Empirical Study of a Spatial Analysis for Prone Road Traffic Accident Classification based on MCDM Method

Anik Vega Vitianingsih[1]
Informatics Departments
Universitas Dr. Soetomo
Faculty of Information and Communication Technology
Universiti Teknikal Malaysia Melaka
Surabaya, Indonesia. Melaka, Malaysia

Zahriah Othman[2], Safiza Suhana Kamal Baharin[3]
Faculty of Information and Communication Technology
Universiti Teknikal Malaysia Melaka, Melaka, Malaysia

Aji Suraji[4]
Department of Civil Engineering
University of Widyagama Malang, Malang, Indonesia

***Abstract*—Spatial analysis techniques are widely used as an effective approach for prone road traffic accident classification. This paper will present the results of empirical behavioral testing on the spatial analysis for prone road traffic accident classification using the Multicriteria Decision Making (MCDM) method. The performance of MCDM is compared on arterial and collector road types processed with multicriteria parameters. MCDM was chosen because it can be used as a decision making based on an alternative selection with many criteria. Empirical tests of the MCDM method used include Weighted Sum Model (WSM), Weighted Product (WP), Simple Additive Weighting (SAW), Weighted Product Model (WPM), Multi-Attribute Utility Theory (MAUT), Technique for Others Reference by Similarity to Ideal Solution (TOPSIS), and Analytical Hierarchy Process (AHP). The multicriteria parameter weight values are based on expert judgment and the Fuzzy-AHP method (EJ-AHP), which comprises volume-to-capacity ratio (VCR), international roughness index (IRI), vehicle type, horizontal alignment, vertical alignment, design speed, and shoulder. Then, the performance of the models was compared to determine the value of accuracy, precision, recall, and F1-score as decision-making on the prone road traffic accident classification using Multicriteria Evaluation Techniques (MCE). The empirical test results on arterial roads show that the SAW and TOPSIS methods have the same performance and are superior to other methods, with an accuracy value of 63%. However, the results on the collector road type show that the accuracy value of the AHP method outperforms other methods with an accuracy value of 70%.**

Keywords*—Spatial Analysis; GIS; prone road traffic accident; MCDM Model; WSM; WP; SAW; WPM; MAUT; TOPSIS; AHP

I. Introduction

The rate of road traffic accidents (RTA) that results in deaths increases every year. Data for 2004-20130 states that RTA is the leading cause of death, which is ranked 9th in the world; WHO estimates that in 2030 the RTA will increase to the 5th rank if there are no efforts to overcome this problem [1]. The number of deaths due to RTA annually reached 1.35 million worldwide in 2016 [2].

The case study for the spatial analysis for prone road traffic accident classification in the discussion of this paper is on the type of arterial and collector roads in the Province of East Java, Indonesia, which is one of the areas with a very high accident-prone. The Global Status Report on Road Safety 2018 states that in Indonesia, with a population number is 261,115,456 people in 2016, the number of deaths due to RTA reached 31,282 million people [2]. The accident factors include 69.70% due to the human factor, 21.21% due to road facilities, and 9.09% due to road infrastructure factor (Komite Nasional Keselamatan Transportasi, 2016). In 2010 the United Nations General Assembly declared a Decade of Action for Road Safety year 2011-2020 aimed at stabilizing the level of fatality of global casualties by increasing activities undertaken at national, regional, and global scales [1][3][4] The spirit of the Road Safety Action Declaration 2011-2020 is in line with the mandate of Law Number 38 of 2004 [5], Number 34 of 2006 [6] concerning roads, and Law Number 22 the Year 2009 concerning road traffic and transportation [7] to prepare a National General Plan for Road Safety 2011-2035 in Indonesia [8] as outlined in and Regulation of the President of the Republic of Indonesia Number 2 of 2012 concerning the national transportation safety committee [9].

Spatial data modeling (GIS-spatial analysis) is a part of multicriteria decision-making (GIS-MCDM). Spatial analysis in geographic information systems (GIS) is the process of developing artificial intelligence (AI) formulations by combining geo-referenced data (spatial data) with multicriteria parameters as value assessment attribute data (decision-makers preferences and uncertainty) to obtain the appropriate information in georeferencing-based decision making. GIS is commonly regarded as a technology capable of integrating, storing, manipulating, analyzing, and displaying spatial data and attribute data for decision-making and decision-supporting operations [10]. The MCDM method provides a collection of procedures and AI algorithms for formulating decision-making problems, designing, evaluating, and prioritizing alternative decisions [11][12]. This empirical study aims to analyze the sensitivity of the methods tested through spatial data modeling with MCE. MCE evaluates the methods [13] by testing the extent to which the values of accuracy, precision, recall, and F1 scores when multicriteria parameters systematically vary on various interests.

The characteristic of GIS-Spatial MCDM is to determine the weighting of the spatial datasets used for spatial data modeling. The literature study [14]–[18] summarizes several issues related to GIS-spatial relationship modeling for prone-roads classification traffic accidents (PRTA) using the MCDM method. First, GIS-Spatial relationship modeling using multicriteria decision-making methods (GIS-Spatial MCDM) is a spatial analysis process to perform spatial data modeling that involves multicriteria parameters from the expert judgment in spatial decision making. The spatial analysis involves multicriteria parameters in building software GIS for spatial decision-making based on combined theory, methods, and measurement tools from expert judgment [19][20]. An expert judgment is required to validate the spatial dataset parameters used [19].

Many researchers give parameter-weighted values only from the point of view of expert judgment. The expert judgments give a subjective and objective risk and bias in the evaluation process for weighting and parameter priority scale [21]. The weighted value given to the multicriteria parameter will impact the accuracy of the spatial data modeling results for PRTA classification [22][23]. Many researchers claim that the multicriteria parameters used are effective and capable of determining the PRTA classification [24]–[32].

Secondly, classification techniques are needed to produce accurate spatial data modeling without overlapping interests and to avoid overfitting problems, and the deep-neuro-fuzzy classification method is used for road weight measurement [33]. The analytical hierarchy process (AHP) method can improve the road safety audit technique used to identify and prioritize black spots in the absence of statistical data of accidents that were not recorded correctly. [34]. The AHP method is used to perfect the weighting value generated from literature studies and expert assessments [31][35] by determining the priority scale ranking of the parameters using the random forest (RF) [31] method, preference ranking organization method for enrichment of evaluations (PROMETHEE), and VlseKriterijuska Optimizacija I Komoromisno Resenje (VIKOR) method [35].

The literature study shows that the MCDM method such as SAW and AHP methods and Fuzzy AHP for determining the weight can help decisions process in Road Safety Analysis (RSA) such as road management prioritization and provide mitigating actions against the most vulnerable to accidents. In another literature study, the TOPSIS classification model is used to manage road safety to reduce the number of traffic accidents by knowing the position of a road safety study based on various quantitative and qualitative criteria [36]. Besides that, the simple ranking (SR) method and the empirical Bayes (EB) combine the type and severity of the accident to the data series in Australia, then proposed to evaluate alternative indicators with multiple criteria parameters for the identification of accident-prone road (blackspot). The SR and EB method is used to calculate the value of accidents by type of case, societal cost of any accident, and the crash prediction models using data series [37].

GIS is region specific [38][39], where 96% [14] of research uses private spatial datasets with small dataset characteristics. The challenge for small datasets is at the data pre-processing stage to produce optimal performance on the AI method used. The GIS-Spatial MCDM approach is proposed in this study based on the characteristics of the MCDM model based on multi-criteria parameter weighting. The weighting that has been carried out by expert judgment will be combined with the AHP computational method (EJ-AHP) as a means of measuring the resulting weight value.

Based on the review of these literature studies, however, no research studies specifically for evaluating and comparing through an empirical study approach in the spatial analysis using the MCDM method (WSM, WP, SAW, WPM, MAUT, TOPSIS, and AHP methods) for prone road traffic accidents. The classification of multicriteria parameter weight values is based on expert judgment and the AHP method (EJ-AHP). Therefore, this study proposed a combination of expert judgment and Fuzzy-AHP (EJ-Fuzzy-AHP) to produce weighting values in spatial datasets and provide the appropriate parameter priority scale values. Fuzzy-AHP has a procedure following decisions that involve expert judgment, so it can be used to combine data knowledge in Fuzzy-AHP with expert judgment. Then, these weighting values were used in the MCDM method for GIS-based spatial modeling for PRTA classification based on multicriteria parameters, namely speed design, volume/capacity ratio (V/C Ratio) [40], the width of the road, the number of lanes, road shoulders, median strip, horizontal alignment, vertical alignment, road condition [40], and vehicle type.

The discussion structure in this paper includes section II: which discusses multicriteria parameters through the description of spatial datasets, section III: which discusses research methodology; section IV: which describes results and discussion; and section V: which discusses the conclusion and future works directions.

II. Spatial Datasets

The spatial dataset parameters in this study were obtained from private data sources, so a decision-making model using GIS-Spatial MCDM is proposed. The GIS decision-making system is used for specific regions case studies in which 96% of researchers use private data types on specific regional case studies [14] with multicriteria parameters from the expert judgments. The MCDM method is applied to making decisions through management priority ranking related to existing or specific region-specific planning policies [41]. The MCDM method is one of the right approaches to deal with the problem of the PRTA classification because it uses several road and environmental criteria, both quantitative and qualitative; MCDM is related to the results of decision making for planning that involves stakeholders [42].

According to the Republic of Indonesia Law No.38 of 2004 article 8, the type of road based on its function is divided into 4 (four), namely arterial roads, collector roads, local roads, and environmental roads. The arterial road is a public road with the number of access roads is limited efficiently, works to serve the main transport which connects provincial capitals with the characteristics of long-distance travel dan high average speeds [43]–[45]. The collector road is a public road that works to serve vehicle which connects between regency capitals with medium distance travel characteristics, medium average speed, and a limited number of entrances [43][44]. Local roads are public roads with an unlimited number of access roads that serve local transport, which link the sub-district cities with short-distance travel and low average speeds [43][44]. The environmental road is a public road that serves environmental transportation between villages with short-distance travel characteristics and low average speed [43][44].

