ISSN 2278-3091
Volume 8, No.3, May - June 2019
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse08832019.pdf
https://doi.org/10.30534/ijatcse/2019/08832019

Developing Composite Indicators for Flood Vulnerability Assessment: Effect of Weight and Aggregation Techniques

D.Y.C. Liew[1,2], F. Che Ros[2], A. N. Harun[2]
[1]Department of Irrigation and Drainage, Malaysia, daniel@water.gov.my
[2] Disaster Preparedness and Prevention Centre, Malaysia-Japan International Institute of Technology, University Teknologi Malaysia, Malaysia, crfaizah@utm.my

ABSTRACT

The assessment of flood vulnerability involves multidimensional and complex interactions between environment, social and economic dimensions. Indicator-based vulnerability assessment such as the Flood Vulnerability Index (FVI) is widely used in vulnerability studies to summarise complexity and multidimensionality issues to gauge the level of vulnerability. To assess the various factors of vulnerability, we employed a set of 21 environmental, social and economic indicators to quantitatively assess the three factors of vulnerability, namely exposure, susceptibility and resilience to flood at the subnational level. The construction of the vulnerability index involves sequential steps, including selecting the indicators, their normalisation, weightage and aggregation to a final index. In this study, we looked into which weighting and aggregation technique as the most suitable to develop the final composite indicator. The weighting techniques employed are equal weight, unequal weight and principal component analysis. Two different aggregation techniques, namely additive mean and geometric mean were used to aggregate the indicators to a single index. This study employed reliability and sensitivity analyses to evaluate the robustness of the FVI constructed using various techniques. This study shows the wider application of the equal weighting and additive mean techniques to develop composite indicator for flood vulnerability assessment.

Key words : Equal weight, flood vulnerability index, principal component analysis, unequal weight.

1. INTRODUCTION

Composite indicators are widely used to encapsulate a range of complex and multidimensional issues [1]. Birkmann [2] reviewed in his paper a few approaches to measure risk and vulnerability using indicators, namely the Disaster Risk Index (DRI), the Hotspots Project, the Americas Indexing Programme and the Community-Based Risk Index which aimed to quantitatively and qualitatively measure risk and vulnerability. Vulnerability indicators method is commonly used in flood vulnerability studies [3]–[5] and are preferred by policy makers due to its ability to provide a simple model of the complex measurement of vulnerability. This method uses available data to provide an operational representation of the characteristic of a place and information of an element at risk [6]. The depiction of vulnerability over space provides policy makers the means to prioritise strategy and measures to manage disaster risk in specified region [7].

As vulnerability covers a multidimensional characteristic of risk such as environment, social and economic, the composite indicators method is able to provide an assessment of flood vulnerability in particular geographical region. With respect to the main issues and problems pertaining to the construction of a composite index, [8]–[10] recognised that important considerations should be given to proper conceptual framework and methodology. A framework defines the phenomenon to be measured and guides the selection of indicators [10]. An indicator or a set of indicators is characterised by its inherent characteristic to estimate quantitatively the condition of a system. Specific vulnerability indicators for flood have been developed to measure the factors contributing to vulnerability, namely exposure, susceptibly and resilience [3], [11], [12].

In the case of flood, vulnerability can be defined as "the extent which a system is susceptible to flood due to its exposure to a disturbance and its capacity or incapacity to be resilient, to cope, recover or adapt" [3]. Figure 1 shows the framework of flood vulnerability as a function of exposure, susceptibility and resilience.

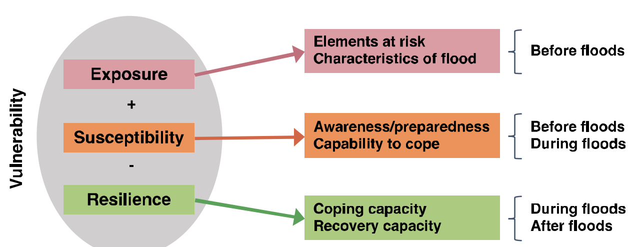

Figure 1: Vulnerability as a function of exposure, susceptibility and resilience [3]

The formulation of the flood vulnerability index (FVI) equation can be expressed as in (1), where E is exposure, S is susceptibility, and R is resilience. These are known as factors influencing vulnerability. Resilience is expressed as a negative connotation, in which a higher score causes the vulnerability to be higher.

$$FVI = E + S - R \quad (1)$$

In brief, exposure is the geographical predispositions of a system to be disrupted by a flooding event [3]. Susceptibility refers to the extent to which elements within the system are exposed that influences the odds of being damaged at times of hazardous floods [3]. Resilience is the capability of a system to endure any disturbance while maintaining fundamental efficiencies in its social, economic, and physical environmental dimensions [3].

This paper aims to show the differences between weighting and aggregation techniques and their influences on the resulting rankings by applying a methodological process of constructing indicators for flood vulnerability assessment. Malaysia is a country in south-east Asia, with a total surface are of 330,345 km^2 and population of 31.6 million [13]. In a 2003 national study, approximately 29,800 km^2 (9%) of the landmass of the country is prone to flood with an estimated 4.8 million (21%) population affected by flood [14]. Figure 2 illustrates the flood prone areas in Malaysia.

Figure 2: Flood prone areas in Malaysia [14]

2. METHODOLOGY

For this paper, three weighting techniques and two aggregation techniques were examined to study their effect and to determine the preferred method to construct composite indicators for flood vulnerability assessment in Malaysia. The FVI for the 14 states of Malaysia was computed using equal weight and aggregation method [15]. A set of 21 environmental, social and economic indicators was selected to quantitatively assess the three factors of flood vulnerability, namely exposure, susceptibility and resilience at the subnational level. As shown in Figure 3, the indicators were grouped to make 3 composite indicators and finally combined as a single index. The null hypothesis is there is no significant agreement among the rankings of the states based on the FVI.

