



Techno Science Africana Journal

Journal homepage: www.technoscienceafricana.com



FLOOD RISK ASSESSMENT USING RAINFALL AND STREAM FLOW AT SEGMAT RIVER BASIN, MALAYSIA

ADAMU MUSTAPHA

Department of Geography, Faculty of Earth and Environmental Studies, Kano University of Science and Technology, Wudil,

ARTICLE INFO

Keywords:

Mann-Kendal test, Flood risk, Segamat River, Non parametric statistics, ARIMA model

Correspondence author:

e-mail:

amustapha494@gmail.com

Phone:

ABSTRACT

This study presents non parametric statistical analyses for flood forecasting using rainfall and stream flow data of monitoring stations at Segamat River, Malaysia with the following objectives (i) to develop time series flood forecasting model (ii) to investigate the correlation between selected rainfall stations and stream flow and (iii) to obtain a trend analysis of the rainfall and stream flow data. Autoregressive Integrated Moving Average (ARIMA) modeling was computed to get a suitable ARIMA model for the dataset. Spearman's correlation coefficients test was used to test the correlation between rainfall and stream flow. Trend analysis was conducted using Mann-Kendall trend test to detect the monotonic trend direction and magnitude of change. Results revealed that ARIMA (0, 1, and 2) model was the most suitable for all the data obtained at three rainfall and stream flow stations, it was suitable to forecast flood events at Segamat River using monthly rainfall and stream flow data. The finding of ARIMA (0, 1, and 2) model was verified by the minimum value of Akaike Information Criteria (AIC), Mean Absolute Percentage Error (MAPE) and Pearson Correlation Coefficients tests between predicted and observed data. Spearman's correlation coefficient showed a significant relationship between rainfall and flooding. Mann-Kendall trends test showed that there were significant upward and downward trend detected for rainfall station at the study area ($S= 1372, p < 0.05$), $S= -886, p < 0.05$).

1. INTRODUCTION

Flood is a natural disaster which often causes property damage, disease and loss of life. Efficient flood warning is required to apply mitigations and to enable evacuation from the risky area (Sene, 2010). Types of flood occur in Malaysia are seasonal flood during monsoon season and sometimes flash flood occurs in cities.

Malaysia experienced North-East Monsoon season between November to March while Southwest Monsoon between May to September (Billa et al., 2004; Razi et al., 2010).

Flood forecasting and flood mitigation are crucial because flood events caused by continuous heavy rainfall is unavoidable.

According to Billa *et al.*, (2004), Malaysia had applied various flood forecasting and warning systems based on an advanced hydraulic model, but an advanced hydraulic model was not the best model to predict impending floods. Various flood mitigation also had been practiced such as widening and deepening of major rivers, embankments, construction of dam and retention ponds (Chan, 1997). The latest flood mitigation had been successfully implemented in Malaysia was a diverting flood system, namely Storm Water Management and Road Tunnel (SWMART). This tunnel had reduced flood problem in Kuala Lumpur city centre. Factors that can lead to flood event are heavy rainfall, large concentration of runoffs and low infiltration that will exceed the river capacity (Razi *et al.*, 2010; Sene, 2010). Chan (1997) reported that developments on flood plain also contributed to flood occurrences. Mohd *et al.*, (2006) reported that 6.3% of total land uses on floodplain in Segamat are town area, 42.3% was rubber plantation and 43.4% was palm oil plantation. Increasing of flood events may due to changes of land use (Gasim *et al.*, 2006; Adnan and Atkinson, 2011). Changes of land use also will change the factors that can lead to flood event in an area. Gasim *et al.*, (2010) had identified the main causes of flooding in the study area were extra ordinary heavy rainfall, low and undulating topography of Segamat area.

Major flood occurs in Johor on 2006 and 2011 are disastrous, for example flood in 2006 caused 16 death and 104 people were evacuated. The latest flood events were on January and February 2011. On 31st January 2011, two people died and 28,932 were evacuated in Johor floods (The Straits Times, 2011). On 1st February 2011, 42 300 flood victims were evacuated from Segamat, Muar, Kota Tinggi, Kluang and Ledang, the flood and landslides also forced Keretapi Tanah Melayu (KTM) to

cancel several train services to the south (The Straits Times, 2011).

Flood forecasting is very important for better flood risk management. The efficient early-warning system can help us in order to improve emergency rescue plan and to avoid loss of life. Statistical analysis such as trend analysis is required to help the managers and local authority in understanding the behaviour of flood pattern and its risk potential in order to develop accurate flood forecasting model. Flood occurs at Segamat area during North-East Monsoon season between Novembers to March. Monsoon is the main factor of heavy rainfall at Segamat (Suhaila *et al.*, 2010). However, Rainfall can be affected by complex relationship between wind, temperature, evaporation and topography of an area (Grist and Nicholson, 2001). The impact of temperature and wind on rainfall can be observed before the flood event in Segamat on December 2006 (Gasim *et al.*, 2010). Cold air from the high pressure area in Siberia had strengthened the velocity of northeastern wind at the South China Sea thus brought heavy rainfall to Johor State which caused the major flood (Fredolin *et al.*, 2008; Gasim *et al.*, 2010).

Lana-Renault *et al.*, (2011) had reported that different type of catchment area such as agriculture or forest area could affect stream flow. Based on a study conducted by Juahir *et al.*, (2010), downward trend of a stream flow was affected by increasing of urban development and decreasing of vegetation. Decreasing of stream flow also might be affected by dam in-filling at the upstream (Juahir *et al.*, 2010). Segamat River had a dam located at the upstream, Juasseh Dam. Decreasing trend of stream flow probably may be affected by the Juasseh Dam. Ground water storage and soil moisture within the basin also might influence the decreasing trend of the stream flow. Douglas *et al.*, (2000)

reported that capacity of water contains in ground water storage and moisture capacity in the soil within the basin may be impacted by climate change and rainfall trend.

The objectives of this study are (i) to develop time series flood forecasting model; (ii) to investigate the correlation between selected rainfall station and stream flow; (iii) to obtain the trend of the rainfall and stream flow data.

2. MATERIALS AND METHODS

2.1 Description of Study area and hydrological data

Segamat River is a tributary of Muar River which flows through Segamat Town. It is located at about $102^{\circ} 49''$ E and $2^{\circ} 30.5''$ N. The river has a drainage length of 23 km with average width of 40 m with elevation of 14 m above sea level. About 70% of the catchments is highland rising up to 1000 m and the remainder another 30% are undulating highland and swampy areas. According to Segamat District Council (2010), Segamat district consists of 269,709 people which formed 8.6% of the total population Johor State.

