

# Estimating Generalised Flood Sew Coefficients for Peninsular Malaysia

(Date received: 4/9/2015/Date accepted: 9/2/2016)

David Hong Jer Lang<sup>1</sup>, Amirah Hanim bt. Mohd Puad<sup>2</sup>, Intan Shafillah bt. Abdul Azia<sup>3</sup> and Hong Kee An<sup>3</sup>

<sup>1</sup>School of Computing and Information Technology, Taylor's University, Malaysia.

<sup>2</sup>ZHL Engineers Sdn. Bhd.

<sup>3</sup>Hong & Associates

## ABSTRACT

Flood estimations based on fitting the frequency of occurrence of annual peak discharges using the Log-Pearson Type 3 distribution are commonly used but they are sensitive to the skew coefficients of the gauging stations. The estimation accuracy can be improved by using a weighted average population skew coefficient calculated from the sample station skew and the generalised unbiased skew. The U.S. Water Resources Council (WRC) has documented guidelines for estimating the generalised skew coefficients and published a map of generalised skew values for the United States. The map shows isolines of skew coefficient values and the average skew coefficient for each 1-degree quadrangle of latitude and longitude for the United States. Following the WRC guidelines, many of the state authorities in the US have developed the generalised skew coefficients separately on a state/regional basis. In Malaysia, the Log Pearson Type 3 distribution has been widely used for flood peak estimation but there are no guidelines available for estimating the generalised skew coefficients for use in conjunction with the distribution and as more accurate results can be obtained if these data are available, it is clear that a regional flood skew study is needed. With the regional skew data available, the peak flow can be simply and easily calculated with the aid of a software such as HEC-SSP. The aim of this paper is to use the WRC guidelines to derive the generalised skew coefficients using the peak annual discharge data of Peninsular Malaysia for general use.

The WRC recommended several techniques for estimating and evaluating generalised skew of the Log-Pearson Type 3 distribution for the annual peak discharges. Station skews (skew coefficients computed from gauging station records) and unbiased and weighted skews derived from these station skews are to be used to develop these techniques. In this study, peak discharge records at 66 gauging stations having 16 or more annual peak discharges in Peninsular Malaysia were selected for computing station skews. Station skew values ranged from  $-0.831(\log_{10} \text{ unit})$  to  $1.475(\log_{10} \text{ unit})$ .

The three techniques recommended by WRC used for estimating the generalised skew of annual peak discharges were adopted for this study. These methods are: (1) An isoline map, (2) a prediction equation (3) a regional mean skew. Attempts to develop a prediction equation were unsuccessful. An error analysis showed that the regional mean skew method has a lower MSE (mean square error) than that obtained from the state wide generalised skew coefficient contour map. As a result, the mean station skew for the selected gauging stations can be used to estimate the generalised skew for any gauging site in the peninsula. The mean skew is  $-0.022(\log_{10} \text{ unit})$  and the associated mean square error is  $0.05(\log_{10} \text{ unit})$ .

## 1.0 INTRODUCTION

Accurate estimates of the magnitude and frequency of floods are often required in the planning and design of water resources projects and flood managements as well as the cost effective design of structures like dams and levees on and along rivers and streams. The Interagency Advisory Committee on Water Data (1982) of U.S.A. recommends using the Log-Pearson Type 3 probability distribution to define the frequency of occurrence of annual peak discharges. The Log-Pearson Type 3 distribution is defined by the mean, standard deviation and the skew coefficient of the logarithms of the peak discharges. Estimates of the peak discharges for low probabilities using the distribution are sensitive to the skew coefficient.

Estimates of the skew coefficient for a single gauging station are biased and subject to large sampling errors, especially when computed from short periods of streamflow records. The accuracy of the station skew can be improved by weighting the station skew with a generalized skew value that represents

pooled skew coefficient data from nearby stations with long records. The U.S. Water Resources Council (WRC) has published a map of generalized skew values for the United States (U.S. Water Resources Council, 1976). The map shows isolines of skew coefficient values and the average skew coefficient for each 1-degree quadrangle of latitude and longitude. Following the guidelines of WRC, several states in the U.S. have developed the generalized skew coefficients separately on a regional /state level. Typical examples are: Judd *et al.*, (1996) for Texas, Lumia *et al.*, (2000) for New York, Reuben (1984) for Hawaii, Croskey *et al.*, (1983) for Michigan, Lorenz (1997) for Minnesota, Oberg *et al.*, (1987) for Illinois.

This study adopts the procedures of the U.S. Water Resources Council to derive the generalized skew coefficients for Peninsula Malaysia. The recommended procedures require the use of 40 stations or all stations within a 100 mile (160 kilometer) radius. The stations used should have 25 or more years of record. It is recognized that in some locations a relaxation of these criteria may be necessary.

Station skews used to develop the generalised skew map were calculated using procedures recommended by the U.S. Water Resources Council. In constructing the skew contour map, the skew coefficients were plotted at the gauging sites instead of the centroids of the catchments. The gauge location was used because data are available at the gauging station, and the gauge location has been used by the U.S. Water Resources Council for determining the generalised skew coefficient from the WRC (1982) map. The HEC-SSP (2009) program was used for computing the station skew coefficients. Low outliers were rejected and excluded for further calculation but high outliers were retained unless there was strong hydrological and statistical evidence for their rejection.