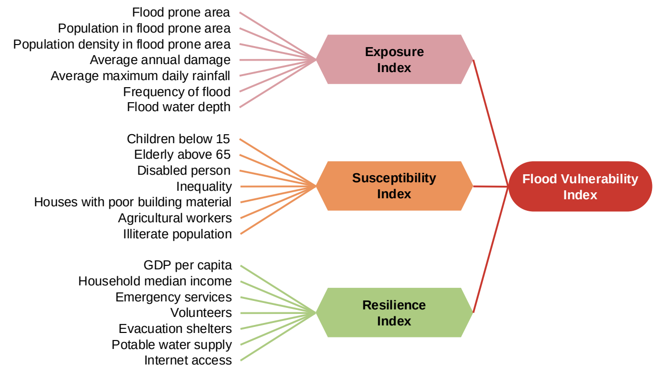

Figure 3: Indicators and their respective vulnerability factors

2.1 Overview of Indicators

Table 1 contains the indicators in different measurement units under the three factors of flood vulnerability. Consideration of vulnerability indicators undertaken in this study is based on literature review while the selection of indicators depends on its relevancy and the availability of secondary data. More importantly, these indicators can describe the relationship between the three factors of vulnerability.

Table 1: Selected indicators for flood vulnerability

Factor	Indicators	Functional Relationship
Exposure	Flood prone area (%)	(+)
	Population in flood prone area (%)	(+)
	Population density in flood prone area (persons/km^2)	(+)
	Average annual damage (RM)	(+)
	Average maximum daily rainfall (mm)	(+)
	Frequency of flood	(+)
	Flood water depth (m)	(+)
Susceptibility	Children under 15 (%)	(+)
	Elderly above 65 (%)	(+)
	Disabled persons (%)	(+)
	Inequality (Gini coefficient)	(+)
	Houses with poor building material (%)	(+)
	Agricultural workers (%)	(+)
	Illiterate population (%)	(+)
Resilience	Household median income (RM)	(–)
	GDP per capita (RM)	(–)
	Emergency services (hospital bed per 10,000 persons)	(–)
	Volunteers (per 10,000 persons)	(–)
	Evacuation shelters (capacity per 10,000 person)	(–)
	Potable water supply (%)	(–)
	Internet access (%)	(–)

2.2 Normalisation

Since the indicators is measured in various units, they needed to be normalised to ensure comparability. The min-max normalisation is a process used to normalise the actual value into 0 to 1 scale. This normalisation technique has been used irrespective of study domain [16]–[19]. Before the values could be normalised, it was necessary to assign functional relationship to the indicators. There are two types of functional relationship; either vulnerability increases with an increase of the value of the indicator or a decrease of the value of the indicator. The functional relationships of the indicators were determined from previous studies. The lists of selected indicators and their functional relationship identified with flood vulnerability are shown in Table 1.

In the case where vulnerability increases corresponding to the value of the indicator (the indicator has "+" functional relationship), normalisation was done using (2), as in

$$y_{ij} = \frac{x_{ij} - x_{min}}{x_{max} - x_{min}} \tag{2}$$

where y_{ij} denotes the normalised value of indicator (j) with respect to state (i), x_{ij}, x_{min} and x_{max} are the actual, minimum and maximum values, respectively, of indicator (j) among all the states (i).

On the contrary, where vulnerability decreases with an increase in the value of the indicator (the indicator has "–" functional relationship), normalisation was carried out using (3), as in

$$y_{ij} = \frac{x_{max} - x_{ij}}{x_{max} - x_{min}} \tag{3}$$

2.3 Weighting

The weights are required in composite indicator construction. Weights obtained from mathematical algorithms or models are known as objective weighting method as it does not involve subjective judgement from decision makers [20]. As there is no consensus on the method used in deriving weights [8], this study used equal weighting method as the baseline approach. In addition to that, weights derived from an inverse variance method proposed by [21] and principal component analysis (PCA) were also employed to investigate the implication of the use of weighting system.

A. Equal Weight

Most composite indicators depend on on equal weighting (EW) to assign the same weight to the indicators [8]. This method of averaging ensures an equal importance to all the indicators which may not be quite correct but may be acceptable when no other available means of weighting are known [22]. The various justification to use equal weighting includes simplicity of construction and lack of theoretical structure to justify a differential weighting scheme [23]. Similarly, [18], [19], [24] have employed equal weighting in their approach to construct vulnerability index.

In this study, considering the straightforwardness and simplicity of equal weighting over other methods, the weights of the indicators are assumed equal. Using equal weighting (EW) method, the same weightage is assigned for each indicator using (4), as in

$$w_j = 1/k \tag{4}$$

where w_j is a weight for k indicators (j = 1, …, k) for each factor.