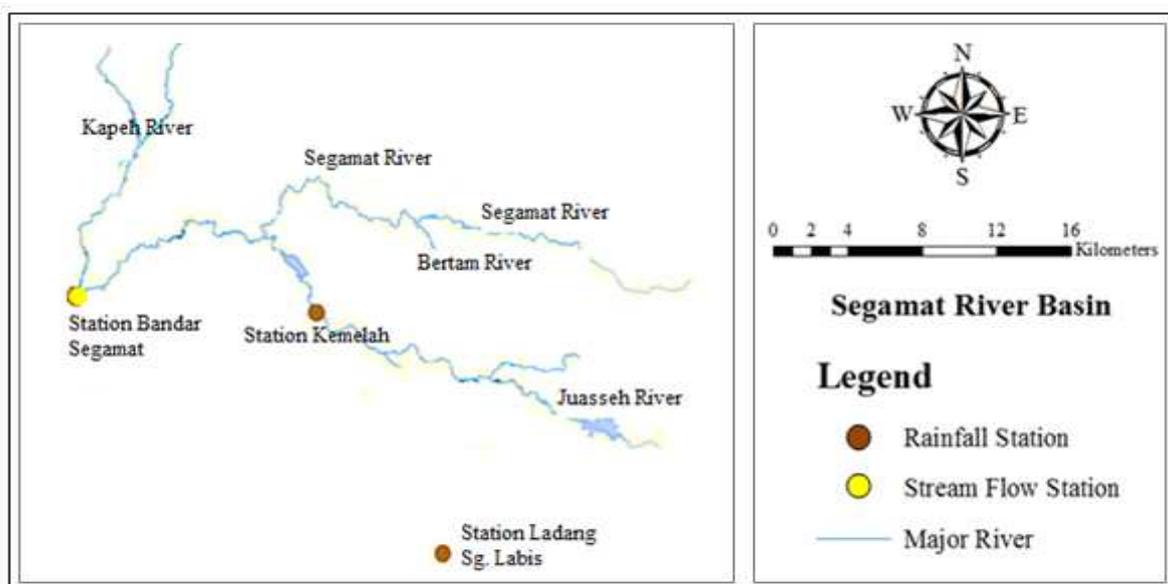


Fig.1 Map of the study area showing Segamat River Basin Malaysia

2.2 Methods

Daily and monthly rainfall data of December 2008, March 2009, July 2009, October 2009, January 2010, May 2010, August 2010, November 2010 and February 2011 at three rainfall stations covering Segamat River Basin at Kemelah, Ladang Sungai Labis and Bandar Segamat stations (Figure 1) were obtained from Department

of Irrigation and Drainage (DID). These three rainfall stations were selected because rainfall events at the stations can give rise to stream flow at Bandar Segamat. Stream flow station was chosen at Sungai Segamat covering from 2008 to 2011 likewise. Table 1 shows detail information of stream flow and rainfall stations selected in this study.

Table 1 Stream flow and rainfall stations at Segamat River Basin

Station Number	Station Name	Latitude	Longitude	Recording Type
2529001	Kemelah	02° 30 00 N	102° 56 53 E	Rainfall, Manual
2330009	Ladang Sg. Labis	02° 23 05 N	103° 01 00 E	Rainfall, Manual
2528002	Bandar Segamat	02° 30 30 N	102° 49 05 E	Rainfall, Manual
2528414	Sugai Segamat	02° 30 25 N	102° 49 05 E	Stream Flow, Automatic

The raw data were treated by logarithmic transformation (log scaling), column scaling, and column auto-scaling. All log transformed variables were also z-scale standardized to minimize the effects of different units and variance of variables and equally to transform the data into dimensionless format. The data were analyzed using descriptive statistics, non parametric statistics and were interpreted into tables, histograms. All statistical analyses were conducted using statistical packages for social sciences (SPSS) version 21 and XLSTAT Excel for window.

2.2.1 ARIMA modeling approach

Autoregressive integrated moving average (ARIMA) was a widely used time series model (Ömer Faruk, 2010). ARIMA model may autoregressive moving average or autoregressive integrated moving average (Chattopadhyay and Chattopadhyay, 2010; Wu and Chau, 2010). The model can be written as AR (p) for autoregressive, MA (q) for moving average, ARIMA (p,q) for autoregressive moving average and ARIMA (p,d,q) for autoregressive integrated moving average, where p was the autoregressive order, d was integrated order and q was moving average order (Chattopadhyay and Chattopadhyay, 2010 ; Wu and Chau, 2010). The p, d and q values vary from the range of 0 to 12.

Firstly, autocorrelation function (ACF) and partial correlation function (PACF) of the time series were investigated to get suitable Auto Regressive (p) value (Gemitzi and Stefanopoulus, 2011; Chattopadhyay and Chattopadhyay, 2010; Weisang and Awazu, 2008). ACF and PACF

plot provide lag number of the time series. The highest spike of lag number indicates p value of a model (Weisang and Awazu, 2008). Another two orders, d and q can be determined by generating ARIMA model. An ARIMA (p,d,q) process is given as in Equation 1, 2 and 3 (Brockwell and Davis, 2002; Box et al., 1994).

$$\phi(B)\Delta^d z_t = \theta_0 + \theta(B)a_t \tag{Eq.1}$$

Here, B is the backward shift operator;

$$\phi(B) = 1 - \sum_{k=1}^p \phi_k B^k \tag{Eq. 2)}$$

$$\theta(B) = 1 - \sum_{k=1}^q \theta_k B^k \tag{Eq.3}$$

Minimum value of Akaike Information Criteria (AIC) and Mean Absolute Percentage Error (MAPE) indicates the best fit model (Ömer Faruk, 2010). Chattopadhyay and Chattopadhyay (2010) stated that the best fit model can be determined by the minimum value of AIC which is a screening statistics.

In this study, Pearson correlation test between observed and predicted data computed to verify the best ARIMA model. AIC, MAPE and Pearson correlation mathematical formulae were given as Equation 4, 5 and 6 (Habib et al., 2001; Ömer Faruk, 2010).

$$AIC = n \left[\ln \left(\frac{2\pi \sum R^2}{n} \right) + 1 \right] + 2x$$

(4)

Where $x = (p + q)$ is the number of terms estimated in the model and R^2 is squared residuals.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{p_i - o_i}{p_i} \right| 100$$

(5)

Where p_i and o_i are the predicted and observed data and n is the number of data.

$$r(X, Y) = \frac{\overline{XY} - \bar{X}\bar{Y}}{\sqrt{\sum(X - \bar{X})^2 \cdot \sum(Y - \bar{Y})^2}}$$

(6)

Where X and Y denotes rainfall data and stream flow data and n represent the sample size.

