ORIGINAL PAPER

Long-term variation analysis of a tropical river's annual streamflow regime over a 50-year period

Mohammed Seyam · Faridah Othman

Received: 5 August 2013 / Accepted: 7 July 2014 / Published online: 20 July 2014 © The Author(s) 2014. This article is published with open access at Springerlink.com

Abstract Studying the long-term changes of streamflow is an important tool for enhancing water resource and river system planning, design, and management. The aim of this work is to identify the long-term variations in annual streamflow regime over a 50-year period from 1961 to 2010 in the Selangor River, which is one of the main tropical rivers in Malaysia. Initially, the data underwent preliminary independence, normality, and homogeneity testing using the Pearson correlation coefficient and Shapiro-Wilk and Pettitt's tests, respectively. The work includes a study and analysis of the changes through nine variables describing the annual streamflow and variations in the yearly duration of high and low streamflows. The analyses were conducted via two time scales: yearly and sub-periodic. The sub-periods were obtained by segmenting the 50 years into seven sub-periods by two techniques, namely the change-point test and direct method. Even though analysis revealed nearly negligible changes in mean annual flow over the study period, the maximum annual flow generally increased while the minimum annual flow significantly decreased with respect to time. It was also observed that the variables describing the dispersion in streamflow continually increased with respect to time. An obvious increase was detected in the yearly duration of danger level of streamflow, a slight increase was noted in the yearly duration of warning and alert levels, and a slight decrease in the yearly duration of low streamflow was found. The perceived changes validate the existence of long-term changes in annual streamflow

M. Seyam · F. Othman (🖂)

Civil Engineering Department, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia e-mail: faridah ce@yahoo.com

F. Othman e-mail: faridahothman@um.edu.my

M. Seyam e-mail: mohseyam@yahoo.com regime, which increase the probability of floods and droughts occurring in future. In light of the results, attention should be drawn to developing water resource management and flood protection plans in order to avert the harmful effects potentially resulting from the expected changes in annual streamflow regime.

1 Introduction

Streamflow is a river basin's integrated reaction to rainfall and other relevant hydrological parameters under certain climatological conditions. Factors known to affect a basin's reaction to precipitation are temperature, wind speed, humidity, soil type, land cover, and land use (Hodgkins et al. 2003; Rakhshandehroo et al. 2010). Streamflow regimes can be characterized by some such features as rate, magnitude, duration, timing, and fluctuations over time scales including hourly, daily, monthly, annual, and multi-annual streamflow (Krasovskaia and Gottschalk 2002; Morán-Tejeda et al. 2011). These characteristics serve in describing the whole streamflow regime and related phenomena such as drought periods and flood events. Besides, by describing streamflow through these characteristics, the long-term changes in a streamflow regime can be investigated explicitly (Poff et al. 1997; Richter 1996; Yang et al. 2005).

Since streamflow records vary in period scales of hours, days, weeks, months, seasons, years, and multi-years, long streamflow period observations are necessary to analyze and describe a river's streamflow regime (Post and Jones 2001). To gain complete understanding of long-term streamflow regime changes, related statistical analyses should be conducted using extensive records—preferably longer than 50 years; trends quantified from short records may be part of climate fluctuations (Gautam and Acharya 2012; Kundzewicz and Robson 2004; Opitz-Stapleton and Gangopadhyay 2011). There are several reasons why changes in streamflow take place, for instance climate variability, human activities, and

geomorphic changes, which are perhaps the principal causes of streamflow change (Sang et al. 2013). Variations in streamflow evolve gradually, at a slow pace, and such changes are not only linked by local but also regional weather patterns via local/regional climate dynamics. Over the past century, a significant decrease in annual streamflow has been observed in approximately a quarter of the world's rivers (Descroix et al. 2012; Walling and Fang 2003; Yang et al. 2005; Yue et al. 2003; Zhang et al. 2000).

The streamflow regime of rivers is not only significant from a hydrologic viewpoint but also from both socioeconomic and natural perspectives. In fact, ecosystems are extremely sensitive to hydrological variations in a river's regime because they depend on river flow to conserve their composition and structure (Richter 1996). Investigating the variations in a streamflow regime is a key to understanding the influence of climatic variability and fluctuations on river hydrologic systems. The enhanced awareness of streamflow regimes is essential for the development of river system and surface water resource planning, design, and management, as well as for the improvement of flood protection. Furthermore, decision makers will benefit for future planning, while water resource planners will be better able to devise adaptive management strategies to deal with such probable changes. Better streamflow regime analysis also helps to forecast the expected timing of low and high flow over certain periods as well as identify future research directions and requirements (Gautam and Acharya 2012; Xu et al. 2012).