B. Unequal Weight

Iyengar and Sudarshan [21] developed a statistically sound and well-suited method to create a composite indicator from multivariate data in order to rank the economic performance of the districts in India. This method has been used by [4], [25], [26]. Let y_{ij}, represent the normalised value of the j^{th} indicator and the i^{th} state (i = 1, 2, …, m; j = 1, 2, …, k), the indicator I_i according to their respective factors, are then summed using (5), as in

$$I_i = (EI; SI; RI) = \sum_{j=1}^{k} w_j y_{ij} \tag{5}$$

where the w's ($0 < w_j < 1$ and $w_1 + w_2 + \ldots + w_j = 1$) are the weights. I_i denotes exposure index (EI), susceptibility index (SI) and resilience index (RI) for each i^{th} state respectively. The weights are assumed to vary inversely as the variance over the states in the respective indicators and the weight w_j is determined by (6) while the normalising constant, c, is determined by (7).

$$w_j = c/\sqrt{var(y_{ij})} \tag{6}$$

$$c = \left|\sum_{j=1}^{k} 1/\sqrt{var(y_{ij})}\right|^{-1} \tag{7}$$

This weighting method prevents peculiarly large variation in any of the indicators from overshadowing the contribution of the rest of the indicators, and that improvements in an indicator can compensate the deficiencies of the other indicators [21], [25].

C. Principal Component Analysis

Principal component analysis (PCA) is a statistical approach to capture the highest variance possible in the original indicators by retaining as few components as possible. It has been used by the majority of the more recent vulnerability indices studies in vulnerability studies [19], [27]–[31]. This method has seemingly gained interest as a 'data-driven technique' to derive weight [23].

In this study, we used the Kaiser criterion to select the principal component with eigenvalues more than one which accounts for the maximum variance [8]. Kaiser [32] retained the factors with eigenvalue greater than one as there are as many reliable factors which implies reliability. For the q^{th} factor with eigenvalue greater than 1, the weight of each indicator is computed by dividing the explained variance with the total variance given by (8). Varimax rotation was then performed to maximise the sum of variances of the squared loadings.

$$w_j = \sum_{q=1}^{p} a_{jq}^2 / \sum_{q=1}^{p} \lambda_q \tag{8}$$

where, w_j is the weight of the j^{th} indicator, λ_q is the eigenvalue of the q^{th} factor and a_{jq} is the loading value of the j^{th} indicator on q^{th} factor.

2.4 Aggregation

Aggregation based on addition of components using equal weights is used comprehensively [33]. As in previous studies, [27], [34], defended that the additive mean method model is more appropriate compared to geometric mean method, as the former do not make *a priori* assumption which allows for the weights of each factor to vary with different importance. Thereby, each factor is viewed as contributing equally to the state's overall vulnerability. This is the best method in the absence of a defensible method for assigning weights to each factor.

The geometric mean method, on the other hand, is used to address the shortcoming of additive aggregation in which a high value in one indicator can compensate a low value in another. This behaviour is known as compensability [33]. Geometric aggregation technique is a nonlinear approach and is calculated as the product of weighted indicators. Hence, the conceptual FVI equation can then be written as the additive mean in (9) and the geometric mean in (10) respectively.

$$FVI_a = \sum_{j=1}^{n} w_j I_j = \frac{1}{3}(EI + SI + RI) \tag{9}$$

$$FVI_g = \prod_{j=1}^{n} I_j^{w_j} = (EI \times SI \times RI)^{\frac{1}{3}} \tag{10}$$

where FVI_a and FVI_g are the flood vulnerability index derived using additive and geometric means respectively, w_j is the weight of the factor indicator I_j, and n is the number of factor indicators. I_j denotes EI, SI and RI for the exposure index, susceptibility index and resilience index respectively. Both additive and geometric approaches yield quantitative index score.

2.5 Sensitivity and Reliability Analysis

The null hypothesis is there is no significant agreement among the ranking of the states based on the FVI scores.

A. Average Rank Shift

Average shift in rank is a sensitivity analysis technique used to assess the robustness of the composite indicator [16], [22], [27]. The stability of the FVI and the resulting rank of a given state, $Rank(FVI_i)$, indicates the robustness of the estimation. The rank shift, R_s, measures the uncertainty of each input factor and the mean value is computed as the differences between the respective state's ranking and the reference ranking over the total number of states (m) in (11), as in

$$\bar{R}_s = \frac{1}{m}\sum_{i=1}^{m} \left| Rank_{ref}(FVI_i) - Rank(FVI_i) \right| \tag{11}$$

where the reference ranking is perceived as the median rank.

B. Spearman's Rank-Order Correlation

This study used the Spearman's rank-order correlation, a non-parametric statistic test, to test the reliability of the rankings. Spearman's has a value from -1 to +1, where +1 signifies the ranks are perfectly positive correlated and -1 signifies a perfect negative relationship between ranks while 0 indicates no correlation between ranks. Taking into consideration of tied ranks, the Spearman's rank-order correlation coefficient, ρ, is given as (12) [35], as in

$$\rho = \frac{12\left[\sum_{i=1}^{n} R_i S_i - n(n+1)^2/4\right]}{\{[n(n^2-1) - 12t'][n(n^2-1) - 12u']\}^{1/2}} \tag{12}$$

where $R_i = \text{rank}(X_i)$, $S_i = \text{rank}(Y_i)$ and the expression u' denotes the summation over all set of u tied in the additive mean method while t' the corresponding sum for the geometric mean method for n sample pairs. The formula is given in (13) and (14) respectively, as in

$$t' = \sum t(t^2 - 1) \tag{13}$$

$$u' = \sum u(u^2 - 1) \tag{14}$$

C. Cronbach's Alpha

We used the Cronbach's alpha, α, developed by Cronbach [36] to measure the reliability or consistency of the rankings. The Cronbach's alpha is a function of the number of ranking methods and the average inter-correlation among the ranking methods. An alpha value greater than 0.9 means the rankings has excellent consistency while a value below 0.7 is questionable. The Cronbach's alpha equation is given in (15).

$$\alpha = \frac{n \times \bar{c}}{\bar{v} + (n-1)\bar{c}} \tag{15}$$

where n is the number of ranking methods, $\bar{c}$ is the average covariance between item-pairs and $\bar{v}$ is the average variance.