In this study, 8 years of rainfall and stream flow data from 2001 to 2008 were used. The other 2 years data for the year 2009 and 2010 were used to calibrate the obtained best fit ARIMA model by comparing between predicted and observed data.

2.2.2 Spearman Rank Correlation Coefficient

Correlation coefficient is a measure of the strength of the linear association between variables. Points lying precisely along a straight line with positive slope will have a correlation of +1, while points lying along a negative slope will have correlation of -1. Points that are randomly scattered on the graph will have a correlation close to zero. Strong correlation coefficient is indicated by > 0.7 , moderate correlation by $0.5-0.7$ and weak correlation by < 0.5 . The close inspection of correlation matrix was very useful, because it can point out associations between variables and

indicate the participation of individual in the model.

Spearman's correlation Coefficient test was computed to test the relationship between rainfall data and stream flow at Segamat River Basin. Spearman's test is the best correlation test adopted for non-parametric data and is given in Equation 7 (Siegel, 1957; Kendall, 1975). Where; X and Y denotes rainfall and stream flow, X_i and Y_i denotes the rank of both variables and n denotes the size of the sample. Correlation coefficient of 1 and -1 indicates the variables are perfectly positively and negatively correlated respectively. Correlation coefficient range from 0.5 to 1 indicates strong positive correlation and value range from 0.1 to 0.5 indicates weak positive correlation. While the correlation coefficient of 0 indicates the variables were not correlated. The sign in front of the correlation coefficient value determine the direction of the relationship. A plus sign denotes positive relationship and a minus sign denotes negative correlation. The correlation (r) provides a standardized measure of the linear association between two variables as given in Equation 7 and 8.

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n - 1) s_x s_y}$$

(Eq.

7)

Where x and y are the bivariate to be correlated, s_x and s_y are the sample standard deviations of variable x and y respectively.

$$r_s(X, Y) = 1 - 6 \frac{\sum(X_i - Y_i)^2}{n^3 - n}$$

(Eq.

9)

2.2.3 Mann-Kendall non- parametric test of trend.

Temporal trends of rainfall and stream flow data were analyzed using Mann-Kendall trend analysis. Mann-Kendall trend test (Kendall, 1975; Mann, 1945) is a non parametric rank-based statistical test which is suitable to data sets with seasonality and missing values (Hirsch et al., 1982). The test also can provide positive or negative trends with linear or non linear regression (Van Gelder et al., 2007; Chattopadhyay et al., 2011). According to a study conducted by Yue et al. (2002), Mann-Kendall trend test is a powerful tool to detect significant trends in hydro-meteorological time series and the test was dependent on distribution type, sample size, magnitude of trend and variation of the time series. The difference between observed values ($y_j - y_i$), where $j > i$, assign the integer values of 1, 0, or -1 to positive differences, no differences, and negative differences (Juahir et al., 2010). The Mann-Kendall test statistic, S was given as in Equation 9. Positive S value indicates upward trend, negative S value indicates downward trend and zero S value indicates no trend. Significant trend is indicated by p value which less than 0.05.

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^n \text{sign}(x_j - x_k) \tag{Eq. 9}$$

Where S is the Mann-Kendall test value, x_j and x_k are the sequential data values and j, k and n are the length of the data. S value is assumed to be 0 (e.g. no trend) if the value of S shows a positive value (incremented) whereas a negative value of S shows decremental as in equation 10.

$$\text{sign}(x_j - x_k) = \begin{cases} (x_j - x_k) = 1 & \text{if } x_j - x_k > 0 \\ (x_j - x_k) = 0 & \text{if } x_j - x_k = 0 \\ (x_j - x_k) = -1 & \text{if } x_j - x_k < 0 \end{cases} \tag{Eq. 10}$$

A very high positive value of S is an increasing trend, and a very low negative value indicates a decreasing trend. In situation where n is greater than 10, the standard normal variate is computed to determine the trend in the data sets. The normalized test statistics Z is computed using the equation 11.

$$Z = \begin{cases} \frac{S - 1}{\sqrt{\text{VAR}(S)}} & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ \frac{S + 1}{\sqrt{\text{VAR}(S)}} & \text{if } S < 0 \end{cases} \tag{Eq. 11}$$

Where Z is the normalized test statistics, S is Mann-Kendall statistics and VAR(S) is the Mann-Kendall test of variance. The variance of S can be computed using the equation 12.

$$\text{VAR}(S) = \frac{1}{18} \left[n(n-1)(2n+5) - \sum_{p=1}^g t_p(t_p-1)(2t_p+5) \right] \tag{Eq.12}$$

Where n indicates the number of data points, g indicates the number of tied groups, and t_p indicates the number of data points in the pth group.

3. RESULTS AND DISCUSSION

3.1 ARIMA Modeling

From the ACF and PACF plot of the rainfall and stream flow data, the highest spike of lag number was zero. This means the suitable p value for ARIMA model for these data was zero. After generating different model of ARIMA (0, p, q), the fit model for the three rainfall stations and a stream flow station was ARIMA (0, 1, 2). ARIMA (0, 1, 2) process can be written as in Equation 13 after Brock well and Davis, 2002; Box et al., 1994.

$$Y(t) = \mu - \theta_1 \epsilon(t-1) - \theta_2 \epsilon(t-2) \tag{Eq. 13}$$

Where $Y(t) = Y(t) - Y(t-1)$, $e(t)$ = forecast error at period t , $\theta_1 = MA(1)$ coefficient and $\theta_2 = MA(2)$ coefficient.

The ARIMA (0, 1, 2) model was tested using rainfall and stream flow data set. Table 2 produces three verifications used in the model: AIC, MAPE and Pearson correlation tests. The lowest value of AIC

and MAPE indicated that the model is fit. To verify correlation between observed and predicted value, Spearman Rank correlation test was computed and the predicted values of the models showed significant correlation with observed value at alpha 0.05.