In this study, the spatial analysis datasets include the types of primary arterial networks and primary collector networks since both types of roads are the main roads that supply sufficient datasets for this study. The case study involves the research objects in Indonesia with data retrieval from the National Road Development Center, Police Corps and Traffic Police of Indonesia, Traffic Corps National Police, and Transportation Department. Descriptions of the spatial datasets on the multicriteria parameters used are shown in Table I were to the range and score for Spatial Datasets from:

- The Directorate General of Highways Standard Specifications for Geometric Design of Urban Roads. Ministry of Public Works, Directorate General of Highways, Jakarta 1992.
- Indonesian Highway Capacity Manual (IHCM), 1997
- TRB Highway Capacity Manual. Transportation Research Board Special Report 209; Washington D.C. USA 1985. Revised 1994.
- SNRA Manual on Calculation of Capacity, Queues, and Delay in Traffic Facilities (in Swedish). Swedish National Road Administration Report TV 131, 1977.

Geospatial Datasets comprised spatial data needs for a base map (layer) and attribute data requirements for multicriteria parameters that were utilized for spatial analysis of PRTA classification. The data requirements used in this study used private data types from the National Road Implementation Center, East Java Bali, Indonesia. Spatial datasets include:

1) The base map: consists of attributes road number, suffix, road names, length of roads (km), and road function.

a) Arterial primary networks

b) Collector primary networks

2) Multicriteria parameters

a) Volume-to-capacity ratio (VCR): to measure the overall service quality provided. If the VCD is high, it indicates a high risk of accidents.

b) International Roughness Index (IRI). Condition of the pavement. If the IRI is heavily damaged, the likelihood of an accident increases significantly.

c) Vehicle Type: vehicle types 2/1 UD, 2/2 UD, 4/2 UD, 4/2 D, and 6/2 can pass through arterial or principal collector roads.

d) Horizontal alignment (HA). Projection of the axis of the road for roads without a median or the projection of the inner edge of the pavement for roads with a median. If the horizontal alignment is sharp, the potential for accidents is high.

e) Vertical alignment (VA): the intersection of the vertical plane with the road pavement surface through the road axis for 2-speed 2-way roads or through the inner edge of each pavement for roads with a median. If the vertical alignment is high, the potential for accidents is high.

f) Design speed (Vr): The vehicle speed can be achieved safely when running without interruption. If the speed is high, then the accident potential is high.

g) Shoulder: The lane is located side by side with the traffic lane. If the shoulder there isn't, then the potential for an accident is high.

TABLE I. SPATIAL DATASETS PARAMETERS

Arterial Road			
Parameters	**Range**	**Description**	**EJ Scoring**
VCR (%)	VCR ≥ 0.85 && VCR < 1.00	The condition reaches capacity with 2000 units of passenger cars (pcu/hour), 2 directions.	5
	VCR ≥ 0.70 && VCR < 0.85	The conditions approach unsteady flow with traffic volume reaching 85% of the capacity, namely 1700 units of passenger cars (pcu/hour), 2 directions.	4
	VCR ≥ 0.45 && VCR < 0.70	Traffic flow conditions are still stable, with traffic volume reaching 70% of capacity (pcu/hour), 2 directions	3
	VCR ≥ 0.20 && VCR < 0.45	The start of a stable flow condition with traffic volume reaching 45% of the capacity is 900 units of passenger cars (pcu/hour), 2 directions.	2
	VCR < 0.20	The conditions free flow with traffic volume reaching 20% of the capacity, namely 400 units of passenger cars (pcu/hour), 2 directions.	1
IRI (m/km)	IRI ≥ 12	Heavy Damage	4
	IRI ≥ 8 && IRI < 12	Light Damage	3
	IRI ≥ 4 && IRI < 8	Moderately	2
	IRI < 4	Good	1
HA(rad/km)	HA ≥ 3.50	Poor	3
	HA ≥ 0.25 && HA < 3.50	Fair	2
	HA < 0.25	Good	1
VA	VA ≥ 45	Poor	3

Arterial Road			
Parameters	**Range**	**Description**	**EJ Scoring**
(m/km)	VA ≥ 5 && VA < 45	Fair	2
	VA < 5	Good	1
Vr (km/jam)	Vr ≥100	Traffic speed more than 100 kilometres per hour.	6
	Vr ≥ 80 && Vr < 100	Traffic speed is more than 80 kilometers per hour.	5
	Vr ≥ 65 && Vr < 80	Traffic speed more than 65 kilometres per hour.	4
	Vr ≥ 60 && Vr < 65	The speed limit is reduced to 60 kilometers per hour.	3
	Vr ≥ 50 && Vr < 60	The average pace of traffic is approximately 50 kilometers per hour.	2
	Vr < 50	Traffic moving at a speed of fewer than 50 kilometers per hour.	1
Road Type	2/2 UD	The traffic road is a two-lane two-way without a median (2/2 UD)	5
	4/2 UD	The traffic road is a four-lane two-way without a median (4/2 UD)	4
	4/2 D	The traffic road is four lanes two-way with a median (4/2 D).	3
	6/2 D	The traffic road has six two-way lanes with a median (6/2 D).	2
	2/1 UD	The traffic road has two lanes with no median (2/1 UD).	1
Shoulder	No	No, there is no roadside shoulder.	2
	Yes	Yes, the roadside shoulder is available.	1

Collector Road			
Parameters	**Range**	**Description**	**EJ Scoring**
VCR (%)	VCR ≥ 0.90 && VCR < 1.00	The condition reaches capacity with 2000 units of passenger cars (pcu/hour), 2 directions.	5
	VCR ≥ 0.75 && VCR < 0.90	The conditions approach unstable flow with traffic volume reaching 90% of the capacity, namely 1800 units of passenger cars (pcu/hour), 2 directions.	4
	VCR ≥ 0.50 && VCR < 0.75	Traffic flow conditions are still stable, with traffic volume reaching 75% of capacity (pcu/hour), 2 directions	3
	VCR ≥ 0.30 && VCR < 0.50	The start of a stable flow condition with traffic volume reaching 50% of the capacity is 1000 units of passenger cars (pcu/hour), 2 directions.	2
	VCR < 0.30	The conditions free flow with traffic volume reaching 30% of the capacity, namely 600 units of passenger cars (pcu/hour), 2 directions.	1
IRI (m/km)	IRI ≥ 12	Heavy Damage	4
	IRI ≥ 8 && IRI < 12	Light Damage	3
	IRI ≥ 4 && IRI < 8	Moderately	2

Arterial Road			
Parameters	**Range**	**Description**	**EJ Scoring**
	IRI < 4	Good	1
HA (rad/km)	HA ≥ 3.50	Poor	3
	HA ≥ 0.25 && HA < 3.50	Fair	2
	HA < 0.25	Good	1
VA (m/km)	VA ≥ 45	Poor	3
	VA ≥ 5 && VA < 45	Fair	2
	VA < 5	Good	1
Vr (km/jam)	Vr ≥100	Traffic speed more than 100 kilometres per hour.	6
	Vr ≥ 90 && Vr < 100	Traffic speed is more than 80 kilometers per hour.	5
	Vr ≥ 75 && Vr < 90	Traffic speed is more than 65 kilometers per hour.	4
	Vr ≥ 60 && Vr < 75	The speed limit is reduced to 60 kilometers per hour.	3
	Vr ≥ 50 && Vr < 60	The average pace of traffic is approximately 50 kilometers per hour.	2
	Vr < 50	Traffic moving at a speed of fewer than 50 kilometers per hour.	1
Road Type	2/2 UD	The traffic road is a two-lane two-way without a median (2/2 UD)	5
	4/2 UD	The traffic road is a four-lane two-way without a median (4/2 UD)	4
	4/2 D	The traffic road is four lanes two-way with a median (4/2 D).	3
	6/2 D	The traffic road has six two-way lanes with a median (6/2 D).	2
	2/1 UD	The traffic road has two lanes with no median (2/1 UD).	1
Shoulder	No	No, there is no roadside shoulder.	2
	Yes	Yes, the roadside shoulder is available.	1

III. RESEARCH METHODOLOGY

The proposed MCDM experiment procedure in Fig. 1 has major differences from the existing framework [46]–[48]. Fig. 1 describes the proposed MCDM experiment procedure, namely:

- The requirement gathering a primary data set as a base map to determine the category of roads to be studied. This research uses private data types. The base maps used include primary arterial and primary collector networks.
- Attribute data for the multicriteria parameters used is based on an assessment by expert judgment, including VCR, IRI, vehicle type, horizontal alignment, vertical alignment, design speed, and shoulder. The data requirements are described in Table II.
- Conduct a literature study related to the multicriteria parameters used in each road category based on expert judgment assessment with the results in Table II.

- Mathematical modeling for spatial data analysis through empirical study for PRTA classification based on the MCDM method using WSM, WP, SAW, WPM, MAUT, TOPSIS, and AHP methods. In this case, the data pre-processing process will be carried out for the classification analysis process, that is:
  - Determine the priority weight of the parameters using the AHP method.
  - Determine the multiclass classification range obtained from the final value of the results of mathematical modeling on the MCDM method using the Guttman Scale. This process is carried out because there is no standardized assessment from expert judgment regarding the value of the PRTA classification range based on the multicriteria parameters used.
- The results of the multiclass classification will be validated through the value of accuracy and F1 score, which the F1 score is derived from the precision and recall.

A. *The Priority Weight of the Parameters*

Spatial decision-making based on multicriteria parameters is almost always faced with the problem of determining the level of importance or influence between parameters. Decision-makers will weigh each parameter based on the importance or influence between these variables, which is usually done by expert judgment. The AHP method can solve the complex multicriteria parameter problems into a hierarchical unit. The hierarchy represents a complex problem in a multilevel structure, where the first level is the goal, followed by the factors level, criteria, sub-criteria, and the last level of alternatives. With a hierarchy, complex problems can be described in groups which are then arranged into a hierarchical form so that problems will appear more structured and systematic.

How to overcome the biases from the weighting given by expert judgment overdue of various factors of interest, then the decision-maker can perform parameter weighting using the AI method. The AHP is a pairwise comparison method through an analytic hierarchy process. The parameter weights are determined by normalization through the eigenvectors associated with the maximum eigenvalues in the unit ratio matrix. The weighting between parameters in this study is accomplished by the AHP method approach based on flow depicted in Fig. 2.