Results indicate that our findings are consistent with findings of other studies, which were carried out mainly in the U.S.A. For example, the generalised skew and mean square error obtained from our study are comparable to the values derived from basins of Texas, Michigan, New York and Hawaii. The table shown below are our results compared to those of U.S. basins.

*Results of generalised skew coefficient from various studies.*

State/Country	Generalised Skew Coefficient	MSE	Method
Peninsular Malaysia	-0.022	0.05	Regional
Peninsular Malaysia	-	0.19	Isoline
New York	-0.2 to 0.7	0.16	Isoline
Texas	-0.2 to 0.3	0.35	Isoline
Hawaii	-0.14	-	Regional
Michigan (UPPER PART)	0.12	0.2	Regional
Michigan (southwest)	0.081	-	Isoline
Michigan (lower)	-0.017	-	Isoline
Minnesota	-0.5 to 0.2	0.182	Isoline
Illinois	-0.16	0.2	Regression

## 2.0 PURPOSE AND SCOPE

This paper describes the results of a study to determine an accurate technique for estimating generalised flood skew in Peninsular Malaysia.

Station skews were computed using guidelines of Bulletin 17B (WRC 1982). In addition, unbiased skew is calculated using a biased correction factor suggested by Tasker and Stedinger (1986).

Previous investigators have developed equations to estimate the variance of station skew for weighting the computed skew. Tasker and Stedinger (1986) used the inverse variance of the station skew to give a weight to the unbiased skew.

Some researchers prefer to apply a weighting factor for correcting the skew coefficient based on the station record length because the number of years of data varied greatly between stations. The weighting factor in this context is the number of years of annual peak discharge record at each station, divided by the average number of recorded annual peak discharges at all stations in the region.

The techniques for estimating generalised skew evaluated are: (1) an isoline map of skew coefficients for the study area, (2) a prediction equation relating skew coefficients to catchment characteristics, (3) the mean station skew coefficient for the study area.

The techniques were applied separately to the station, unbiased, and the two weighted unbiased skew data sets.

It is more appropriate to describe methods used to select gauging stations, compute station skews, and estimate generalised skew first. Results of station skew computations and a description of the techniques selected to estimate generalised skew are presented next. The main idea presented in the paper and conclusions based on the study are summarised in the final section.

## 3.0 MATERIALS AND METHODS

### 3.1 Gauging Station Selection

Selection of gauging stations used in this study was based on several criteria. Gauging records of more than 16 years from recording stations that were not significantly affected by regulation or diversion were selected. The unregulated records for stations where streamflow has been both unregulated and regulated for certain periods of time were also used.

### 3.2 Computation of Station Skew

The Log-Pearson Type 3 frequency distribution is recommended by the Hydrology Subcommittee of the IACWD (1982) for defining the frequency of occurrence of floods in an annual flood series. The Log-Pearson Type 3 distribution is defined by three parameters—the mean, standard deviation, and the skew coefficient—which can be calculated by the method of moments. The mean,  $\bar{X}$ , is:

$$\bar{X} = \frac{\sum x}{N} \quad (1)$$

the standard deviation,  $S$ , is:

$$S = \left( \frac{\sum [x - \bar{X}]^2}{N - 1} \right)^{1/2} \quad (2)$$

And the skew coefficient  $G$  is:

$$G = \frac{N \sum (x - \bar{X})^3}{(N - 1)(N - 2)S^3} \quad (3)$$

Where  $x$  = common logarithm of annual peak discharge, and  
 $N$  = number of annual peak discharges

The skew coefficient is a measure of the asymmetry of the frequency distribution. The Log-Pearson Type 3 distribution is equivalent to a Log-Normal distribution when the skew coefficient is zero. The discharge corresponding to small exceedence probabilities (say less than 0.1) will be larger for a positively skewed distribution and smaller for a negatively skewed distribution when compared to a Log-Normal distribution.

Estimates of station skew are sensitive to extremely high and low peak discharges in the annual flood series. Methods used in this study to identify outliers are the same as those described by the Hydrology Committee of the IACWD (1982). High outliers are examined for errors and were compared with peaks at nearby sites. Suspect discharges were eliminated from the annual series and the peak discharge statistics (mean, standard deviation

and skew coefficient) were recalculated. When low outliers were detected, the peak discharge statistics were recomputed without the outlier and a conditional probability adjustment was performed (IACWD 1982). No historical events were considered in this study.

Station skew,  $G$  in Equation 3, is a biased estimate of the population skew coefficient,  $g$  (Wallis *et al.*, 1974). An unbiased estimate of the population skew coefficient,  $G_g$ , can be obtained by using an unbiased correction factor (Tasker and Stedinger 1986), i.e.

$$G_g = (1+6/N)*G \quad (4)$$

$G_g$  is hereafter referred to as unbiased skew

Several investigators have developed equations to estimate the variance of station skew coefficients, Tasker and Stedinger (1986) estimated the station variance as:

$$V_s = \frac{6N(N-1)(1+\frac{6}{N})^2}{(N-2)(N+1)(N+3)} \quad (5)$$

A station's unbiased skew is weighted in inverse proportion to the estimated station variance ( $V_s$ ), and the weight given to the unbiased station skew is:

$$W=1/V_s \quad (6)$$

This weighted unbiased skew using the above approach is termed weighted unbiased skew using station inverse variance.