Table 2 ARIMA (0, 1, 2) Model Results

Station	Location	Value		Verification measures		
		MA(1)	MA(2)	AIC	MAPE	Spearman Rank's r*
Rainfall	Kemelah	-0.986	0.069	542.34	108.68	0.615
	Ladang Sg. Labis	-0.964	0.09	511.52	143	0.952
	Bandar Segamat	-0.795	-0.205	478.95	166	0.995
Stream flow	Bandar Segamat	-0.866	-0.134	851.09	408.41	0.991

AIC = Akaike Information Criteria, MAPE = Mean Absolute Percentage Error, MA = Moving Average

Figure 2, 3, 4 and 5 showed scatter plot of observed and predicted rainfall and stream flow ARIMA (0, 1, 2). This model showed the lowest AIC of 542.34 for Kemelah Station, 511.52 for Station Ladang Sg. Labis, 478.95 for Station Bandar

Segamat and 851.09 for stream flow station at Bandar Segamat. MAPE for this model were 108.68 for Station Kemelah, 143.0 for Station Ladang Sg. Labis, 166.0 for Station Bandar Segamat and 408.41 for stream flow station.

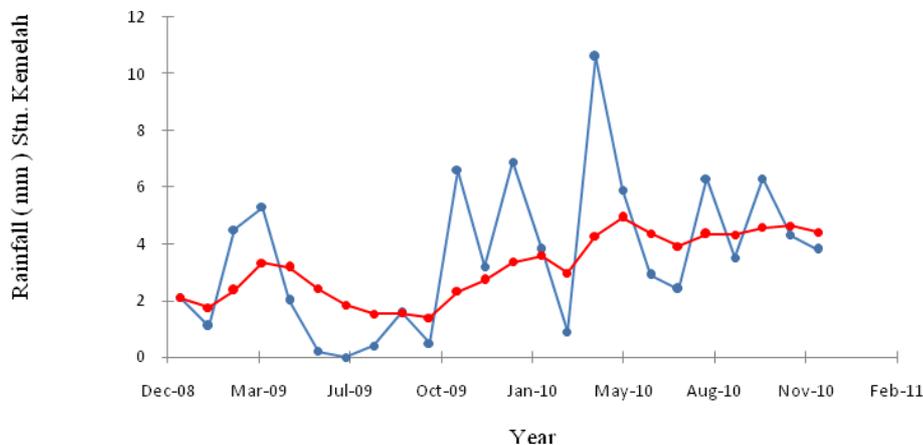


Figure 2 Scatter plot of observed and predicted values of rainfall event at Station Kemelah

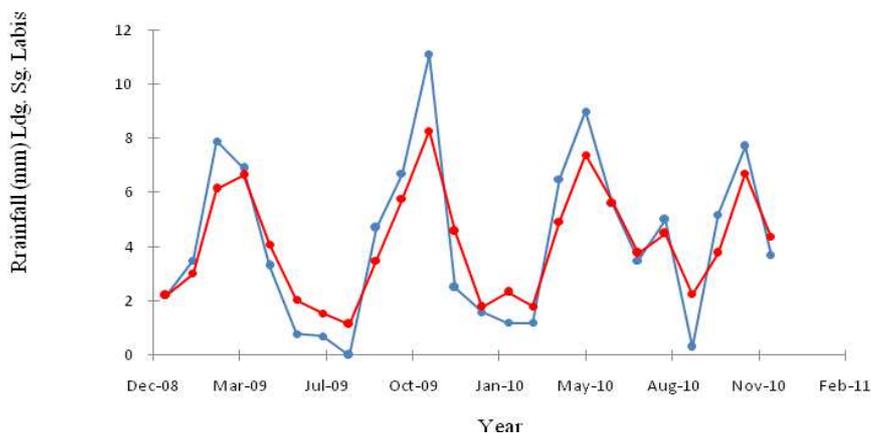


Figure 3 Scatter plot of observed and predicted values of rainfall event for Station Ladang Sg. Labis

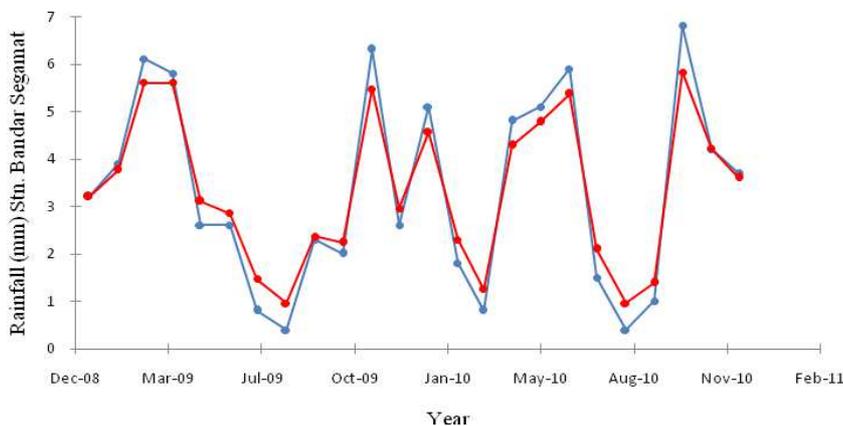


Figure 4 Scatter plot of observed and predicted values of rainfall event at Station Bandar Segamat.

