ISSN: 2754-6659

Journal of Artificial Intelligence & Cloud Computing

SCIENTIFIC Research and Community

Review Article

Open Access

Traffic Behaviour Analysis using Logistic Regression Method (LRM) and Structural Equation Modelling (SEM)

Fajaruddin Mustakim[1]*, Mohammad Nazir Ahmad[1], Rabiah Abdul Kadir[1], Azlan Abdul Aziz[2] and Othman Che Puan[3]

[1]Senior Lecturer, Faculty of Engineering and Technology, Kebangsaan University, Malaysia

[2]RMC, Faculty of Engineering and Technology, Malaysia Multimedia University, Malaysia

[3]School of Civil Engineering, University Technology, Malaysia

ABSTRACT

This study aims to determine traffic behaviour at the selected unsignalized intersection and the development of right-turn motorists (RTM) by adopting the logistic regression method (LRM) and structural equation modelling (SEM). In the early stage of the study, we analysed the traffic behaviour focusing on traffic volume and turning volume at the field site. This study involves five unsignalized intersections (UI), and it observes three types of turning volume: right turn volume (RTV) from a minor road onto a major road, left turn volume (LTV) from a minor road onto a major road, and right turn volume (RTV) from a major road onto a minor road. Although the SEM approach is among the popular scientific analysis and wisely applied in various fields of study, there is less attention to traffic behaviour and road safety. An SEM model for right-turn motorists using 812 datasets was developed and variables that influenced the decision of right-turn motorists (RTM) were identified. Out of the six variables analysed in this statistical model, we identified gap, motorcycle rider, conflict lane change and the traffic signal to be significant.

***Corresponding author**

Fajaruddin Mustakim, Senior Lecturer, Faculty of Engineering and Technology, Kebangsaan University, Malaysia.

Received: August 22, 2023; **Accepted:** August 28, 2023; **Published:** September 04, 2023

Keywords: Structural Equation Method, Logistic Regression Method, Traffic Behaviour

Introduction

Programming Package, namely Linear Structural Relationship LISRELhas widely spread and become extremely popular among researchers [1-4]. The existence of these statistical methods has encouraged scholars to adopt them in various fields.

SEM stands for Structural Equation Modelling, which is a statistical method used to model complex relationships between observed and unobserved (latent) variables. It is a multivariate statistical approach that allows researchers to examine the direct and indirect relationships between variables, as well as test theoretical models and hypotheses. Hair et al interpreted SEM as a multiple regression equation, estimating a group of datasets interdependently and simultaneously by using a structural model [5].

SEM typically has two functions: first, explain the relationship between two latent variables, endogenous and exogenous. Consequently allows the estimation of both relationships and identifies the significant factors among these parameters. Second, analyse the relationship between the latent and the observed variables. Equation 1, defined the relationship between an observed an unobserved variable:

$$x=l.Y+e \quad (1)$$

Where x is the observed variable, the loading l is a regression coefficient measuring connectivity between x and Y, and e represent random measurement error.

Analysis of Moment Structures (AMOS), which is introduced by IBM Corporation, and the Structural Equation Program EQS provides SEM analysis using visualisation and syntax techniques [4, 6]. Besides dataset analysis, this software can present over 20 model fit indexes to test and validate the model [7].

Previous study by F M et al was analysed crash prediction model using non-linear regression method [8]. Accident frequency and accident point weighting has been adopting in the Models as dependent variable and found motorcycle crossing from unsignalized intersection, approach speed and vehicle gap was statistical significance factor contributing the accident rate.

The study conducted by F M et al reveals the percentage of accident reduction by changing the measures of every parameter [9]. One access point per kilometre reduction can decrease crashes by 16%. A 5 kph reduction in approach speed equated to a 27% drop in a crash. Conversely, providing one traffic signalised at the unsignalized intersection would reduce the number of casualties by 44%. By using regression and logistic regression approach F M et al analysed accident causality and traffic behaviour at black spot locations on Malaysia Federal Route [10]. The study revealed that the most unsafe turning manoeuvre is the movement of a right-turning motorcycle from a minor to a major road and a left-turning motorcycle from a minor road.