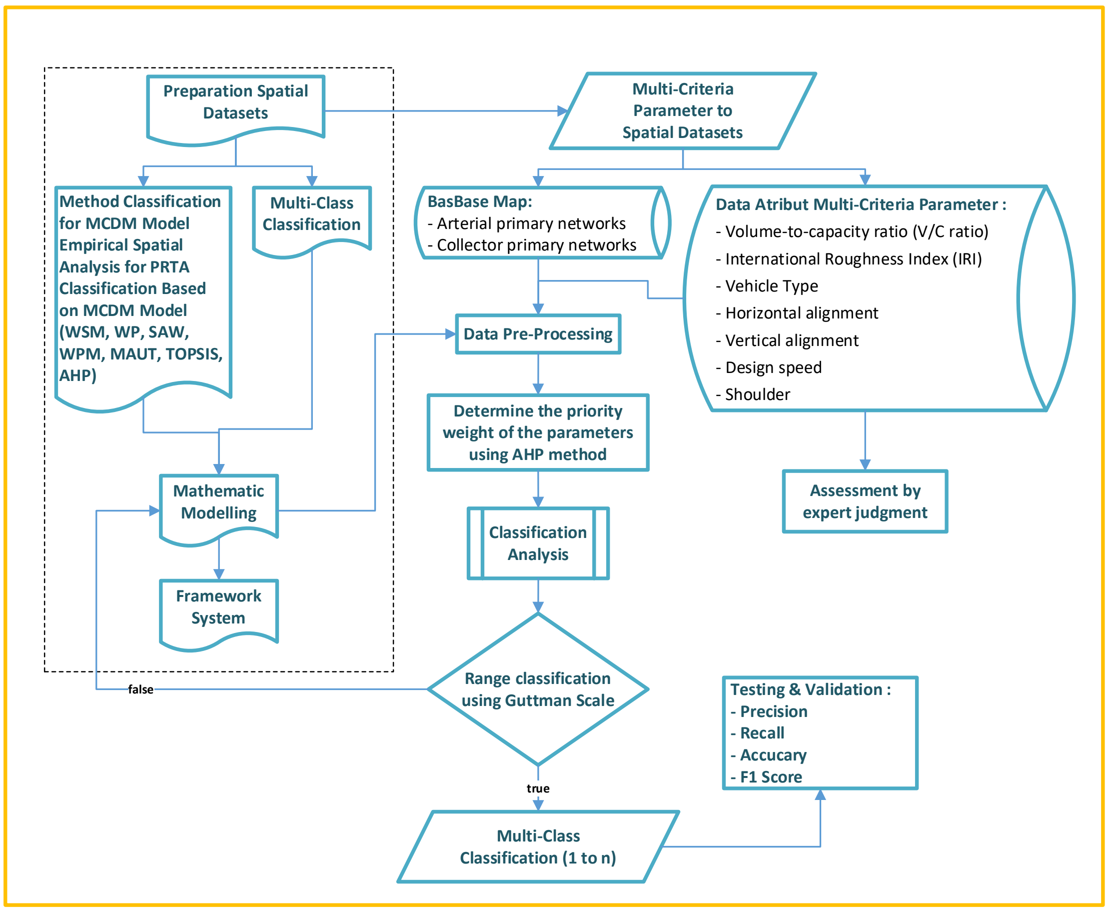

Fig. 1. Proposed MCDM Experiment Procedure.

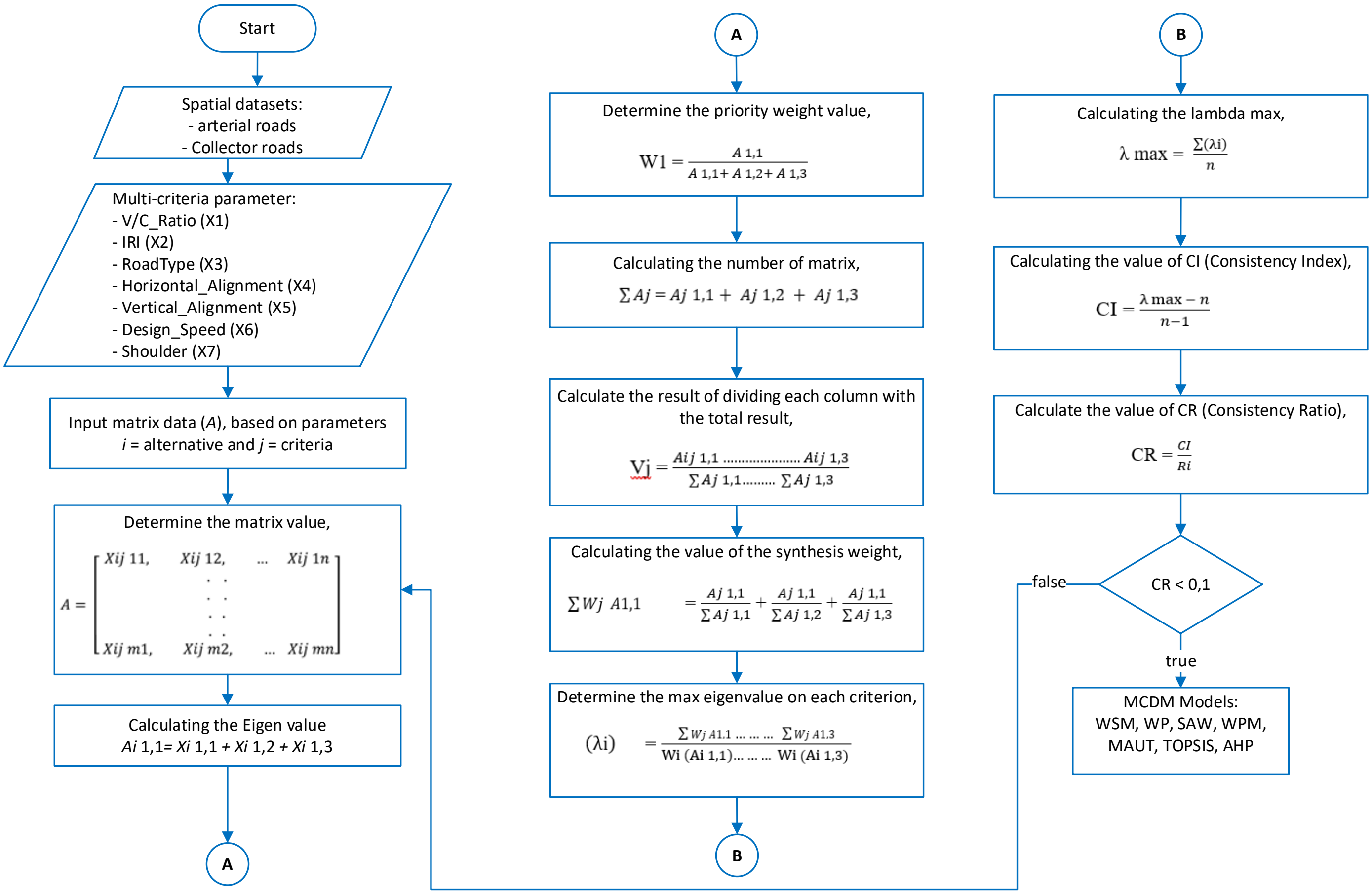

Fig. 2. The Priority Weight of the Parameters using the AHP Method.

Which consists of the following:

Step 1: Input the spatial datasets based on the base map used, namely the arterial and collector roads network.

Step 2: Input data multicriteria parameters, provide the labeling of the parameters used, namely VCR (X1), IRI (X2), Road Type (X3), HA (X4), VA (X5), Vr (X6), and Shoulders (X7).

Step 3: Input matrix data (A), based on parameters. Where i variable is an alternative, and j variable is criteria.

Step 4: Determine the multicriteria parameters matrix value (A) using pairwise comparison based on Eq. (1). A value is assigned to each criterion in accordance with the specifications of the hierarchical structure based on the number of multicriteria parameters present.

$$A = \begin{bmatrix} Xij\ 11, Xij\ 12, \dots\ Xij\ 1n \\ \cdot\cdot \\ \cdot\cdot \\ \cdot\cdot \\ \cdot\cdot \\ Xij\ m1, Xij\ m2, \dots\ Xij\ mn \end{bmatrix} \quad (1)$$

The recommended values for creating a pairwise comparison matrix, where the values referring to Table III [49].

TABLE II. THE RECOMMENDED VALUES FOR CREATING A PAIRWISE COMPARISON MATRIX

Value	Description
1	Equally important (equal)
3	A little more important (slightly)
5	More importantly, with a strong type(strongly)
7	More importantly, with a very strong type (very strong)
9	More important to the extreme (extreme)

Step 5: Calculate the eigenvalues of each element in each pairwise comparison matrix. The eigenvalues are the weight of each element used to determine the priority of items in each hierarchical structure. Operate for adding values to each column in question to obtain the normalization of the matrix, based on Eq. (2).

$$Ai\ 1{,}1 = Xi\ 1{,}1 + Xi\ 1{,}2 + Xi\ 1{,}3 \dots Xi\ 1{,}n \quad (2)$$

Step 6: Calculate the priority weight value of the parameter (Wi) using Eq. (3), adding up each column's values in the pairwise comparison matrix, then dividing each value in the column by the total number of related columns obtain a normalized matrix. Where, $\sum$ Aij is the number of matrices,

$$Wi = \frac{Ai\ 1{,}1 \dots\dots\dots Ai\ 1{,}5}{\sum Aij} \quad (3)$$

Step 7: Calculate the result of divide each column by the result of the total number (Vj) using Eq. (4), where $\sum \text{Aj} = \text{Aj } 1{,}1 \ \dots\dots\dots \ \text{Aj } 1{,}5$. Addition the values of each row and divide them with the number of elements to get the average value.

$$\text{Vj} = \frac{\text{Aij } 1{,}1 \ \dots\dots\dots\dots \ \text{Aij } 1{,}5}{\sum \text{Aj } 1{,}1 \dots\dots\dots \sum \text{Aj } 1{,}5} \tag{4}$$

