

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Temporal Water Quality Assessment of Langat River from 1995-2006

Zalina Mohd Ali¹, Noor Akma Ibrahim², Kerrie Mengersen³,
Mahendran Shitan², Hafizan Juahir⁴ and
Faridatul Azna Ahmad Shahabuddin¹

¹*Faculty of Science and Technology, University Kebangsaan,*

²*Faculty of Science, University Putra,*

³*Faculty of Science and Technology, Queensland University of Technology,*

⁴*Faculty of Environmental Studies, University Putra,*

^{1,2,4}*Malaysia*

³*Australia*

1. Introduction

Water quality is generally described according to biological, chemical and physical properties (Coke et al 2005). Based on these properties, the quality of water can be expressed via a numerical index (i.e. Water Quality Index, WQI) by combining measurements of selected water quality variables. The index is important in evaluating the water quality of different sources and in observing the changes in the water quality as a function of time and other influencing factors (Sarkar and Abbasi 2006). The time when samples are taken is one of the contributing factors that can influence the concentration of a particular water quality variable (Coke et al 2005). Thus, temporal assessment is a good indication in determining the presence or absence of trend and seasonality to which water quality is responding to changes in the catchment and time.

However, the assessment on the temporal effect of water quality sub indices variables and WQI are rarely carried out by the Malaysian Department of Environment and Malaysian Department of Irrigation and Drainage (DID). Several studies by past researchers investigated the water quality assessment in Langat River especially in spatial assessment such as evaluating the polluting effects from various land use pattern (Suki et al 1988), relationship between water quality and sewage discharge and location Lee et al (2006) and on spatial variations of water quality variables (Juahir et al 2010a). Most of the studies did not consider the temporal assessment in details. Therefore, the influence of time on selected water quality variables and water quality index of Langat River are studied by using box plot to examine the annual and quarterly pattern. Then, regression time series and decomposition analysis are carried out on normally distributed variables with no outliers at particular stations along the Langat River. Both methods are helpful in evaluating the changes with time and in determining the best fitted models of the selected variables.

2. Description of study area

2.1 Background

Pollution prevention improvement programme has been introduced by the Malaysian Department of Environment (DOE) from 2001 to improve the condition of polluted rivers in Malaysia. Langat River which is situated in the state of Selangor, Peninsular Malaysia with a total catchment area of approximately 1,815km² is chosen for this programme. The catchment area is shown in Fig. 1. Data used in the analysis for Langat River were collected from six monitoring stations as shown in Table 1 and Fig.1.

DOE station number	Station number	DOE Station Code	Distance from estuary (km)	Grid reference	Location
2814602	1	IL01	4.19	2°52.027' 101°26.241'	Air Tawar Village
2815603	2	IL02	33.49	2°48.952' 101°30.780'	Telok Datuk, near Banting town
2817641	3	IL03	63.43	2°51.311' 101°40.882'	Bridge at Dengkil Village
2918606	4	IL04	81.14	2°57.835' 101°47.030'	Near West Country Estate
2917642	5	IL05	86.94	2°59.533' 101°47.219'	Kajang Bridge
3118647	6	IL07	113.99	3°09.953' 101°50.926'	Bridge at Batu 18

Table 1. DOE sampling station at the study area

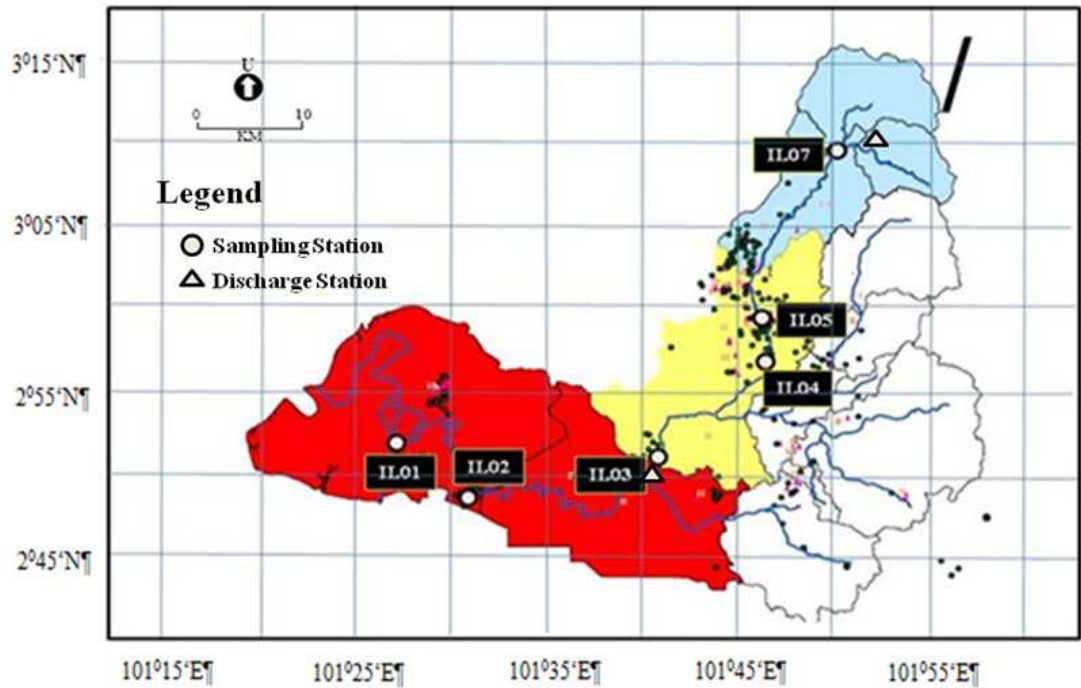


Fig. 1. Location of the selected sampling stations

2.2 Land use changes

The land use of Langat catchment consists of mainly agriculture, forest, urban areas (commercial and residential) and water bodies. There are 3 types of forest in Langat River catchment area such as dipterocarp, peat swamp and mangrove. Agriculture is the dominant land use, followed by forest, urban areas and water bodies of Langat catchment as presented by Table 2 and Fig. 2. However, Langat River as a tropical catchment area is experiencing rapid urbanization (Amini et al 2009) where the urban expansion occurred since 1981 (see Table 2 and Fig. 2). The gain in size of urbanised area was also reported by Jaafar et al (2009) and the urban development which occurred in the Langat River catchment was due to extensive land exchange from agriculture to urban-industrial-commercial use.

Year	Forest	Developed Area/Urban	Agriculture	Water bodies
1981	41.42	2.59	54.79	1.19
1988	30.32	9.61	57.97	2.10
1991	29.89	9.72	58.21	2.18
1996	27.47	10.56	60.22	1.75
2001	24.11	17.56	56.21	2.12

Table 2. Land use composition (in %) for Langat river basin (Source : Idrus et al 2003)

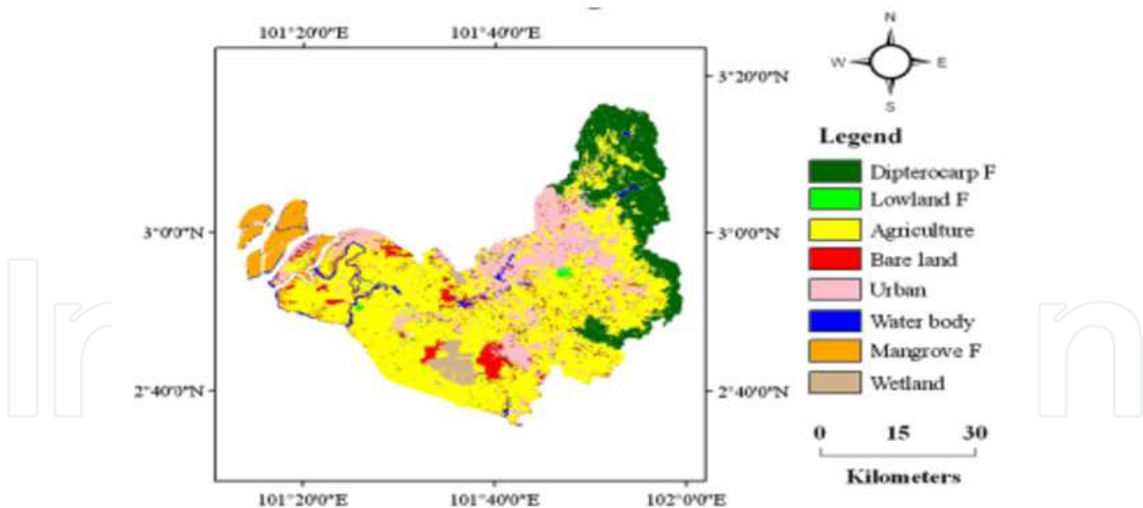


Fig. 2. GIS map of land uses in Langat River Basin for 2005 (Source : Mokhtar et al 2011)

2.3 Population change

Selangor is one of the most populated states in Malaysia with rapid growth from 1980 to 2010. Based on the Malaysian government census reports in 1980, 1991, 2000 and 2010, the recorded population of Selangor was 1.4 million in 1980, 2.3 million in 1990, close to 3.9 million in 2000 and rise to 5.4 million in 2010. In Selangor, a large portion of Langat River

Basin is located in three administrative districts of Selangor, namely Hulu Langat, Sepang and Kuala Langat as shown in Fig. 3. The rest of the basin covers part of Negeri Sembilan, Klang, Petaling and Federal Territory of Kuala Lumpur.

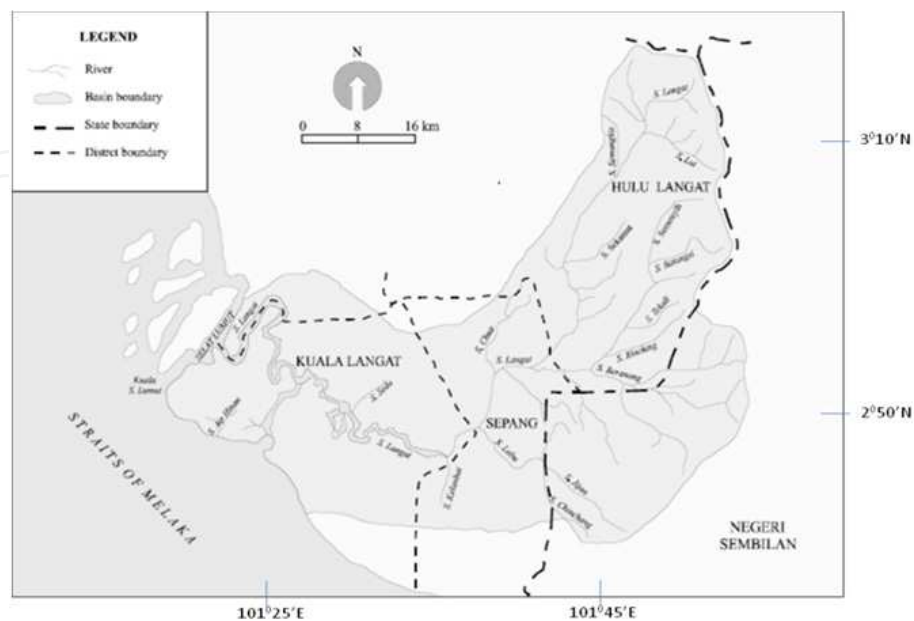


Fig. 3. Langat River Basin

Table 3 and Fig. 4 show the district population in the Langat River Basin for Kuala Langat, Sepang and Hulu Langat. The population of these three districts increased from 1980 to 2010 and the population increased was consistent with the urban development discussed above.