F M et al Determination of critical gap for passenger car and motorcyclist at unsignalized intersection by applying raff method and logic approach have found the same result [11]. In addition, the study has identified right turn vehicle is motorcyclist likely to accept the shorter gaps as the waiting time at the head of queues increase.

F M et al have developed rider and passenger car models by applying binary regression [12]. However, it does not concentrate on traffic volume, turning behaviour, and implementation of structural equation modelling. F M et al have established for the right-turning motorist, and clearly described a gap sequence at unsignalized intersections. Recently F M et al have developed lane change serious conflict model by using Logistic Regression Method (LRM) and Artificial Neuoron Network (ANN) [10, 13, 14]. The study found that the combination between both scientific approach may sustainable hydrodynamic in describing the outcomes result.

SEM Fit Index

SEM programming comprises standard error with each calculated loading and related t-test. If the sample size for SEM estimation is larger than 200, the Chi-square value can also be determined, which usually results in high statistical power. Therefore, significant loading is desirable, and even if it is not attained, the model is still meaningful. Typically, the chi-square ($\chi 2/\mathrm{df}$) statistical ratio recommended ranges as low as 2.0 and up to 5.0 [15].

Steiger introduced the root-mean-square error of approximation (RMSEA) as the second fit index applied in the LISREL program [16]. It provides information about how well the model fits the data, considering the number of parameters estimated in the model. If the model provides a good fit to the population covariance matrix, it suggests that the model accurately represents the relationships between the variables in the population [17]. The cut-off values of the RMSEA should be between 0 to 0.08 [18].

Besides the Chi-Square test, Jöreskog and Sorbom have created the Goodness of Fit Index (GFI) to measure the quantity of variance in the model by the estimated amount of covariance [15]. This statistical index ranges from 0 to 1. Typically, the larger sample increases its value and the bigger number of variables influences the increase of GFI [19]. Mile and Shevlin have performed simulation studies and recommend the cut-off of GFI 0.95 is more appropriate when the factor of loading and sample sizes are low. Adjusted Goodness of Fit Index (AGFI) is a modification of the Goodness of Fit Index (GFI) that takes into account the degrees of freedom in the model [20]. AGFI is calculated by adjusting GFI for the expected fit under the null hypothesis of a completely saturated model, given the degrees of freedom in the actual model. Like the GFI, the AGFI ranges between 0 and 1, and a value of 0.90 or higher shows a well-fitting model, according to Tabachnick and Fidell (2007).

The root mean square residual (RMR) is a measure of the discrepancy between the hypothesised covariance model and the sample residuals of the covariance matrix, and its scale depends on the variables being measured. According to Hu and Bentler (1999), an acceptable RMR index value is equal to or less than 0.08.

The Norm Fit Index (NFI) compares the χ^2 value of the model with the χ^2 value of the null model, which assumes that all tested variables are unrelated and represents the worst-case scenario. This statistical index ranges from 0 to 1. Bentler et al suggesting an index value greater than 0.90 represents a good fit model. Meanwhile Hu and Bentler recommend the cut-off value NFI equal to or greater than 0.95 would be appropriate [21, 22].

Bentler discovered the comparative fit index (CFI) [23]. It is a reviewed form of the NFI which concentrates on sample size. The index performs well even though the sample size is smaller [15]. Like NFI, the value of CFI range between 0 and 1. An index value that is close to 1 shows a good fit. Hu and Bentler emphasised a CFI value equal to or greater than 0.95 is recognised as a good index [21].

Too many models' fit indexes in SEM create phenomena of reporting results because of complexity, issues, and the acceptable margin of the fit indexes. Describing and evaluating a hypothesised model does not require all fit indices to be included. The choice of the fit index to use typically depends on the purpose of the study, as noted by Mousa Alavi et al [23].