Some researchers prefer to use a weighting factor to give more weight to stations with longer periods of record than to stations with shorter periods of record. This weighting factor is the number of years of annual peak discharge record at a station, divided by the average number of recorded peak discharges at all stations in the region. In this study, the unbiased skew was also corrected using this weighting factor based on station record length.

The weighted skew derived using the above method is referred to as weighted unbiased skew using station record length.

### 3.3 Estimation of Generalised Skew

Use of a generalised skew is desirable because station skew is sensitive to extreme events. Accurate estimates of station skew are also difficult from gauging stations with short periods of record lengths. Therefore, the IACWD (1982) recommends using gauging stations having 25 or more annual peak discharges to minimise time-sampling errors. The three techniques mentioned earlier for estimating generalised skew for annual peak discharge records of Peninsular Malaysia were evaluated in detail.

The accuracy of the three techniques was evaluated using a split-sampling approach (Tasker 1982). Split sampling is the reservation of a portion of available data to obtain an independent measure of the accuracy of an estimating technique. It is an effective way to compare different techniques for estimating the same variable. The entire set of gauging stations was split into two sets of approximately equal size, having similar geographic distribution and similar ranges in catchment characteristics and station skew values. One set, the estimation set, was used in the application of each technique for estimating generalised skew;

the other, the prediction set, was used to compute the MSE (mean square error) of each technique.

The MSE is a measure of the accuracy of an estimating technique. It is computed by dividing the sum of the squared differences between generalised skew,  $\bar{G}$ , and station skew  $G$ , by the number of stations in the prediction set,  $N$ , as shown in Equation 7:

$$MSE = \frac{\sum(\bar{G}-G)^2}{N} \quad (7)$$

Values for  $\bar{G}$  were determined from each of the three techniques used for estimating generalised skew.

The most accurate technique, the one with the lowest MSE can be used for estimating the generalised skew. The methods used to develop each technique for estimating generalised skew are presented in the following sections.

### 3.4 Isoline Map

The following procedure was used to develop a contour map of generalised skew based on station skews from the estimation set. Mean skews were computed for all stations within the 100-mile (160 kilometer) radius of grid points. These grid points were located at the intersection of every 12 minutes interval of latitude and longitude across the peninsula. These mean skews were plotted on the map and lines of equal mean skew constructed to obtain the contour map. To assess the accuracy of this method, the mean skew for a station in the prediction set was first calculated using the station skew of the stations in the estimation set within the 100 mile (160 kilometer) radius of the station. The difference between the mean skew and the station skew for all the stations in the prediction set were then used to calculate the MSE. The mean skews for the stations in the prediction set can also be obtained from the contour map.

The above procedure can also be used to calculate the difference between mean and station skew for unbiased and weighted skew sets. For simplicity, the mean station, unbiased, and weighted skews derived from the estimation set for the grid point which is nearest to a station in the prediction set were used as the generalised station skew, unbiased and weighted skews for the station and these values were also used for calculating the mean square error.

### 3.5 Prediction Equation

A prediction equation for estimating generalised skew was developed by performing multiple linear regression analysis to relate skews to catchment and climatic characteristics. The characteristics considered are drainage area, main channel slope, stream length, rainfall intensity and mean runoff. Station and unbiased and weighted skews were plotted against each catchment and climatic characteristics. Each of the characteristics was also transformed using common logarithms and plotted against station and unbiased, and weighted skews. These plots were examined for evidence of correlation between station skew and unbiased and weighted skew and catchment characteristics. The catchment or climatic characteristics most related with station skew or unbiased or weighted skew were considered in the regression analysis.

### 3.6 Regional Mean Skew

The study area was not divided into sub-regions in order to

include more gauging stations used in the calculation of regional mean skew. The regional mean of station skew, unbiased and weighted skews were calculated for all the gauging stations in the peninsula.

#### 4.0 RESULTS

The results of this study are presented in the following sections. In the first section, the results of gauging station selection and station skew computations are presented. In the second section, the split-sampling approach and the results of the three techniques for estimating generalised skew are discussed. Values of generalised skew and the MSE's associated with each technique are also presented.

#### 5.0 STATION SKEWS

Peak discharge records from gauging stations located in the peninsula were evaluated for use in estimating generalised skew. Records of 16 or more annual peak discharges through year 2010 are available at 66 of the gauging stations. The locations of the gauging stations are shown in Figure 1. Reference number and name of the gauging station together with DID station number, record length and the peak discharge statistics computed using Equations 3,4,5,6,7 are also presented in Table 1. Station skew ranged from -0.831( $\log_{10}$  unit) to 1.475 ( $\log_{10}$  unit).