3. RESULTS AND DISCUSSION

3.1 Derived Weights of the Indicators

The derived weights of the indicators using various weighting techniques are plotted in a spider diagram as shown in Figure 4. Using equal weight method as a baseline, each indicator is assigned a weightage of 1/7. The weights derived from unequal weight (Iyengar and Sudarshan) method and principal component analysis do not follow any pattern. Table 2 shows the derived weights of the normalised indicators using equal weight (EW), unequal weight (UW) and principal component analysis (PCA).

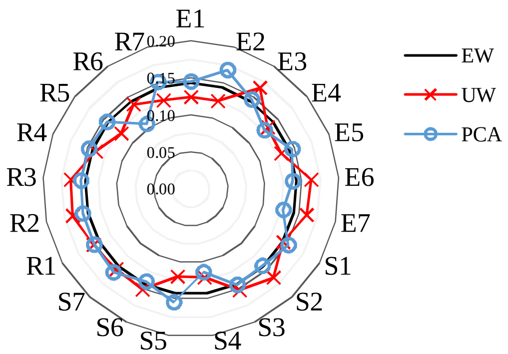

Figure 4: Spider diagram of the derived weights

Table 2: Weights obtained from different weighting methods

Indicators	Weight		
	EW	UW	PCA
Flood prone area (E1)	0.143	0.124	0.145
Population in flood prone area (E2)	0.143	0.124	0.168
Population density in flood prone area (E3)	0.143	0.165	0.146
Average annual damage (E4)	0.143	0.133	0.127
Average maximum daily rainfall (E5)	0.143	0.131	0.147
Frequency of flood (E6)	0.143	0.163	0.138
Flood water depth (E7)	0.143	0.160	0.128
Children below 15 (S1)	0.143	0.144	0.152
Elderly above 65 (S2)	0.143	0.164	0.142
Disabled person (S3)	0.143	0.152	0.144
Inequality (S4)	0.143	0.121	0.114
Houses with poor building material (S5)	0.143	0.120	0.155
Agricultural workers (S6)	0.143	0.151	0.139
Illiterate population (S7)	0.143	0.148	0.154
Household median income (R1)	0.143	0.152	0.151
GDP per capita (R2)	0.143	0.164	0.150
Emergency services (R3)	0.143	0.163	0.149
Volunteers (R4)	0.143	0.138	0.148
Evacuation shelters (R5)	0.143	0.120	0.145
Potable water supply (R6)	0.143	0.138	0.106
Internet access (R7)	0.143	0.125	0.151

The reliability of each indicator due to the effect of weights was tested using the Cronbach's alpha and average rank shift method. We investigated the reliability of the exposure index, susceptibility index and resilience index.

3.2 Reliability of Exposure Index

Table 3 presents the outcome of the exposure index computed using three different weighting methods, namely equal weighting (EW), unequal weighting (UW) and principal component analysis (PCA). All three weighting methods consistently ranked the exposure for Kelantan (1st), Pahang (2nd), Terengganu (3rd), Selangor (4th), Kuala Lumpur (8th), Johor (9th), Perak (10th) Kedah (11th) and Melaka (12th). It is observed that the rank varies within 1 position for the state of Sarawak (5th - 6th) and Perlis, and varies within 2 positions for the state of Pulau Pinang (5th – 7th), Sabah (5th – 7th) and Negeri Sembilan (12th – 14th).

Table 3: Exposure index (EI) and ranks using various weighting methods

State	EI Score			EI Rank		
	EW	UW	PCA	EW	UW	PCA
Johor	0.21	0.19	0.21	9	9	9
Kedah	0.13	0.14	0.13	11	11	11
Kelantan	0.60	0.58	0.61	1	1	1
Melaka	0.05	0.05	0.05	12	12	12
N. Sembilan	0.04	0.05	0.04	13	12	14
Pahang	0.51	0.47	0.53	2	2	2
Perak	0.18	0.19	0.18	10	10	10
Perlis	0.04	0.04	0.04	13	14	13
Pulau Pinang	0.32	0.30	0.32	6	7	5
Sabah	0.31	0.30	0.30	7	5	7
Sarawak	0.32	0.30	0.31	5	5	6
Selangor	0.42	0.41	0.42	4	4	4
Terengganu	0.49	0.45	0.50	3	3	3
Kuala Lumpur	0.28	0.29	0.28	8	8	8

Cronbach's alpha shows the exposure index ranking has an excellent reliability of 0.99. The average shift in rank was calculated to assess the reliability of the rank against the median rank. There was no shift in rank using equal weighting method, while unequal weight method resulted in a shift of 0.36 points and principal component analysis method resulted in a shift of 0.21 points.

3.3 Reliability of Susceptibility Index

Table 4 shows the susceptibility index values using various weighting methods. Ten states were consistently ranked while 4 states saw their ranks shifting within 1 position. The states which were ranked consistently are Sarawak (1st), Kelantan (2nd), Sabah (3rd), Perak (6th), Perlis (7th), Melaka (10th), Johor (11th), Pulau Pinang (12th), Kuala Lumpur (13th) and Selangor (14th). Whereas, Kedah and Terengganu alternate between the 4th and 5th position while Negeri Sembilan and Pahang alternate between the 8th and 9th position. The Cronbach's alpha value of 0.99 indicates the consistency of the rank among the three different weighting methods. Adopting the median rank as reference rank, the average shift in rank using equal weighting and unequal weight is zero respectively, while principal component analysis is 0.29 position.