Calculate the value of the synthesis weight (Wj) to j using Eq. (5). Where $\sum$ Wj is the result of adding Vj to calculate the weight value of the synthesis.

$$\sum \text{Wj} = \text{Vj } 1{,}1 \ \dots\dots\dots.. \text{Vj } 1{,}5 \tag{5}$$

Step 8: Determine the max eigenvalue (λi) on each criterion using Eq. (6). Where, $\sum$(λi) is (λi) A1,1 … … … (λi) A1,5.

$$(\lambda \text{i}) = \frac{\sum \text{Wj A1,1} \dots\dots\dots \sum \text{Wj A1,5}}{\text{Wi (Ai 1,1)} \dots\dots\dots \text{Wi (Ai 1,5)}} \tag{6}$$

Calculate the Lambda max (λ max) using Eq. (7). Where, $\sum$(λi) is the total value of the sum of the eigenmax max, and n variable is the number of criteria.

$$\lambda \text{ max} = \frac{\sum(\lambda \text{i})}{\text{n}} \tag{7}$$

Step 9: Check the consistency of the hierarchy by calculating the value of the consistency ratio using the consistency index (CI) using Eq. (8), where:

- If the CI value 10%, then the consistency ratio is correct
- If the CI value is > 10%, then the consistency ratio is wrong, so data assessment must be corrected and reviewed.

$$\text{CI} = \frac{\lambda \text{ max} - \text{n}}{\text{n} - 1} \tag{8}$$

The consistency ratio (CR) value can be calculated using Eq. (9). Where Ri is a random index value determined by the hierarchy structure, as described in Table III, where:

a) If CR < 0.1 , then the level of consistency shown is quite rational in the pair comparison matrix.

b) If CR > 0.1, it indicates an inconsistent assessment of the pair comparison matrix.

$$\text{CR} = \frac{\text{CI}}{\text{Ri}} \tag{9}$$

B. Weighted Sum Model (WSM)

WSM is a simple method that is extensively used in decision-making on single-dimensional problems. Attribute normalization is done by altering the value of the numeric column in the data set to the same scale to obtain a balance on the overall attribute value. Spatial data modeling with the WSM method is an approach to determining the weight of the priority value of each parameter of the attribute parameter, then multiplying with the data of each attribute to take a high alternative value as a solution [50]–[53].

The following sequence of the steps spatial analysis process for the PRTA classification using the WSM method:

Step 1: Follow the steps in the flow in Fig. 2 to define the criteria used as a benchmark for solving the problem and determine the priority of parameter weights.

Step 2: Calculate the priority values for each layer dataset by using the matrix of Eq. (10) [51].

$$\text{A}_\text{i}^{\text{WSM}-\text{score}} = \sum_{\text{j}=1}^{\text{n}} \text{W}_\text{j}\text{X}_{\text{ij}}; \text{ for i} = 1,2,3, \dots \text{m} \tag{10}$$

Where, $\text{A}_\text{i}^{\text{WSM}-\text{score}}$ is potential WSM score, X_{ij} a variable is an alternative to the i data score based on the j relative weight criterion, and W_j a variable is the j relative weight criterion

Step 3: Determine the range of PRTA classification values using the Guttman scale based on Eq. (11).

$$\text{I} = \text{R} / \text{K} \tag{11}$$

Where the I variable is the interval range, the R variable is the result of the calculation of the highest scores value of A_i minus the lowest score value of A_i, and the K variable is the number of alternatives. The alternative assessment criteria for the PRTA classification are obtained from the result of the calculation of the highest scores value of A_i minus the value of I variable as shown in the result of Eq. (12).

$$\begin{cases} \text{PRTA, if } \text{A}_\text{i} \geq \text{the scale range} \\ \text{Non\_PRTA, if } \text{A}_\text{i} < \text{the scale range} \end{cases} \tag{12}$$

C. Weighted Product (WP)

The WP method is a decision support system that connects attribute ratings through multiplication operations to be raised to the power of the appropriate attribute weights. The normalization process to handle different units of measurement is done through multiplication operations on attribute ratings [54].

The following sequence of the steps spatial analysis process for the PRTA classification using WP method:

Step 1: Follow the steps in the flow in Fig. 2 to define the criteria used as a benchmark for solving the problem and determine the priority of parameter weights. Determine the initial and final input to change the name of the input into a rating value and determine the weight of each criterion. Improve the weights of each criterion by adding up the weights of each criterion, followed by dividing the result of the sum of the weights of the criteria by the starting weight of each criterion divided by the result of the sum of the weights of the criteria.

Step 2: Calculate the normalization value using Eq. (13) to get the alternative preference value of each criterion represented by the vector S_i. Where the S_i variable is the value of alternative preference, X_{ij} is the variable value of the alternatives on each attribute. The W_j variable is the value of the weight of the criteria, the n variable is the number of criteria, the i variable is an alternative value 1,2,..m, and the j variable is the criterion value.

$$\text{S}_\text{i} = \prod_{\text{j}=1}^{\text{n}} \text{x}_{\text{ij}}^{\text{w}_\text{j}} \tag{13}$$

Step 3: Determine the range of PRTA classification values using the Guttman scale based on Eq. (11). Where the I variable is the interval range, the R variable is the result of the calculation of the highest scores value of S_i minus the lowest score value of S_i , and the K variable is a number of alternatives. The alternative assessment criteria for the PRTA classification are obtained from the result of the calculation of the highest scores value of S_i minus the value of I variable as shown in the result of Eq. (14) .

$$\begin{cases} \text{PRTA, if } S_i \geq \text{the scale range} \\ \text{Non\_PRTA, if } S_i < \text{the scale range} \end{cases} \tag{14}$$

D. Simple Additive Weighting (SAW)

The SAW method is a multi-process method in spatial decisions making with multicriteria parameters. The SAW method performs a weighted summation of the performance ratings on each alternative attribute. The process of normalizing the decision matrix (X) to a scale that can be compared with all existing alternative ratings [55]. The advantage of the SAW method compared to the decision support system method that involves other multicriteria parameters lies in its ability to make a more precise assessment because it is based on the criteria value and the weight of the level of importance required.

The following sequence of the steps spatial analysis process for the PRTA classification using the SAW method:

Step 1: Follow the steps in the flow in Fig. 2 to define the criteria used as a benchmark for solving the problem and determine the priority of parameter weights.

Step 2: Perform normalization using Eq. 15 for each alternative value on each attribute by calculating the performance rating value.

$$r_{ij} = \begin{cases} \frac{x_{ij}}{\text{Max}_i\, x_{ij}}, \text{if j is benefit attribute} \\ \frac{\text{Min}_i\, x_{ij}}{x_{ij}}, \text{if j is cost attribute} \end{cases} \tag{15}$$

where r_{ij} variable is the normalized performance rating of alternative A_i on attributes C_j and j, Max X_{ij} variable is the greatest value of each criterion i, and Min X_{ij} variable is the smallest value of each criterion i, X_{ij} variable is the attribute values that each criterion has. If the largest value is the best, then it is included in the benefit attribute category. If the smallest value is the best, then it is included in the cost attribute category.

Step 3: Calculate the value of preference weight on each alternative (V_i) using Eq. (16). Where the V_i variable is the ranking for each alternative. the W_j variable is the ranking weight value of each criterion, and r_{ij} variable is the normalized performance rating value.

$$V_i = \sum_{j=1}^{n} w_j\, r_{ij} \tag{16}$$

Step 4: Determine the range of PRTA classification values using the Guttman scale based on Eq. (11). Where the I variable is the interval range, the R variable is the result of the calculation of the highest scores value of V_i minus the lowest score value of V_i , and the K variable is a number of alternatives. The alternative assessment criteria for the PRTA classification are obtained from the result of the calculation of the highest scores value of V_i minus the value of I variable as shown in the result of Eq. (17).

$$\begin{cases} \text{PRTA, if } V_i \geq \text{the scale range} \\ \text{Non\_PRTA, if } V_i < \text{the scale range} \end{cases} \tag{17}$$

E. Weighted Product Model (WPM)

Spatial data modeling with the WPM method is a process to determine the weight of the priority value on each attribute parameter criterion, perform weighting by dividing the attribute weights by the weight of all attributes to get the total value equal to 1, determining the total vector value S to produce the vector V in produce the highest value that will be used as an alternative selection [51]–[53] The WPM method can be used for MCDM single or multi-dimensional categories [56].

The spatial data modeling process for the PRTA classification using the WPM method. The following sequence of the steps is as follows:

Step 1: Follow the steps in the flow in Fig. 2 to define the criteria that will be used as a benchmark for solving the problem and determine the priority of parameter weights.

Step 2: Calculate the normalized decision matrix value using Eq. (18) to get the alternative preference value of each criterion represented by the vector S_i. Where the S_i variable is the value of alternative preference, X_{ij} is the variable value of the alternatives on each attribute, and the W_j variable is the value of the weight of the criteria, the n variable is the number of criteria, the i variable is an alternative value 1,2,..m, the j variable is the criterion value.

$$S_i = \prod_{j=1}^{n} x_{ij}^{w_j}, i = 1,2,3,..,m \tag{18}$$

The normalized decision matrix value is calculated in order to obtain the x_{ij} value by providing the *i-th* alternative performance rating value on the *j-th* sub-criteria in the normalized decision matrix value computation. Furthermore, the value of the performance rating is elevated to the relative weight value (w_j), where w_j will be positive for the benefit attribute and negative for the cost attribute, depending on the attribute being evaluated. The sum of the w_j values for each sub-criteria on the same criteria will be worth 1. The value of w_j is calculated using Eq. 4.19.

$$w_j = \frac{w_j}{\sum w_j} \tag{19}$$

Step 3: Calculate the relative preference value of each alternative V_i using Eq. 20. Where, V_i variable is the relative preference of each *i-th* alternative. x_{ij} variable is the criteria value for each alternative to the *i-th* and the criteria *j-th*. w_j variable is the weight of the criteria or sub-criteria and the *n* variable is the number of criteria.

$$V_i = \frac{\prod_{j=1}^{n} x_{ij}^{w_j}}{\prod_{j=1}^{n} x_j^{*\,w_j}}, i = 1,2,3, \dots, m \tag{20}$$

Step 4: Determine the range of PRTA classification values using the Guttman scale based on Eq. (11). Where the *I* variable is the interval range, the *R* variable is the result of the calculation of the highest scores value of V_i minus the lowest score value of V_i, and the *K* variable is a number of alternatives. The alternative assessment criteria for the PRTA classification are obtained from the result of the calculation of the highest scores value of V_i minus the value of *I* variable as shown in the result of Eq. (21).

$$\begin{cases} PRTA, if\ V_i \geq the\ scale\ range \\ Non\_PRTA, if\ V_i < the\ scale\ range \end{cases} \quad (21)$$

F. Multi-Attribute Utility Theory (MAUT)

Spatial data modeling with the MAUT method is to determine the value of $U(A_i)$With the weight value on each sub-criteria parameter and the priority value of each attribute parameter's interest, calculate the number of criteria in each attribute [57][58]. the more value of sub-criterion of every single parameter, the obtained value will end up with a high value $U(A_i)$ [59].