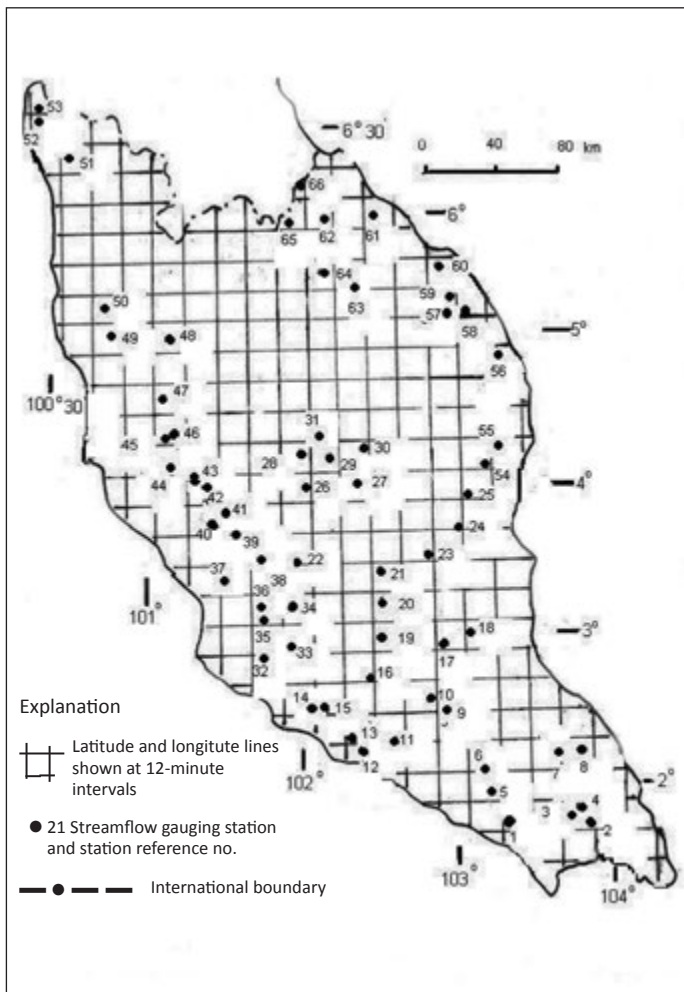


Figure 1: Gauging Stations in Peninsular Malaysia Used in this Analysis.

Table 1: Annual Maximum Streamflow Statistics for Selected Stations in Peninsular Malaysia