Model Modification and Fitness

Although SEM involves complex statistical methods, there are software tools available that make it relatively easy to apply. The scientific tool can analyse single and multiple models concurrently and estimate effectively. Modification indices (MI) found under analysis properties in the SEM provide information on modification value. The change of covariance between the variables, by increasing the network or removing errors in the MI, might improve fitting models such as Chi-Square, NFI, RMSEA, RMR, GFI, AGFI, NFI, RFI, IFI, CFI and RMSEA [4, 24, 25]. When attempting to improve the model fit of an SEM, it is important to consider both the theoretical aspects of the model and the fit indices. We should not disregard theoretical considerations in favour of only improving the model fit [26]. Before implementing the model in SEM, this study used the logistic regression method to examine and validate all model parameters carefully to ensure that they align with the theoretical foundation of the model.

Data Collection

We performed this study on Malaysia's Federal Route 50. It has four lanes of a two ways partially divided road. The total stretch of the roadway from Batu Pahat to Ayer Hitam is about 40 km. The existing road consists high density of access roads or unsignalized intersections. It is because the location of the route crosses several housing areas, industrial hubs, and commercial buildings. In the year 2022, it has a capacity providing approximately 80,102 veh/day and up to 7,949 veh/hr. The design speed for this route is around 100 kph. Meanwhile, the method applies in this study includes site investigation, video recording, traffic behaviour analysis, critical gap, speed study, gap pattern, development of right turn model and the conflict model. We have implemented binary regression or logistic regression and structural equation modelling in the models.

Video cameras were at selected unsignalized intersections (UI) and data collection was concentrated on all traffic manoeuvre behaviour as illustrated in Figure 2. Data of vehicles, classified by types, like cars, motorbikes, lorries, and public transport were gathered from hourly traffic volume (disaggregated according to every type of motor vehicle), occurring on every chosen unsignalized intersection (UI2, UI8, UI9, UI10 and UI20). All selected UI is in the urban area except UI 20 in the suburban region.

Data on conflict situations, approach speeds, vehicle flow, and pedestrian crossings were simultaneously collected. Once the recording finished, all the video cams were brought to the laboratory for further microscope analysis. The selection of unsignalized intersection (UI) was based on two aspects; first, the accident's blackspot ranking recorded, and second, the road safety facilities provided on that UI [12]. UI2 was a three-leg unsignalized junction. It has concrete dividers, right-turn channelisation, and traffic lights approximately 100 meters from the intersection in the middle of the mainstream road. UI9 was a three-leg junction and traffic signal located around 100 meters from the intersection. UI9 didn't equip concrete road median and right turn channelisation. Meanwhile, UI8 and UI10 were a three-leg junction connecting four lanes on a major road with a two lanes minor road as illustrated in Figure 1. It has no traffic safety facilities such as traffic signals, concrete road medians and right-turn channelisation. UI 20 is in a suburban area, which has 4 leg unsignalized intersections and four lines on the major road. In this study, five UI are involved in turning volume behaviour and only three UI selected (UI2, UI9 and U10) focus on right-turning motorist analysis and development of structural equation modelling.

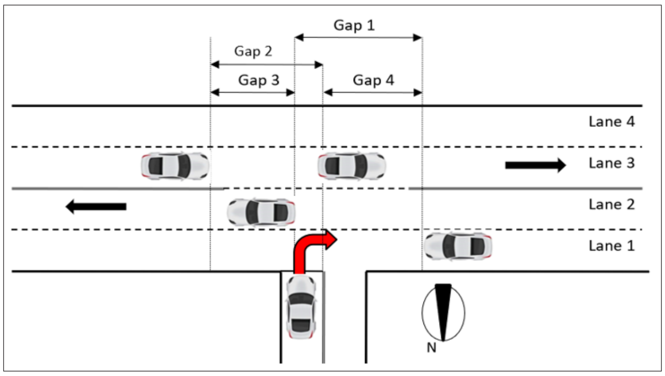

Figure 1: Right Turning At A Three-Leg Unsignalized Intersection, The Four Potential Gap Patterns For Motor Vehicle Acceptance Or Rejection (UI)