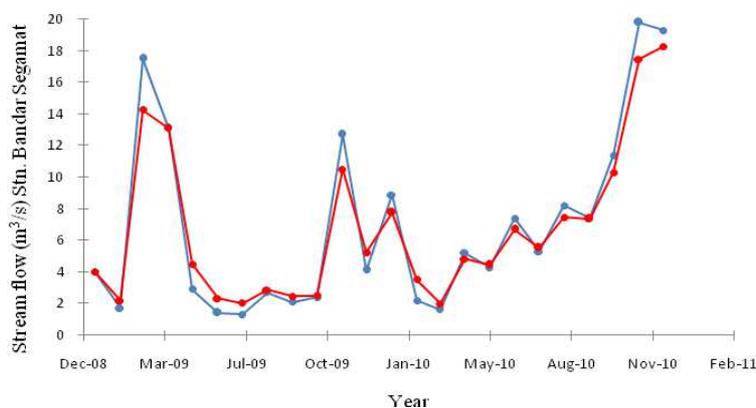
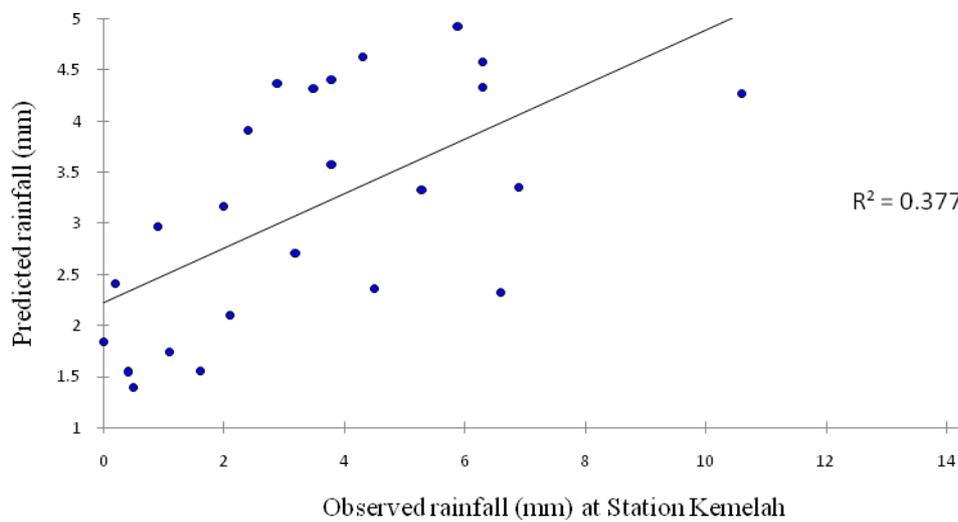


Figure 5 Scatter plot of observed and predicted values of stream flow at Station Bandar Segamat.

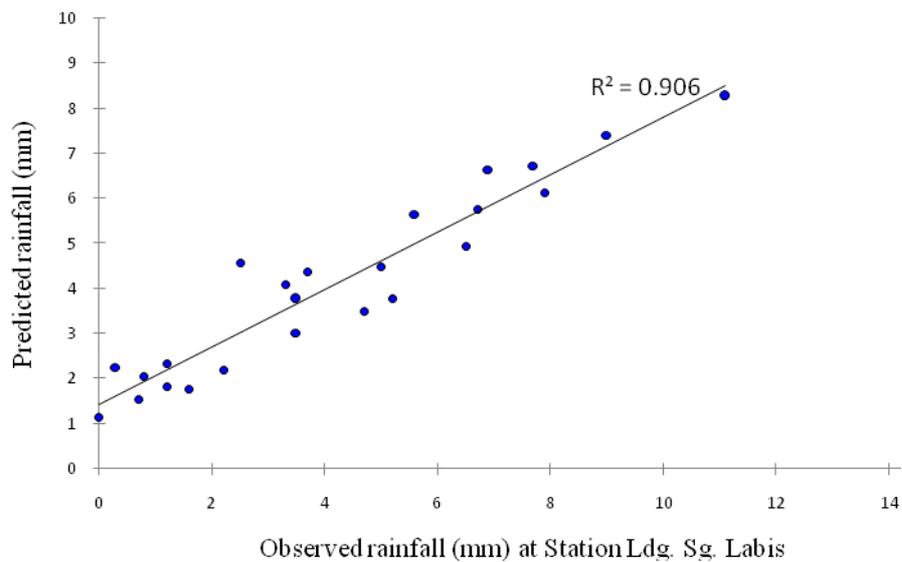
Note : — Observed value,
 — : ARIMA predicted value

Spearman Rank Correlation test conducted verified that there were significant correlations between observed and predicted values of rainfall and stream flow. Correlation coefficient value at Rainfall Station was: Kemelah ($r = 0.615$), Ladang Sg. Labis ($r = 0.952$), and Bandar Segamat ($r = 0.995$) and for stream flow station at Bandar Segamat was ($r = 0.991$). Correlation coefficients were significant at

alpha 0.05, this means that, there were significant relationship between the observed and predicted values of rainfall and stream flow events. Figure 6 (a,b,c and d) shows scatter plot generated by Spearman Rank correlation Coefficient. This revealed that, the ARIMA prediction model fit the mean monthly data of rainfall and stream flow.

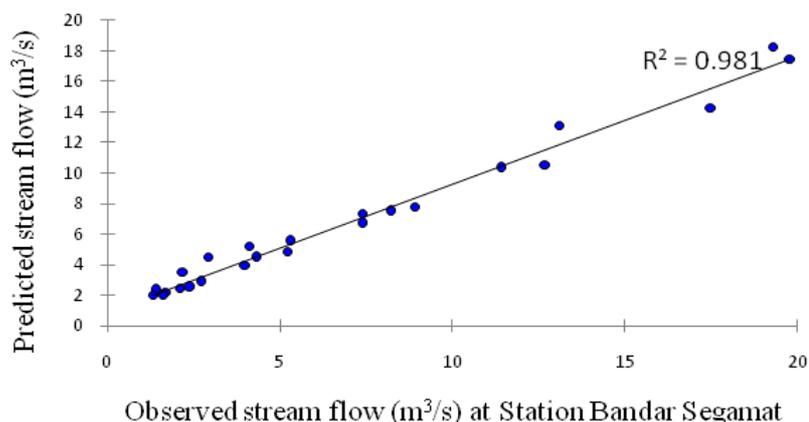
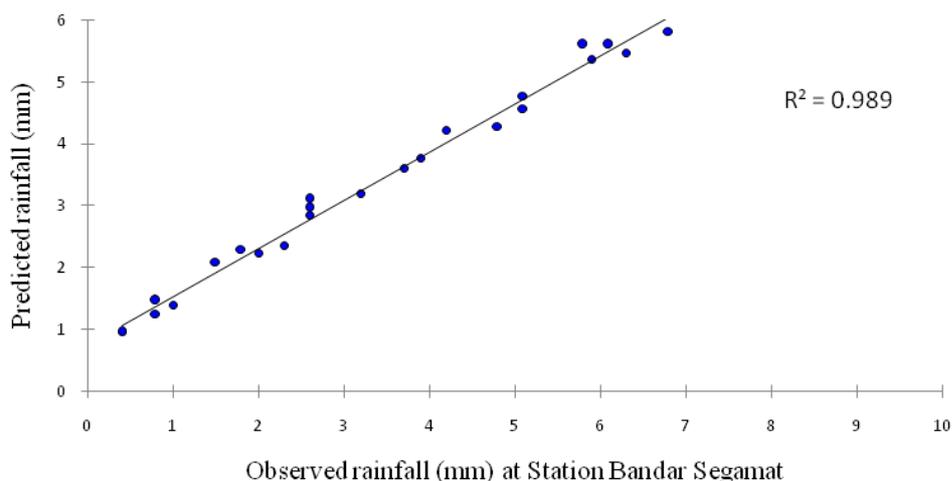


a.



b.