District	1980	1991	2000	2010
Kuala Langat	101578	130,090	192176	224648
Sepang	46025	54671	108640	211361
Hulu Langat	177877	413900	864451	1156585

Table 3. Number of population by districts along Langat River

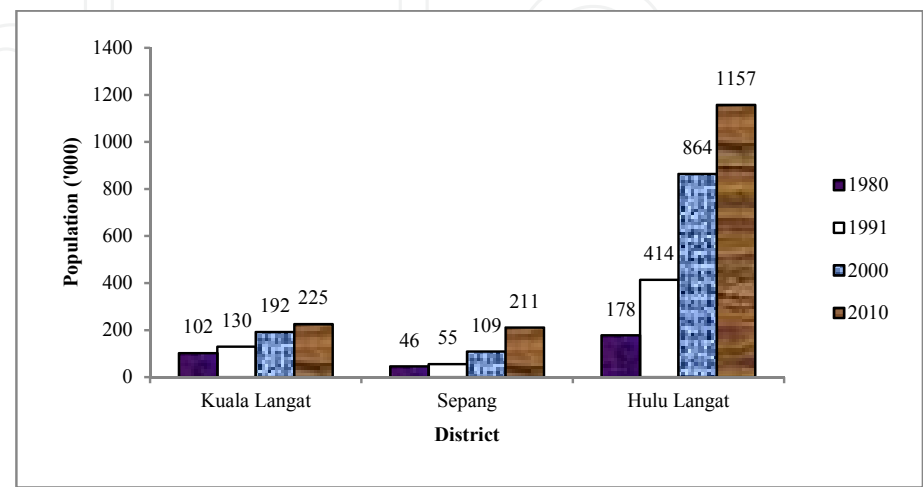


Fig. 4. Number of population along Langat River Basin, by district from 1980-2010.

2.4 Sources of pollutions

Langat River is one of the most important raw water resources for drinking water, recreation, industry, fishery and agriculture. The river flows from the highest peak of 1493 meter of Gunung Nuang across Langat Basin to Kuala Langat and land use activities along the river banks contribute to deterioration of river water quality (Charlie 2010). The sources of the Langat River pollution are identified as industrial discharge (58%), domestic sewage from treatment plants (28%), construction projects (12%) and pig farming (2%) (Khairuddin et al 2002). A study from Juahir et al (2010a) showed that major sources of surface water quality variations in Langat River come from industrial effluents, wastewater treatment plants, domestic and commercial areas.

The declining quality of the river water is caused by two main sources of water pollution, i.e. point sources and non-point sources. Point source (PS pollution) is single and identifiable source that discharge pollutants into the environment such as discharge from manufacturing and agro-based industries, sewage treatment plants and animal farms. On the other hand, non-point sources pollution (NPS pollution), also known as polluted runoff is pollution where sources cannot be traced to a single point. NPS is defined as pollution originating from diffused sources such as agricultural activities and surface runoffs which contributed by storm runoff, e.g. from rainfall, snowmelt, or irrigation over land surfaces into the drainage system (Sapari et al 2009). The increase of PS and NPS pollution loading such as discharges of surface runoff, domestic sewage, ship wastes and industrial discharges into coastal waters may resulted by rapid urbanization along the river. The NPS pollution is seen as the main contributor to the pollution load in Langat River compared to PS pollution (UPUM 2002). It is also obvious that the Langat River ecosystem is under stress from the discharge of effluents particularly domestic sewage (Lee et al 2006). Table 4 indicates some of the major source of pollution in the study area (Department Of Environment, DOE 2007).

Point Source (PS)	Non-Point Sources (NPS)
Industrial Estate	Agriculture
Landfills	Construction
Piggeries	Forest
Private Sewage Treatment Plants	Residential
Public Sewage Treatment Plants	Residential and Commercial
Restaurants	Urban
Sand Mining Activities	
Wet Markets	
Workshops	

Table 4. Sources of Pollution in Sub-catchment of Langat River

Due to rapid urbanization and changes from undeveloped to developed area, Langat River experienced changes of discharges and direct runoff volume. A study done by Juahir et al (2010b) has showed that there is a relationship between land use and discharge or flow rate and runoff in Langat River. The annual mean discharge and direct runoff of Langat River at two selected gauging stations, i.e. Dengkil Station (station number 2816441) and Lui Station (station number 2917401) are shown in Fig. 5 and Fig. 6. The Dengkil gauging station is located at the downstream of Langat River (2°51'20"N, 101°40'55"N) and Lui gauging

station is located at the upstream of Langat River ($3^{\circ}10'25''\text{N}$, $101^{\circ}52' 20''\text{E}$). The increment trend of discharge and direct runoff at Dengkil Station compared to Lui Station is consistent with the increasing trend of urban development and the decrease of agriculture and forest areas within the region (Juahir et al 2010b).

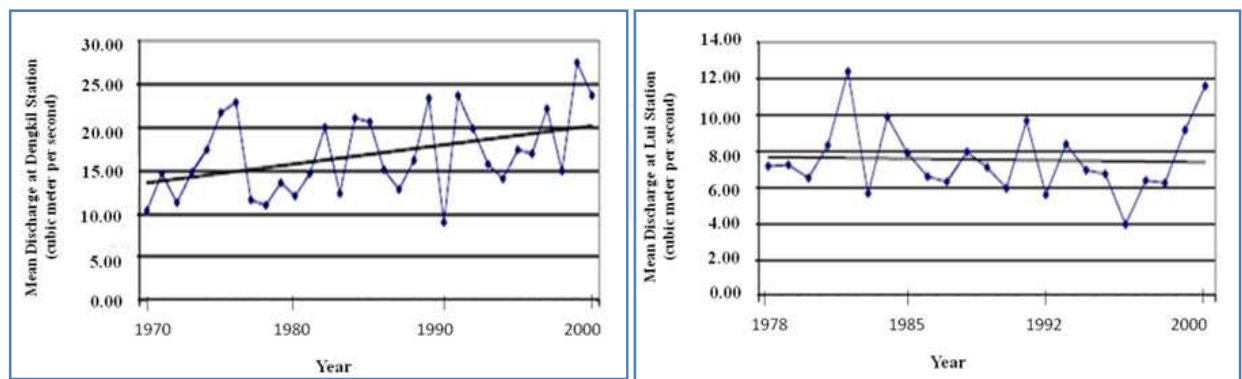


Fig. 5. Mean discharge at Dengkil Station and Lui Station (adapted from Juahir et al 2010)

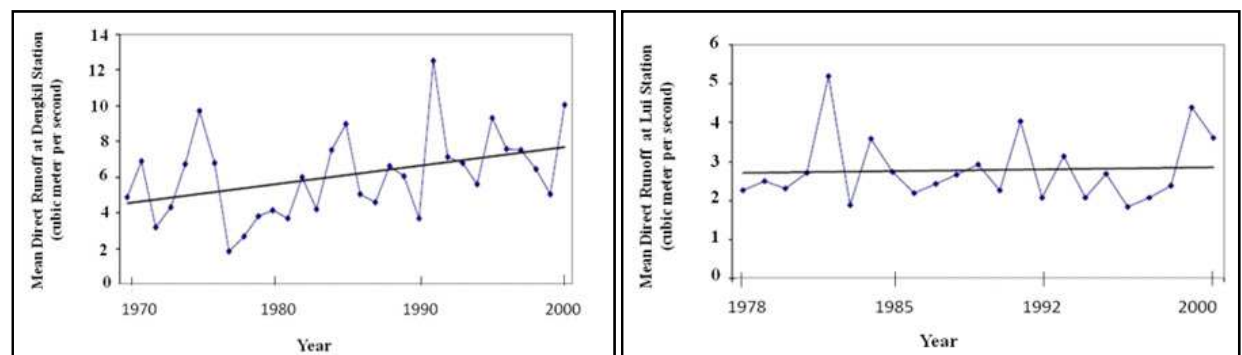


Fig. 6. Mean direct runoff at Dengkil Station and Lui Station (adapted from Juahir et al 2010)

3. Water quality data

Water quality data used in this study were obtained from the Malaysian Department of Environment (DOE). The data obtained however, were not collected at regular time intervals and to facilitate the analysis, quarterly data was used instead. Time series data from September 1995 until December 2007 for selected parameters and stations were used in the present study. Since quarterly data is used in the analysis, the data that represent the first quarter is taken from the last month of that quarter i.e. data from the month of March. Similarly, for the second quarter, data from the month of June will represent the data for that second quarter. However, if the last month of a quarter does not contain any data, then data from either the first or second month of that quarter will be considered . For example, for quarter 1, the data that represent quarter 1 will either be from the month of February or January, likewise for quarter 2 the data from the month of May or April will be taken to represent the second quarter data.

The six selected water quality variables used in this study are Suspended Solids (SS), Biochemical Oxygen Demand (BOD), Ammoniacal Nitrogen (AN), Chemical Oxygen Demand (COD), Dissolved Oxygen (DO) and pH. These variables were selected by the

panel of experts as the variables that when calculated and use collectively will give some indication on the water quality level or water quality index of a river (DOE 1997). According to the best-fit relationship for each six parameters, the new variables of the 6 sub indices (SI) were determined and the overall trend for Langat River were obtained using the formula given by

$$WQI = 0.22\ SIDO + 0.19\ SIBOD + 0.16\ SICOD + 0.16\ SISS + 0.15\ SIAN + 0.12\ SIpH \tag{1}$$

where,

WQI = Water quality index; SIDO = Sub-index of Dissolved Oxygen; SIBOD = Sub-index of Biological Oxygen Demand; SICOD = Sub-index of Chemical Oxygen Demand; SIAN = Sub-index of Ammonical Nitrogen; SISS = Sub-index of Suspended Solid; SIpH = Sub-index of pH value.

Generally, WQI is a unitless number varies between 0 and 100. Measurements of each of these parameters are taken and compared to a classification table (see Table 5), where the water is identified as excellent, good, fair, poor or very poor (DID 2009).

Status	Polluted	Slightly Polluted	Clean
WQI value	0-59	60-80	81-100

Table 5. WQI categorisation scheme

4. Methods

4.1 Graphical analysis

Graphical analysis is very useful in data analysis and helps the researcher in seeing pattern, trends and other features not easily apparent using numerical summaries. Box-Whisker plot, normal probability plot and scatter plot were used to analyze data graphically in this study. Displaying data using graphs allow for more effective visualization and presentation of large data sets in a small space (Cooke et al 2005). By using graphical analysis, we can visualize any gaps in the data, relationship between variables and trends that might exist in the selected water quality data.