Ref. No.	Station Name	Station No.	Years of Record	Station Skew	Unbiased Weighted	Unbiased Weighted Skew / Inverse / Variance	Unbiased Weighted Skew / Record Length
1.	MARINODO	1732401	17	-0.278	-0.376	-0.680	-0.192
2.	JOHOR	1737451	29	0.104	0.126	0.458	0.109
3.	SAYONG	1836402	33	0.532	0.629	2.696	0.624
4.	PENGEI	1836403	23	0.3395	0.498	1.352	0.345
5.	SEMBRONG	1931423	21	-0.476	-0.612	-1.474	-0.387
6.	BEKOK	2130442	19	-0.125	-0.164	-0.346	-0.094
7.	KAHANG	2235401	30	0.325	0.390	1.486	0.352
8.	LENGGOR	2237471	38	0.527	0.610	3.106	0.698
9.	SEGAMAT	2528414	28	0.613	0.744	2.601	0.627
10.	MUAR B KASAP	2527411	44	0.340	0.386	2.341	0.511
11.	KESANG	2224432	48	0.445	0.501	3.359	0.723
12.	D TUNGGAL	2322415	35	0.340	0.398	1.835	0.419
13.	MELAKA	2322413	48	0.100	0.113	0.755	0.162
14.	LINGGI S B	2519421	48	0.446	0.502	3.367	0.724
15.	PEDAS	2520423	31	0.115	0.137	0.545	0.128
16.	KEPIS	2723401	22	-0.216	-0.275	-0.704	-0.182
17.	KERATONG	2928401	16	0.938	1.290	2.142	0.621
18.	ROMPIN	3030401	18	-0.715	-0.953	-1.865	-0.516
19.	SERTING	3024443	21	0.778	1.000	2.409	0.632
20.	TRIANG K A	3224433	34	-0.322	-0.379	-0.1685	-0.387
21.	PAHANG TEMERLOH	3424411	45	-0.213	-0.241	-1.502	-0.327
22.	BENTONG	3519426	40	-0.266	-0.306	-1.656	-0.368
23.	MENTIGA	3329401	16	-0.392	-0.539	-0.895	-0.259
24.	LEPAR JAB GELUGOR	3629403	37	-0.366	-0.425	-2.096	-0.473
25.	KUANTAN	3930401	33	-0.746	-0.882	-3.780	-0.875
26.	LIPIS	4019462	45	0.626	0.709	4.414	0.960
27.	PAHANG S Y	4023412	37	-0.467	-0.543	-2.675	-0.604
28.	JELAI K M	4218416	27	0.070	0.086	0.286	0.069
29.	JELAI J B	4219415	30	-0.346	-0.415	-1.582	-0.375
30.	TEMBILING	4223450	19	-0.168	-0.221	-0.465	-0.126
31.	KECAU	4320401	26	-0.478	-0.588	-1.871	-0.460
32.	LANGAT D'KIL	2816441	50	0.256	0.287	2.017	0.431
33.	SEMENYIH R'RING	2918401	29	0.570	0.688	2.513	0.600
34.	LUI	3118445	44	0.556	0.632	3.829	0.836
35.	GOMBAK	3116433	28	0.398	0.483	1.689	0.407
36.	BATU	3116434	24	0.416	0.520	1.492	0.375
37.	SELANGOR R PAN'G	3414421	49	0.389	0.437	3.001	0.644
38.	SELANGOR RASA	3516244	32	-0.685	-0.813	-3.358	-0.783
39.	BERNAM T MALIM	3615412	44	0.390	0.443	2.686	0.587
40.	BERNAM SKC	3813411	47	0.395	0.445	2.916	0.630
41.	SLIM S RIVER	3814416	39	-0.362	-0.418	-2.193	-0.490
42.	SUNGKAI	3913458	48	-0.355	-0.399	-2.680	-0.577
43.	BIDOR TIN MINE	4012401	29	-0.278	-0.336	-1.225	-0.293
44.	BTG PADANG	4111455	45	-0.333	-0.377	-2.348	-0.511
45.	KINTA WEIR	4310401	31	-0.378	-0.451	-1.791	-0.421
46.	KAMPAR	4311464	31	-0.422	-0.504	-1.999	-0.470
47.	PARI	4610466	46	-0.533	-0.603	-3.847	-0.834
48.	PLUS K LINTANG	4911415	41	-0.269	-0.308	-1.719	-0.380
49.	KURAU	5007421	49	-0.245	-0.275	-1.890	-0.405
50.	KRIAN	5206423	46	0.554	0.626	3.998	0.867
51.	KULIM	5405421	47	1.475	1.663	10.890	2.352
52.	ARAU	6402435	25	0.045	0.056	0.169	0.042
53.	PELARIT	6502431	20	0.496	0.645	1.455	0.388
54.	CHERIL	4131453	26	-0.257	-0.316	-1.006	-0.247
55.	KEMAMAN	4232453	30	-0.665	-0.798	-3.041	-0.720
56.	DUNGUN	4831441	37	-0.831	-0.966	-4.760	-1.075
57.	TELEMONG	5129437	25	-0.737	-0.914	-2.764	-0.687
58.	TERENGGANU	5130432	23	-0.392	-0.494	-1.342	-0.342
59.	NERUS	5229436	28	0.324	0.393	1.375	0.331
60.	CHALOK	5428401	30	0.216	0.259	0.988	0.234
61.	BESUT	5724441	24	-0.641	-0.801	-2.299	-0.578
62.	KELANAN	5721442	50	-0.288	-0.323	-2.270	-0.485
63.	LEBIR	5222452	32	-0.653	-0.775	-3.201	-0.746
64.	GALAS	5320443	33	-0.221	-0.261	-1.120	-0.259
65.	LANAS	5718401	19	-0.309	-0.407	-0.856	-0.232
66.	GOLOK R PJANG	6019411	35	-0.104	-0.122	-0.561	-0.128

Note: Skew coefficient in  $\log_{10}$  unit

### 5.1 Techniques for Estimating Generalised Skew

Peak discharge statistics and catchment characteristics computed for every gauging station were split into two sets, the estimation and the prediction sets. The estimation set has 35 gauging stations, and the prediction set has 31 gauging stations.

Data in the estimation set were used to develop the three techniques for estimating generalised skew in Peninsular Malaysia. Data in the prediction set were used to determine the accuracies of these techniques.

Box and whisker plots in Figure 2 show that station skew, unbiased and weighted skews have almost similar statistical properties for both sets of data. The ends of the whiskers represent the range of values, the ends of the box represent 25 and 75% quartiles and the line in the middle of each box represents the median value. The median values of the station skew for the estimation and the prediction set are almost the same as those of unbiased skew and the weighted unbiased skew using station record length. However, the median values of the weighted unbiased skew using inverse station variance are much smaller. The variance of the unbiased and weighted skew sets are greater than those of the station skew sets. This is shown in Figure 2, where the ends of the whiskers and the 25 and 75% quartiles are further apart for the unbiased and weighted skews than for the station skew. The larger variance in these skew sets is due to the factors used to correct and weight the station skew for bias.