C.



d.

Figure 6 Scatter plot of predicted rainfall (mm) against observed rainfall (mm) at a) Station Kemelah, b) Station Ladang Sg. Labis, c) Station Bandar Segamat and d) Scatter plot predicted stream flow (m^3/s) against observed stream flow (m^3/s) at Station Bandar Segamat.

3.2 Correlation test between rainfall and stream flow data.

Table 3 shows correlation coefficient (r) and p-value for Spearman’s correlation test for rainfall stations at Kemelah, Ladang Sg. Labis and Bandar Segamat with the stream flow.

Table 3 Correlation coefficient and p-value of Spearman test for each rainfall station

Variable (Rainfall)	Correlation Coefficient	p value
Kemelah	0.289	0.001
Ladang Sg. Labis	0.433	<0.0001
Bandar Segamat	0.350	<0.0001

Note: The values are significant at $\alpha=0.05$.

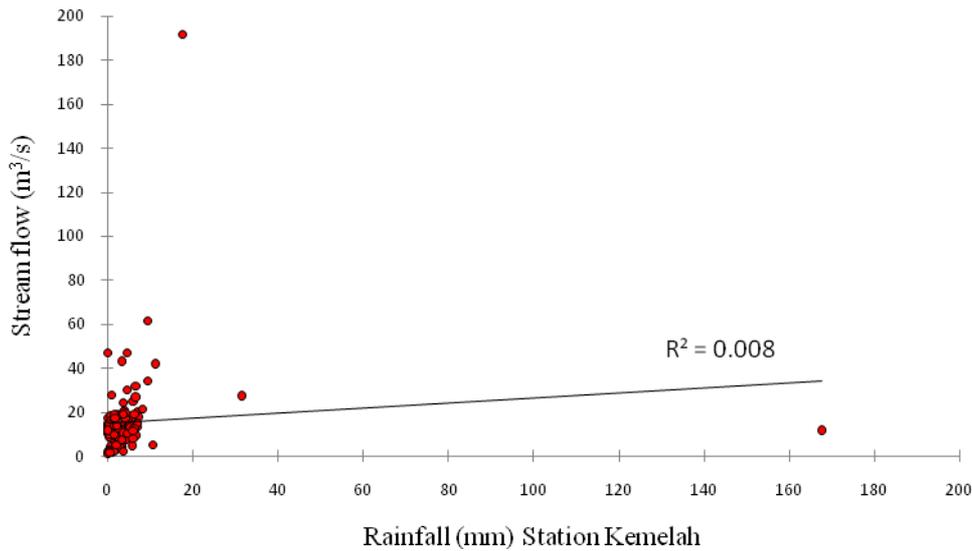
All the three rainfall stations were significantly correlated to stream flow with significant level less than the alpha

0.05 and correlation coefficient for each station at Kemelah, Ladang Sg.

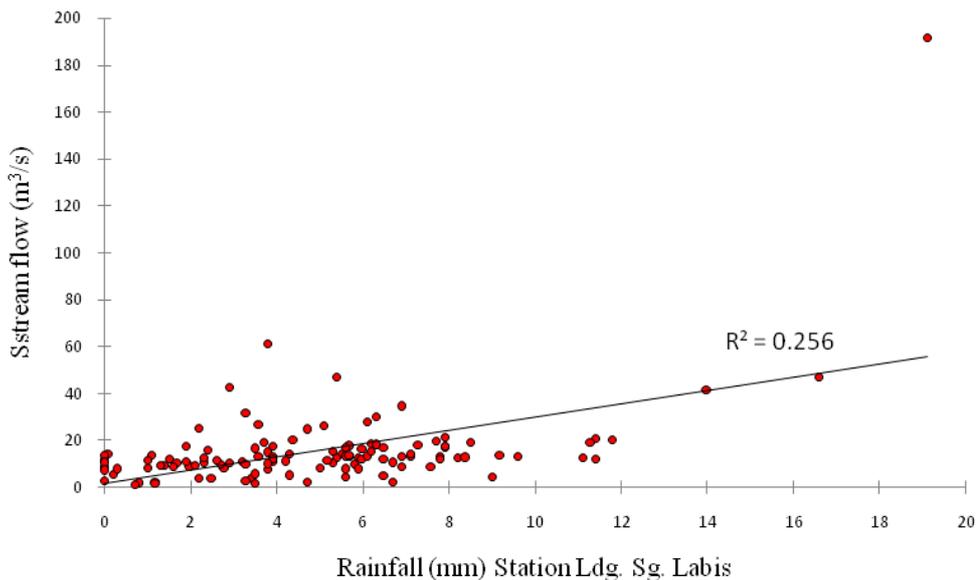
Labis and Bandar Segamat were $r = 0.289^*$, $r = 0.433^*$ and $r = 0.350^*$ respectively.. Rainfall at Ladang Sg. Labis station showed the strongest correlation coefficient to stream flow compared to Kemelah and Bandar Segamat stations. The weakest relationship between rainfall

and stream flow was found at Kemelah Station. Figure 7 showed the scattered plots with predicted line clearly for all the three rainfall stations. Based on the scatter graph analysis, rainfall has weak positive correlation with stream flow within the study area.

a)



b)



c)

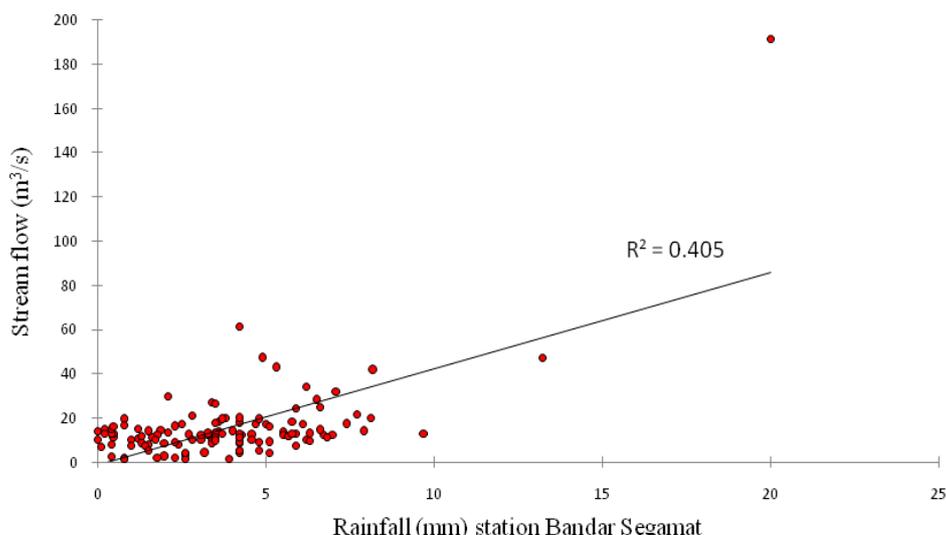


Figure 7. Scatter plot between stream flow against rainfall; a) at Kemelah, b) at Ladang Sg. Labis, c) at Bandar Segamat