Investigations into the long-term changes of streamflow regime have witnessed growing interest over the last years. Several studies have been conducted worldwide to examine the possible long-term changes in streamflow regime. Sun et al. (2013) studied the changes in annual streamflow using climate elasticity in Poyang Lake basin, China. They examined streamflow datasets at four hydrological stations from 1961 to 2000. A minor change has been observed in the timing of the mass center of the streamflow. The annual streamflow increases at different rates, with a statistically significant trend at the 5 % level. Chen et al. (2013) quantified the effects of climate variability and human activities on streamflow for Kaidu River basin in an arid region of northwest China over the period from 1960 to 2009. Aberrant increases have been observed in annual streamflow. Ahmad et al. (2013) analyzed the annual maximum streamflow in Terengganu, Malaysia using TL-moments approach. Their study revealed that the generalized Pareto and generalized logistic distributions are appropriate for describing the annual maximum streamflow series in Terengganu, Malaysia. Hannaford and Buys (2012) evaluated the trends in river flow regime in the UK for the 1969-2008 period. Trends were characterized for four seasons. It was found that runoff and high flow increased in winter and autumn, flow decreased in spring, and no changes occurred in summer. Gautam and Acharya (2012) presented an initial widespread trend detection system of streamflow in Nepal. They applied Mann-Kendall and Sen's trend tests on two data sets of streamflow. It was revealed that 24 % of variables had trends, of which 41 % were descending and 59 % ascending. Miao et al. (2012) aimed to explore the reasons behind streamflow changes in the Yellow River. Their results showed that streamflow presented a downward trend from 1956 to 2008. Morán-Tejeda et al. (2011) looked at variability in the Duero River basin's regime (northern Spain) for the 1961-2006 period. They categorized streamflow regime types by applying the principal component analyses. A noticeable drop in river flow was observed in winter and spring. They also noted that river regime fluctuations occurred in both the timing and magnitude directions. Korhonen and Kuusisto (2010) studied the streamflow regime fluctuations in Finland for the 1912-2004 period. Generally, no variations in mean annual flow were detected, but remarkable variations in seasonal means of streamflow were observed. The mean monthly streamflow in winter and spring increased, but there were no changes in autumn. Burn et al. (2010) examined the changes in extreme high and low flow events in Canada. They applied the Mann-Kendall test on 68 streamflow stations over at least 50 years. The results revealed a decrease in the magnitude of annual maximum flow. Kumar et al. (2009) evaluated the streamflow trend for 31 streamflow stations using records from more than 50 years. Trends were calculated using the Mann-Kendall test for streamflow characteristics including low, mean, and high annual streamflows. They observed a decreasing trend in low and mean flow conditions in roughly 70 % of stations. Ma et al. (2008) investigated the changes in annual streamflow for eight sub-basins in the Shiyang River basin over 50 years. They applied the Kendall and Pettitt tests to detect variations in streamflow. It was revealed that the annual streamflow in five of the eight sub-basins significantly decreased. Birsan et al. (2005) studied the trends from the daily streamflow data of 48 basins in Switzerland using the Mann-Kendall test over three sub-periods (1931-2000, 1961-2000, 1971-2000). An increase in annual runoff in about 60 % of the stations was recorded. Burn et al. (2004) studied the hydrologic trends of the Liard River basin in Canada. According to results, the river's hydrologic trends are associated to changes in both climatological and largescale atmospheric progression. Using a trend test, Xiong and Guo (2004) explored the variations in the annual streamflow of the Yangtze River. They reached the conclusion that there was no significant trend in annual maximum flood, but there was a decreasing trend in the annual minimum and mean streamflows.

The abundance of research works including the abovementioned, which mainly explore the long-term changes of streamflow, along with additional international research works reveal the great importance of gaining better understanding of river flow regimes which is a fundamental tool for river systems and surface water resource planning, design, and management. Thus, it is the interest of the authors to investigate the long-term changes in the streamflow regime of the Selangor River as one the main rivers in Malaysia and which generally represents tropical rivers in the south eastern Asia region.

The main objective of this study is to identify the long-term variations in the annual streamflow regime of the Selangor River over a 50-year period from 1961 to 2010. This is accomplished by investigating the changes in nine hydrological variables describing the annual streamflow and changes in yearly duration of high and low streamflows. The analyses were conducted via two time scales: yearly changes and subperiodic changes. The sub-periods were derived from the segmentation of the 50-year period into seven sub-periods in two ways, namely the change-point test and direct technique. The paper is organized in five sections beginning with the introduction. The study area is described in the second part. The third section comprises the data and research methodology, while the fourth section is devoted to results and discussion. The conclusions are specified in the fifth and final part.

2 Study area

The study area for this research work is the Selangor River basin. The Selangor River is one of the rivers flowing along the Malaysian Peninsula's west coast. It is the main river in the state of Selangor. The catchment area is nearly 1,960 km² covering approximately 25 % of Selangor state. The main tributary of the Selangor River starts at the border between the states of Selangor and Pahang at an elevation of 1,700 m. The Selangor River streams nearly 110 km from the northeast to the southwest, flowing in the state of Selangor through three districts: Kuala Selangor, Hulu Selangor, and Gombak. The southern border of the river basin touches the outskirts of Kuala Lumpur City (Hassan et al. 2004; Lee 2002; Samsudin et al. 2011). The Selangor River basin is the main water source for Selangor and Kuala Lumpur. About 50 % of the water consumed in these two states is obtained from the Selangor River (Subramaniam 2004). Figure 1 presents the location map of the Selangor River basin in peninsular Malaysia.

2.1 Climate and rainfall in the Selangor River basin

The Selangor River basin has a humid, tropical climate much like Peninsular Malaysia. The characteristic climate feature here is uniform temperature with minimal variation throughout the year. On average, the temperature rises during the day up to 32°C and drops to 23°C at night. The highlands are slightly cooler. The average annual rainfall varies between

2,000 and 3,000 mm throughout the basin. Open water evaporation ranges from 1,600 to 1,800 mm, while the relative humidity is 80 % on average each year (Breemen 2008; Shafie 2009; Zin et al. 2013).

2.2 Hydrology of the Selangor River basin

The main tributaries of the Selangor River are Rening, Kerling, Batang Kali, and Guntong. Several minor branches join the major ones or the river itself. The upstream branches are fast-flowing streams through mountain with granite and sedimentary bedrock, and they are at some points interrupted by rapids and waterfalls. In the lower basin area, the river enters the fluvial plain and becomes a low-gradient, meandering river. The river bed slope in the last 30 to 40 km is roughly zero, and at some sites, the slope is negative (Breemen 2008). The river's average streamflow is 57 m³/s. Seasonal variations in rainfall cause flow to exceed 122 m³/s or to fall below 23 m³/s around 10 % of the time (Nelson 2002). Figure 2 presents the mean annual streamflow of the Selangor River over a 50-year period between 1961 and 2011.

3 Data and methodology

3.1 Data

Data is considered the backbone of studies on changes in hydrological systems, thus, high quality data is vital. Data must undergo quality control prior to conducting hydrological analyses such as studying long-term variations in streamflow regime. For complete understanding of long-term variations in streamflow regime, related statistical analyses should be conducted using records from lengthy periods, preferably longer than 50 years, as changes quantified from short periods may be components of climate fluctuations (Kundzewicz and Robson 2004; Walling and Fang 2003).

Research data was sourced from hydrological stations located in the Selangor River basin. In the Selangor River basin, there are two streamflow (SF) stations, seven water level (WL) stations, and more than 30 rainfall (RF) stations. The two streamflow stations are Rantau Panjang and Ampang Pecah. Unfortunately, only the Rantau Panjang SF station has sufficient streamflow records of hourly streamflow data from 1961 to 2012 to study the long-term changes in streamflow regime. This station is located in the downstream area of the Selangor River. All major tributaries join the Selangor River before this station; for this reason, the streamflow at the Rantau Panjang station is considered the best indicator of streamflow in the study area (Seyam and Othman 2014). Figure 3 presents the location of the two



Fig. 1 The location map of Selangor River basin

streamflow stations and the main tributaries in the Selangor River basin.