The MAUT method will change from several parameters of importance to a numerical value with a scale of 1-5, where a scale of 1 is the worst choice, and a scale of 5 is the best choice. The results of the MAUT method will provide a ranking order of alternative evaluations that describe the choices of policymakers. The spatial data modeling process for the PRTA classification using the MAUT method. The following sequence of the steps is as follows:

Step 1: Follow the steps in the flow Fig. 2 to define the criteria used as a benchmark for solving the problem and determine the priority of parameter weights.

Step 2: Make the normalized matrix using Eq. (1). Where the U(x) variable is the normalized alternative weight, the x is the alternative weight, the x_i^- is the minimum weight of the *x*-th criterion, and the x_i^+ is the maximum weight of the *x*-th criterion using Eq. (22).

$$U_{(x)} = \frac{x - x_i^-}{x_i^+ - x_i^-} \quad (22)$$

Step 3: Calculate the evaluation value of each alternative $V_{(x)}$ by multiplying utility $U_{(x)}$ by weight using Eq. (23).

$$V_{(x)} = \sum_{i=1}^{n} w_j * x_{ij} \quad (23)$$

Where the $V_{(x)}$ variable is the evaluation value of each alternative of the PRTA classification for the *i*-th data, the value of the division between the parameter weighting value and the number of sub-criteria on each parameter then multiplied by the weight of the attribute priority value at each parameter criteria. The w_k variable is the weight of the attribute sub-criterion on each parameter of the parameter until the *k*-th data and $u_k(x_{ik})$ is the parameter of the *k*-th data multiplied by the priority value of each parameter x_{ik}. The A_i variable is the weighting value of multicriteria parameters.

Step 4: Determine the range of PRTA classification values using the Guttman scale based on Eq. (11). Where the I variable is the interval range, the R variable is the result of the calculation of the highest scores value of V_x minus the lowest score value of V_x, and the K variable is a number of alternatives. The alternative assessment criteria for the PRTA classification are obtained from the result of the calculation of the highest scores value of V_x minus the value of the I variable as shown in the result of Eq. (24).

$$\begin{cases} PRTA, if\ V_x \geq the\ scale\ range \\ Non\_PRTA, if\ V_x < the\ scale\ range \end{cases} \quad (24)$$

G. Technique for Others Reference by Similarity to Ideal Solution (TOPSIS)

The TOPSIS method is a decision-making method that involves multicriteria parameters used to overcome alternative problems due to uncertainty/inconsistency [60]–[62] The TOPSIS method also determines the distance of the ideal solution to smaller and larger before making the determination of alternative value with the result of alternative calculation has the final value < 1 [50]–[53] The concept of selecting the best alternative in the TOPSIS method is that the best-selected alternative consists of alternatives with the shortest distance from the positive ideal solution and the longest distance from the negative ideal solution.

The spatial data modeling process for the PRTA classification using the MAUT method used the following sequence of the steps is as follows:

Step 1: Follow the steps in the flow in Fig. 2 to define the criteria used as a benchmark for solving the problem and determine the priority of parameter weights.

Step 2: Calculate a normalized decision matrix using Eq. (25) [63]. Where the r_{yx} variable is the normalized value for each y-th alternative to the x-th criteria with i=1,2,...,m and j=1,2,...,n.

$$r_{yx} = \frac{x_{yx}}{\sqrt{\sum_{i=i}^{m} x_{yx}^2}} \quad (25)$$

Step 3: Calculate a weighted normalized decision matrix using Eq. (26). Multiply the weight of the parameter criteria with the value of each attribute.

$$v_{yx} = W_{yx} * r_{yx} \quad (26)$$

Where v_{yx} variable is the weighted normalized value, the variable w_{yx} is the weight of each criterion, and the variable r_{yx} is the normalized value of each alternative against the j-th criterion with i=1,2,...,m and j=1,2,...,n.

Step 4: Calculate the ideal solution based on the maximum value of A^+ using Eq. (27) [51] [64] and the negative ideal solution based on a minimum value of A^- using Eq. (28) [51] [64].

$$A_x^+ = \{v_{y1}^+, v_{y2}^+, v_{y3}^+\}$$

$$A^+ = \{max\ y_1, max\ v_2, max\ v_3, \dots, max\ v_n\}$$

Where,

$$v_x^+ = \begin{cases} \max v_{yx};\ if\ yx\ is\ benefit\ attribute \\ \min v_{yx};\ if\ yx\ is\ cost\ attribute \end{cases} \quad (27)$$

$A_x^- = \{v_{y1}^-, v_{y2}^-, v_{y3}^- \}$

$A^- = \{min\ v_1, min\ v_2\ , min\ v_3\ , \dots\ , min\ v_n\ \}$

Where,

$$v_x^- = \begin{cases} \max v_{yx}\ ;\ if\ yx\ is\ benefit\ attribute \\ \min v_{yx}\ ;\ if\ yx\ is\ cost\ attribute \end{cases} \quad (28)$$

Step 5: Calculate the positive and negative ideal solution spacing, as referenced in Eq.s (29) and (30) [51] [64]. In this research, the ideal positive solution is based on the maximum value of $\mathrm{D}+$ of the Eq. (29) [51] [64] and the ideal negative solution distance based on a minimum value of $\mathrm{D}-$ of Eq. (30) [51] [64].

$$D_y^+ = \sqrt{\sum_{y=1}^{n}\left(v_{yx} - A_x^+\right)^2} \quad (29)$$

$$D_y^- = \sqrt{\sum_{y=1}^{n}\left(v_{yx} - A_x^-\right)^2} \quad (30)$$

Where the D_y^+ variable is used to calculate the maximum ideal solution distance as much as the *y*-th data. The D_i^- variable is used to calculate the minimum ideal solution distance as much as the *y*-th data.

Step 6: Calculate the preference value for each alternative to be generated by Eq. (31) [51] [64].

$$V = \frac{D_i^-}{D_i^- + D_i^+} \quad (31)$$

where D_i^- is the ideal minimal solution distance value of the *i*-th data dan D_i^+ is the maximum ideal solution distance value as much as a number of *i* data.

Step 7: Determine the range of PRTA classification values using the Guttman scale based on Eq. (11). Where the I variable is the interval range, the R variable is the result of the calculation of the highest scores value of V_i minus the lowest score value of V_i , and the K variable is a number of alternatives. The alternative assessment criteria for the PRTA classification are obtained from the result of the calculation of the highest scores value of V_i minus the value of I variable as shown in Eq. (32) .

$$\begin{cases} PRTA, if\ V_i\ \geq the\ scale\ range \\ Non\_PRTA, if\ V_i < the\ scale\ range \end{cases} \quad (32)$$

H. Analytical Hierarchy Process (AHP)

The Analytical Hierarchy Process (AHP) is a pairwise comparison method through an analytic hierarchy process, where the parameter weights are determined by normalization through the eigenvectors associated with the maximum eigenvalues in the unit ratio matrix. The weighting between parameters in this study is accomplished by the AHP method approach based on flow depicted in Fig. 2. Which consists of the following:

Step 1-9: Use the process in section III subsection A for the priority weight of the parameters using the AHP method.

Step 10: Determine the range of PRTA classification values using the Guttman scale based on Eq. (11). Where the I variable is the interval range, the R variable is the result of the calculation of the highest scores value of CR variable minus the lowest score value of CR variable, and K variable is a number of alternatives. The alternative assessment criteria for the PRTA classification are obtained from the result of the calculation of the highest scores value of CR variable minus the value of I variable, as shown in the result of Eq. (33).

$$\begin{cases} PRTA, if\ CR\ \geq the\ scale\ range \\ Non\_PRTA, if\ CR < the\ scale\ range \end{cases} \quad (33)$$

I. Multicriteria Evaluation techniques (MCE)

This research method evaluation uses the accuracy, and F1 score approaches. The F1 score is obtained from the values of precision and recall. A confusion matrix [65] is used in this evaluation technique, consisting of two positive classes and a negative class to compare actual data and classification data [66]. Multi-class classification [65] is used in the discussion of this paper: prone road traffic accident (PRTA), and non-prone road traffic accident (Non-PRTA). The precision and recall value is calculated with the average value in each class.