Traffic Volume

We collected a traffic count at five unsignalized intersections UI 2, UI 9, UI 10, UI 8 and UI20 using video cameras. Traffic data conducted in this paper were based on hourly traffic volume and focused on three peak hours, namely morning (8:00-10:00), midday (12:00-14:00) and afternoon (16:00-18:00). Figure 2, illustrates the highest traffic volume was during the afternoon at (17:00-18:00) stated 4,804 veh/hr, 4,500 veh/hr, 4,368 veh/hr, 4,142 veh/hr and 2218 veh/hr for UI 2, UI 8, UI 9, UI20 and UI 10 respectively. The second highest traffic volume was during the morning (8:00-9:00) recorded at 3,887 veh/hr, 3,632 veh/hr, 3,072 veh/hr, 2,760 veh/hr and 2,707 veh/hr for UI8, UI2, UI 10, UI20 and UI9 respectively. Conscience, the third highest traffic volume was during midday at (13:00-14:00) got 3,628veh/hr, 3,612 veh/hr, 2,822 veh/hr, 2,551 veh/hr and 2,168 veh/hr for UI20, UI 2, UI8, UI 9 and UI 10, respectively.

Although UI 2 achieved the highest traffic volume during (17:00-18:00= 4804 veh/hr, however in right turning volume from a minor road onto a major UI2 received the lowest (17:00-18:00= 34veh/hr).

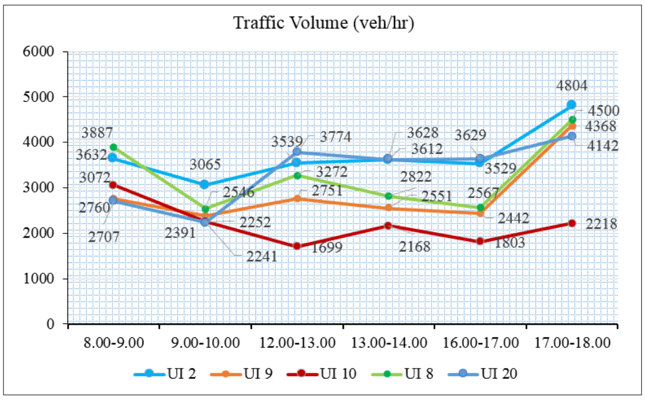

Figure 2: Traffic Volume from Minor Road Fluctuation Over Six Hours Survey During A Typical Weekday

Right Turn Volume from Minor Road

Right turning volume in this section define as right turn motorist from a minor road onto a major road at the selected intersection (Figure 2). Five unsignalized intersections (UI) were involved in right turn volume analysis, which are UI 2, UI 8, UI 9, UI 10 and UI 20. As mentioned before in the previous section, the traffic count was based on hourly traffic turning volume (veh/hr) at three peak hours (8:00-10:00), (12:00-14:00) and (16:00-18:00). Figure 3 shows turning volume over six hours' duration. Right turning flow at UI 10 recorded in uniform trend in early stage recorded (8:00-9:00 = 41 veh/hr), (9:00-10:00 = 39 veh/hr), (12:00-13:00 = 44 veh/hr), (13:00-14:00 = 55 veh/hr), (16:00-17:00 = 42 veh/hr) but drastically increase at (17:00-18:00 = 244 veh/hr). Other turning flows at UI 2, UI 8, UI9 and UI20 represent a stable turning manoeuvre and less fluctuation with a minimum range of 34 veh/hr and a maximum of 92 veh/hr, compared with UI 10.

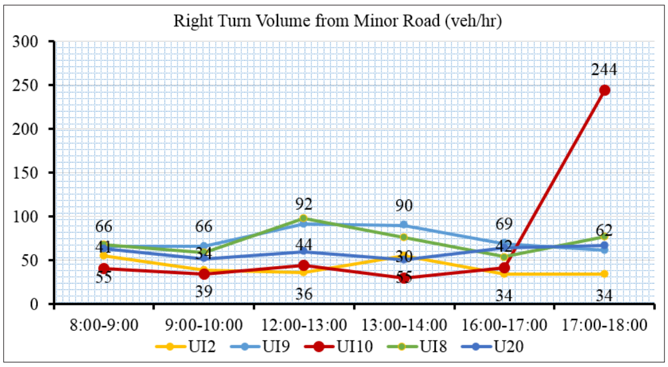

Figure 3: Right Turning Volume from Minor Road Fluctuation Over Six Hours Survey During A Typical Weekday