4.1.1 Box-whisker plot

Box-Whisker plot is a powerful exploratory data analysis tool. It is also called the five-number summary (Tukey, 1977). To plot, the given sub index data are ranked from smallest to largest value. Then, the five-number summary which include the smallest and largest values, the median (a measure of central tendency that is more robust and not sensitive to outlying values, indicates the 50th percentile), and the lower and upper hinges (i.e. 0.25 and 0.75 quartiles respectively) are obtained. This information is then represented by the Box-Whisker plot. The box represents the inter-quartile range and the whiskers are lines that extend from the box to the highest and lowest values, excluding outliers. The outliers are individual points with values beyond the highest and lowest limits and are plotted with asterisks. A line across the box indicates the median of the

data. If the median lie in the middle of the box with upper and lower whiskers of similar length, it shows that data for any given year or quarter are symmetric. The lacks of symmetry suggest departure from normality. Apart from that, the plot can also give immediate visuals about the center, the spread, and the overall range of distribution. Additionally, the median confidence intervals can be plotted along the boxes in Minitab as shown in Fig. 8-Fig. 10 and are useful in offering a rough guide to determine the differences of medians. If the two boxes do not overlap, this offers evidence of a statistically significant difference between the medians (Mcgill et al 1978). In this study, the plots are used to evaluate several things, for example to ascertain how the sub indices and water quality index time series data are distributed, to demonstrate outlier and to evaluate normality of the data. It also can be used to visualize the median differences and to track the annually and quarterly changes to the water quality data.

4.1.2 Normal probability plot

Normal probability plot is a graphical method for testing normality. If our data follows the hypothesized normal distribution, then the plotted points fall approximately along a straight line. There are various test statistics for normality and Anderson-Darling statistic is a widely used test (Montgomery et al 2008). If the p -value is smaller than the critical value (usually 0.05), the underlying population is not normal.

4.1.3 Scatter plot

The sub indices water quality parameters with normal distribution will be selected based on the results from Box-Whisker plot and probability plot. The scatter plot of the selected parameters was constructed to see any trend and seasonal patterns. Scatter plot is used to show the relationship between dependent variable (sub indices water quality parameters) and independent variable (time). Each value of dependent variable is plotted against its corresponding time. If the sub indices for water quality parameters values tend to increase or decrease in a straight line fashion as the time increases, and if there is a scattering of the (time, sub index) points around the straight line, then it is reasonable to describe the relationship between the sub index and time by simple linear regression model. This could help to identify whether the differences in water quality variables are due to the actual trend or due to the changes in the water quality variables.

4.1.4 Analysis framework

Fig. 7 shows the analysis framework to be used in this study. The first step is to assess the temporal variations of the sub indices and water quality variables. The assessments were performed using Minitab 15.1. Then, exploratory assessment is carried out to determine the annual and quarterly changes of the selected water quality variables. From this analysis, if there seem to be some apparent trends or oscillations in the selected water quality variables, it will be evaluated. To assess the statistical significance of the changes and to examine the uncertainty about the possible trend and seasonality, regression and additive decomposition analysis were performed. To continue with these analyses, the variables considered are those without any outliers and are normally distributed. Models fitted for the selected variables at certain stations can be used to predict future values.

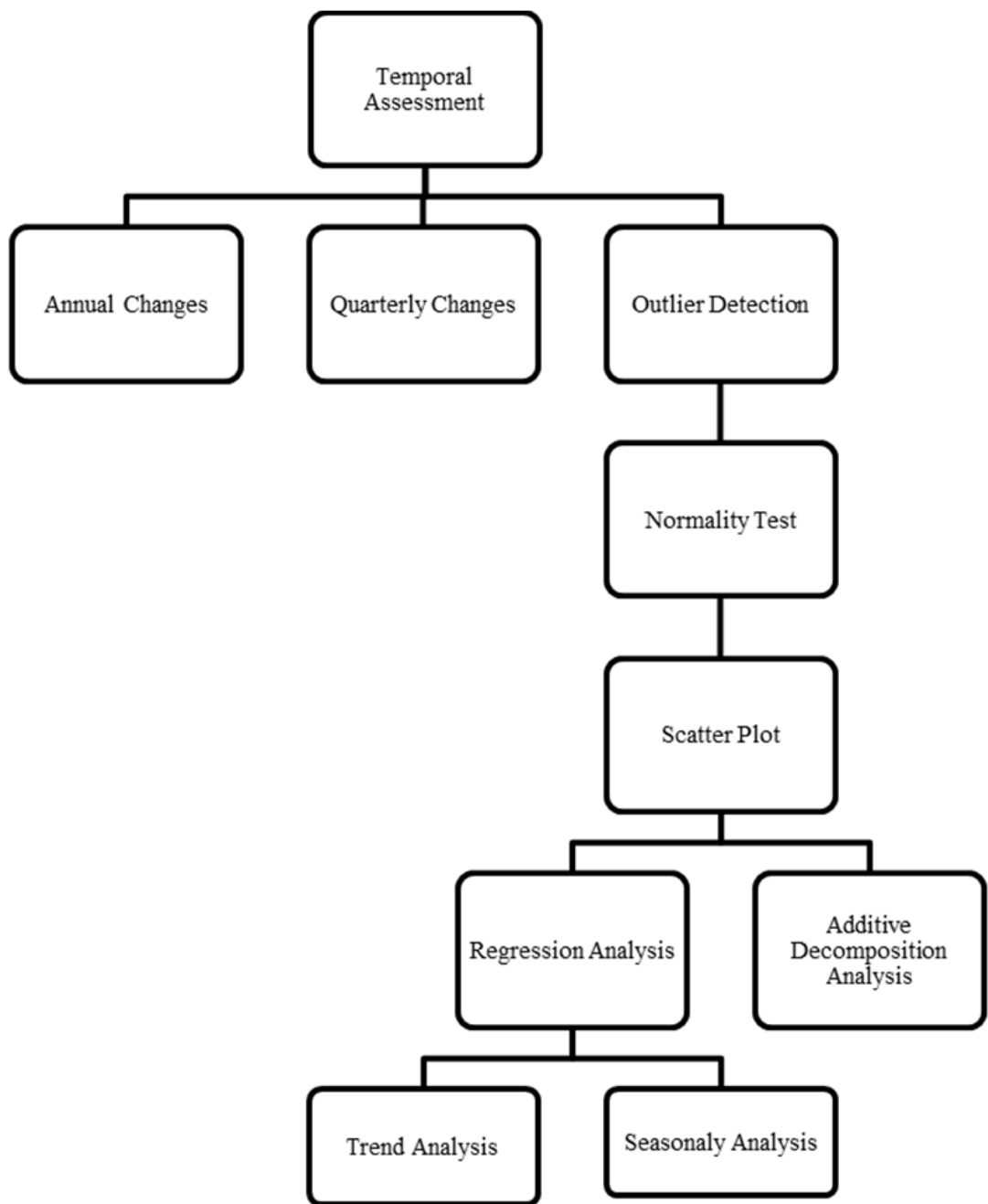


Fig. 7. Analysis Framework

4.2 Linear regression analysis

Linear regression analysis is an important parametric method to identify the monotonic trend in a time series. It is useful to describe the relationship between variables. The method is often performed to determine the slope of selected variables. In this study, the regression analysis was used to investigate and to model the relationship between the selected water quality variables versus time. The slope indicates the mean temporal change of the variables. Positive values of the slopes show increasing trends in the mean temporal change while negative values of the slopes indicate decreasing trends.

4.2.1 Trend analysis

The trend model

$$y_t = TR_t + e_t \quad (2)$$

$$\text{or } y_t = \beta_0 + \beta_1 * time + e_t \quad (3)$$

where,

y_t = variable values, TR_t = trend and e_t = error term, in time period t .

This model explains that the time series y_t can be represented by an average level (denoted $\mu_t = \beta_0 + \beta_1 * time$) that changes over time according to the equation $TR_t = \mu_t$ and by the error term, e_t .

4.2.2 Quarterly analysis

The quarterly model

$$y_t = \beta_0 + \beta_1 * time + \beta_2 Q1 + \beta_3 Q2 + \beta_4 Q3 + \beta_5 Q4 + e_t \quad (3)$$

where,

$Q1$ = quarter 1, $Q2$ = quarter 2, $Q3$ = quarter 3 and $Q4$ = quarter 4.

4.3 Additive decomposition analysis

The additive model is useful when the seasonal variation is relatively constant over time. In the present study, the selected parameters were separated into linear trend, seasonal and error components.

$$y_t = TR_t + SN_t + CL_t + e_t \quad (4)$$

where,

TR_t = trend, SN_t = seasonal, CL_t = cyclical and e_t = error term in time period t .

Basic steps in decomposition method are:

1. Estimate the trend using centered moving average.
2. De-trend the series by subtracting the trend estimated in (1) from the original series.
3. Estimate seasonal factors. This entails an effect for each quarter by using the average of the de-trended values for a particular season. However, Minitab uses median rather than mean.
4. Final step is to determine the random component. The random component could be summarized using measures such as mean absolute percentage error (MAPE, %), mean absolute deviation (MAD, expressed in the same unit as the variables) and mean squared deviation (MSD). MAPE and MAD measure the accuracy of fitted time series values, \hat{y} and the MSD is more sensitive to unusually large forecast errors (outliers). The formulas are as follows :

$$MAPE = \frac{\sum |(y_t - \hat{y}_t)/y_t|}{n} \times 100 \quad (5)$$

$$MAD = \frac{\sum |y_t - \hat{y}_t|}{n} \quad (6)$$

$$MSD = \frac{\sum |y_t - \hat{y}_t|^2}{n} \quad (7)$$

5. Results and discussion

5.1 Annual changes

Box plots provide a visual impression of the location and shape of the underlying distributions (Vega et al 1998). In this section, the presence or absence of trends over time was examined graphically through Fig. 8. From Fig. 8, SIDO and WQI showed a positive trend except for certain particular years for example in 1997 and 1999 for SIDO and 1997 for WQI where the medians were clearly lower than the median for the rest of the data. For SICOD and SIpH, the median were fairly stable without many variations for most of the years selected. However for SIBOD, the trend highly fluctuates from 1995 up to 1999 and after 2000 the fluctuation stabilised. The trend for SISS, however, showed a decreasing trend for three consecutive years from 1998 to 2000 and then in 2001 the trend increased. Similar decreasing pattern was observed again in 2004. The decreasing and increasing trend in SISS was also observed in SIAN but the decreasing trend started in 1995. The changes that occurred to the annual trend of water quality variables were influenced by many factors for example, certain hydrological events and/or developments in the river basin (Ravichandran 2003).

5.2 Quarterly changes

To plot the graphs in Fig.9, all quarterly data for each parameter and each station were combined from all selected years from 1995-2006. For example, quarterly data for SIDO for all selected years were combined for all stations i.e. Station 1 to Station 6. The distribution of each parameter and each quarter can then be examined as shown in Fig. 9. From Fig. 9, SIDO, SIBOD, SIAN, SISS and WQI showed that the quarterly median values did not differ much. The median values for SIDO in quarter 3 and quarter 4 were fairly equal. However, SISS showed a decreasing trend in quarter 4. The quarterly medians in SICOD and SIpH were generally similar for all quarters except in quarter 4 where SICOD showed a slight increase in the quarterly median.

5.3 Outliers detection

The box plots in Fig. 10 show that most of the variables depart from normality in their skewness. Many variables also have outliers and extreme values. To ease our analysis, variables with no outliers were selected i.e. SIDO at Station 1-5, SIBOD at station 5, SICOD at station 1 and 5, SIAN at station 2-5, SISS at station 2-3, SIpH at station 2 and WQI at station 1,4 and 5.