3.2 Methods

Prior to analyzing the long-term variations in streamflow regime over 50 years from 1961 to 2010 in the Selangor River, the streamflow data over the study period from the Rantau Panjang station was preliminarily examined. The initial examination entailed testing the data with the assumptions of independence, normality, and homogeneity tests using the

Fig. 2 Mean annual streamflow (m³/s) in the Selangor River over a 50-year period from 1961 to 2010

Pearson correlation coefficient and Shapiro–Wilk and Pettitt's tests, respectively. The hydrological variables which best represent the long-term change of streamflow were then determined. Nine hydrological variables were selected to represent the long-term changes in streamflow regime over the study phase. The analyses additionally included an investigation of the changes in the yearly duration of high and low streamflows throughout 50 years and testing for any trend in changes.

The analyses of the long-term changes in hydrological variables and yearly duration of high and low streamflows





Fig. 3 The locations of the streamflow stations and the main tributaries in the Selangor River basin

were conducted based on two time scales, yearly changes, and sub-periodic changes. The sub-periods were obtained by dividing the 50-year interval into seven sub-periods. The segmentation process involved two techniques: the change-point test and direct method.

3.2.1 Preliminary data analysis

The independence test The hydrological time series are generally highly persistent, and even measurements far apart in time can exhibit strong interdependence (Montanari et al. 2000). The independence test was applied to decide whether any significant correlation exists through the annual streamflow records of the Rantau Panjang station over 50 years from 1961 to 2010. This test was conducted by computing the correlation coefficient (r) between the time series. There are different means of calculating r with the most widely used being the Pearson correlation coefficient. It is obtained by dividing the covariance of the two variables by the product of their standard deviations as described in Eq. 1.

$$r = \frac{\sum_{i=1}^{n} \left(X_{\text{obs},i} - \overline{X}_{\text{obs}} \right) \cdot \left(X_{\text{model},i} - \overline{X}_{\text{model}} \right)}{\sqrt{\sum_{i=1}^{n} \left(X_{\text{obs},i} - \overline{X}_{\text{obs}} \right)^2 \cdot \sum_{i=1}^{n} \left(X_{\text{model},i} - \overline{X}_{\text{model}} \right)^2}}$$
(1)

When the absolute values of correlation coefficient for different lag times calculated for a time series data consisting of *n* observations are not larger than the typical critical value (i.e., $1.96/\sqrt{n}$, corresponding to the 5 % significance level), the observations in this time series data can be accepted as independent from each other (Douglas et al. 2000; Noguchi et al. 2011).

The normality test The normality test is a statistical test used to examine whether the data is fine-modeled by a normal distribution or not. It is also used to calculate the likelihood that the studied variables will be normally distributed (Coin 2008; Tenreiro 2011). Normality test is extensively utilized in statistical analyses (Hwang and Wei 2006). About 40 normality tests are available in statistical literature (Coin 2008; Declercq and Duvaut 1999; Lesaffre 1983; Önder and Zaman 2005; Razali and Wah 2011; Zhang and Wu 2005). The Shapiro–Wilk test is considered one of the most popular tests for checking normality in scientific and hydrologic research (Jurečková and Picek 2007; Kottegoda 1984).

In this research, the normality of average annual flow from 1961 to 2010 was verified with the Shapiro–Wilk test. This test was published in 1965 by Samuel Shapiro and Martin Wilk, and it was later modified by Royston in 1992 and again in 1995 (Razali and Wah 2011; Royston 1992). The output of

streamflow Number Variable Definition Unit

Table 1 The hydrological variables used to describe the annual

Nullibei	variable	Deminition	Ullit
1	SF_1	Mean annual flow	m ³ /s
2	SF_2	Maximum annual flow	m^3/s
3	SF ₃	Minimum annual flow	m^3/s
4	SF_4	Maximum monthly flow per year	m^3/s
5	SF_5	Minimum monthly flow per year	m^3/s
6	RA	The variance between maximum and minimum annual streamflow	m ³ /s
7	SD	Standard deviation	m^3/s
8	CV	Coefficient of variation	Ratio
9	PR	The pluviometric ratio	Ratio

the Shapiro-Wilk test consists of two values: W and P. The value of W lies between 0 and 1. High W values lead to acceptance of normality whereas smaller W values lead to rejection. When W is 1, it indicates full data normality. For the *P* value, if it is higher than the selected significance level, normality will be accepted (Razali and Wah 2011). Another method of testing the normality is a graphical technique of comparing the frequency distribution (histogram) of streamflow records of the Rantau Panjang station over 50 years from 1961 to 2010 and the normal probability curve. The data is deemed to be normally distributed if its histogram resembles the normal distribution of data.

The homogeneity test The homogeneity test is a statistical test for detecting data variability and checking whether a data series has been sourced from homogeneous or heterogeneous records. This is a fundamental test in hydrological statistical analysis (Kang and Yusof 2012; Qin et al. 2009). In general, when the hydrological data series is homogeneous, it means that data records were recorded with similar instruments, techniques, and environments (Kang and Yusof 2012). Literature offers several methods of testing the homogeneity of hydrological time series data such as the standard normal homogeneity test, Buishand's test, and Pettitt's test (Buishand 1982; Chiu and Wang 2009; Firat et al. 2010; Pettitt 1979; Rocha et al. 2012; Wang et al. 2007). In this research, Pettitt's test was applied to check the homogeneity of streamflow records from the Rantau Panjang station over 50 years from 1961 to 2010.