Accuracy in the measurement of a method is used to determine the accuracy value in clarifying the results of classification data with actual data with Eq. (34) [65]. Precision describes the amount of positive-valued data divided by total positive-valued data in Eq. (35) [65]. The recall describes the percentage of data in the positive category classified by the system with the calculation in Eq. (36) [65]. Results of precision and recall values are used to calculate F1-score, as in formula (37) [65]. The accuracy of the data generated in classification is known from the percentage after testing between the actual data in the form of an analog map of the classification of the watershed erosion zone and prediction data with MAUT, WPM, WSM, and TOPSIS methods. Performance value classification with categories 91% – 100% is very good classification, 81% – 90% is good classification, 71% – 80% is fair classification, 61% – 70% is poor classification, and values below 60% are false classification [67].

$$Accuracy_M = \frac{\sum_{i=1}^{l}\frac{TP_i+TN_i}{TP_i+FN_i+FP_i+TN_i}}{l} \quad (34)$$

$$Precision_M = \frac{\sum_{i=1}^{l}\frac{TP_i}{TP_i+FP_i}}{l} \quad (35)$$

$$Recall_M = \frac{\sum_{i=1}^{l}\frac{TP_i}{TP_i+FN_i}}{l} \quad (36)$$

$$F-score_M = \frac{(\beta^2+1)Precision_M Recall_M}{\beta^2 Precision_M + Recall_M} \quad (37)$$

Where TP_i is the amount of data + which when the classification is true by the method used for the *i*-th class. TN_i is the amount of data. When the classification is true by the method used for the *i*-th class. FP_i is the amount of + data that is classified as false by the method used for the *i*-th class. FN_i is the amount of data - which when the classification is false by the method used for the *i*-th class. l is the number of classification classes. Average accuracy is the average value of method accuracy in all classification class. $Precision_\mu$ is the precision value of each classification class. $Precision_M$

represents the average value of the precision in all classification classes. $Recall_{\mu}$ is the recall value of each classification class. $Recall_{M}$ represents the average value of recalls in all classification classes. $Fscore_{\mu}$ is a performance matrix to calculate the average of precision and recall values in each classification class. $F-score_{M}$ is a performance matrix to calculate the average of precision and recall values in all classification classes.

IV. Result And Discussion

Based on the private spatial datasets and quantitative attribute data explained in section III, the results of this study are discussed in the following subsections.

A. Parameter Priority Weight

In each of the methods used in the MCDM, the weight of each parameter priority value in this study uses the opinion of EJ (score) and mathematical calculations using AHP based on the score given by EJ (EJ-AHP). Tables IV and V are the results of mathematical calculations of the pairwise comparison matrix of the AHP method to produce the priority weight of the parameters based on the flow in Fig. 2 with the process in Section III for sub-section A.

TABLE III. Parameter Priority Value Weighting Results

Parameters Symbol	**AHP Weight**
VCR (X1)	0.02
IRI (X2)	0.06
HA (X3)	0.10
VA (X4)	0.14
Vr (X5)	0.18
Road Type (X6)	0.22
Shoulder (X7)	0.27
Total:	**1.00**

TABLE IV. Sub-parameter Priority Value Weighting Results

Parameters	**Multicriteria Parameters**	**EJ Scoring**	**EJ-AHP Weight**
VCR (%)	VCR ≥ nm0.85 && VCR < 1.00	5	0.34
	VCR ≥ 0.70 && VCR < 0.85	4	0.26
	VCR ≥ 0.45 && VCR < 0.70	3	0.24
	VCR ≥ 0.20 && VCR < 0.45	2	0.12
	VCR < 0.20	1	0.04
IRI (m/km)	IRI ≥ 12	4	0.43
	IRI ≥ 8 && IRI < 12	3	0.35
	IRI ≥ 4 && IRI < 8	2	0.17
	IRI < 4	1	0.05
HA (rad/km)	HA ≥ 3.50	3	0.54
	HA ≥ 0.25 && HA < 3.50	2	0.37
	HA < 0.25	1	0.09
VA (m/km)	VA ≥ 45	3	0.54
	VA ≥ 5 && VA < 45	2	0.37
	VA < 5	1	0.09
Vr (km/jam)	Vr ≥100	6	0.33
	Vr ≥ 80 && Vr < 100	5	0.25
	Vr ≥ 65 && Vr < 80	4	0.17
	Vr ≥ 60 && Vr < 65	3	0.16
	Vr ≥ 50 && Vr < 60	2	0.06
	Vr < 50	1	0.03
Road Type	2/2 UD	5	0.36
	4/2 UD	4	0.29
	4/2 D	3	0.24
	6/2 D	2	0.08
	2/1 UD	1	0.04
Shoulder	No	2	0.83
	Yes	1	0.17

Parameters	**Range**	**EJ Scoring**	**EJ-AHP Weight**
VCR (%)	VCR ≥ 0.90 && VCR < 1.00	5	0.34
	VCR ≥ 0.75 && VCR < 0.90	4	0.26
	VCR ≥ 0.50 && VCR < 0.75	3	0.24
	VCR ≥ 0.30 && VCR < 0.50	2	0.12
	VCR < 0.30	1	0.04
IRI (m/km)	IRI ≥ 12	4	0.43
	IRI ≥ 8 && IRI < 12	3	0.35
	IRI ≥ 4 && IRI < 8	2	0.17
	IRI < 4	1	0.05
HA (rad/km)	HA ≥ 3.50	3	0.54
	HA ≥ 0.25 && HA < 3.50	2	0.37
	HA < 0.25	1	0.09
VA (m/km)	VA ≥ 45	3	0.54
	VA ≥ 5 && VA < 45	2	0.37
	VA < 5	1	0.09
Vr (km/jam)	Vr ≥100	6	0.33
	Vr ≥ 90 && Vr < 100	5	0.25
	Vr ≥ 75 && Vr < 90	4	0.17
	Vr ≥ 60 && Vr < 75	3	0.16
	Vr ≥ 50 && Vr < 60	2	0.06
	Vr < 50	1	0.03
Road Type	2/2 UD	5	0.36
	4/2 UD	4	0.29
	4/2 D	3	0.24
	6/2 D	2	0.08
	2/1 UD	1	0.04
Shoulder	No	2	0.83
	Yes	1	0.17

B. The Guttman Scale to Determine the Classification of Accident Prone Roads

The Guttman scale [68] is used to measure the generated classification values in this paper. This scale is used to draw conclusions from qualitative data [69]. It is also used to estimate the value of the classification resulting in an intervention value that is still ambiguous due to uncertainty [70]. It is possible to assess the uncertainty factor of a variable class defined using the Guttman scale [71] in Eq. (11) for a dataset that employs a weight in the analysis process and delivers a value.

The test data consisted of 180 primary arterial roads and 201 primary collector roads, where the data is categorized as a small-scale dataset. The value of the scale on the SAW, WP, SAW, WPM, MAUT, TOPSIS, and AHP methods using Eq. (12), (14), (17), (21), (24), (32), and (33) based on the process calculations in section III sub-sections B to H, respectively.

TABLE V. THE SCALE TO DETERMINE THE CLASSIFICATION OF ACCIDENT PRONE ROADS

MCDM Models	Arterial Road Scale	Collector Road Scale
WSM	$\begin{cases} PRTA, if\ A_i \geq 0{,}2729 \\ Non\_PRTA, if\ A_i < 0{,}2729 \end{cases}$	$\begin{cases} PRTA, if\ A_i \geq 0{,}2886 \\ Non\_PRTA, if\ A_i < 0{,}2886 \end{cases}$
WP	$\begin{cases} PRTA, if\ S_i \geq 0{,}2367 \\ Non\_PRTA, if\ S_i < 0{,}2367 \end{cases}$	$\begin{cases} PRTA, if\ S_i \geq 0{,}2579 \\ Non\_PRTA, if\ S_i < 0{,}2579 \end{cases}$
SAW	$\begin{cases} PRTA, if\ V_i \geq 0{,}5753 \\ Non\_PRTA, if\ V_i < 0{,}5753 \end{cases}$	$\begin{cases} PRTA, if\ V_i \geq 0{,}6037 \\ Non\_PRTA, if\ V_i < 0{,}6037 \end{cases}$
WPM	$\begin{cases} PRTA, if\ V_i \geq 0{,}0060 \\ Non\_PRTA, if\ V_i < 0{,}0060 \end{cases}$	$\begin{cases} PRTA, if\ V_i \geq 0{,}0055 \\ Non\_PRTA, if\ V_i < 0{,}0055 \end{cases}$
MAUT	$\begin{cases} PRTA, if\ V_x \geq 0{,}2886 \\ Non\_PRTA, if\ V_x < 0{,}2886 \end{cases}$	$\begin{cases} PRTA, if\ V_x \geq 0{,}4723 \\ Non\_PRTA, if\ V_x < 0{,}4723 \end{cases}$
TOPSIS	$\begin{cases} PRTA, if\ V_i \geq 0{,}4073 \\ Non\_PRTA, if\ V_i < 0{,}4073 \end{cases}$	$\begin{cases} PRTA, if\ V_i \geq 0{,}4157 \\ Non\_PRTA, if\ V_i < 0{,}4157 \end{cases}$
AHP	$\begin{cases} PRTA, if\ CR \geq 0{,}00229 \\ Non\_PRTA, if\ CR < 0{,}00229 \end{cases}$	$\begin{cases} PRTA, if\ CR \geq 0{,}000080 \\ Non\_PRTA, if\ CR < 0{,}000080 \end{cases}$

C. Model Performance Evaluation

The MCDM spatial analysis model was developed to assist the decision-making process by selecting alternatives in the multi-class classification [52][72]. The steps in the MCDM model are to determine the multicriteria parameter that will be an alternative to the mult-iclass classification, to describe the quantitative data requirements that will be processed to have an impact on the alternatives of the multi-class being processed, then to process the numerical values on the qualitative data to determine the rating on each of the multicriteria parameters.

Table VI results from multicriteria evaluation techniques using a confusion matrix based on the process in section III sub-section I. The accuracy values in the experimental test of arterial road type data using the WSM and TOPSIS methods were 63% superior to other methods, followed by the MAUT, SAW, WP, and WPM methods, and the AHP method, namely 59%, 58%, 54%, 43%, respectively. However, the AHP method is superior in the collector road type experiment with an accuracy value of 70%, followed by the TOPSIS and WSM, SAW, MAUT, WP, and WPM methods, namely 58%, 57%, 53%, 41%, respectively.

Fig. 3 is a sampling test of the spatial analysis results of accident-prone road classification using the WSM method on the North Rim Probolinggo arterial road type, Indonesia. Calculate the weight value for each multicriteria parameter using Flow in Fig. 2 with the results in Tables IV and V. Perform the calculation process based on section III subsection B, then obtain the value of the variable A_i is 0.3593, referring to Table VI, the road is included in the PRTA classification.