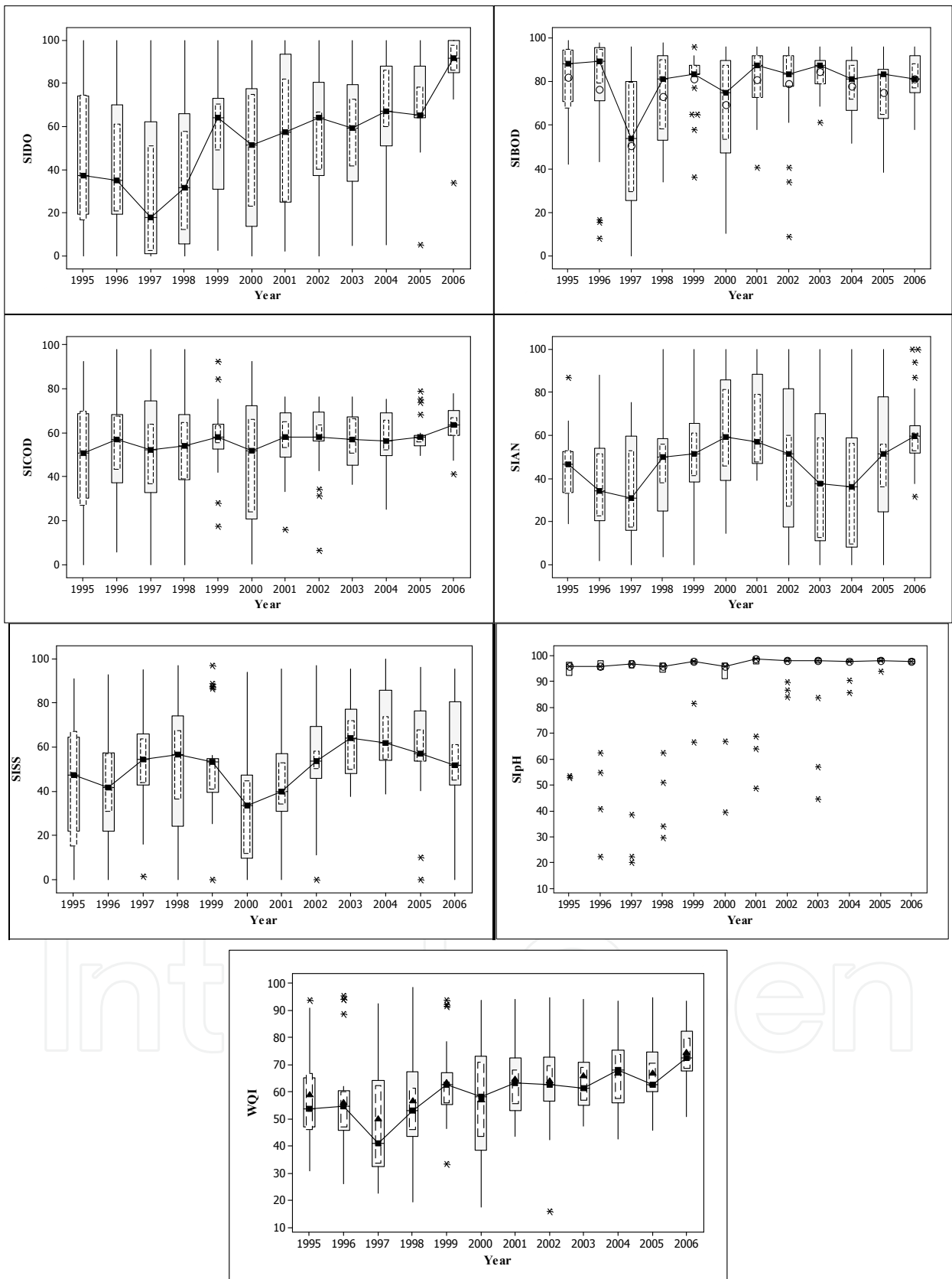


Fig. 8. Box plots of measured values of sub indices water quality parameters and water quality index in Langat River during 1995-2006

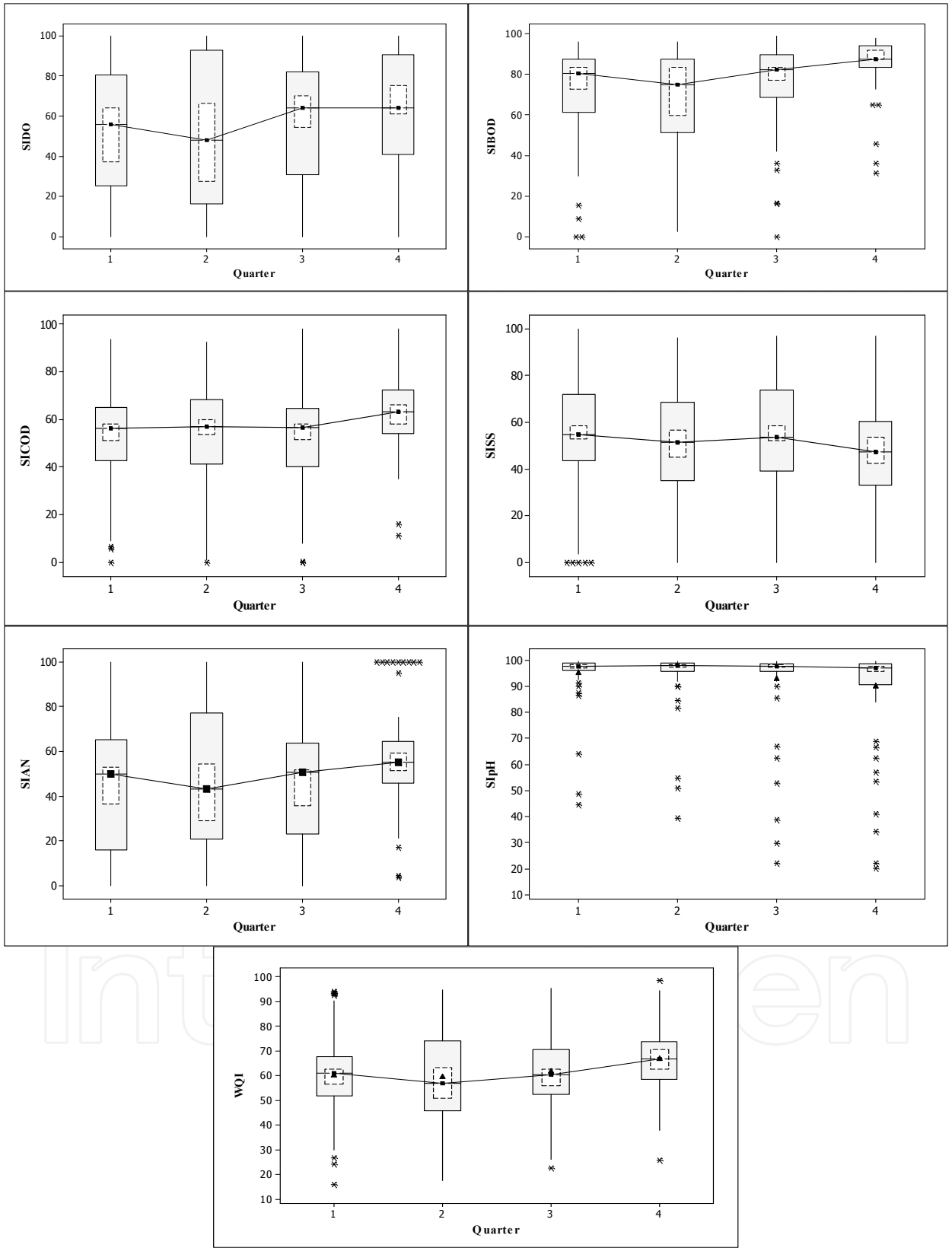


Fig. 9. Box plots of measured values of sub indices water quality parameters and water quality index in Langat River 1995-2006

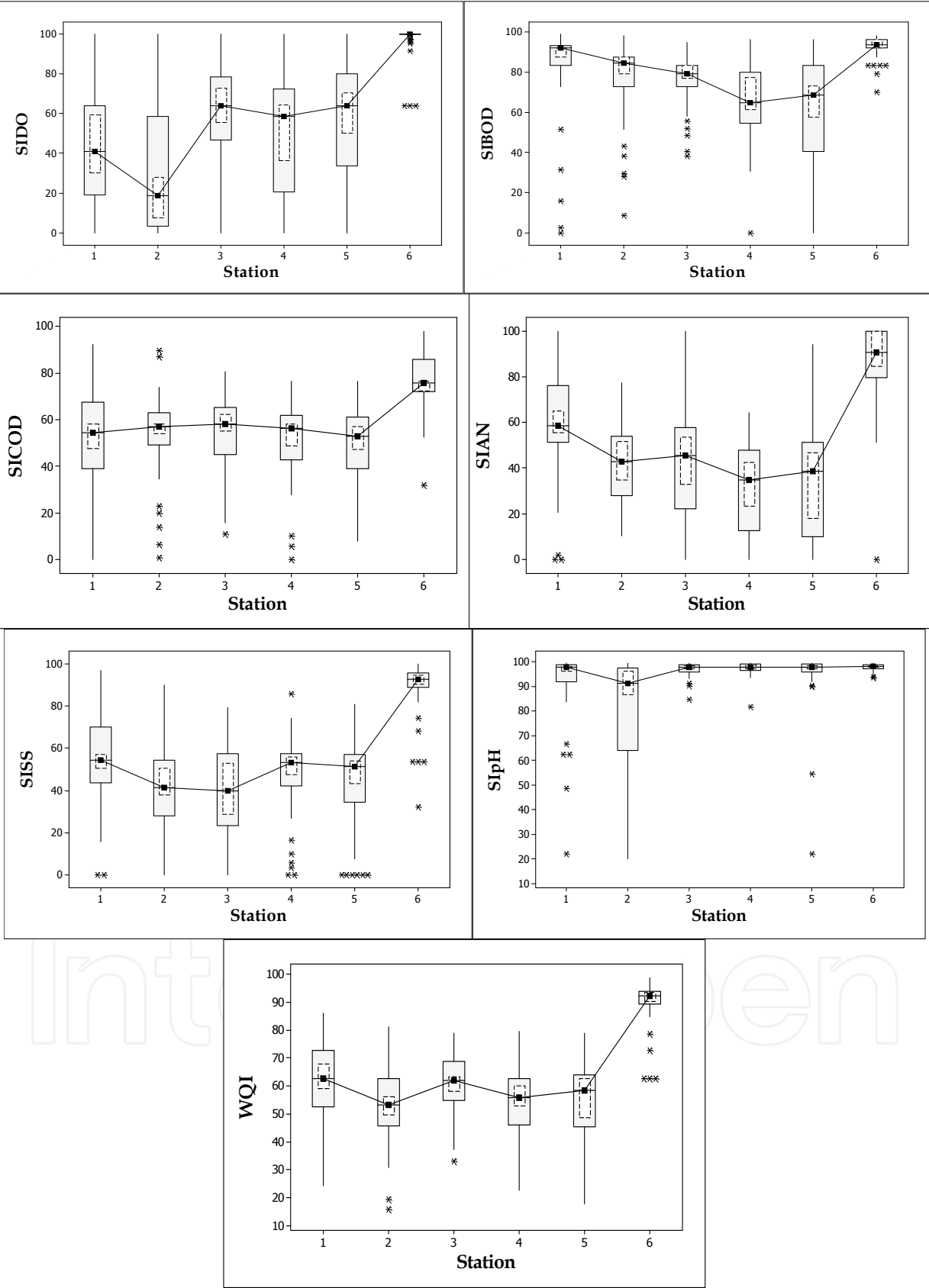


Fig. 10. Box plots of measured values of sub indices water quality parameters and water quality index at 6 main stations of the Langat River.