Left Turning Volume from Minor Road

Left turning volume in this section is defined as a left turn motorist from a minor road onto a major road at the selected intersection. Figure 4 shows the turning volume over six hours' duration. Left turning volume at UI 9 recorded in less fluctuation trend in morning and midday, was (8:00-9:00 = 89 veh/hr), (9:00-10:00 = 65 veh/hr), (12:00-13:00 = 78 veh/hr), (13:00-14:00 = 67 veh/hr), (16:00-17:00 = 97 veh/hr) however sharply rise during the afternoon (17:00-18:00 = 213 veh/hr). Meanwhile, UI 10 has a contrariwise situation pattern during midday (12:00-13:00 = 115veh/hr) achieving the highest left-turning volume and second highest during the afternoon (17:00-18:00= 184 veh/hr). Subsequently, other turning volumes at UI 2, UI 8 and UI 20 represent a uniform turning manoeuvre and less fluctuation with a minimum range of 8 veh/hr and a maximum of 63 veh/hr.

Citation: Fajaruddin Mustakim, Mohammad Nazir Ahmad, Rabiah Abdul Kadir, Azlan Abdul Aziz, Othman Che Puan (2023) Traffic Behaviour Analysis using Logistic Regression Method (LRM) and Structural Equation Modelling (SEM). Journal of Artificial Intelligence & Cloud Computing. SRC/JAICC-130.
DOI: doi.org/10.47363/JAICC/2023(2)121

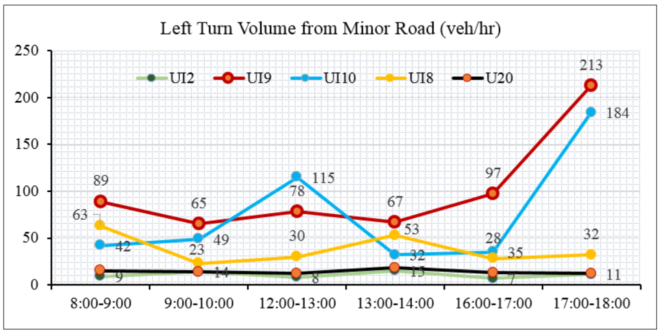

Figure 4: Left Turning Volume from Minor Road Fluctuation Over Six Hours Survey During A Typical Weekday

Right Turn Volume from Major Road (RTVmr)

Right turning volume from a major road in this section is defined as the right turn motorist from a major road onto a minor road at the selected intersection. Figure 5 shows the turning volume over six hours' duration. Right turn volume from the major road has a different turning pattern compared with other right turning volumes from the minor road (Figure 3) and left turning volume from the minor road (Figure 4). Concisely, the right turning volume at UI 8 in the morning was (8:00-9:00 = 70 veh/hr) and the number of traffic plumped (9:00-10:00 = 40 veh/hr). Meanwhile, the volume rises drastically at 121 veh/hr and 123 veh/hr during midday (12:00-13:00) and (13:00-14:00) respectively. However, the volume dropped suddenly with 46 veh/hr in the afternoon (16:00-17:00) and increase doubled to 83 veh/hr at (17:00-18:00). Briefly, UI 20 demonstrate active variation volume flow from morning stated (8:00-9:00 = 46 veh/hr) drop at (9:00-10:00 = 26 veh/hr) sharply rise in midday (12:00-13:00 = 58 veh/hr) and (13:00-14:00 = 95 veh/hr), slightly reduce in the afternoon (16:00-17:00 = 86 veh/hr) before rocketed (17:00-18:00 = 111 veh/hr).

UI 2 turning volume performed a gradually increasing trend from early morning (8:00-9:00=18 veh/hr) and (9:00-10:00= 19 veh/hr), continuing during midday (12:00-13:00 = 21 veh/hr) and (13:00-14:00= 28 veh/hr), slightly rise during the afternoon was (16:00-17:00 = 46 veh/hr) and (17:00-18:00 = 56 veh/hr).