TABLE VI. THE MCDM MODEL PERFORMANCE

MCDM Models	*Accuracy*	*F1-Score*
Arterial Roads		
WSM	63%	54%
WP	54%	55%
SAW	58%	52%
WPM	54%	55%
MAUT	59%	53%
TOPSIS	63%	54%
AHP	43%	53%
Collector Roads		
WSM	58%	49%
WP	41%	45%
SAW	57%	48%
WPM	41%	45%
MAUT	53%	47%
TOPSIS	58%	49%
AHP	70%	47%

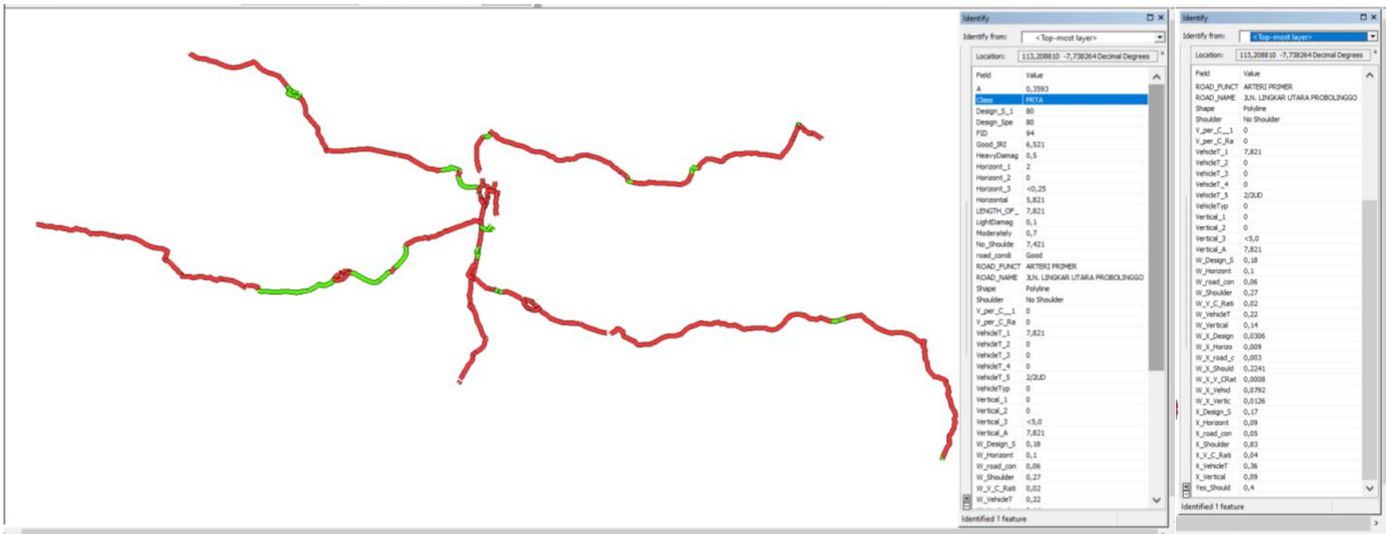

Fig. 3. Results of Spatial Analysis of Accident Prone Roads using the WSM Method on Arterial Road Types.

V. Conclusion and Future Works

This paper presents an empirical study to determine the ability of the MCDM model based on multi-criteria parameters that combine the weight values given by expert judgment and mathematical computation using the AHP method (EJ-AHP). MCDM model is a method that depends on the value of weights and the priority scale of parameter values that depends on expert judgment. Weight analysis on the MCDM model by combining EJ-AHP can bridge the difference between subjective and objective risks that are biased in the evaluation process for weights and parameter priority scales between expert judgments. The labeling results in this research can be used as a labeling basis for further research on the category of the private dataset types with small dataset scales in determining the category of PRTA or Non-PRTA classification based on multi-criteria parameters.

The parameter weight values generated in the EJ-AHP computation process will be used as the basis for the empirical study for the spatial analysis of the PRTA classification based on the MCDM model using a comparison of the WSM, WP, SAW, WPM, MAUT, TOPSIS, and AHP methods to measure the performance of the method. The performance evaluation results of this method will be used as a reference for whether or not a method is feasible to be developed further. The accuracy value in the whole process of the MCDM model performance is below 71% (Table VI), where each method produces a different rating value, and this concludes that a new alternative method can be applied to produce a high accuracy value. Therefore, it is essential to further research using machine learning (ML) by applying several alternative scenarios through performance tests on ML single classifier, ML parameter tuning, and ML hybrid ensemble learning to improve the performance of the resulting classification values.

Acknowledgment

The results of this study are part of a thesis research in the Universiti Teknikal Malaysia Melaka (UTeM), Malaysia. The results from a study funded by the Indonesian Directorate General of Strengthening Research & Development of Research, Technology, and Higher Education Ministry in 2015-2016 and the Universitas Dr. Soetomo, Indonesia.

References

[1] A. A. Hyder, N. Paichadze, T. Toroyan, and M. M. Peden, “Monitoring the Decade of Action for Global Road Safety 2011–2020: An update,” Glob. Public Health, vol. 12, no. 12, pp. 1492–1505, 2017.

[2] World Health Organization, Global Status Report on Road Safety 2018. World Health Organization 2018, 2018.

[3] G. Plan, D. Of, A. For, and R. Safety, Global Plan for the DECADE OF ACTION FOR ROAD SAFETY, 2011-2020, Version 3 Global Plan Decade of Action for Road. 2011, pp. 2011–2020.

[4] F. Wegman, “The future of road safety: A worldwide perspective,” IATSS Res., vol. 40, no. 2, pp. 66–71, 2017.

[5] Presiden Republik Indonesia, Undang-Undang Republik Indonesia Nomor 38 tahun 2004 Tentang Jalan. 2004.

[6] Republik Indonesia, “Peraturan Pemerintah Republik Indonesia Nomor 34 Tahun 2006 Tentang Jalan.” 2006.

[7] Republik Indonesia, Undang-Undang Republik Indonesia Nomor 22 Tahun 2009 Tentang Lalu Lintas dan Angkutan Jalan. 2009.

[8] Republik Indonesia, “Rencana Umum Nasional Keselamatan (RUNK) Jalan 2011 - 2035.” 2011.

[9] Republik Indonesia, “Peraturan Presiden Nomor 2 Tahun 2012 tentang KNKT.” 2012.

[10] H. Briassoulis, D. Kavroudakis, and N. Soulakellis, The practice of spatial analysis. Springer, 2019.

[11] G. Haseli, R. Sheikh, and S. S. Sana, “Base-criterion on multi-criteria decision-making method and its applications,” Int. J. Manag. Sci. Eng. Manag., vol. 00, no. 00, pp. 1–10, 2019.

[12] A. Aljuhani, “Multi-Criteria Decision-Making Approach for Selection of Requirements Elicitation Techniques based on the Best-Worst Method,” Int. J. Adv. Comput. Sci. Appl., vol. 12, no. 11, pp. 732–738, 2021.

[13] M. G. Delgado and J. B. Sendra, “Sensitivity analysis in multicriteria spatial decision-making: A review,” Hum. Ecol. Risk Assess., vol. 10, no. 6, pp. 1173–1187, 2004.

[14] A. V. Vitianingsih, Z. Othman, S. Suhana, and K. Baharin, “Spatial Analysis for the Classification of Prone Roads Traffic Accidents: A Systematic Literature Review,” Int. J. Adv. Trends Comput. Sci. Eng., vol. 10, no. 2, pp. 583–599, 2021.

[15] A. V. Vitianingsih, N. Suryana, and Z. Othman, “Spatial analysis model for traffic accident-prone roads classification: A proposed framework,” IAES Int. J. Artif. Intell., vol. 10, no. 2, pp. 365–373, 2021.

[16] A. V. Vitianingsih, D. Cahyono, and A. Choiron, “Web-GIS Application using Multi-Attribute Utility Theory to Classify Accident-Prone Roads,”

J. Telecommun. Electron. Comput. Eng., vol. 10, no. 2–3, pp. 83–89, 2018.

[17] A. V. Vitianingsih and D. Cahyono, "Geographical Information System for Mapping Road Using Multi-Attribute Utility Method," in International Conference on Science and Technology-Computer (ICST), 2016, pp. 0–4.

[18] A. V. Vitianingsih and D. Cahyono, "Geographical Information System for mapping accident-prone roads and development of new road using Multi-Attribute Utility method," in Proceedings - 2016 2nd International Conference on Science and Technology-Computer, ICST 2016, 2017, pp. 66–70.

[19] N. Mahmoody Vanolya, M. Jelokhani-Niaraki, and A. Toomanian, "Validation of spatial multicriteria decision analysis results using public participation GIS," Appl. Geogr., vol. 112, no. November, pp. 1–16, 2019.

[20] W. Alkhadour, J. Zraqou, A. Al-Helali, and S. Al-Ghananeem, "Traffic Accidents Detection using Geographic Information Systems (GIS): Spatial Correlation of Traffic Accidents in the City of Amman, Jordan," Int. J. Adv. Comput. Sci. Appl., vol. 12, no. 4, pp. 484–494, 2021.

[21] S. G. and N. D. Bekir Çetintav, Gözde Ulutagay, "A New Approach Of Combining Expert Judgment And Data Knowledge In Multi-Attribute Decision Making," in Uncertainty Modelling in Knowledge Engineering and Decision Making, 2016, pp. 112–118.

[22] T. Chen, C. Zhang, and L. Xu, "Factor analysis of fatal road traffic crashes with massive casualties in," Adv. Mech. Eng., vol. 8, no. 4, pp. 1–11, 2016.

[23] A. C. L. Vieira, M. D. Oliveira, and C. A. Bana e Costa, "Enhancing knowledge construction processes within multicriteria decision analysis: The Collaborative Value Modelling framework," Omega (United Kingdom), vol. 94, no. July, pp. 1–36, 2020.

[24] H. Fang and Z. Guo, "Vessel Collision Accidents Analysis based on Factor Analysis and GA-SVM *," in IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), 2017, pp. 191–195.

[25] E. Turunen, "Using GUHA Data Mining Method in Analyzing Road Traffic Accidents Occurred in the Years 2004–2008 in Finland," Data Sci. Eng., vol. 2, no. 3, pp. 224–231, 2017.