Table 4: Susceptibility index (SI) and ranks using various weighting methods

State	SI Score			SI Rank		
	EW	UW	PCA	EW	UW	PCA
Johor	0.33	0.33	0.32	11	11	11
Kedah	0.49	0.48	0.47	4	4	5
Kelantan	0.62	0.60	0.63	2	2	2
Melaka	0.34	0.35	0.34	10	10	10
N. Sembilan	0.38	0.38	0.37	8	8	9
Pahang	0.37	0.38	0.38	9	9	8
Perak	0.42	0.44	0.42	6	6	6
Perlis	0.39	0.40	0.40	7	7	7
Pulau Pinang	0.26	0.27	0.26	12	12	12
Sabah	0.56	0.53	0.55	3	3	3
Sarawak	0.71	0.71	0.71	1	1	1
Selangor	0.22	0.21	0.21	14	14	14
Terengganu	0.47	0.46	0.49	5	5	4
Kuala Lumpur	0.24	0.24	0.22	13	13	13

3.4 Reliability of Resilience Index

The resilience index and ranks using various weighting methods is tabulated in Table 5. The ranks for ten states were consistent, Sabah (1st), Johor (4th), Sarawak (5th), Perak (6th), Pulau Pinang (6th), Negeri Sembilan (10th), Melaka (11th), Perlis (12th), Terengganu (13th), Kuala Lumpur (14th). Four states saw their ranks shifting within 1 position, namely Kedah and Kelantan alternating between 2nd and 3rd position while Pahang and Selangor between 7th and 8th position.

Table 5: Resilience index (RI) and ranks using various weighting methods

State	RI Score			RI Rank		
	EW	UW	PCA	EW	UW	PCA
Johor	0.63	0.64	0.66	4	4	4
Kedah	0.74	0.75	0.78	2	3	2
Kelantan	0.74	0.76	0.73	3	2	3
Melaka	0.47	0.49	0.49	11	11	11
N. Sembilan	0.48	0.50	0.50	10	10	10
Pahang	0.52	0.54	0.54	8	7	8
Perak	0.55	0.55	0.58	6	6	6
Perlis	0.44	0.46	0.46	12	12	12
Pulau Pinang	0.50	0.51	0.52	9	9	9
Sabah	0.76	0.77	0.78	1	1	1
Sarawak	0.63	0.63	0.64	5	5	5
Selangor	0.53	0.54	0.55	7	8	7
Terengganu	0.43	0.46	0.45	13	13	13
Kuala Lumpur	0.27	0.24	0.27	14	14	14

Similar to the exposure index and susceptibility index, the calculated Cronbach's alpha for the various weighting method for the resilience index is 0.99 indicating high reliability. The average shift in rank is tested for the equal weighting, unequal weighting and principal component analysis and was found to be 0.00, 0.29 and 0.00 respectively.

3.5 Computation of FVI

Having tested the reliability of the three factor indices, the next step involved aggregating the factor indices to an overall composite index. As explained in the methodology, the additive mean and geometric mean were used to aggregate the exposure index (EI), susceptibility index (SI) and resilience index (RI) to compute the flood vulnerability index (FVI).

Table 6 presents the computed FVI scores based on two aggregation methods and three weighting methods. Using the additive mean, each factor is seen to equally contributing to the state's overall vulnerability, unlike the geometric mean which lowers the overall vulnerability index when one of the factor indices has a low score.

Table 6: FVI scores of the various computation methods

State	FVI Score					
	Additive mean			Geometric mean		
	EW	UW	PCA	EW	UW	PCA
Johor	0.39	0.39	0.40	0.35	0.34	0.36
Kedah	0.45	0.45	0.46	0.36	0.36	0.36
Kelantan	0.66	0.65	0.66	0.65	0.64	0.65
Melaka	0.29	0.29	0.29	0.20	0.20	0.20
N. Sembilan	0.30	0.31	0.30	0.20	0.21	0.20
Pahang	0.47	0.46	0.48	0.46	0.46	0.48
Perak	0.38	0.39	0.39	0.35	0.35	0.35
Perlis	0.29	0.30	0.30	0.19	0.19	0.20
Pulau Pinang	0.36	0.36	0.37	0.35	0.35	0.35
Sabah	0.55	0.53	0.54	0.51	0.50	0.51
Sarawak	0.55	0.55	0.55	0.52	0.51	0.52
Selangor	0.39	0.39	0.39	0.37	0.36	0.36
Terengganu	0.46	0.45	0.48	0.46	0.45	0.48
Kuala Lumpur	0.26	0.26	0.26	0.26	0.26	0.26

Table 7 shows the FVI rankings based on various computation methods. Kelantan ranks the first, followed by Sarawak (2nd) and Sabah (3rd) under all scenarios. The difference in ranking changes between 1 to 2 positions for Pahang (from 4th to 5th), Terengganu (from 4th to 5th), Kedah (from 5th to 7th), Perak (from 8th to 10th), Pulau Pinang (from 9th to 10th), Melaka (from 12th to 13th) and Perlis (from 12th to 14th). The biggest difference in rankings is observed for Selangor (from 6th to 9th), Johor (from 7th to 10th), Negeri Sembilan (from 11th to 14th) and Kuala Lumpur (from 11th to 14th).