Rainfall can reach a river by three pathways which are; directly into the stream channel, runoff on surface with low infiltration capacities and subsurface flow pathways (Renshaw *et al.* 2003). The correlation between rainfalls data at Station Ladang Sg. Labis with stream flow data at Bandar Segamat is ($r = 0.433, p < 0.05$) which showed a relationship is existing between rainfall and stream flow compared to Station Kemelah ($r = 0.289, p < 0.05$) and Station Bandar Segamat ($r = 0.350, p < 0.05$) although the rainfall stations are located far away from the stream flow station. Station Ladang Sg. Labis was located at agricultural area with high vegetation and low covered soil which had high infiltration capacity. The downpour probably reaches Segamat River by subsurface flow pathways.

Theoretically, rainfall at Bandar Segamat has a highest correlation with a stream flow due to covered soil at the town area which would lower the infiltration capacity, hence increase the runoff to the river basin. However, correlation between stream flow and rainfall station at Bandar Segamat ($r = 0.350, p < 0.05$) was lower than correlation between the stream flow

and rainfall station at Ladang Sg. Labis ($r = 0.433, p < 0.05$).

There are factors that affect the stream flow other than rainfall such as geological and topographical conditions (Xing and Davis, 2002; Vidon and Hill, 2004; Frei *et al.* 2010). If a drainage basin is located at a steep slope topography, runoff from the surface will reach the stream faster and the speed of the stream flow will increase quickly, whereas if drainage located at a lower area of flat slope, the runoff will reach the stream slower and much quantity of the runoff may be disappeared by evaporation and infiltration into the soil. These can be observed on correlation between rainfall at Station Bandar Segamat and the stream flow ($r = 0.350, p < 0.05$). Rainfall at the nearby stream flow station is very low and this might be cause by the low topography of the stream area. Station Kemelah had lowest correlation coefficient value ($r = 0.289, p < 0.05$) to stream flow probably due to high amount of runoff disappear by evaporation and infiltration into the soil. The increased stream flow after the rainfall also might be homogenized along the river before it reaches the stream flow detector at Station Bandar Segamat.

3.3 Trend analysis using Mann-Kendall test

Table 4 presented the result of trend test for the rainfall and stream flow stations using Mann-Kendall test.

Table 3 Mann-Kendall test of trend for rainfall and stream flow during 2000-2010.

Station	Location	S	p value	Trends
Rainfall	Kemelah	1372	0.002	↑
	Ladang Sg. Labis	262	0.554	↑
	Bandar Segamat	-886	0.045	↓
Stream flow	Bandar Segamat	-1446	0.001	↓

Note: ↑-upward trend, ↓-downward trend, bolded values were significant at alpha=0.05.

Rainfall at Kemelah stations showed significant upward trend with p-value 0.002 (less than alpha=0.05) and positive S value, 1372. Rainfall at Ladang Sg. Labis showed slightly upward trend with S value 262. However, the trend was not significant at alpha 0.05 because p-value of 0.554 was greater than alpha 0.05. Rainfall station at Bandar Segamat revealed a significant downward trend with p-value of 0.045 (less than 0.05) and negative S value, -886. Lastly, stream flow at Station Bandar Segamat revealed downward trend with negative S value of -1446 with significant trend p-value of 0.001 which was less than the alpha 0.05. Based on the trend test, downward trend of stream flow was consistent with downward rainfall trend at Station Segamat. However, based on correlation test, rainfall at Station Segamat shows the weakest correlation to stream flow compared to the other two stations.

5. CONCLUSION

In this study, the best fit ARIMA model for monthly mean data of rainfall and stream flow is ARIMA (0, 1, 2). The suitability of this model was verified by minimum value of AIC and MAPE. Correlation test between observed and predicted data also proved that ARIMA (0, 1, 2) is the best model. The correlation test shows the ARIMA (0, 1, 2) predicted values at rainfall Station Kemelah ($r = 0.615, p < 0.05$),

Station Ladang Sg. Labis ($r = 0.952, p < 0.05$), Station Bandar Segamat ($r = 0.995, p < 0.05$) and stream flow station at Bandar Segamat ($r = 0.991, p < 0.05$) were closed to observed data. Spearman correlation proved that rainfall at Station Kemelah ($r = 0.289, p < 0.05$), Station Ladang Sg. Labis ($r = 0.433, p < 0.05$) and Station Bandar Segamat ($r = 0.350, p < 0.05$) could give rise to stream flow. Rainfalls at the three stations were not the only factor affecting the stream flow. The stream flow might be affected by other factor such as temperature, geological and topographical condition. Mann-Kendall trend analysis showed that there were significant upward trend at Station Kemelah ($S = 1372, p < 0.05$). Rainfall station at Bandar Segamat ($S = -886, p < 0.05$) shows significant downward trend same as stream flow station at Bandar Segamat ($S = -1446, p < 0.05$). There was slightly upward trend detected for rainfall at Station Ladang Sg. Labis ($S = 262, p > 0.05$). The study identified that monsoon climate were the main factor affecting the rainfall trends, but other factor such as wind, temperature and topography of the area also could affect the rainfall variability. Factors influencing downward trend of stream flow might be due to changes of land use, dam in-filling at the upstream and climate change.