[26] X. Qu, W. Wang, W. fu Wang, and P. Liu, "Real-time rear-end crash potential prediction on freeways," J. Cent. South Univ., vol. 24, no. 11, pp. 2664–2673, 2017.

[27] M. Juhász and C. Koren, "Getting an Insight into the Effects of Traffic Calming Measures on Road Safety," Transp. Res. Procedia, vol. 14, pp. 3811–3820, 2016.

[28] D. Nilsson, M. Lindman, T. Victor, and M. Dozza, "Definition of run-off-road crash clusters—For safety benefit estimation and driver assistance development," Accid. Anal. Prev., vol. 113, no. November 2017, pp. 97–105, 2018.

[29] M. Bassani, L. Rossetti, and L. Catani, "Spatial analysis of road crashes involving vulnerable road users in support of road safety management strategies," in Transportation Research Procedia, 2020, vol. 45, pp. 394–401.

[30] H. Zhang, X. Wang, J. Cao, M. Tang, and Y. Guo, "A multivariate short-term traffic flow forecasting method based on wavelet analysis and seasonal time series," Appl. Intell., vol. 48, no. 10, pp. 3827–3838, 2018.

[31] R. Al-Ruzouq, K. Hamad, S. Abu Dabous, W. Zeiada, M. A. Khalil, and T. Voigt, "Weighted Multi-attribute Framework to Identify Freeway Incident Hot Spots in a Spatiotemporal Context," Arab. J. Sci. Eng., vol. 44, no. 10, pp. 8205–8223, 2019.

[32] B. Pradhan and M. Ibrahim Sameen, "Review of Traffic Accident Predictions with Neural Networks," in Urbanization and Its Impact in Contemporary China, Springer, Cham, 2020, pp. 97–109.

[33] R. Goel, "Modelling of road traffic fatalities in India," Accid. Anal. Prev., vol. 112, no. October, pp. 105–115, 2018.

[34] M. Keymanesh, H. Ziari, S. Roudini, and A. N. Ahangar, "Identification and Prioritization of 'Black Spots' without Using Accident Information," Model. Simul. Eng., vol. 2017, 2017.

[35] Ö. Kaya, A. Tortum, K. D. Alemdar, and M. Y. Çodur, "Site selection for EVCS in Istanbul by GIS and multi-criteria decision-making," Transp. Res. Part D Transp. Environ., vol. 80, no. February, p. 102271, 2020.

[36] M. S. Fatemeh Haghighat, "Application of a Multi-Criteria Approach To Road Safety Evaluation in the Bushehr Province , Iran," Traffic Plan. Prelim. Commun., vol. 23, no. 5, pp. 341–352, 2011.

[37] S. Da Costa, X. Qu, and P. M. Parajuli, "A Crash Severity-Based Black Spot Identification Model," J. Transp. Saf. Secur., vol. 7, no. 3, pp. 268–277, 2015.

[38] F. Torrieri and A. Batà, "Spatial multi-Criteria decision support system and strategic environmental assessment: A case study," Buildings, vol. 7, no. 4, 2017.

[39] S. Kumar and D. Toshniwal, "A data mining approach to characterize road accident locations," J. Mod. Transp., vol. 24, no. 1, pp. 62–72, 2016.

[40] T. Sipos, "Spatial statistical analysis of the traffic accidents," Period. Polytech. Transp. Eng., vol. 45, no. 2, pp. 101–105, 2017.

[41] R. Mosadeghi, J. Warnken, R. Tomlinson, and H. Mirfenderesk, "Comparison of Fuzzy-AHP and AHP in a spatial multi-criteria decision making model for urban land-use planning," Comput. Environ. Urban Syst., vol. 49, pp. 54–65, 2015.

[42] F. Yakar, "A multicriteria decision making–based methodology to identify accident-prone road sections," J. Transp. Saf. Secur., pp. 1–15, 2019.

[43] A. G. Macbeth and R. Eng, "Road Classification Systems – Christchurch and Toronto," in 2001 Traffic Management Workshop Auckland, 2001, no. 03, pp. 1–12.

[44] E. M. Setton, P. W. Hystad, and C. P. Keller, "Road Classification Schemes – Good Indicators of Traffic Volume?," in Spatial Sciences Laboratories Occasional Papers, 2005, vol. i, pp. 1–11.

[45] A. Suraji, L. Djakfar, and A. Wicaksono, "Analysis of bus performance on the risk of traffic accidents in East Java-Indonesia," EUREKA, Phys. Eng., vol. 2021, no. 3, pp. 111–118, 2021.

[46] H. Pilko, S. Mandžuka, and D. Barić, "Urban single-lane roundabouts: A new analytical approach using multi-criteria and simultaneous multi-objective optimization of geometry design, efficiency and safety," Transp. Res. Part C Emerg. Technol., vol. 80, no. July, pp. 257–271, 2017.

[47] M. Rosić, D. Pešić, D. Kukić, B. Antić, and M. Božović, "Method for selection of optimal road safety composite index with examples from DEA and TOPSIS method," Accid. Anal. Prev., vol. 98, no. January, pp. 277–286, 2017.

[48] A. Ait-Mlouk, T. Agouti, and F. Gharnati, "Mining and prioritization of association rules for big data: multi-criteria decision analysis approach," J. Big Data, vol. 4, no. 1, pp. 1–21, 2017.

[49] T. L. Saaty, "A scaling method for priorities in hierarchical structures," J. Math. Psychol., vol. 15, no. 3, pp. 234–281, 1977.

[50] E. Triantaphyllou and S. H. Mann, "An examination of the effectiveness of multi-dimensional decision-making methods: A decision-making paradox," Decis. Support Syst., vol. 5, no. 3, pp. 303–312, 1989.

[51] E. Triantaphyllou, Multi-Criteria Decision Making Methods: A Comparative Study. Springer, Boston, MA, 2000.

[52] E. Mulliner, N. Malys, and V. Maliene, "Comparative analysis of MCDM methods for the assessment of sustainable housing affordability," Omega (United Kingdom), vol. 59, pp. 146–156, 2016.

[53] V. Maliene, R. Dixon-Gough, and N. Malys, "Dispersion of relative importance values contributes to the ranking uncertainty: Sensitivity analysis of Multiple Criteria Decision-Making methods," Appl. Soft Comput. J., vol. 67, pp. 286–298, 2018.

[54] C. H. Yeh, "A Problem-based Selection of Multi-attribute Decision-making Methods," Int. Trans. Oper. Res., vol. 9, no. 2, pp. 169–181, 2002.

[55] E. Boltürk, A. Karaşan, and C. Kahraman, "Simple additive weighting and weighted product methods using neutrosophic sets," in Studies in Fuzziness and Soft Computing, vol. 369, 2019, pp. 647–676.

[56] Z. Chourabi, F. Khedher, A. Babay, and M. Cheikhrouhou, "Multi-criteria decision making in workforce choice using AHP, WSM and WPM," J. Text. Inst., vol. 110, no. 7, pp. 1092–1101, 2019.

[57] J. S. Dyer, "MAUT — Multiattribute Utility Theory," in Muiltiple Criteria Decision Analysis: State of the Art Surveys, 2005, pp. 265–292.

[58] M. Bystrzanowska and M. Tobiszewski, "How can analysts use multicriteria decision analysis?," TrAC - Trends Anal. Chem., vol. 105, pp. 98–105, 2018.

[59] M. Cinelli, S. R. Coles, and K. Kirwan, "Analysis of the potentials of multi criteria decision analysis methods to conduct sustainability assessment," Ecol. Indic., vol. 46, pp. 138–148, 2014.

[60] K. Mela, T. Tiainen, and M. Heinisuo, "Comparative study of multiple criteria decision making methods for building design," Adv. Eng. Informatics, vol. 26, no. 4, pp. 716–726, 2012.

[61] S. Başaran and F. El Homsi, "Mobile Mathematics Learning Application Selection using Fuzzy TOPSIS," Int. J. Adv. Comput. Sci. Appl., vol. 13, no. 2, pp. 270–282, 2022.

[62] N. A. M. Zulkefli, M. Madanan, T. M. Hardan, and M. H. M. Adnan, "Multi-Criteria Prediction Framework for the Prioritization of Council Candidates based on Integrated AHP-Consensus and TOPSIS Methods," Int. J. Adv. Comput. Sci. Appl., vol. 13, no. 2, pp. 352–359, 2022.

[63] L. Vargas, Why the Analytic Hierarchy Process Is Not Like Multiattribute Utility Theory, vol. 356. 1989.

[64] T. Vulevic, N. Dragovic, S. Kostadinov, S. Belanovic Simic, and I. Milovanovic, "Prioritization of Soil Erosion Vulnerable Areas Using Multi-Criteria Analysis Methods," Polish J. Environ. Stud., vol. 24, no. 1, pp. 317–323, 2015.

[65] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," Inf. Process. Manag., vol. 45, no. 4, pp. 427–437, 2009.

[66] Max Bramer, Principles Data Mining. London-Spinger: http://www.springer.com/series/7592, 2007.

[67] Florin Gorunescu, Data Mining: Concept, Models, Techniques. Springer, 2011.

[68] L. Guttman, "The Determinacy Of Factor Score Matrices With Implications For Five Other Basic Problems Of Common-Factor Theory," Br. J. Stat. Psychol., vol. 7, no. 2, pp. 65–81, 1955.

[69] L. Guttman, "A Basis for Scaling Qualitative Data," Am. Sociol. Rev., vol. 9, no. 2, p. 139, 1944.

[70] R. E. Tractenberg, F. Yumoto, P. S. Aisen, J. A. Kaye, and R. J. Mislevy, "Using the Guttman scale to define and estimate measurement error in items over time: The case of cognitive decline and the meaning of 'points lost,'" PLoS One, vol. 7, no. 2, pp. 1–8, 2012.

[71] A. Stegeman, "A new method for simultaneous estimation of the factor model parameters, factor scores, and unique parts," Comput. Stat. Data Anal., vol. 99, no. July, pp. 189–203, 2016.

[72] M. R. Asadabadi, "The stratified multi-criteria decision-making method," Knowledge-Based Syst., vol. 162, no. December, pp. 115–123, 2018.