3.2.2 Segmentation of the study period

To study and analyze the long-term changes in streamflow regime over long periods such as 50 years, sometimes, yearly changes are not sufficient to describe the trend of long-term changes. For this reason, lengthy periods could be subdivided into short sub-periods, i.e., 7, 10 years. The short sub-periods are generally formed by sequential years with similar hydrological characteristics (Descroix et al. 2012). The segmentation process thus simplifies analysis and helps present the long-term changes over long periods. A number of segmentation techniques exist such as the Hidden Markov model, the Hubert model, and the change-point test (Gedikli et al. 2010; Hubert 2000; Kehagias 2004; Kehagias et al. 2006; Petersen-Øverleir and Reitan 2005). The change-point is a statistics test applied to identify the date (s) at which an abrupt change occurs in a data series. The identified dates show a shift in the mean or variance potentially due to other contemporaneous events like a concomitant shift in other climatic variables or major anthropogenic interventions, e.g., dam construction for regulating streamflow (Beaulieu et al. 2009; Rougé et al. 2012; Villarini et al. 2011).

In literature, many methods have been applied to check for the presence of change-points in the data of long periods such as Bayesian inference, moving t test, and Pettit's test (Descroix et al. 2012; Ma et al. 2008; Rougé et al. 2012; Xiong and Guo 2004; Zheng et al. 2007). In this research, the 50-year period 1961-2010 was divided into seven subperiods by two methods. The first was Pettit's test, which was applied to verify the presence of change-points over the 50year phase from 1961 to 2010. The second technique was direct segmentation of the study period into seven equal subperiods.

3.2.3 Determining the hydrological variables

The long-term changes in streamflow regime could be explored via streamflow characteristics like magnitude, rate, frequency, duration, timing, and rate of change. These characteristics are usually applied in three main conditions: average flow, low flow, and high flow (Poff and Ward 1989). A wide range of hydrological variables can help assess the changes in the characteristics of streamflow (Moliere et al.

Table 2 The independence test of the annual streamflow data

Lag times (year)	r	r	Ν	$1.96/\sqrt{n}$	Test result
lag-1	0.20	0.20	50	0.28	Independence
Lag-2	0.06	0.06	49	0.28	Independence
Lag-3	0.04	0.04	48	0.28	Independence
Lag-4	0.14	0.14	47	0.29	Independence
Lag-5	-0.13	0.13	46	0.29	Independence
Lag-6	-0.10	0.10	45	0.29	Independence
Lag-7	0.16	0.16	44	0.30	Independence
Lag-8	-0.13	0.13	43	0.30	Independence
Lag-9	-0.32	0.32	42	0.30	Dependence
Lag-10	-0.39	0.39	41	0.31	Dependence
Lag-11	0.14	0.14	40	0.31	Independence

Fig. 4 The frequency

annual streamflow



2009). In this research, the changes in streamflow were studied based on streamflow rate (discharge), which is the amount of water passing through a specified site per unit of time (Poff et al. 1997; Richter 1996; Yang et al. 2005).

Nine hydrological variables describing streamflow were chosen to study long-term modifications of the Selangor River's streamflow regime. The variables are mean annual flow (SF1), maximum annual flow (SF2), minimum annual flow (SF3), maximum monthly flow per year (SF4), minimum monthly flow per year (SF5), the range between maximum and minimum streamflows (RA), standard deviation (SD), coefficient of variation (CV), and the pluviometric ratio (PR). SF2 represents the maximum annual flow throughout the sub-period; SF3 is the minimum annual flow over the sub-period; SF4 signifies the maximum monthly flow over a single year; SF5 corresponds to the minimum monthly flow over a single year; SD and CV are statistical measures

Table 3 The sub-periods obtained via two segmentation techniques

Sub-periods	Segmentation technique			
	Change-point	Direct		
1	1961-1972	1961–1967		
2	1973–1978	1968–1974		
3	1979–1986	1975–1981		
4	1987-1991	1982–1988		
5	1992-1995	1989–1995		
6	1996-2004	1996–2002		
7	2005-2010	2003-2009		

of dispersion in a data series around its mean; and the CV of annual streamflow data series represents the ratio of standard deviation to the mean annual flow. The CV is valuable in matching the degree of dispersion and variation among data series (Albrecher et al. 2010; Boik and Shirvani 2009). PR corresponds to the ratio between maximum and minimum streamflow values, and it provides initial indication of seasonal variability. When the PR value is close to 1, seasonal variability is small, but when the value is above 1, seasonality increases and seasonal variability is high (Laraque et al. 2007). SD, CV, RA, and PR are calculated twice based on two time scales: yearly and sub-periodic. The values of all variables were derived from the hourly streamflow records at the Rantau Panjang station over a 50-year period (1961-2010). Table 1 presents the nine hydrological variables and their measurement units.

3.2.4 The long-term change analyses

The long-term change analyses include an investigation of the variations in nine hydrological variables (Table 1) over the 50-year period from 1961 to 2010 along with testing the trend of changes. Analysis was conducted based on two time scales. The first is the evaluation of yearly changes and the second depends on the subperiodic changes. The sub-periods were obtained by segmenting the 50 years into seven sub-periods via the direct and change-point test techniques.

The work also takes into account an investigation of the changes in the yearly duration of high and low





streamflows from 1961 to 2010 and testing the trend of changes. For the high and low streamflows, the assessment was generally with respect to their duration and magnitude. Both features play an essential role in analysis (Mirabbasi et al. 2012). Analysis is based on the two previously mentioned time scales. Meanwhile, the

study of yearly high streamflow duration comprises three levels, danger level, warning level, and alert level. The yearly duration of low streamflow analysis was conducted at a single level that when the streamflow fell below 14.5 m³/s, which represents about 25 % of the average streamflow over the study period.



Fig. 6 The changes in hydrological variables over the 50-year period. a Standard deviation (SD). b The range between maximum and minimum streamflows (RA). c Coefficient of variation (CV). d The pluviometric ratio (PR)



Fig. 7 The changes in the hydrological variables over the sub-periods obtained by change-point test technique. **a** mean, maximum, and minimum annual flows. **b** Standard deviation and variance between maximum and minimum annual streamflows. **c** Coefficient of variation and pluviometric ratio

4 Results and discussion

4.1 Preliminary data analysis

The preliminary data analysis includes the independence test, normality test, and homogeneity test. The results of these tests prove that the annual streamflow data is considered independent, normally distributed, and homogeneous.