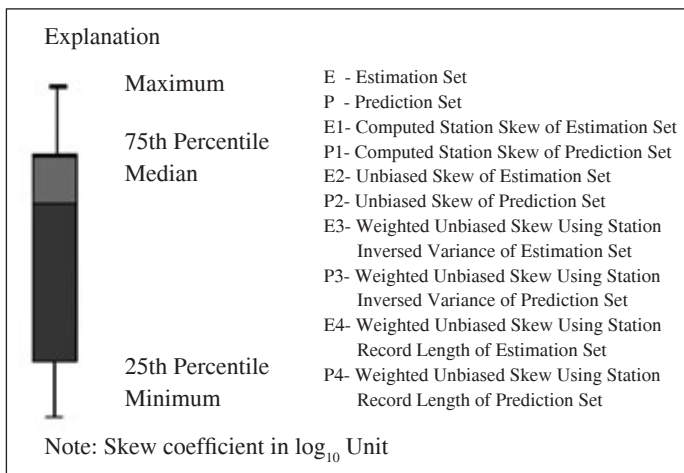
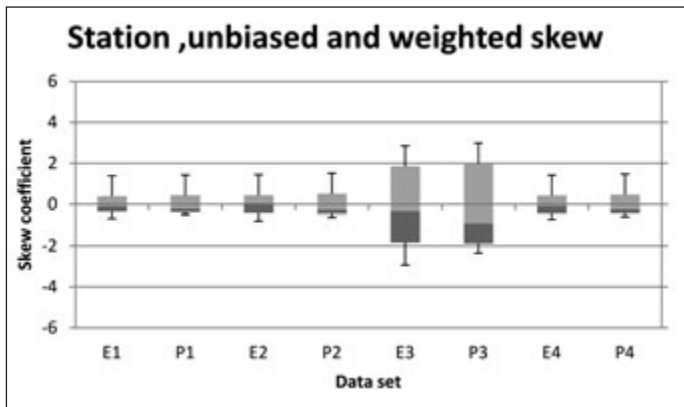


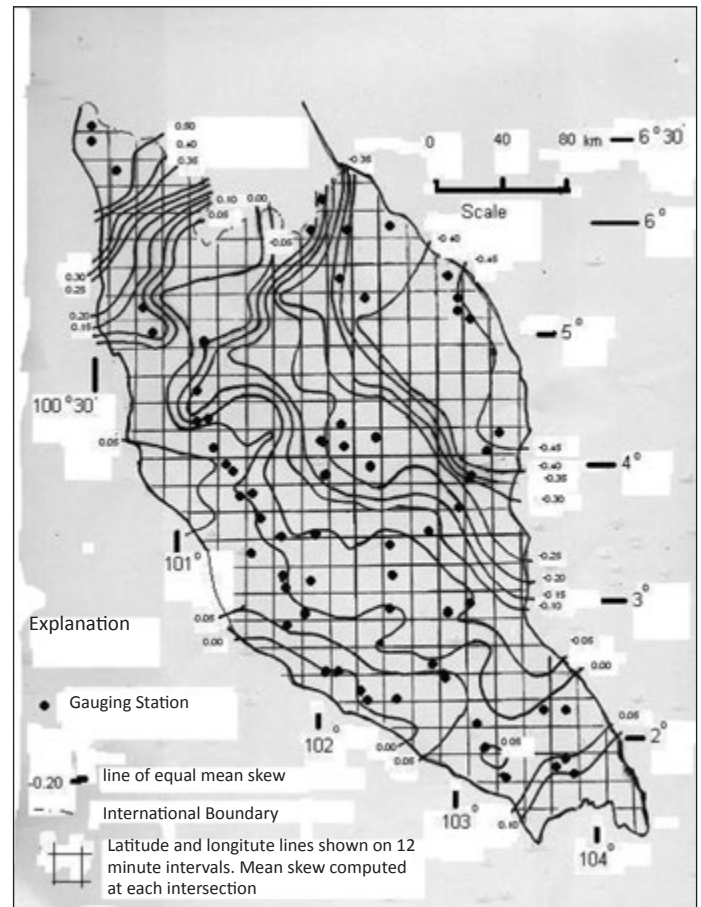
Figure 2: Schematic Plot of Selected Characteristics in the Estimation and Prediction Sets

Figure 3 shows the isoline map constructed using the method described in “Materials and methods”. As stated in “isoline map”, the accuracy for this technique was evaluated for station skew, unbiased and weighted skew data sets using the station skew, unbiased and weighted skews of each station in the prediction set and the mean station skew, unbiased and weighted skews obtained from the estimation set of the grid point located nearest to the station. The mean square error values obtained for the station skew, unbiased, weighted unbiased skews are:

Data	Mean square error ( $\log_{10}$ unit)
Station Skew	0.19
Unbiased Skew	0.29
Weighted Unbiased Skew/ Inverse Variance	5.11
Weighted Unbiased Skew/ Record Length	0.26

It can be seen that the station skew data sets give the lowest MSE and thus if the isoline map technique is used, the generalised skew coefficient should be obtained from Figure 3 and the MSE value of 0.19 ( $\log_{10}$  unit) should be adopted.

For the regression method, the station skew, unbiased and weighted skews are not well correlated with any of the catchment and climatic factors. Attempts to correlate the independent and dependent variables using multiple linear regression were also unsuccessful. This technique was dropped for further consideration.



