322 | 17.4  | 0.000 |
| 4-harmonic | 2  | 6.33   | 0.087 | 3.21    | 0.246 | 4.9  | 0.142 | 3.86  | 0.285 |
| 5-harmonic | 2  | 1.42   | 0.577 | 2.46    | 0.368 | 22.02 | 0.000 | 18.54 | 0.000 |
| Residual  | 354| 457.04 | -     | 486.35  | -     | 589.8 | -     | 636.63 | -     |

The results have been plotted in the figure 2(a) together with the observed values of the mean rainfall received per rainy day and it shows that the observed values are fitted well by the three harmonics.

Figure 2. Observed and fitted amount of rainfall received at (a) Kuching, (b) Sibu, (c) Bintulu and (d) Miri

The results for both Sibu and Miri stations indicate that three harmonics also are adequate to the model since no further harmonics reduce the deviance significantly. The results have been plotted in figure 2(b) and figure 2(d) along with the observed values and it shows that the three harmonics of
Fourier series fitted the data very well. Thus, the model for Sibu and Miri can be expressed as in equation 8 and equation 9, respectively.

\[ g(t) = 2.826 - 0.066\sin(t) + 0.047\sin(2t) - 0.007\cos(t) - 0.107\cos(2t) - 0.055\cos(3t) \]  
\[ (8) \]

\[ g(t) = 2.832 + 0.111\sin(t) - 0.018\sin(2t) - 0.055\sin(3t) - 0.042\cos(t) + 0.077\cos(2t) - 0.027\cos(3t) \]  
\[ (9) \]

However, when one harmonic was fitted to the mean amount of rainfall received for Bintulu station, the deviance values were very large indicating that the model was inadequate. By applying two harmonics into the model, it shows significant reduction in deviance (p-value is less than 0.05). Therefore, two harmonics is adequate to represent the amount of rainfall received in Bintulu and the model is expressed as follows:

\[ g(t) = 2.970 + 0.05\sin(t) + 0.012\sin(2t) - 0.082\cos(t) + 0.072\cos(2t) \]  
\[ (10) \]

The observed and fitted mean rainfall received per rainy day for Bintulu station is shown in figure 2(c) and the curve seems to fit the observed data adequately.

The analysis is then continued with determination of the wettest and the driest period at the four stations. As shown in figure 3, the extremely high mean amount of rainfall received per rainy day is recorded at Kuching, followed by Bintulu with fitted values of 31 and 23 mm. The other stations record a fairly same extremely high mean value which is 20 mm per day. Meanwhile, the extremely low mean amount of rainfall received per rainy day is observed at Miri with 13 mm per day, followed by Kuching and Sibu with 14 mm per day.

Through the observation, extreme values may vary for one station to other stations. Therefore, in order to identify the peaks for each of the studied stations, the lower and upper quartile of thresholds are calculated. These threshold values are showed as dash line in figure 3 used to project the time frame of the wettest and driest period over the stations. This study uses the 2nd percentile to identify the driest period, whereas the 98th percentile to identify the wettest period.

Figure 3(a) and figure 3(b) shows a single peak can be observed for Kuching and Sibu stations. The northeast monsoon leads to the highest mean rainfall amount received for both stations during the period of middle to the late January. The highest peak observed was on 27th January and 18th January with 31.07 mm and 20.16 mm per day respectively. In contrast, the lowest mean rainfall amount recorded in the late July for the Kuching station, whereas the late June to the early July for the Sibu station. These periods considered as the driest period for both stations, and it was strongly affected by the southwest monsoonal flow.

A similar rainfall patterns also can be observed for the other two stations in Sarawak. As shown in the figure 3(c), the period of the late December can be considered as the wettest period for the Bintulu station with the highest mean rainfall amount was observed of 22.74 mm on 28th December. Meanwhile, the driest period was occurred in between the middle and the late April, with the lowest mean rainfall amount was observed on 19th April with values of 16.98 mm per rainy day. The rainfall pattern at this station had been influenced by the northeast monsoon and the first inter-monsoon seasons.

However, figure 3(d) shows that the period of the late December to the early January can be considered as the wettest period for the Miri station with the maximum peak of 19.66 mm per rainy day, which was observed on the 2nd January. Whereas, the driest period, in term of the lowest mean rainfall amount was recorded on the middle March. The rainfall pattern at this station also had been influenced by the northeast monsoon and the first inter-monsoon seasons.
According to the results, the two and three harmonics are best described by unimodal pattern. These results are consistent with the earlier studies that had been done by [4]. The study found that two harmonics are significant to model the rainfall amount received in Sandakan, Kota Kinabalu and Labuan stations and the two harmonics are also best described with unimodal pattern. Another study conducted by [21] in Johor found that the Hospital Kota Tinggi station shows that three harmonics is sufficient to model the rainfall amount at the station. The result shows the three harmonics also depict a unimodal pattern.