The correlation coefficient of the annual streamflow data for multi-years lagged records was applied to test the independency. Regarding the calculated *r* of lag-1 year to lag-11 year of the annual streamflow, it is accepted as independent data at the 5 % significance level since the value of *r* is smaller than the typical critical value $1.96/\sqrt{n=1.96}/\sqrt{49=0.28}$ as described in Table 2. As an example, the *r* of lag-2 year is 0.06, and it is smaller than 0.28. Table 2 presents the correlation coefficient of lag-1 year to lag-11 year of the annual streamflow.

The Shapiro–Wilk test was applied to check the normality of streamflow data. Based on the test results W(0.976) and P(0.411), annual flow time series data are considered to be normally distributed. The normality was also checked by

graphical technique by exploring the resemblance between the frequency distribution of the average annual streamflow data (histogram) and normal probability curve. Figure 4 demonstrates a very high resemblance between the frequency distribution and normal probability curve. Thus, the average annual flow time series is considered normally distributed.

Pettitt's test was applied to verify the homogeneity of the annual streamflow data. The P value (two-tailed), computed using 10000 Monte Carlo simulations, is 0.161. According to the calculated P value which is higher than the significance level P value of 0.05, the average annual flow time series is considered homogeneous time series data.

4.2 Segmentation of the streamflow data

The 1961 to 2010 time frame was separated into seven subperiods by means of two techniques. The primary method is direct segmentation in which the study period is divided into seven identical 7-year sub-periods. The subsequent technique entails determining multiple change-points, as calculated using Pettitt's test. This second technique leads to the subdivision of the study period into seven, non-identical subperiods. The sub-periods obtained in the two ways are presented in Table 3.

4.3 The changes in hydrological variables over 50 years

4.3.1 The yearly changes

An investigation of the differences in the nine variables representing annual streamflow in the Selangor River over 50 years (1961–2010) was carried out based on a yearly time scale to provide an overall view of the long-term changes in streamflow regime. Figure 5 demonstrates the three time series of mean annual flow (SF1), maximum monthly flow per single year (SF2), and minimum monthly flow per single year (SF3). There is no apparent significant trend in the three time series. Although Fig. 5 reveals nearly negligible changes and a minor trend in the three time series, it is noted that the time series values get farther from the mean values with respect to time. The time series of SD, CV, RA, and PR over 50 years (from 1961 to 2010) are presented in Fig. 6. There is no clear trend in the SD, CV, RA, and PR variables to describe the dispersion of annual streamflow data. The results of the variable changes based on a yearly time scale justify the need to study long-term changes of streamflow based on a subperiodic time scale as obtained via two segmentation techniques explained previously.

4.3.2 The sub-periodic changes

The changes over the sub-periods obtained via the changepoint technique The hydrological variable changes over the sub-periods obtained via the change-point technique are shown in Fig. 7. Although the analysis revealed nearly negligible changes in mean annual flow throughout the sub-period as presented in Fig.7a, the maximum annual flow progressively increased, especially from the fourth sub-period. According to this result, the incidence of high flow increased substantially in the latter sub-periods. Such changes lead to the formation of suitable hydrological conditions for floods to occur. As a direct result, flood events will probably happen more frequently in future. Although the maximum annual flow generally increased, it is noticed that the minimum annual flow decreased significantly, particularly from the fifth sub-period. Obviously, low flow occurrence increased substantially in the latter sub-periods, leading to suitable hydrological conditions for droughts to take place. As such, drought periods may occur





Fig. 8 The changes in hydrological variables over the sub-periods obtained by direct technique. a Mean, maximum and minimum annual flow. b Standard deviation and variance between maximum and minimum annual streamflows. c Coefficient of variation and pluviometric ratio



Fig. 9 The yearly duration of high and low streamflows over 50 years. a Danger level. b warning level. c Alert level. d Low streamflow

more often in future. The analysis in Fig.7b indicates that the difference between maximum and minimum yearly streamflows increases sharply with respect to time. Figure 7b also shows a continuous increment of standard deviation with respect to time. In other words, the annual streamflow gets farther from the mean annual flow, another indication of the increasing probability of high and low flows occurring.

In the early sub-periods, the PR values are close to 1 (Fig. 7c), demonstrating that the dispersion in annual streamflow data is very low. In the later sub-periods, there is a continuous increase in pluviometric ratio. The ratio reaches 2.89 and 2.35 in the last two sub-periods, meaning that annual variability of streamflow data is becoming very high. This is an additional indicator of increased probability of high and low flow incidences. Figure 7c indicates a constant increase in the CV with respect to time. The CV values in the last two subperiods are double compared to the first two subperiods. Clearly, the dispersion of annual flow around the mean becomes very high in later sub-periods. These analyses therefore prove the presence of large changes in the annual streamflow regime of the Selangor River along with the fact that these changes can create

suitable hydrological conditions for the increased probability of floods and droughts occurring in future.

The changes over the sub-periods obtained via the direct technique The changes of hydrological variables over the sub-periods obtained by direct technique are illustrated in Fig. 8. Generally, the hydrological variable changes over the sub-periods obtained by direct method are similar to the changes over the sub-periods obtained by change-point test. This similarity emphasizes the results regarding the long-term streamflow regime variations of the Selangor River and provides additional evidence that the observed changes in hydrological variables can produce suitable hydrological conditions for the increased probability of floods and droughts happening in future.

4.4 The changes in high and low streamflow durations

The yearly duration of high and low streamflows was investigated over 50 years (1961–2010). The work includes an assessment of the changes in high and low streamflow durations throughout the 50 years along with testing for any trend in changes as described earlier.

The high streamflow duration was looked at from three levels: danger level, when the streamflow is higher than 250 m³/s; warning level, when the streamflow is above 180 m³/s; and alert level, when the streamflow is more than 160 m^3/s . The three levels are as determined by the Department of Irrigation and Drainage (DID). The low streamflow analysis was run in a single level that is when the streamflow drops below $14.5 \text{ m}^3/\text{s}$ (which represents roughly 25 % of the average annual streamflow). Figure 9 illustrates the yearly duration of high and low streamflows over 50 years. There is an apparent significant increase in the danger level duration, especially in the last decade while minor change in the duration of warning and alert levels takes place. A slight change in the yearly duration of the low streamflow was also noticed through the study period. Three-year moving average of yearly duration of the high and low streamflows was applied to study the overall trend of changes in the study interval. Figure 10 presents the three-year moving average of the yearly duration of high and low streamflows over 50 years and the trend line of yearly data. The three moving average confirms the evident considerable increase in the duration of danger level. Figure 10 also indicates a

slight increase in the duration of warning level, no changes in the duration of alert level and a slight decrease in the duration of low streamflow. Such changes in the yearly duration of streamflow danger level lead to the formation of suitable hydrological conditions for floods to occur. As a direct result, flood events would probably happen more frequently in future.