Note: Skew coefficient in  $\log_{10}$  Unit

Figure 3: Lines of Equal Mean Skew at Stations Within 100 Mile (160 Kilometer) Radius of Grid Points

In the regional mean skew technique, the peninsula is considered as a single region in order to include sufficient stations for estimating a regional mean skew. Regional mean skew for station skew, unbiased, weighted unbiased skew data sets were calculated using the respective estimation data sets. The MSE for the data sets were estimated using the regional mean skew from the estimation data sets and the station skew, unbiased and weighted skews for all stations in the prediction data sets. Results are:

Data	Regional Mean Square ( $\log_{10}$ unit)	Mean Square Error ( $\log_{10}$ unit)
Station Skew	-0.022	0.05
Unbiased Skew	-0.036	0.07
Weighted Unbiased Skew/Inverse Variance	0.192	1.97
Weighted Unbiased Skew/Record Length	0.030	0.10

From the summary shown above, the station skew estimation and prediction sets give the lowest MSE.

The MSE calculated from the above techniques are shown in Figure 4.

As previously mentioned, unbiased and weighted skews have greater variance than station skews due to the use of biased correction and weighting factors. The accuracy of techniques developed using station skew corrected and weighted is less than that of techniques using station skew. The MSE computed from unbiased and weighted skews were larger than the MSE calculated from station skew.

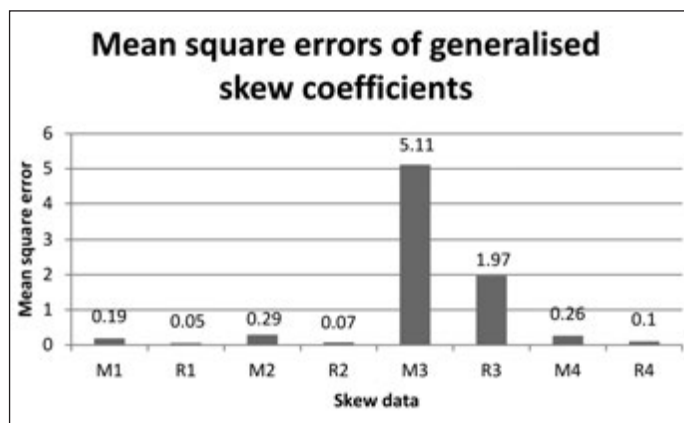
As the smallest MSE is obtained from the station skew and regional mean, the most appropriate generalised skew and MSE to use for gauging stations in Peninsular Malaysia are -0.022 ( $\log_{10}$  unit) and 0.05( $\log_{10}$  unit).

## 6.0 COMPARISON OF RESULTS WITH OTHER STUDIES

Results indicate that our findings are consistent with findings of other studies, which were carried out mainly in the U.S.A. For example, the generalised skew and mean square error obtained from our study are comparable to the values derived from basins of Texas, Michigan, New York and Hawaii. The table shown below are our results compared to those of U.S. basins.

*Results of generalised skew coefficient from various studies.*

State/Country	Generalised Skew Coefficient	MSE	Method
Peninsular Malaysia	-0.022	0.05	Regional
Peninsular Malaysia	-	0.19	Isoline
New York	-0.2 to 0.7	0.16	Isoline
Texas	-0.2 to 0.3	0.35	Isoline
Hawaii	-0.14	-	Regional
Michigan (UPPER PART)	0.12	0.2	Regional
Michigan (Southwest)	0.081	-	Isoline
Michigan (Lower)	-0.017	-	Isoline
Minnesota	-0.5 to 0.2	0.182	Isoline
Illinois	-0.16	0.2	Regression



Explanation

M1 - Station skew and generalised skew from nearest map grid point

R1 - Station skew and regional mean skew

M2 - Unbiased skew and generalised unbiased skew from nearest map grid point

R2 - Unbiased skew and unbiased regional mean

M3 - Weighted unbiased skew/ Inverse variance and generalised weighted skew/Inverse variance from nearest map grid point

R3 - Weighted unbiased skew/Inverse variance and weighted regional mean skew/Inverse variance

M4 - Weighted unbiased skew/ Record length and weighted generalised mean skew/ Record length

R4 - Weighted unbiased skew/ Record length and weighted regional mean/ Record length

Note: Skew coefficient in  $\log_{10}$  unit

**Figure 4: Mean Square Error for Generalized Skew Coefficients, As Computed From the Isoline Map and Regional Mean Technique**

## 7.0 SUMMARY AND CONCLUSIONS

Gauging stations having more than 16 annual peak discharges that were not significantly affected by regulation or diversion were considered for use in estimating generalised skews, Station skews were computed for each gauging station according to the guidelines published by the Hydrology Subcommittee of the Interagency Advisory Committee on Water Data. Unbiased and weighted skews were also computed using correction factors.

A split-sampling approach was used to evaluate three techniques for estimating generalised skew. The gauging stations used in this analysis were split into two sets of approximately equal size. One set, the estimation set, was used to compute the generalised skew; the other, the prediction set, was used to compute the mean square error, MSE, of each technique. The most accurate technique, the one with the lowest MSE, was considered to be most appropriate for the estimation of generalised skew for Peninsular Malaysia.

Three techniques for estimating generalised skew for Peninsular were evaluated. The techniques are: (1) an isoline map of skew for the study area, (2) a prediction equation relating station skew to catchment and climatic characteristics, (3) the regional mean skew for Peninsular Malaysia.

Records of 66 gauging stations were selected for use in estimating generalised skew. Station skew values computed for each station ranged from -0.831( $\log_{10}$  unit) to 1.475 ( $\log_{10}$  unit).