5.4 Normality test

Normality tests were applied to selected variables in Table 6 at particular sampling stations. From Fig. 11- Fig. 15, normal probability plots showed that SIDO at Station 2, 4 and 5, SIBOD,SICOD,SIAN at Station 5, SISS at Station 2, SICOD at Station 1 and SIAN at Station 4 were not normally distributed based on the Anderson-Darling test. Table 6 shows the list of the parameters with the Anderson-Darling test statistics and the corresponding *p*-values in the bracket.

Variables	Station 1	Station 2	Station 3	Station 4	Station 5	Station 6
SIDO	0.456 (0.256)	2.816 (<0.005)	0.521 (0.175)	0.887 (0.021)	1.117 (0.006)	
SIBOD					1.190 (<0.005)	
SICOD	0.840 (0.028)				1.119 (0.006)	
SISS		0.839 (0.028)	0.576 (0.127)			
SIAN		0.745 (0.049)	0.540 (0.157)	1.264 (<0.005)	1.156 (<0.005)	
SIpH						
WQI	0.518 (0.179)			0.289 (0.600)	0.745 (0.049)	

Table 6. Parameters selected with normality test results

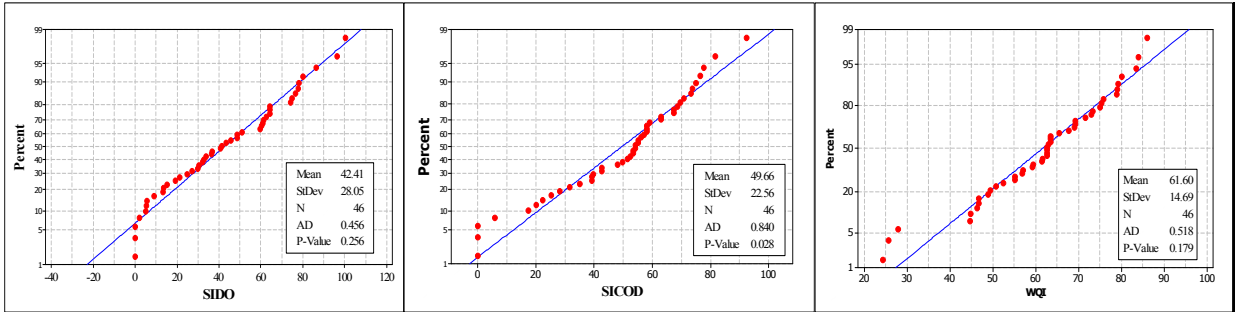


Fig. 11. Normal Probability Plot of SIDO, SICOD and WQI at Station 1

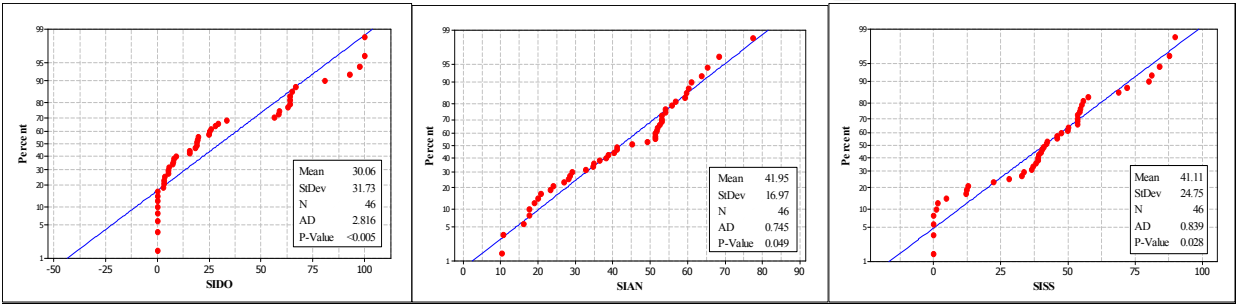


Fig. 12. Normal Probability Plot of SIDO, SIAN and SISS at Station 2

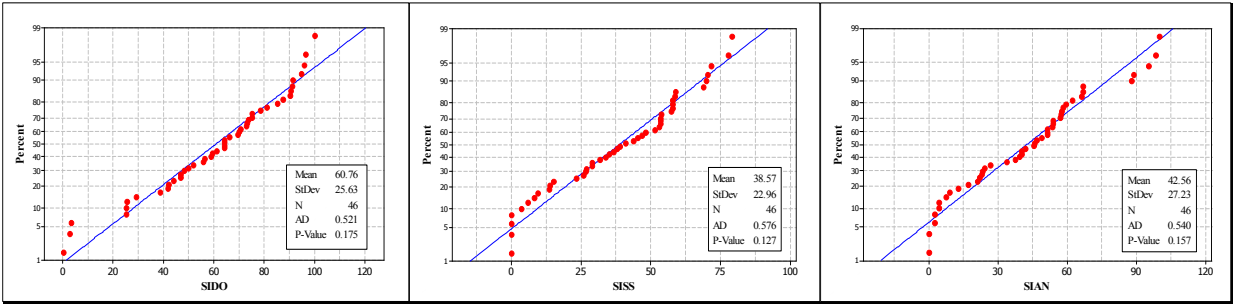


Fig. 13. Normal Probability Plot of SDO, SISS and SIAN at Station 3

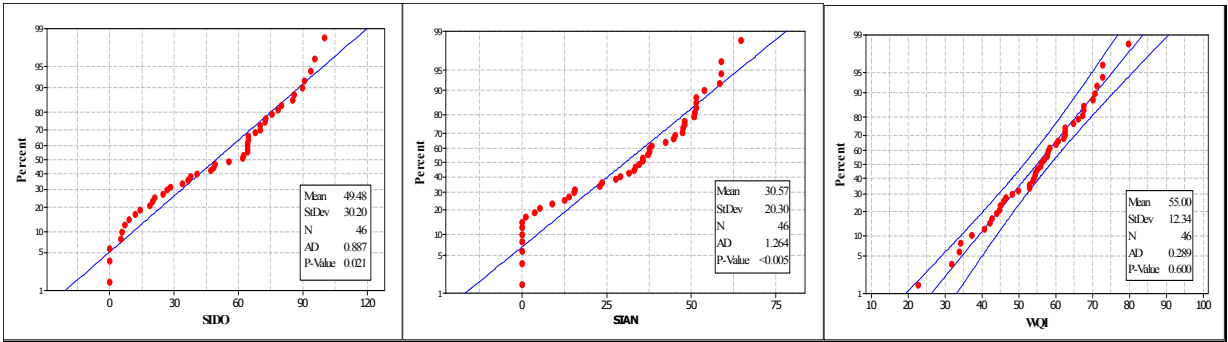


Fig. 14. Normal Probability Plot of SDO, SIAN and WQI at Station 4

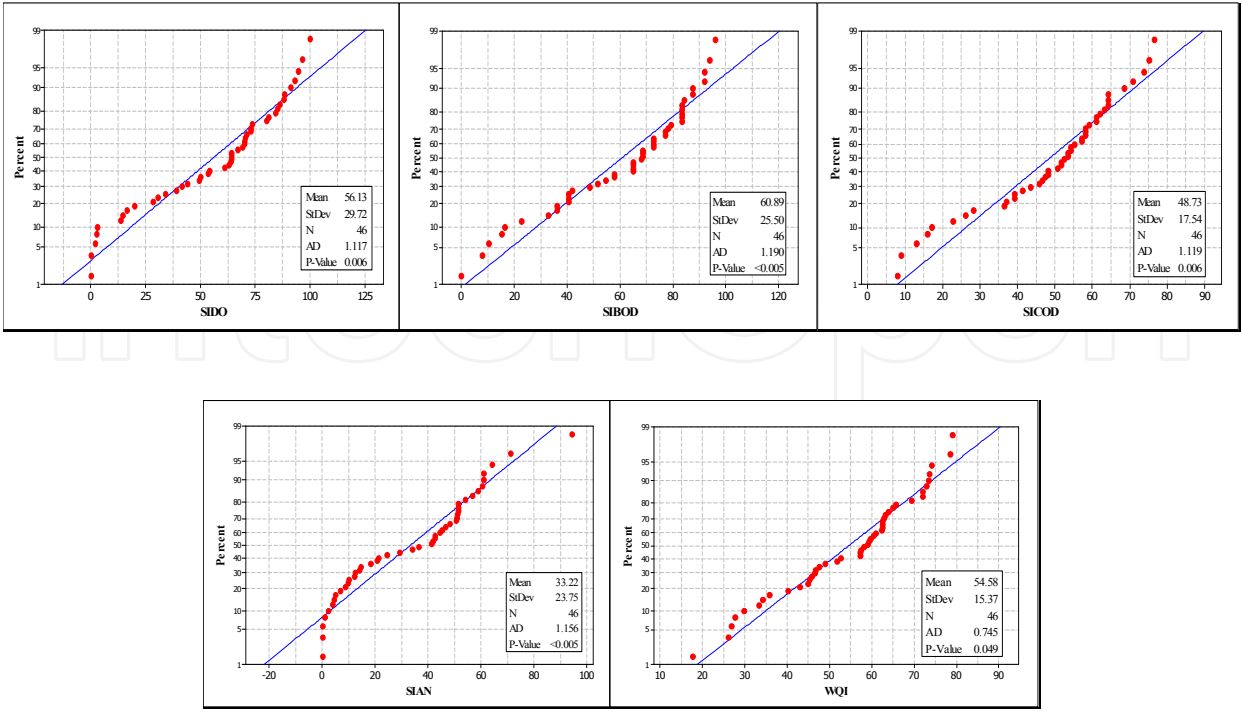


Fig. 15. Normal Probability Plot of SDO, SIBOD, SICOD, SIAN and WQI at Station 5

5.5 Trend and seasonal analyses

Trend and seasonal analyses using scatter plot were performed for selected stations in Fig. 16 i.e. SIDO at Station 1 and 3, SIAN at Station 2 and 3, SISS at Station 3 and WQI at Station 1, 4 and 5. Scatter plots in Fig.16 showed positive trend patterns in SIDO and WQI as well as quarter-to-quarter variations; hence, the model should include both trend and seasonal variations. The magnitudes of the seasonal variations were fairly constant around the level of the series, so an additive model is appropriate. To check on the significance of trend and seasonality, regression analysis and decomposition analysis were carried out. The trend-only models in regression analysis exclude the seasonal variation for both variables. Notice that the R^2 values for the trend-only models in Table 7 are between 16.70 and 38.20. To improve the forecast accuracy, seasonal variations for both variables were taken into account. Four indicator variables which represent quarter 1 to quarter 4 were used to model the seasonal variations and to test whether the seasonality were statistically significant. In Minitab, the last indicator variable is removed because it is highly correlated with the first three indicator variables for both SIDO and WQI. The results in Table 8 show that the R^2 values for the trend-and-seasonal models were increased between 39.70 and 50.30. Even though the increase was not substantial, but it is acceptable for quarterly time series data.