5 Conclusions

A new advanced analytical approach to investigating the long-term changes of the annual streamflow regime in the Selangor River over a 50-year period from 1961 to 2010 has been developed in this paper. The long-term change analysis includes a study of the changes in nine hydrological variables that describe annual streamflow, as well as an investigation of the modifications in the yearly duration of high and low streamflows. The analyses were conducted based on two time scales, namely yearly changes and sub-periodic changes. The subperiods were the result of study period segmentation into seven sub-periods by two techniques: the changepoint test and direct technique. Prior to conducting the



Fig. 10 Three-year moving average of the yearly duration of high and low streamflows. a Danger level. b Warning level. c Alert level. d Low streamflow

long-term change analyses, the streamflow data was first tested with the assumption of independence, normality, and homogeneity tests using the Pearson correlation coefficient and Shapiro–Wilk and Pettitt's tests, respectively. The tests' results proved that the annual streamflow data can be considered independent, normally distributed, and homogeneous.

Even though analysis revealed nearly negligible changes in mean annual flow over the study period, the maximum annual flow generally increased while the minimum annual flow significantly decreased with respect to time. It was also observed that the variables describing the dispersion in streamflow data continually increased with respect to time.

An obvious increase was detected in the yearly duration of danger level of streamflow, a slight increase was noted in the yearly duration of warning and alert levels, and a slight decrease in the yearly duration of low streamflow was found. According to these results, the main conclusions drawn are as follows:

- Apparent changes were observed in the hydrological variables that describe annual streamflow and yearly duration of high streamflow over the 1961–2010 study period; these changes prove that there are long-term changes in the annual streamflow regime of the Selangor River
- The observed long-term changes in annual streamflow regime may potentially result in the development of suitable hydrological conditions that can increase the probability of flood and drought events occurring in future.

These results draw attention to the necessity to improve water resource management in light of the increasing probability of future droughts. Also, awareness is drawn to the development of flood protection plans in response to the increasing probability of future flood events. This work is an initial step toward reaching comprehensive knowledge about the changes and variations in the streamflow regime of the Selangor River basin. It raises awareness of the real need for higher understanding of the hydrology in the Selangor River basin, which would enhance water resource management and flood protection in the Selangor River basin to avoid the harmful effects that could ensue from the probable changes in annual streamflow regime.

Acknowledgments The authors appreciate the support rendered by the Hydrology and Water Resources Division of the Department of Irrigation and Drainage (DID), Malaysia. Special thanks are due to Malaysian International Scholarship (MIS, MOHE) and Asian Core Program (ACP) for their support in the study. The authors would like to acknowledge the financial support from University of Malaya through the grant FL026-2012D, FL001-13SUS, and RG244-12AFR. Special thanks are due to the reviewers whose supportive comments and ideas have led to a weighty enhancement of the manuscript.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Ahmad U, Shabri A, Zakaria Z (2013) An analysis of annual maximum streamflows in Terengganu, Malaysia using TL-moments approach. Theor Appl Climatol 111:649–663. doi:10.1007/s00704-012-0679-x
- Albrecher H, Ladoucette SA, Teugels JL (2010) Asymptotics of the sample coefficient of variation and the sample dispersion. J Stat Plan Infer 140:358–368. doi:10.1016/j.jspi.2009.03.026
- Beaulieu C, Seidou O, Ouarda TBMJ, Zhang X (2009) Intercomparison of homogenization techniques for precipitation data continued: Comparison of two recent Bayesian change point models Water Resources Research 45
- Birsan M-V, Molnar P, Burlando P, Pfaundler M (2005) Streamflow trends in Switzerland. J Hydrol 314:312–329. doi:10.1016/j. jhydrol.2005.06.008
- Boik RJ, Shirvani A (2009) Principal components on coefficient of variation matrices Statistical. Methodology 6:21–46. doi:10.1016/j. stamet.2008.02.006
- Breemen MTJ (2008) Salt intrusion in the Selangor Estuary in Malaysia. University of Twente, The Netherlands
- Buishand TA (1982) Some methods for testing the homogeneity of rainfall records. J Hydrol 58:11–27. doi:10.1016/0022-1694(82) 90066-x
- Burn DH, Cunderlik JM, Pietroniro A (2004) Hydrological trends and variability in the Liard River basin/Tendances hydrologiques et variabilité dans le basin de la rivière. Liard Hydrol Sci J 49:53–67. doi:10.1623/hysj.49.1.53.53994
- Burn DH, Sharif M, Zhang K (2010) Detection of trends in hydrological extremes for Canadian watersheds. Hydrol Process 24:1781–1790
- Chen Z, Chen Y, Li B (2013) Quantifying the effects of climate variability and human activities on runoff for Kaidu River basin in arid region of northwest China. Theor Appl Climatol 111:537–545. doi:10. 1007/s00704-012-0680-4
- Chiu SN, Wang L (2009) Homogeneity tests for several Poisson populations. Comput Stat Data Anal 53:4266–4278. doi:10.1016/j.csda. 2009.05.017
- Coin D (2008) A goodness-of-fit test for normality based on polynomial regression. Comput Stat Data Anal 52:2185–2198. doi:10.1016/j. csda.2007.07.012
- Declercq D, Duvaut P (1999) Hermite normality tests. Signal Process 69: 101–116. doi:10.1016/s0165-1684(98)00093-0
- Descroix L, Genthon P, Amogu O, Rajot J-L, Sighomnou D, Vauclin M (2012) Change in Sahelian Rivers hydrograph: The case of recent red floods of the Niger River in the Niamey Region. Glob Planet Chang 98–99:18–30. doi:10.1016/j.gloplacha.2012.07.009
- Douglas EM, Vogel RM, Kroll CN (2000) Trends in floods and low flows in the United States: impact of spatial correlation. J Hydrol 240:90– 105. doi:10.1016/s0022-1694(00)00336-x
- Firat M, Dikbas F, Koc AC, Gungor M (2010) Missing data analysis and homogeneity test for Turkish precipitation series. Acad Proc Eng Sci Sadhana 35:707–720
- Gautam MR, Acharya K (2012) Streamflow trends in Nepal. Hydrol Sci J 57:344–357. doi:10.1080/02626667.2011.637042
- Gedikli A, Aksoy H, Erdem Unal N, Kehagias A (2010) Modified dynamic programming approach for offline segmentation of long hydrometeorological time series. Stoch Env Res Risk A 24:547– 557. doi:10.1007/s00477-009-0335-x
- Hannaford J, Buys G (2012) Trends in seasonal river flow regimes in the UK. J Hydrol. doi:10.1016/j.jhydrol.2012.09.044