6. Conclusion
The study is aimed to fit Fourier series in the daily rainfall data in Sarawak to describe the rainfall patterns over the studied stations. Results from the analysis showed that three harmonics is significant to model Kuching, Sibu and Miri stations, whereas only two harmonics is appropriate for Bintulu station. According to the findings, the two or three harmonics are best described with unimodal pattern. In terms of wettest and driest periods, January and December can be considered as the wettest period for the studied stations, while driest months were June for the Sibu station, July for the Kuching station, April for the Bintulu station and March for the Miri station.

The information gained from this study benefits the Malaysian Meteorology Department so that the occurrence of flood and drought events can be monitored in the area. Continuous monitoring of the events is essential so that appropriate actions can be taken to reduce the impacts. Apart from using the rainfall characteristic of the mean rainfall per rainy day, it is proposed that future study to use other
characteristic such as the probability of wet and dry event to identify the rainfall pattern over the studied stations.

References

[1] Parker D J, Islam N and Chan N W 1997 Chapter 3: Reducing vulnerability following flood disaster: Issues and practices A. Awotona (Ed.) Reconstruction after disaster pp 23-24 (London: Avebury)

[2] Drainage and Irrigation Department 2012 Annual Flood Report of DID for Peninsular Malaysia. Unpublished report (Kuala Lumpur: DID)

[3] Ching Y C, Baharudin Y, Ekhwam T, Maimon M A, Salmijah S and Lee Y H 2013 Impacts of Climate Change on Flood Risk in the Muar River Basin of Malaysia Disaster Advances 6(10) pp 11-17

[4] Amini A, Ali T M, Ghazali A, Aziz A A and Akib S M 2011 Impacts of land-use change on streamflows in the Damansara Watershed, Malaysia Arabian Journal for Science and Engineering 36 pp 713-720

[5] Asmat A, Deni S M and Ujang S 2018 Identifying Rainfall Pattern in Sabah and Labuan, Malaysia. American Institute of Physics 1974 pp 040021(1-8)

[6] Musa Hasni N A 2014 The Peninsular Malaysia Flooding - A Spatio-Temporal Analysis of Precipitation Records (Faculty of Civil Engineering and Earth Resources, Universiti Malaysia Pahang)

[7] Kourgialas N N and Karatzas G P 2011 Flood management and a GIS modelling method to assess flood-hazard areas-A case study Hydrological Sciences Journal 56 (2) pp 212-225

[8] Elias Z, Hamin Z and Othman M B 2013 Sustainable Management of Flood Risks in Malaysia: Some lessons from the legislation in England and Wales Procedia-Social and Behavioral Sciences 105(1) pp 491-497

[9] Balek J 1983 Hydrology and Water Resources in Tropical Regions Developments in Water Science 18

[10] Kong Y Y, Noor Azima Bahrun and Yew K A 2010 Study on the urban flooding

[11] Wiri L and Sibeate P U 2020 A Comparative Study of Fourier Series Models and Seasonal - Autoregressive Integrated Moving Average Model of Rainfall Data in Port Harcourt. Asian Journal of Probability and Statistics 10 (3) pp 36-46

[12] Kreyszig E 1985 Advanced Engineering Mathematics Mohinder Singh Sejwal for wily Eastern limited (Paris)

[13] Iwok I A 2016 Seasonal modelling of Fourier series With Linear Trend International Journal of Statistics and Probability 5 pp 6-10

[14] Omkara C O and Okere O E 2010 Fourier series models through Transformation Global Journal of Mathematical Sciences 18 pp 87-94

[15] Etuk E H and Bartholomew U 2012 Forecasting Nigerian inflation rates by a seasonal ARIMA model Sendra Academic Publishers Burnaby, British Columbia Canadian 6 (3)

[16] Yevjevich V 1992 Probability and Statistics in Hydrology Water Resources Publication (USA)

[17] Afshar N R and Fahmi H 2012 Rainfall Forecasting Using Fourier series Journal of Civil Engineering and Architecture 6 (9) pp 1258-62 ISSN 1934-7359 USA

[18] Kottek M, Grieser J, Beck C, Rudolf B and Rubel F 2006 World Map of the Köppen-Geiger Climate Classification Updated Meteorologische Zeitschrift 15 (3) pp 259-263.

[19] Jamaludin S and Jemain A A 2009 Meteorology Appl. 16 pp 391-401

[20] Kottegoda N T, Natale L and Raiteri E 2004 Journal of Hydrology 296, pp 23-37

[21] Hanisah S and Suhaila J 2014 Genrealized Linear Models (GLMs) Approach in Modelling Rainfall Data over Johor Area AIP Conference Proceedings 1605 pp 979-985