Table 7: FVI rankings of the various computation methods

State	FVI Rank					
	Additive mean			Geometric mean		
	EW	UW	PCA	EW	UW	PCA
Johor	7	9	7	8	10	8
Kedah	6	5	6	7	6	6
Kelantan	1	1	1	1	1	1
Melaka	13	13	13	12	13	12
N. Sembilan	11	11	11	12	12	14
Pahang	4	4	4	4	4	5
Perak	9	7	8	10	8	10
Perlis	12	12	12	14	14	13
Pulau Pinang	10	10	10	9	9	9
Sabah	3	3	3	3	3	3
Sarawak	2	2	2	2	2	2
Selangor	8	8	9	6	7	6
Terengganu	5	5	5	4	5	4
Kuala Lumpur	14	14	14	11	11	11

The difference between aggregation using additive mean and geometric mean is substantial. The compensability is constant in additive aggregation, whereas under the geometric aggregation, the lower the values of the factor indices, the lower the vulnerability. Test of sensitivity and reliability could help with interpretation of the results.

3.6 FVI Sensitivity and Reliability Analysis

Table 8 shows the average shift in vulnerability rankings from the median rank. The statistics comprise the relative shift in the position of all states in a single number. The value of $\bar{R}_s$ closer to zero means the more similar is the ranking to the median ranking. The use of additive mean using equal weighting indicates the smallest difference from the median rank. The average shift in rank using geometric mean is higher, this is probably caused by the partial compensability of the geometric mean which prevents high value of one vulnerability factor from compensating the very low value of vulnerability of the other factors.

Table 8: FVI rankings of the various computation methods

Aggregation method	Additive mean			Geometric mean		
	EW	UW	PCA	EW	UW	PCA
Average rank shift, $\bar{R}_s$	0.36	0.50	0.43	0.71	0.50	0.79

Table 9 presents the Spearman rank order correlation between the different methods of weighting and aggregation method. The correlation coefficient ranges between 0.92 to 0.99 indicating very high positive correlation among the different methodology. There is significant agreement among the ranking of the states on the FVI since all coefficients are significant at $p < 0.01$. Lastly, Cronbach's alpha shows the flood vulnerability ranking has an excellent reliability of 0.99.

Table 9: Spearman rank-order correlation for various weighting and aggregation methods

Correlation matrix of derived FVI rankings						
	EW_a	UW_a	PCA_a	EW_g	UW_g	PCA_g
EW_a	1					
UW_a	0.98	1				
PCA_a	0.99	0.98	1			
EW_g	0.95	0.92	0.93	1		
UW_g	0.94	0.96	0.94	0.97	1	
PCA_g	0.94	0.92	0.92	0.98	0.96	1

Note: Coefficients are all significant at the p < 0.01 level (two tailed). EW – equal weight, UW – unequal weight, PCA – weight derived from principal component analysis, a – additive mean, g – geometric mean

Within these analyses, the robustness of the ranks was tested through a methodological approach. Overall, the ranks determined using the equal weighting and additive aggregation method lies closely to the median rank and the narrow confidence interval implies reliability of those ranks. Figure 5 presents the validation results of the ranks derived using the additive aggregation with equal weight method, median ranks and the 95% confidence interval.

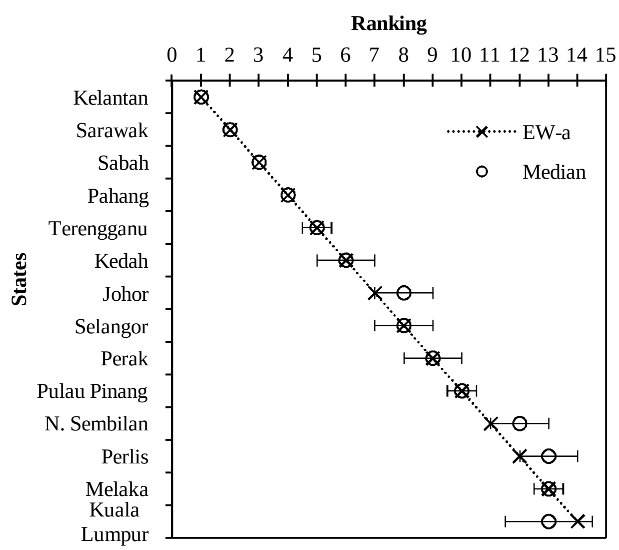

Figure 5: Flood vulnerability ranks with uncertainty consideration

The null hypothesis is there is no significant agreement among the rankings of the states based on the FVI. The null hypothesis was rejected at 99% confidence level and there is statistically significant agreement among the ranks despite different weighting techniques and aggregation methods were employed. The confidence level for the FVI was validated by comparing the ranks with median ranks and also the 95% confidence interval.

4. FUTURE WORKS

We have presented the results of the flood vulnerability index using a methodological approach. While spreadsheet software was used to organise, store and analyse the data, it has limitations in managing the large data sets. Future works concerns improving the database by using Structural Query Language (SQL) which is a standard language for querying and editing information, as in the case of [37] who developed a spread-sheet based user interface to perform data query and update data.

5. CONCLUSION

Our analyses put forward a first attempt to quantify flood vulnerability assessment in Malaysia, hence the selection of indicators may not be comprehensive since indicators were selected based on relevance and data availability. We have quantitatively evaluated the ranking of flood vulnerability of the states in Malaysia through its three factors of exposure, susceptibility and resilience. We assessed the relevance of 21 socio-economic and environmental indicators to reflect these three factors of vulnerability. The flood vulnerability index was constructed using three weighting techniques and two aggregation techniques.

On the basis of the results, we concluded that the FVI scores and ranks do not vary much due to different weighting and aggregation techniques. In this study, we found that the ranking of the states is not unduly affected by assigning equal weightage and additive aggregation method. The results of this study provided valuable knowledge about the current state of vulnerability of the states in Malaysia to flood. The results also provided a baseline for further flood vulnerability assessments with the inclusion of new indicators.