- Hassan AJ, Ghani AA, Abdullah R (2004) Development of flood risk map using GIS for Sg. Selangor Basin. National Hydraulic Research Institute of Malaysia, Malaysia
- Hodgkins GA, Dudley RW, Huntington TG (2003) Changes in the timing of high river flows in New England over the 20th century. J Hydrol 278:244–252
- Hubert P (2000) The segmentation procedure as a tool for discrete modeling of hydrometeorological regimes. Stoch Env Res Risk A 14:297–304. doi:10.1007/pl00013450
- Hwang Y-T, Wei PF (2006) A novel method for testing normality in a mixed model of a nested classification. Comput Stat Data Anal 51: 1163–1183. doi:10.1016/j.csda.2005.11.014
- Jurečková J, Picek J (2007) Shapiro–Wilk-type test of normality under nuisance regression and scale. Comput Stat Data Anal 51:5184– 5191. doi:10.1016/j.csda.2006.08.026
- Kang HM, Yusof F (2012) Homogeneity tests on daily rainfall series in peninsular. Malaysia Int J Contemp Math Sci 7:9–22
- Kehagias A (2004) A hidden Markov model segmentation procedure for hydrological and environmental time series. Stoch Env Res Risk A 18:117–130. doi:10.1007/s00477-003-0145-5
- Kehagias A, Nidelkou E, Petridis V (2006) A dynamic programming segmentation procedure for hydrological and environmental time series. Stoch Env Res Risk A 20:77–94. doi:10.1007/s00477-005-0013-6
- Korhonen J, Kuusisto E (2010) Long-term changes in the discharge regime in Finland. Hydrol Res 41:253–268
- Kottegoda NT (1984) Investigation of outliers in annual maximum flow series. J Hydrol 72:105–137. doi:10.1016/0022-1694(84)90187-2
- Krasovskaia I, Gottschalk L (2002) River flow regimes in a changing climate. Hydrol Sci J 47:597-609. doi:10.1080/ 02626660209492962
- Kumar S, Merwade V, Kam J, Thurner K (2009) Streamflow trends in Indiana: effects of long term persistence, precipitation and subsurface drains. J Hydrol 374:171–183
- Kundzewicz ZW, Robson AJ (2004) Change detection in hydrological records—a review of the methodology/Revue méthodologique de la détection de changements dans les chroniques hydrologiques. Hydrol Sci J 49:7–19
- Laraque A, Ronchail J, Cochonneau G, Pombosa R, Guyot JL (2007) Heterogeneous distribution of rainfall and discharge regimes in the Ecuadorian Amazon basin. J Hydrometeorol 8:1364–1381. doi:10. 1175/2007jhm784.1
- Lee CM (2002) Master plan study on flood mitigation and river management for Sg. Selangor River Basin. vol 2. Drainage and Irrigation Department (DID) Malaysia
- Lesaffre E (1983) Normality tests and transformations. Pattern Recogn Lett 1:187–199. doi:10.1016/0167-8655(83)90061-2
- Ma Z, Kang S, Zhang L, Tong L, Su X (2008) Analysis of impacts of climate variability and human activity on streamflow for a river basin in arid region of northwest China. J Hydrol 352:239–249. doi:10.1016/j.jhydrol.2007.12.022
- Miao CY, Shi W, Chen XH, Yang L (2012) Spatio-temporal variability of streamflow in the Yellow River: Possible causes and implications. Hydrol Sci J 57:1355–1367. doi:10.1080/ 02626667.2012.718077
- Mirabbasi R, Fakheri-Fard A, Dinpashoh Y (2012) Bivariate drought frequency analysis using the copula method. Theor Appl Climatol 108:191–206. doi:10.1007/s00704-011-0524-7
- Moliere DR, Lowry JBC, Humphrey CL (2009) Classifying the flow regime of data-limited streams in the wet-dry tropical region of Australia. J Hydrol 367:1–13. doi:10.1016/j.jhydrol.2008.12.015
- Montanari A, Rosso R, Taqqu MS (2000) A seasonal fractional ARIMA model applied to the Nile River monthly flows at Aswan. Water Resour Res 36:1249–1259. doi:10.1029/2000wr900012
- Morán-Tejeda E, López-Moreno JI, Ceballos-Barbancho A, Vicente-Serrano SM (2011) River regimes and recent hydrological changes

in the Duero basin (Spain). J Hydrol 404:241–258. doi:10.1016/j. jhydrol.2011.04.034