Variables	Station	Regression Analysis			
		Trend Model	$R^2(\%)$	MSD	p -value
SIDO	1	$SIDO = 13.6 + 1.23 * time$ (0.005) (0.000)	34.50	527.00	0.000
	3	$SIDO = 33.03 + 1.18 * time$ (0.000) (0.000)	38.20	415.00	0.000
WQI	1	$WQI = 46.09 + 0.66 * time$ (0.000) (0.000)	36.40	140.40	0.000
	4	$WQI = 46.18 + 0.38 * time$ (0.000) (0.005)	16.70	129.80	0.000
	5	$WQI = 39.96 + 0.62 * time$ (0.000) (0.000)	29.50	170.30	0.000

Table 7. Results of regression trend-only models analysis for SIDO and WQI in selected stations

Result in Table 8 also shows that only the coefficients in WQI model at Station 5 were statistically different from zero for all variables (corresponding p -values in the bracket). Hence, the additive seasonal variations of each quarter can be interpreted relative to Quarter 4. From the model, the WQI in quarter 1 is, on average, 9.71 units lower than WQI in quarter 4. Similarly, quarter 2 and 3 are 18.8 and 13.7 units lower than quarter 4 respectively.

Variables	Station	Regression Analysis			
		Seasonal Model	R ² (%)	MSD	p-value
SIDO	1	$SIDO = 6.73 + 1.23 * time - 4.32 * Q1 + 11.5 * Q2 + 19.4 * Q3$ (0.433) (0.000) (0.635) (0.209) (0.034)	45.90	466.80	0.000
	3	$SIDO = 38.27 + 1.17 * time - 8.26 * Q1 - 6.71 * Q2 - 5.86 * Q3$ (0.000) (0.000) (0.348) (0.445) (0.495)	39.70	434.50	0.000
WQI	1	$WQI = 43.2 + 0.66 * time - 2.06 * Q1 + 5.25 * Q2 + 8.23 * Q3$ (0.000) (0.000) (0.669) (0.280) (0.087)	44.40	131.80	0.000
	4	$WQI = 54.41 + 0.38 * time - 5.98 * Q1 - 17.0 * Q2 - 10.4 * Q3$ (0.000) (0.001) (0.151) (0.000) (0.013)	42.50	96.04	0.000
	5	$WQI = 50.4 + 0.618 * time - 9.71 * Q1 - 18.8 * Q2 - 13.7 * Q3$ (0.000) (0.000) (0.047) (0.000) (0.005)	50.30	128.80	0.000

Table 8. Results of regression trend-and-seasonal models analysis for SIDO and WQI in selected stations

On the other hand, Fig. 17 showed the results from Minitab time series decomposition analysis of WQI showing the original data (labelled ‘actual’) along with the fitted line (“Trend”) and the predicted values (“Fits”) from the additive model which include both the trend and seasonal components. Details of the seasonal analysis were shown in Fig. 18 and Table 9. Estimates of the quarterly variation from the trend line for each season (seasonal indices) are shown in Fig. 18a with box plots of the actual differences shown in Fig. 18b. The percentage of variation by seasonal period was illustrated in Fig. 18c and model residuals by seasonal period in Fig. 18d. Unfortunately, many decomposition methods do not perform significance tests on seasonal indexes. The significance tests on seasonal indexes are important to believe that the seasonal variations exist. The indexes can be tested to confirm that they are statistically different from zero. Since decomposition methods do not perform such tests, the easiest way to test their significance is to create indicator variables as performed in the previous regression analysis. Additional details of the component analysis are shown in Fig. 19. Fig. 19a is the original time series. Fig. 19b is the plot of the time series with the trend removed. Fig. 19c is a plot of the time series with the seasonality removed (should see a trend pattern) and Fig. 19d is a residual plot of the detrended and seasonally adjusted data. The wave-like pattern in Fig. 19d suggests the constant variance over time in WQI at Station 5.

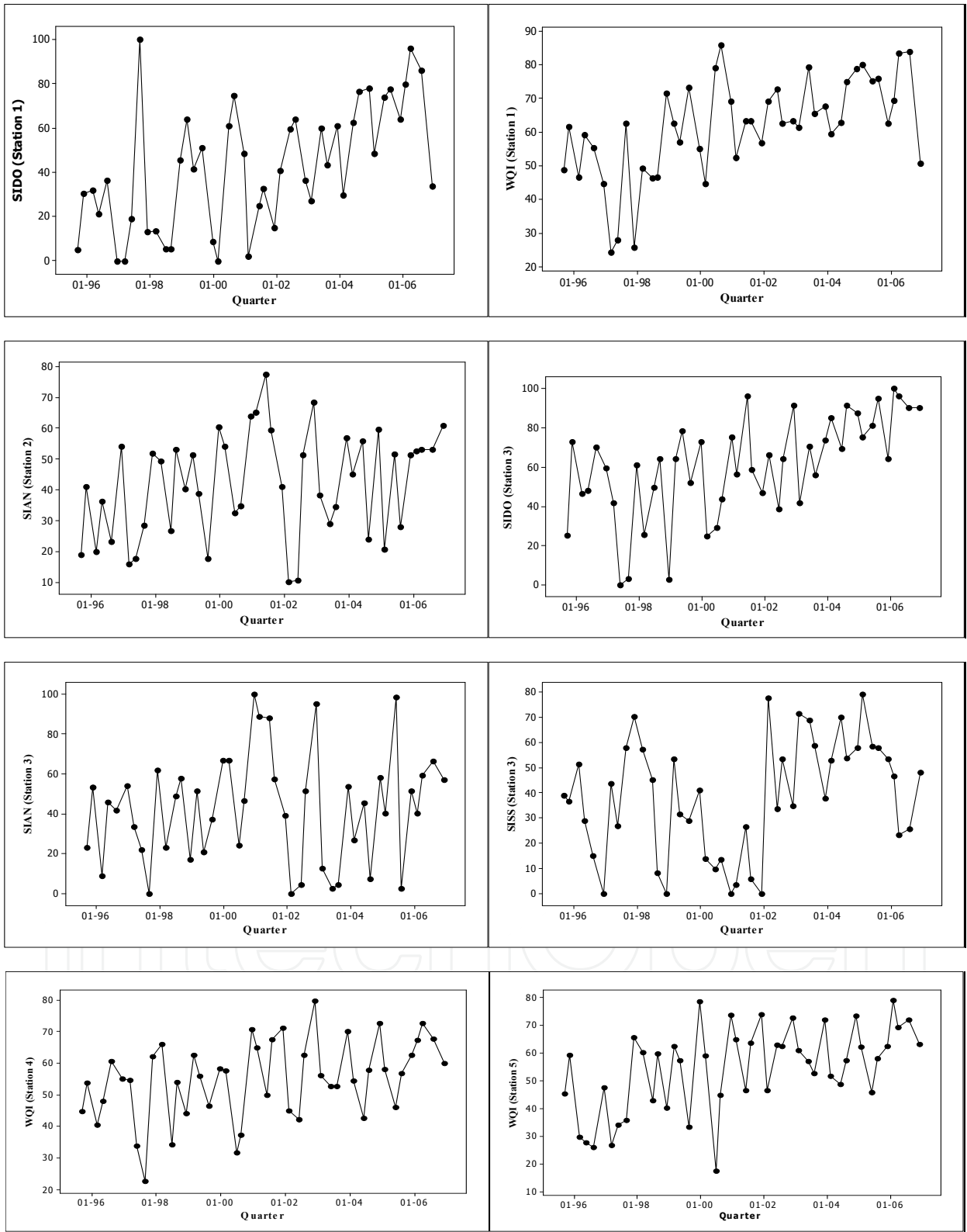


Fig. 16. Scatter plot of selected water quality variables in certain stations (in bracket)