This study provides temporal statistical study on rainfall and stream flow data. Further detail statistical studies in term of

temporal and spatial geological and topography variations

REFERENCES

- Adnan, N. A. and Atkinson, P. M., 2011. Exploring the impact of climate and land use changes on streamflow trends in a monsoon catchment. *International Journal of Climatology*. **31**: 815–831.
- Billa, L., Mansor, S. and Mahmud, A. R., 2004. Spatial information technology in flood early warning systems: an overview of theory, application and latest developments in Malaysia. *Disaster Prevention and Management*. **13** (5): 356 – 363.
- Box, G. E. P., Jenkins, G.M. and Reinsel, G. C., 1994. Time series analysis: forecasting and control. 3rd ed. Englewood Cliffs, NJ:Prentice Hall.
- Brockwell, P. J. and Davis R. A., 2002. *Introduction to Time Series and Forecasting*. 2nd ed. Springer, New York.
- Chan, N. W., 1997. Increasing flood risk in Malaysia: causes and solutions. *Disaster Prevention and Management*. **6** (2): 72 – 86.
- Chattopadhyay, S., and Chattopadhyay, G., 2010. Univariate modelling of summer-monsoon rainfall time series: Comparison between ARIMA and ARNN. *Comptes Rendus Geoscience*. **342**: 100-107.
- Chattopadhyay, S., Jhajharia, D. and Chattopadhyay, G., 2011. Trend estimation and univariate forecast of the sunspot numbers: Development and comparison of ARMA, ARIMA and Autoregressive Neural Network models. *Comptes Rendus Geoscience*. **343**: 433-442.
- Douglas, E. M., Vogel, R. M. and Kroll, C. N., 2000. Trends in floods and low flows in the United States: impact of spatial correlation. *Journal of Hydrology*. **240**: 90-105.
- Fredolin, T. T., Juneng, L., Salimun, E., Vinayachandran, P. N., Seng, Y. K., Reason, C. J. C, Behera, S. K. and Yasunari, T., 2008. On the role of the northeast cold surge, the Borneo vortex, the Madden-Julian Oscillation, and the Indian Ocean Dipole during the extreme 2006/2007 flood in southern Peninsular Malaysia. *Geophysical Research Letter*. **35**: 1-6.
- Frei, S., Lischeid, G. and Fleckenstein, J. H., 2010. Effects of micro-topography on surface–subsurface exchange and runoff generation in a virtual riparian wetland — A modeling study. *Advances in Water Resources*. **33**: 1388–1401.
- Gasim, M. B., Surif, S., Mokhtar, M., Toriman, M. E., Rahim, S. A., and Chong, H. B. 2010. Flood Analysis of December 2006: Focus at Segamat Town, Johor. *Sains Malaysiana*. **39**(3): 353-361.
- Gemitzi, A. and Stefanopoulos, K., 2011. Evaluations of the effects of climate and man intervention on ground waters and their dependent ecosystems using time series analysis. *Journal of Hydrology*. **403**: 130-140.
- Grist, J. P. and Nicholson, S. E, 2001. A study of the dynamic factors influencing the rainfall variability in the West African Sahel. *Journal of Climate*. **14**: 1337-1359.
- Habib, E., Krajewski, W. F. and Ciach, G. J., 2001. Notes and correspondence: Estimation of rainfall interstation correlation. *Journal of Hydrometeorology*. **2**: 621-629.

- Juahir, H., Zain, S. M., Aris. A. Z. Aris, Yusof, M. K., Samah, M. A. A. and Mokhtar, M. 2010. Hydrological trend analysis due to land use changes at Langat River Basin. *Environment Asia*. **3**: 20-31.
- Hirsch, R. M., Slack, J. R. and Smith R. A., 1982. Techniques of trend analysis for monthly water quality data. *Water Resources Research*. **18** (1): 107-121.
- Kendall, M. G., 1975. *Ranks Correlation Methods*. 4th ed. Griffin, London.
- Lana-Renault, N., Latron, J., Karssenber, D., Serrano-Muela, P., Regüés, D. and Bierkens,
- Mann, H. B., 1945. Non-parametric test against trend. *Econometrica*. **13**: 245-259.
- Mohd, M.S., Alias, B. and Daud, D., 2006. *GIS analysis for flood hazard mapping: Case study; Segamat, Johor, West Malaysia*. National GIS Seminar: GIS Application fir Mitigation in Natural Disaster. 1-15.
- Ömer Faruk, D., 2010. A hybrid neural network and ARIMA model for water quality time series prediction. *Engineering Applications of Artificial Intelligence*. **23**: 586-594.
- Razi, M. A. M., Ariffin, J., Tahir, W. and Arish, N. A. M., 2010. Flood Estimation Studies Using Hydrologic Modelling System (HEC-HMS) for Johor River, Malaysia. *Journal of Applied Sciences*. **10** (11): 930-939.
- Renshaw, C. E., Feng, X., Sinclair, K. J. and Dums R. H., 2003. The use of stream flow routing for direct channel precipitation with isotopically-based hydrograph separations: the role of new water in stormflow generation. *Journal of Hydrology*. **273**: 205–216.
- Sene, K. 2010. Chapter 7: Floods. *Hydrometeorology*. DOI 10.1007/978-90-481-3403-8_7. 199-237.
- Siegel, S., 1957. Nonparametric statistics. *The American Statistician*. **11** (3): 13-19.
- Suhaila, J., Deni, S. M., Zin, W. Z. W. and Jemain, A. A., 2010. Trends in Peninsular Malaysia Rainfall Data During the Southwest Monsoon and Northeast Monsoon Seasons: 1975-2004. *Sains Malaysiana*. **39**(4): 533–542.
- Van Gelder, P. H. A. J. M., Wang, W. and Vrijling, J. K., 2007. Statistical estimation methods for extreme hydrological events. *Extreme Hydrological Events: New Concept for Security*. 199-252.
- Vidon, P. G. F. and Hill, A. R., 2004. Landscape controls on the hydrology of stream riparian zones. *Journal of Hydrology*. **292**: 210–228.
- Weisang, G. and Awazu, Y., 2008. Vagaries of the Euro: an introduction to ARIMA modelling. *Case Studies in Business, Industry and Government Statistics (CS-BIGS)*. **2** (1): 45-55.
- Wu, C. L. and Chau, K. W., 2010. Data - driven models for monthly stream flow time series prediction. *Engineering Applications of Artificial Intelligence*. **23**: 1350-1367.
- Xing, J., and Davies, A. M., 2002. Influence of shelf topography upon along shelf flow and across shelf exchange in the region of the Ebro Delta. *Continental Shelf Research*. **22**: 1447–1475.
- Yue, S., Pilon, P. and Cavadias, G., 2002. Power of the Mann-Kendall and Spearman's rho test for detecting monotonic trends in hydrological series. *Journal of Hydrology*. **292**: 254-271.