- Nelson WB (2002) An unusual turbidity maximum. In: Johan CW, Cees K (eds) Proceedings in Marine Science, vol Volume 5. Elsevier, pp 483–497. doi:10.1016/s1568-2692(02)80035-3
- Noguchi K, Gel YR, Duguay CR (2011) Bootstrap-based tests for trends in hydrological time series, with application to ice phenology data. J Hydrol 410:150–161
- Önder AÖ, Zaman A (2005) Robust tests for normality of errors in regression models. Econ Lett 86:63–68. doi:10.1016/j.econlet. 2004.06.008
- Opitz-Stapleton S, Gangopadhyay S (2011) A non-parametric, statistical downscaling algorithm applied to the Rohini River basin. Nepal Theor Appl Climatol 103:375–386. doi:10.1007/s00704-010-0301-z
- Petersen-Overleir A, Reitan T (2005) Objective segmentation in compound rating curves. J Hydrol 311:188–201. doi:10.1016/j.jhydrol. 2005.01.016
- Pettitt AN (1979) A non parametric approach to the change point problem. Appl Stat 28:126–135
- Poff NL, Ward JV (1989) Implications of streamflow variability and predictability for lotic community structure: a regional analysis of streamflow patterns. Can J Fish Aquat Sci 46:1805–1818
- Poff NL et al (1997) The natural flow regime: a new paradigm for riverine conservation and restoration. Bioscience 47:769–784
- Post DA, Jones JA (2001) Hydrologic regimes of forested, mountainous, headwater basins in New Hampshire, North Carolina, Oregon, and Puerto Rico. Adv Water Resour 24:1195–1210. doi:10.1016/s0309-1708(01)00036-7
- Qin Y, Smith B, Lei Q (2009) Test for homogeneity in normal mixtures with unknown means and variances. J Stat Plan Infer 139:4165– 4178. doi:10.1016/j.jspi.2009.06.002
- Rakhshandehroo GR, Vaaghefi M, Shafiee MM (2010) Flood forecasting in similar catchments using neural networks. Turk J Eng Environ Sci 34:57–65
- Razali NM, Wah YB (2011) Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. J Stat Model Anal 2:21–33
- Richter BD (1996) A method for assessing hydrologic alteration within ecosystems. Un Metro Para Evaluar Alteraciones Hidrologicas Dentro de Ecosistemas 10:1163–1174
- Rocha WFC, Nogueira R, da Silva GE B, Queiroz SM, Sarmanho GF (2012) A comparison of three procedures for robust PCA of experimental results of the homogeneity test of a new sodium diclofenac candidate certified reference material. Microchem J. doi:10.1016/j. microc.2012.03.028
- Rougé C, Ge Y, Cai X (2012) Detecting gradual and abrupt changes in hydrological records. Adv Water Resour. doi:10.1016/j.advwatres. 2012.09.008
- Royston P (1992) Approximating the Shapiro-Wilk W-test for nonnormality. Stat Comput 2:117–119. doi:10.1007/bf01891203
- Samsudin R, Saad P, Shabri A (2011) River flow time series using least squares support vector machines. Hydrol Earth Syst Sci 15:1835– 1852. doi:10.5194/hess-15-1835-2011
- Sang Y-F, Wang Z, Liu C, Yu J (2013) The impact of changing environments on the runoff regimes of the arid Heihe River basin, China Theoretical and Applied Climatology:1–9 doi:10.1007/s00704-013-0888-y
- Seyam M, Othman F (2014) The influence of accurate lag time Estimation on the performance of stream flow data-driven based models Water Resour Manage:1–15 doi:10.1007/s11269-014-0628-9
- Shafie A (2009) Extreme flood event: a case study on floods of 2006 and 2007 in Johor, Malaysia. Colorado State University, Fort Collin
- Subramaniam V (2004) Managing water supply in Selangor and Kuala Lumpur vol 22. THE BOARD OF ENGINEERS MALAYSIA, 50580 Kuala Lumpur, Malaysia

- Sun S, Chen H, Ju W, Song J, Zhang H, Sun J, Fang Y (2013) Effects of climate change on annual streamflow using climate elasticity in Poyang Lake basin. China Theor Appl Climatol 112:169–183. doi: 10.1007/s00704-012-0714-y
- Tenreiro C (2011) An affine invariant multiple test procedure for assessing multivariate normality. Comput Stat Data Anal 55:1980– 1992. doi:10.1016/j.csda.2010.12.004
- Villarini G, Smith JA, Baeck ML, Vitolo R, Stephenson DB, Krajewski WF (2011) On the frequency of heavy rainfall for the Midwest of the United States. J Hydrol 400:103–120. doi:10.1016/j.jhydrol.2011. 01.027
- Walling DE, Fang D (2003) Recent trends in the suspended sediment loads of the world's rivers. Glob Planet Chang 39:111–126. doi:10. 1016/s0921-8181(03)00020-1
- Wang XL, Wen QH, Wu Y (2007) Penalized maximal t test for detecting undocumented mean change in climate data series. J Appl Meteorol Climatol 46:916–931
- Xiong L, Guo S (2004) Trend test and change-point detection for the annual discharge series of the Yangtze River at the Yichang hydrological station/Test de tendance et détection de rupture appliqués aux séries de débit annuel du fleuve Yangtze à la station hydrologique de Yichang, Hydrol Sci J 49:99–112. doi:10.1623/hysj.49.1.99.53998
- Xu Y-P, Yu C, Zhang X, Zhang Q, Xu X (2012) Design rainfall depth estimation through two regional frequency analysis methods in

Hanjiang River basin. China TheorAppl Climatol 107:563–578. doi:10.1007/s00704-011-0497-6

- Yang SL, Gao A, Hotz HM, Zhu J, Dai SB, Li M (2005) Trends in annual discharge from the Yangtze River to the sea (1865– 2004)/Tendances et épisodes extrêmes dans les débits annuels du Fleuve Yangtze débouchant dans la mer (1865–2004). Hydrol Sci J 50:836
- Yue S, Pilon P, Phinney BOB (2003) Canadian streamflow trend detection: impacts of serial and cross-correlation. Hydrol Sci J 48:51–63. doi:10.1623/hysj.48.1.51.43478
- Zhang J, Wu Y (2005) Likelihood-ratio tests for normality. Comput Stat Data Anal 49:709–721. doi:10.1016/j.csda.2004.05.034
- Zhang X, Vincent LA, Hogg WD, Niitsoo A (2000) Temperature and precipitation trends in Canada during the 20th century Atmosphere–ocean 38:395–429 doi:10.1080/07055900.2000. 9649654
- Zheng H, Zhang L, Liu C, Shao Q, Fukushima Y (2007) Changes in stream flow regime in headwater catchments of the Yellow River basin since the 1950s. Hydrol Process 21:886–893. doi:10.1002/ hyp.6280
- Zin W, Jemain A, Ibrahim K (2013) Analysis of drought condition and risk in Peninsular Malaysia using standardised precipitation index. Theor Appl Climatol 111:559–568. doi:10. 1007/s00704-012-0682-2