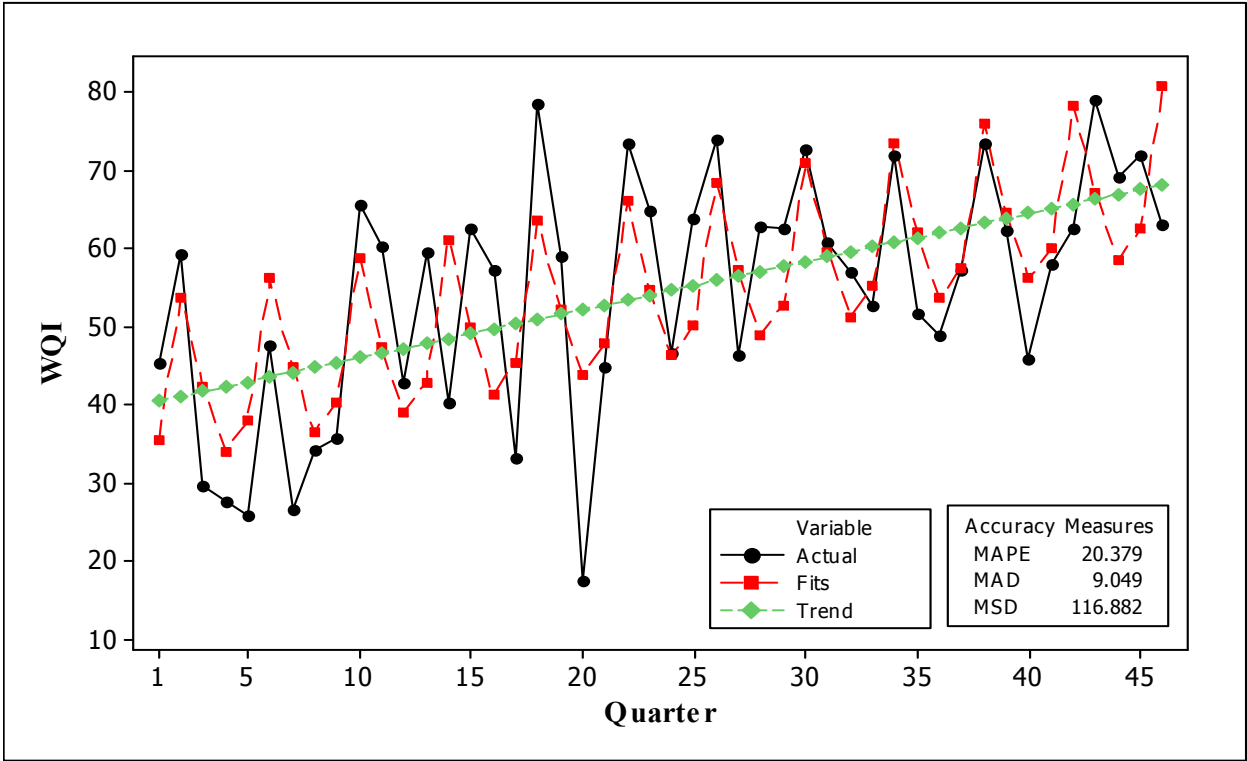


Fig. 17. Additive Time Series Decomposition Plot for WQI in Station 5

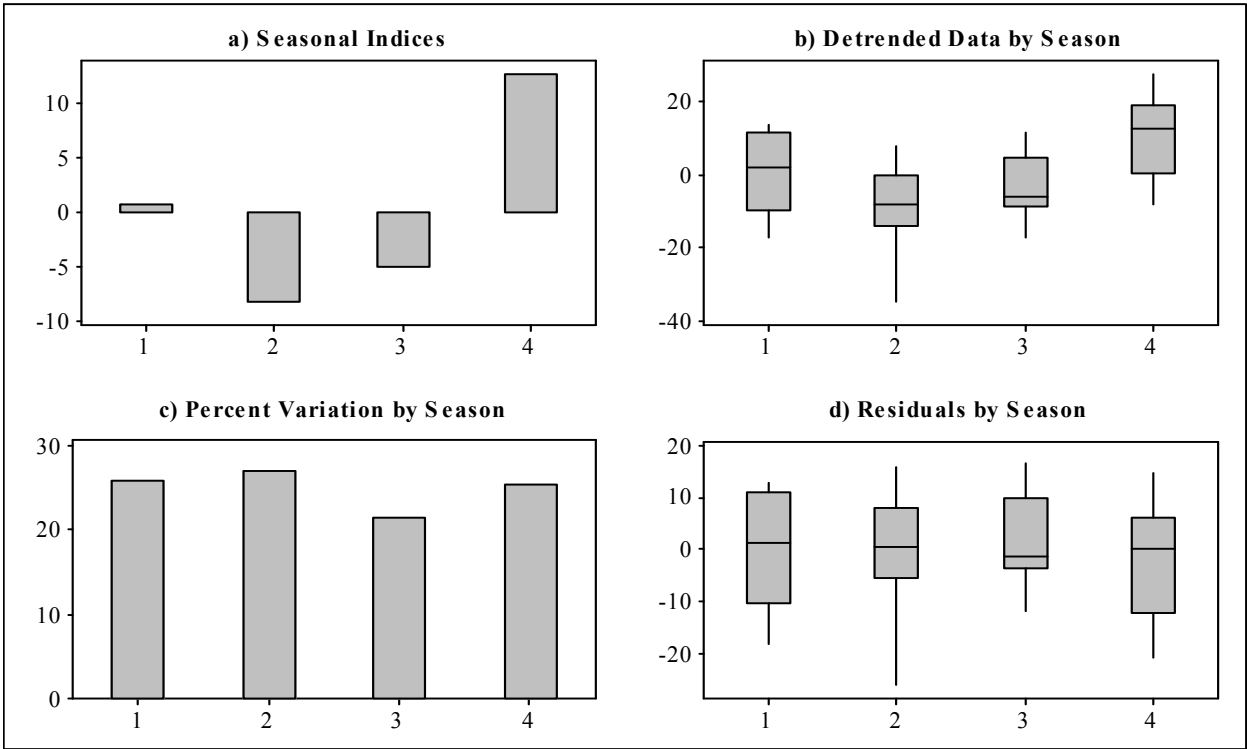


Fig. 18. Seasonal Analysis for WQI at Station 5

Variables	Station	Trend Model	Additive decomposition Models						
			Seasonal Indices				MAPE (%)	MAD	MSD
			Q1	Q2	Q3	Q4			
SIDO	1	$SIDO = 13.28 + 1.23 * time$	-14.30	5.84	16.98	-8.52	104.98	15.69	424.58
	3	$SIDO = 33.13 + 1.17 * time$	-7.00	-0.48	-0.50	7.97	96.93	15.64	393.25
WQI	1	$WQI = 46.08 + 0.658 * time$	-5.44	3.03	3.75	-1.35	16.56	8.20	118.90
	4	$WQI = 45.96 + 0.373 * time$	0.14	-12.09	-0.83	12.79	15.82	7.71	95.12
	5	$WQI = 39.96 + 0.615 * time$	0.70	-8.30	-5.03	12.63	20.38	9.05	116.88

Table 9. Results of additive decomposition model for SIDO and WQI in selected stations

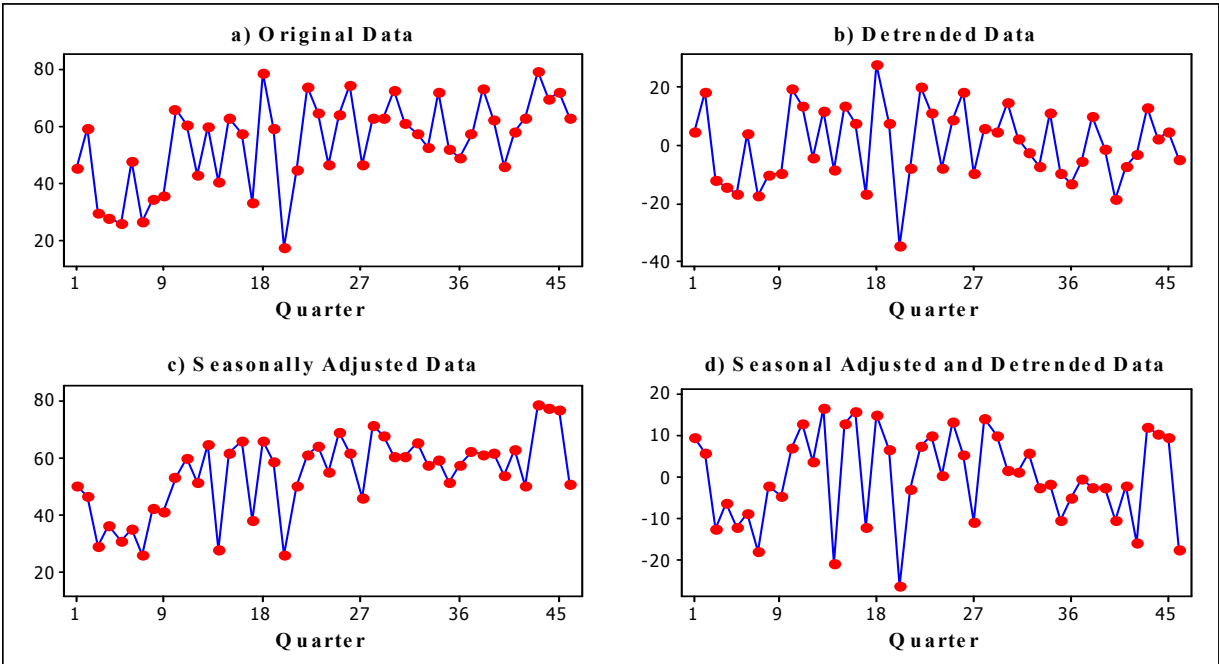


Fig. 19. Component Analysis of WQI in Station 5

5.6 Forecasting model evaluation

The measure of forecast accuracy were evaluated as part of model validation effort (Montgomery et al 2008). To evaluate the resulting model in Section 5.4, the accuracy measures of the model were determined as summarized in Table 9. There were three

measures reported in this study, Mean Average Percentage Error (MAPE), Mean Average Median (MAD) and Mean Square Error (MSE). Focusing on the WQI model at Station 5, the accuracy measures were 20% for MAPE, 9.05 for MAD and 116.88 for MSE. Obviously, small variability in forecast errors is preferred in all forecasting models, but a larger forecast error (i.e. residual) or a relatively small one is very subjective. Therefore, the results of MAPE, MAD and MSE in this case can be reasonably accepted. Further, normality test on the distribution of forecast error was examined. From Fig. 20, the p -value is 0.211, so the hypothesis of normality for the forecast error could not be rejected at the 0.05 level. The forecasts that could adequately model all the structure in the data and the sequence of forecast errors would have no systematic or nonrandom pattern (Montgomery et. al 2008) as shown in Fig. 21. From Fig. 22, the sample autocorrelation function (ACF) shows that all spikes in the sample ACF at lower lags are inside the confidence interval limits. This suggest that there is no pattern in the forecast errors. Therefore, there is strong evidence to support the claim that the residuals are not correlated. Since quarterly data were used, the forecast values in Table 10 with MSD value is considered reasonable for this model.

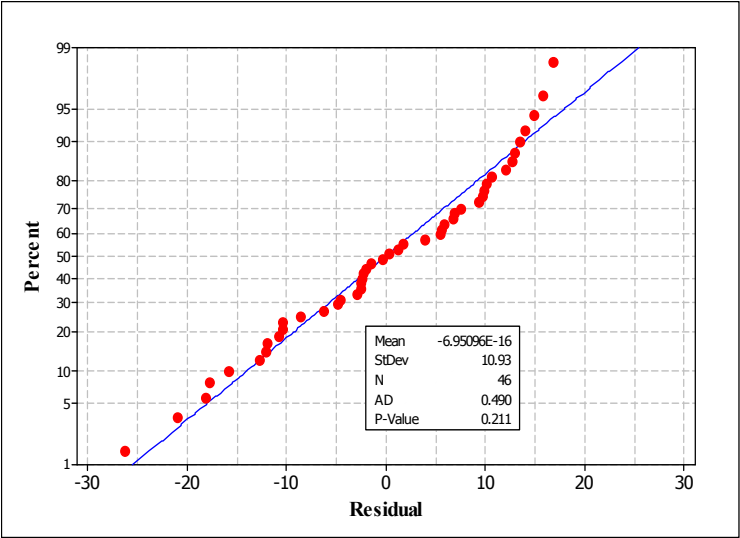


Fig. 20. Normal probability plot of residuals for additive model of WQI in Station 5