Values for unbiased and weighted skews were also computed. Unbiased and weighted station skews have a larger variance due to the correction factors used in computing the respective skew values.

Skew values were split into two sets. Values of station skew, unbiased and weighted skews have almost similar statistical properties for the estimation and prediction sets.

Error analysis shows that the regional mean skew technique using station skew data gives the lowest MSE value and it is considered appropriate to use this technique for estimating generalised skew for Peninsular Malaysia.

Results indicate that our findings are consistent with findings of other studies, which were carried out mainly in the U.S.A. For example, the generalised skew and mean square error for our study are comparable to the values derived from basins of Texas, Michigan, New York and Hawaii.

## 8.0 ACKNOWLEDGEMENT

The permission of the Drainage and Irrigation Department (DID) to use the peak flow data for this study is gratefully acknowledged. ■

## REFERENCES

- [1] Croskey H. M. and Holtschlag D. J. (1983) Estimating generalised flood skew coefficients for Michigan USGS Water Resources Report 83-4194.
- [2] Hydrology Subcommittee of the Interagency Advisory Committee on Water Data(1982) Bulletin 17, Guidelines for determining flood flow frequency U.S. Geological Survey.
- [3] Judd L. J. (1996) Techniques to estimate generalised skew coefficients of annual peak streamflow for natural basins in Texas USGS Water Resources Report 96-4117.
- [4] Lorenz D. L. (1997) Generalised skew coefficients for flood –frequency analysis in Minnesota USGS Water Resources Investigations Report 97-4089.
- [5] Oberg K. A. and Mades D. M. (1987) Estimating generalised skew Of the Log Pearson III distribution for annual peal floods in Illonois USGS water resources investigations Report 86-4008.
- [6] Lee R.(1984) Generalised skew coefficient for flood frequency computations for the state of Hawaii USGS Water Resources Report 84-4027.
- [7] Lumia R. and Baevsky Y. H. (2000) Development of a contour map showing generalised skew coefficients of annual peak discharges of rural, unregulated streams in New York, excluding Long Island USGS Water Resources investigation report 00-4022.
- [8] Tasker G D (1982) Comparing methods of hydrologic regionalisation American Water Resources Association AGU.
- [9] Tasker G D and Stedinger J R (1986) Regional skew with weighted LS regression Jour. Of water resources planning and management ASCE.
- [10] U.S. Water Resources Council(1976) Guidelines for determining flood flow frequency US Water Resources Council Bulletin 17.
- [11] Hydrologic Engineering Centre (2009) HEC SSP-2.0.
- [12] Wallis J R, Matalas N C, and Slack J R, (1974) Just a moment, Water Resources Research AGU.

## PROFILES



**DAVID HONG JER LANG** is currently a lecturer of Taylor’s University, Malaysia. He received his B.Sc degree in computer science from Nottingham University in 2005 and the Ph.D. degree from Monash University in 2010. He is author and co-author of several Association for Computing Machinery/Institution of Electrical and Electronic Engineers (ACM/IEEE) conference papers and is also the author of several papers in automatic wrapper design published in computer science journals and is the co-author of the book “Hydrological Analysis and Design: An approach based on Malaysian Rainfall and Runoff”. David Hong is listed in Who’s Who in the World. E-mail address: jerlang.hong@taylors.edu.my



**AMIRAH HANIM BT. MOHD PUAD** is currently a Civil Engineer of ZHL Engineers Sdn. Bhd., and she is the Chief Engineer of the water resources and hydraulic division. She received her Bachelor Degree (Hons) in Civil Engineering from Universiti Tenaga Nasional in 2008 and a Master’s degree in Civil engineering from the same university in 2012. Email address: amirah.hanim@gmail.com



**INTAN SHAFILIAH BT. ABDUL AZIA** obtained her B.Sc. degree in Civil Engineering from Universiti Tenaga Natiaional (UNITEN). Intan has worked as a research Engineer for UNITEN after graduation and she was actively involved in hydrology and Hydraulic works such as urban hydrology, flood estimation and dam break analysis. She joined ZHL Engineers Sdn. Bhd. in 2011 and she is a modeller and analyst of hydrological and hydraulic works such as flood and drought analysis, and also actively involved in dam inspection, emergency response plan and other related works. Email address: ishafilah@gmail.com



**HONG KEE AN** is presently a consultant in hydrology and water resources and he worked as a hydrologist, water resources planner and modeler for various water supply projects in Malaysia. He received his B.Sc. degree in Agricultural Engineering from National Taiwan University in 1971 and the M.Eng. degree in Water Science from Asian Institute of Technology in 1973. He is the author of several papers published in the journal and bulletins of the Institution of Engineers, Malaysia and Malaysian Water Work Association and the books” A Guide to Flood and Yield Estimation” and “Hydrlogical Analysis and Design: An approach based on Malaysian Rainfall and Runoff”. He is a Fellow of the Institution of Engineers, Malaysia and is listed in Who’s Who in Engineering published by the American Association of Engineering Societies. E-mail address: keeanhong@yahoo.co.uk