APPENDIX

Indicators		Min	Mean	Max	SD
E1	Flood prone area (%)	1.9	9.6	22.6	7.2
E2	Population in flood prone area (%)	4.6	22.1	55.4	17.5
E3	Population density in flood prone area (persons/km^2)	44	1,226	11,935	3,107
E4	Average annual damage (RM million)	2.29	65.36	157.65	50.28
E5	Average maximum daily rainfall (mm)	87	231	537	148
E6	Flood frequency	7	23	86	21
E7	Flood water depth (m)	0.5	1.9	7.0	1.7
S1	Children below 15 (%)	19.7	24.4	30.5	3.0
S2	Elderly above 65 (%)	3.2	6.8	9.9	1.7
S3	Disabled persons (%)	0.7	1.6	2.3	0.4
S4	Inequality (Gini)	0.324	0.363	0.404	0.027
S5	Houses with poor building materials (%)	0.8	20.5	46.0	15.3
S6	Agricultural workers (%)	0.0	7.1	36.1	9.7
S7	Illiterate population (%)	1.2	2.8	7.4	1.7
R1	Household median income (RM)	3,079	4,969	9,073	1,569
R2	GDP per capita (RM)	12,812	36,560	101,420	21,471
R3	Emergency services (hospital beds per 10,000 population)	8.8	15.4	27.3	4.5
R4	Volunteers (per 10,000 population)	476	1,095	1,794	379
R5	Evacuation shelter (capacity/10,000 population)	64	667	1,479	471
R6	Potable water supply (%)	65.4	94.7	100.0	10.0
R7	Internet access (%)	70.0	90.1	79.4	6.4

ACKNOWLEDGEMENT

D.Y.C Liew thanks Public Service Department Malaysia for awarding a scholarship to undertake this master's study and Ministry of Education through Fundamental Research Grant Scheme [FRGS/1/2018/TK01/UTM/02/17].

REFERENCES

1. A. Sumner and S. Tezanos Vázquez. **How has the developing world changed since the late 1990s? A dynamic and multidimensional taxonomy of developing countries**, *Center for Global Development (CGD) Working Paper 375*, Washington, DC, 2014. https://doi.org/10.2139/ssrn.2476558
2. J. Birkmann. **Risk and vulnerability indicators at different scales: Applicability, usefulness and policy implications**, *Environ. Hazards*, vol. 7, pp. 20–31, 2007. https://doi.org/10.1016/j.envhaz.2007.04.002
3. S. Balica and N. G. Wright. **Reducing the complexity of the flood vulnerability index**, *Environ. Hazards*, vol. 9, pp. 321–339, 2010. https://doi.org/10.3763/ehaz.2010.0043
4. A. E. Kissi, G. A. Abbey, K. Agboka, and A. Egbendewe. **Quantitative assessment of vulnerability to flood hazards in downstream area of Mono Basin, South-Eastern Togo: Yoto District**, *J. Geogr. Inf. Syst.*, vol. 7, pp. 607–619, 2015. https://doi.org/10.4236/jgis.2015.76049
5. J. S. Lee and H. I. Choi. **Comparison of flood vulnerability assessments to climate change by construction frameworks for a composite indicator**, *Sustainability*, vol. 10, no. 3, p. 768, Mar. 2018. https://doi.org/10.3390/su10030768
6. J. Birkmann. **Measuring vulnerability to promote disaster-resilient societies : Conceptual frameworks and definitions**, in *Measuring vulnerability to natural hazards: Towards disaster resilient societies*, J. Birkmann, Ed. Tokyo: United Nations University Press, 2006, pp. 9–54.
7. H. Nasiri, M. J. Mohd Yusof, and T. A. Mohammad Ali. **An overview to flood vulnerability assessment methods**, *Sustain. Water Resour. Manag.*, vol. 2, no. 3, pp. 331–336, 2016. https://doi.org/10.1007/s40899-016-0051-x
8. OECD. **Handbook on constructing composite indicators: Methodology and user guide**. Paris:

Organisation for Economic Co-operation and Development, 2008.
9. M. Nardo and M. Saisana. **OECD/JRC Handbook on constructing composite indicators. Putting theory into practice**, in *New Techniques and Technologies for Statistics Conference*, 2009, pp. 1–16.
10. M. Saisana and A. Saltelli. **Rankings and ratings: Instructions for use**, *Hague J. Rule Law*, vol. 3, no. 2, pp. 247–268, 2011. https://doi.org/10.1017/S1876404511200058
11. R. F. Connor and K. Hiroki. **Development of a method for assessing flood vulnerability**, *Water Sci. Technol.*, vol. 51, no. 5, pp. 61–67, 2005. https://doi.org/10.2166/wst.2005.0109
12. S. F. Balica, N. Douben, and N. G. Wright. **Flood vulnerability indices at varying spatial scales**, *Water Sci. Technol.*, vol. 60, no. 10, pp. 2571–2580, 2009. https://doi.org/10.2166/wst.2009.183
13. DOSM. **Statistics yearbook Malaysia 2016**, Deparment of Statistics, Malaysia, 2017.
14. DID. **National register of river basins: Updating of condition of flooding in Malaysia**, Department of Irrigation and Drainage Malaysia, 2003.
15. D. Y. C. Liew, R. Ahmad, M. N. Muhamad Yatim, and F. Che Ros. **A composite indicator aproach to assess flood vulnerability in Malaysia**, in *4th International Conference on Water Resources*, Langkawi, 2018. Unpublished.
16. L. Hudrliková. **Composite indicators as a useful tool for international comparison: The Europe 2020 example**, *Prague Econ. Pap.*, vol. 4, pp. 459–473, 2013. https://doi.org/10.18267/j.pep.462
17. G. Naumann, P. Barbosa, L. Garrote, A. Iglesias, and J. Vogt. **Exploring drought vulnerability in Africa: An indicator based analysis to be used in early warning systems**, *Hydrol. Earth Syst. Sci.*, vol. 18, no. 5, pp. 1591–1604, 2014. https://doi.org/10.5194/hess-18-1591-2014
18. B. Talukder, K. W. Hippel, and G. W. VanLoon. **Developing composite indicators for agricultural sustainability assessment: Effect of normalization and aggregation techniques**, *Resources*, vol. 6, no. 4, p. 66, 2017. https://doi.org/10.3390/resources6040066
19. O. Žurovec, S. Čadro, and B. K. Sitaula. **Quantitative assessment of vulnerability to climate change in rural municipalities of Bosnia and Herzegovina**, *Sustainability*, vol. 9, no. 7, p. 1208, 2017. https://doi.org/10.3390/su9071208
20. N. H. Zardari, K. Ahmed, S. M. Shirazi, and Z. Yusop. **Weighting methods and their effects on multi-criteria decision making model outcomes in water resources management**. Cham: Springer, 2015. https://doi.org/10.1007/978-3-319-12586-2
21. N. S. Iyengar and P. Sudarshan. **A method of classifying regions from multivariate data**, *Econ. Polit. Wkly.*, vol. 17, no. 51, pp. 2047–2052, 1982.
22. M. Nardo, M. Saisana, A. Saltelli, and S. Tarantola. **Tools for composite indicators building**, European Communities, 2005.
23. S. Greco, A. Ishizaka, M. Tasiou, and G. Torrisi. **On the methodological framework of composite indices: A review of the issues of weighting, aggregation, and robustness**, *Soc. Indic. Res.*, 2018. https://doi.org/10.1007/s11205-017-1832-9
24. N. Brooks, W. N. Adger, and P. M. Kelly. T**he determinants of vulnerability and adaptive capacity at the national level and the implications for adaptation**, *Glob. Environ. Chang.*, vol. 15, no. 2, pp. 151–163, 2005. https://doi.org/10.1016/j.gloenvcha.2004.12.006
25. M. A. Fernandez, S. J. Bucaram, and W. Renteria. **Assessing local vulnerability to climate change in Ecuador**, *Springerplus*, vol. 4, no. 1, p. 738, 2015. https://doi.org/10.1186/s40064-015-1536-z
26. M. K. A. Kablan, K. Dongo, and M. Coulibaly. **Assessment of social vulnerability to flood in urban Côte d'Ivoire using the MOVE framework**, *Water*, vol. 9, no. 4, p. 292, Apr. 2017. https://doi.org/10.3390/w9040292
27. S. L. Cutter, B. J. Boruff, and W. L. Shirley. **Social vulnerability to environmental hazards**, *Soc. Sci. Q.*, vol. 84, no. 2, pp. 242–261, 2003. https://doi.org/10.1111/1540-6237.8402002
28. G. A. Gbetibouo, C. Ringler, and R. Hassan. **Vulnerability of the South African farming sector to climate change and variability : An indicator approach**, *Nat. Resour. Forum*, vol. 34, no. 3, pp. 175–187, 2010. https://doi.org/10.1111/j.1477-8947.2010.01302.x
29. L. Wiréhn, Å. Danielsson, and T. S. S. Neset. **Assessment of composite index methods for agricultural vulnerability to climate change**, *J. Environ. Manage.*, vol. 156, pp. 70–80, 2015. https://doi.org/10.1016/j.jenvman.2015.03.020
30. A. Das, A. Khan, P. Daspattanayak, S. Chatterjee, and M. I. Hassan. **Regional model for agricultural imbalances in West Bengal, India**, *Model. Earth Syst. Environ.*, vol. 2, no. 2, p. 58, 2016. https://doi.org/10.1007/s40808-016-0107-9
31. R. Sendhil, A. Jha, A. Kumar, and S. Singh. **Extent of vulnerability in wheat producing agro-ecologies of India: Tracking from indicators of cross-section and multi-dimension data**, *Ecol. Indic.*, vol. 89, pp. 771–780, 2018. https://doi.org/10.1016/j.ecolind.2018.02.053
32. H. F. Kaiser. **The application of electronic computers to factor analysis**, *Educ. Psychol. Meas.*, vol. 20, no. 1, pp. 141–151, 1960. https://doi.org/10.1177/001316446002000116
33. E. Tate. **Uncertainty analysis for a Social Vulnerability Index**, *Ann. Assoication Am. Geogr.*, vol. 103, no. 3, pp. 526–543, 2013. https://doi.org/10.1080/00045608.2012.700616
34. E. Mavhura, B. Manyena, and A. E. Collins. **An approach for measuring social vulnerability in context: The case of flood hazards in Muzarabani district, Zimbabwe**, *Geoforum*, vol. 86, pp. 103–117, 2017. https://doi.org/10.1016/j.geoforum.2017.09.008

35. J. D. Gibbons and S. Chakraborti. **Nonparametric Statistical Inference**, 4th ed. New York: Marcel Dekker, 2003.
36. L. J. Cronbach. **Coefficient alpha and the internal structure of tests**, *Psychometrika*, vol. 16, no. 3, pp. 297–334, 1951. https://doi.org/10.1007/BF02310555
37. S. Chatvichienchai and Y. Kawasaki. **SpreadDB: Spreadsheet-based user interface for querying and updating data of external databases**, *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 7, no. 2, pp. 6-10, 2018. https://doi.org/10.30534/ijatcse/2018/01722018