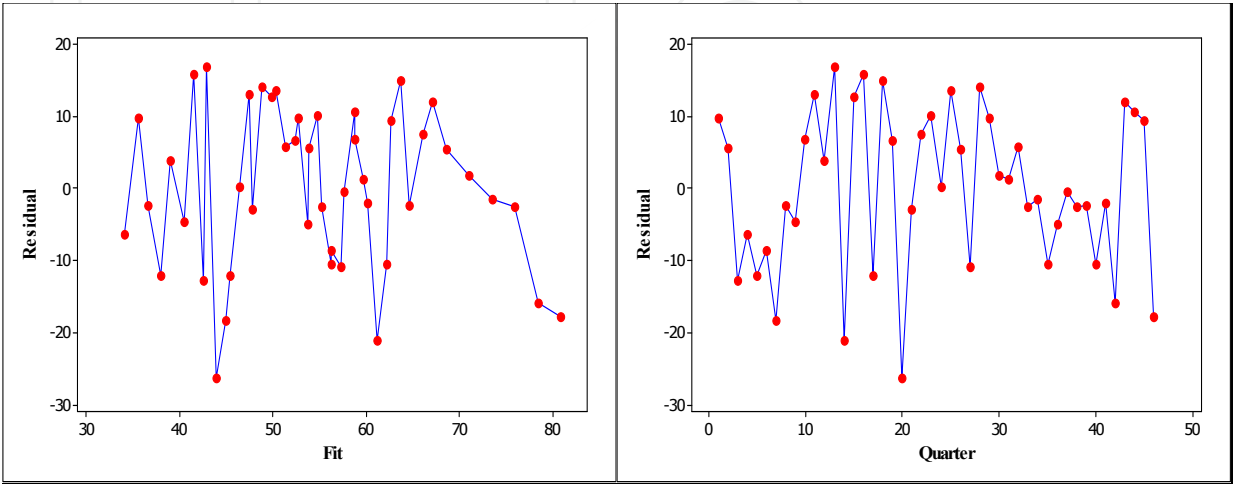


Fig. 21. Scater plot of WQI residual versus WQI fitted value and quarter

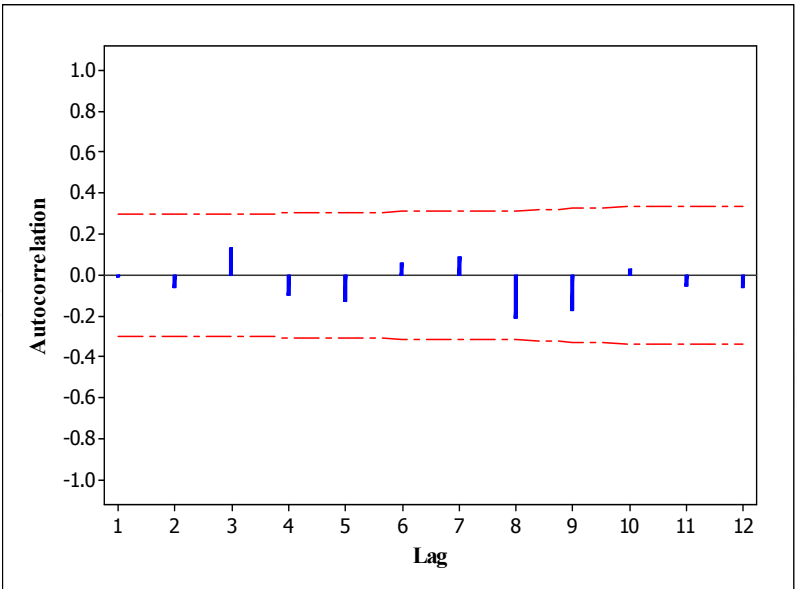


Fig. 22. Sample ACF of forecast error for WQI in Station 5

Variables	Station	Quarter	Forecast	2007	MSD
SIDO	1	1	78.25	61.50	597.91169
		2	90.62	77.50	
		3	66.35	76.86	
		4	61.80	67.68	
	3	1	88.77	91.75	609.25021
		2	89.92	76.99	
		3	99.55	80.96	
		4	85.75	95.11	
WQI	1	1	80.71	55.66	682.64048
		2	82.08	75.47	
		3	77.64	74.89	
		4	74.21	76.28	
	4	1	51.79	65.04	428.17771
		2	63.43	75.64	
		3	77.42	67.66	
		4	65.15	68.07	
	5	1	61.18	63.64	408.44976
		2	65.07	72.47	
		3	83.34	64.73	
		4	72.03	70.96	

Table 10. Forecast value for the selected water quality variables in 2007

6. Conclusion

In this study, various techniques were utilized to evaluate temporal variations in the surface water quality of Langat River. From box plot and median analysis, the annual changes of SIDO and WQI showed an oscillating trend in annual variation with noticeable increment. The annual variation for parameters SICOD, SIBOD, SIAN and SISS showed very little or no trend. There is no trend exhibited by SIpH and its quarterly values do not differ much. The quarterly analyses were examined and the results showed that the median of SIDO, SIBOD, SIAN, SIpH and WQI were lower in quarter 2 than the median of SISS in quarter 4. Significant trends in water quality were found in SIDO at Station 1 and 3 and WQI at Station 1, 4 and 5. However, the effects of quarters appear to be prominent only in WQI at Station 5. WQI is the most significant variable contributing to water quality variations for all quarters. Therefore, further analysis should be carry out to study the relationship between the location of the station (i.e. Station 5) and, sampling measurement time with all the variables that strongly influencing WQI such as urbanization, population density, water shortages and pollution (Cheng et al 2003). In addition, Sapari et al (2009) mentioned that urban NPS pollution has become a growing concern for most major towns in Malaysia due to the serious threat of pollution to river water quality in urban environment. Since this study focus only on the sub indices and index, it would be additionally informative if the same analyses are repeated to all available variables with longer data sets (i.e. monthly data in 1995-2006). Longer data sets could provide clearer indication of trend and seasonal pattern inherent in the time series data. Imputation method should also be considered to overcome the problem of unequal spacing in the measurements.

7. Acknowledgment

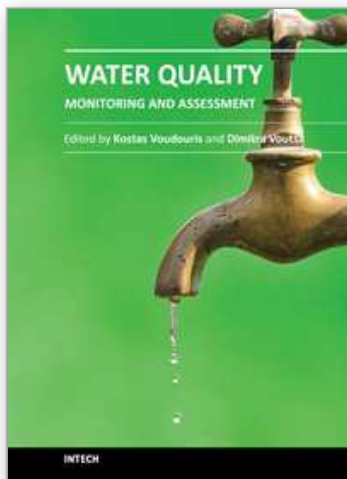
The authors would like to thank Malaysian Department of Environment for supplying the data on which this work was based.

8. References

- Amini, A., Mohammad Ali, T., Ghazali, A.H. & Kim Huat, B. (2009). Adjustment Of Peak Streamflows Of A Tropical River For Urbanization. *American Journal Of Environmental Sciences*. 5 (3), 285-294.
- Bowerman, B.L, O'Connell, R.T. & Koehler, A.B. (2005). *Forecasting, Time Series, And Regression*, ISBN 0-534-40877-6, Thomson Brooks/Cole, U.S.
- Cheng, S. Chan, C.W. & Huang, G.H. 2003, An Intergrated Multi-Criteria Decision Analysis And Inexact Mixed Integer Linear Programming Approach For Solid Waste Management, *Engineering Applications Of Artificial Intelligence*, 543-554.
- Cooke, S.E., S.M. Ahmed & N.D. Macalphine - Revised. (2005). *Introductory Guide To Surface Water Quality Monitoring In Agriculture*. Conservation And Development Branch, Alberta Agriculture, Food And Rural Development. Edmonton, Alberta.
- Department Of Environment Malaysia - DOE (1997). *Malaysia Environmental Quality Reports*. Kuala Lumpur: Ministry Of Science, Technology And Environment.
- Department Of Environment Malaysia - DOE (2007). *Malaysia Environmental Quality Reports*. Kuala Lumpur: Ministry Of Science, Technology And Environment.

- Department Of Irrigation And Drainage - DID (2009). *Study On The River Water Quality Trends And Indexes In Peninsular Malaysia*. Putrajaya : Ministry Of Natural Resources And Environment.
- Idrus, S., Shah, A.H.H. And Mohamed, A.F. (2003). Analysis Of Land Use And Land Cover Changes 1974-2001 In The Langat Basin, Malaysia Using Geographic Information System (GIS). In Mokhtar, M.B., Idrus, S. And Aziz, S. (Eds). *Ecosystem Health Of The Langat Basin, Proceedings Of The 2003 Research Symposium On Ecosystem Of The Langat Basin*, 209-225. LESTARI, UKM, Bangi.
- Jaafar, O., S.A. Sharifah Mastura & Mohd Sood, A. (2009). Land Use and Deforestation Modelling of River Catchments in Klang Valley, Malaysia. *Sains Malaysiana*. 38(5), 655-644.
- Juahir, H., M. Zain, S., Yusoff, M.K., Tengku Hanidza, T. I., Mohd Armi, A. S. Toriman, M.K & Mokhtar, M. (2010a). Spatial Water Quality Assessment Of Langat River Basin (Malaysia) Using Environmetric Techniques. *Environmental Monitoring And Assessment*. Vol. 173, No. 1-4, 625-641.
- Juahir, H., M. Zain, S., Aris, A.Z., Yusof, M.K., Abu Samah, M.A. & Mokhtar, M. (2010b). Hydrological Trend Analysis Due To Land Use Changes At Langat River Basin. *EnvironmentAsia 3* (Special Issue), 20-31.
- Khairuddin M. I, Abd Malek A. 2002. Program Pencegahan Pencemaran Dan Peningkatan Kualiti Air Sungai Langat. Proceeding Simposium Penyelidikan Lembangan Langat 2001. Mazlin M., Shaharuddin I., Ahmad Fariz M. Abdul Hadi H. S. And Sarah A.A.G.A. (Eds). Lestari, UKM, Bangi. 183-189. In Lee, Y.H., Abdullah, M.P., Chai, S.Y., Mokhtar, M. & Ahmad, R. (2006). Development Of Possible Indicators For Sewage Pollution For The Assessment Of Langat River Ecosystem Health. *Malaysia Journal Of Analytical Sciences*. Vo. 10. No. 1. 15-26.
- Lee, Y.H., Abdullah, M.P., Chai, S.Y., Mokhtar, M. & Ahmad, R. (2006). Development Of Possible Indicators For Sewage Pollution For The Assessment Of Langat River Ecosystem Health. *Malaysia Journal Of Analytical Sciences*. Vo. 10. No. 1. 15-26.
- Mokhtar, M., Toriman, M.E. Abraham Hossain, M.A. Kok, W.T. (2011) Institutional Challenges For Integrated River Basin Management In Langat River Basin, Malaysia. *Water And Environment Journal*. 25. 495-503
- Montgomery, D.C., Jennings, C.L. & Kulahci, M. (2008). *Introduction To Time Series Analysis And Forecasting*. Wiley Interscience. ISBN 978-0-471-65397-4.
- Sapari, Nasiman And Adlan, M.N. (2009). *Non-Point Sources (NPS) Pollution Modelling And Water Quality Improvement Using Wetland*. In: International Engineering Convention, Danascus, Syria.
- Ravichandran, S. (2003). Hydrological Influences On The Water Quality Trends In Tamiraparani Basin, South India. *Environmental Monitoring And Assessment*. Vol. 87. (October 2002). 293-309.
- Mcgill, R., Tukey, J.W., & Larsen, W.A. (1978). "Variations Of Box Plots". *The American Statistician* Vol. 32, No. 1 (February 1978), 12-16.
- Sarkar, C. Abbasi, S. A. (2006). QUALIDEX – A New Software For Generating Water Quality Indices. *Environmental Monitoring And Assessment*. Vol. 119, No. 1-3 (August 2006), 201-231.
- Carlie, S.A. 2010. Assessment Of Water Quality Characteristics During Base And Storm Flow Events On Sungai Langat Basin. Master Thesis. Universiti Teknologi Malaysia.

- Suki, A., Yusoff, M.K & Poe, M.K. 1988. Water Quality Profile Of Sg. Langat. *Pertanika*. Vol. 11, No. 2, 273-281.
- Tukey, J.W. 1977. *Exploratory Data Analysis*. Addison-Wesley. ISBN 0-201-07616-0.
- Universiti Malaya Consultancy Unit (UPUM) (2002). *Final Report Program Pencegahan Dan Peningkatan Kualiti Air Sungai Langat*. Kuala Lumpur.
- Vega, M., Pardo, R., Barrado, E. & Deban, L. (1998). Assessment Of Seasonal And Polluting Effects On The Quality Of River Water By Exploratory Data Analysis. *Water Research*. Vol. 32. No. 12, 3581-3592.



Water Quality Monitoring and Assessment

Edited by Dr. Voudouris

ISBN 978-953-51-0486-5

Hard cover, 602 pages

Publisher InTech

Published online 05, April, 2012

Published in print edition April, 2012

The book attempts to covers the main fields of water quality issues presenting case studies in various countries concerning the physicochemical characteristics of surface and groundwaters and possible pollution sources as well as methods and tools for the evaluation of water quality status. This book is divided into two sections: Statistical Analysis of Water Quality Data;Water Quality Monitoring Studies.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Zalina Mohd Ali, Noor Akma Ibrahim, Kerrie Mengersen, Mahendran Shitan, Hafizan Juahir and Faridatul Azna Ahmad Shahabuddin (2012). Temporal Water Quality Assessment of Langat River from 1995-2006, Water Quality Monitoring and Assessment, Dr. Voudouris (Ed.), ISBN: 978-953-51-0486-5, InTech, Available from: <http://www.intechopen.com/books/water-quality-monitoring-and-assessment/temporal-water-quality-assessment-of-langat-river-from-1995-2006>

INTECH
open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](https://creativecommons.org/licenses/by/3.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

IntechOpen

IntechOpen