UI 10 manoeuvre volume trend represents uniform flow in the morning and afternoon recorded (8:00-9:00 = 61 veh/hr, 9:00-10:00 = 69 veh/hr, 12:00-13:00 = 72 veh/hr, 13:00-14:00 = 62 veh/hr), drop slightly in the afternoon (16:00-17:00 = 34 veh/hr) and finally increase (17:00-18:00 = 66 veh/hr). Meanwhile, UI9 turning flow has less traffic volume and fluctuation between 14 veh/hr and 33 veh/hr.

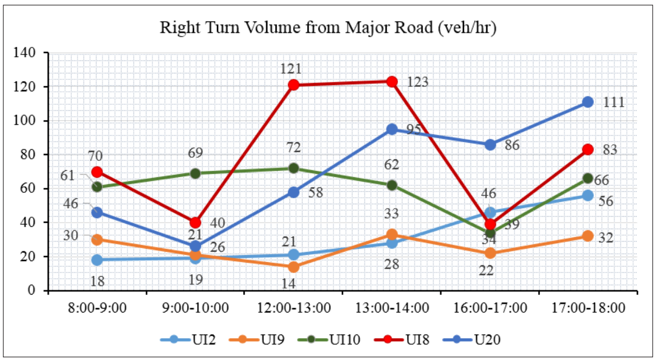

Figure 5: Right Turning Volume from Major Road Fluctuation Over Six Hours Survey During A Typical Weekday

Development of Logistic Regression

812 data points for Right-Turning Motorists (RTM) comprised 351 accepted gaps and 461 rejected gaps utilised in the development of the RTM Models. Subsequently, gap, motorcycle rider, conflict lane change (CLC), channelisation, and traffic signal were set as the independent variables or predictors. Meanwhile, the dependent variable in logistic regression was RTM and set to 1 and 0 if otherwise. Validation of the model was done with SPSS Statistics 26. The significance intervals of 90%, 95% and 99% were determined using a stepwise selection procedure. The description of all dependent and independent variables is explained in Table 1.

Table 1: Attributes of Traffic Behaviour Models

Abbr.	Description
RTM	RTM=1 if motorist turned right at a gap, but 0 if not.
Gap	Gap which is rejected or accepted (sec).
Car	Car=1 if the RTM is car, and 0 if otherwise.
Mc	Mc=1 if the RTM is motorcycle, and 0 if otherwise.
Rider	Rider= 1 if the RTM is rider, and 0 if otherwise.
Van	Van= 1 if the RTM is van, and 0 if otherwise.
Lorry	Lorry= 1 if the RTM is lorry, and 0 if otherwise.
Bus	Bus= 1 if the RTM is bus, and 0 if otherwise.
CLC	Conflict lane change = 1 if CLC occurred, and 0 if otherwise.
Gap1	If the gap was gap pattern 1 in Fig. 4, Gap1=1, but 0 if not.
Gap2	If the gap was gap pattern 2 in Fig. 4, Gap2=1, but 0 if not.
Gap3	If the gap was gap pattern 3 in Fig. 4, Gap3=1, but 0 if not.
Gap4	If the gap was gap pattern 4 in Fig. 4, Gap4=1, but 0 if not.
Gap5	If the gap was gap pattern 5 in Fig. 5, Gap5=1, but 0 if not.
Chanlz	If channelisation facility is in unsignalized intersection, so Chanlz = 1, but 0 if not.
TSignal	If all vehicles are in unsignalized intersection, so TSignal=1, but 0 if not.

In Model 1, three variable which is vehicle gap, conflict lane change (CLC) and traffic signal achieved significance at a 99% level, followed by rider received 95% significance and channelisation stated 90% significance level.

Table 2: Logistic Regression Models for Right Turn Motorists (RTMs)

Attributes	Model 1
	All (detail)
Constant	-5.35
Gap	0.96 (201.97)***
CLC	4.18 (28.95)***
Rider	0.59 (5.73)**
TSignal	-0.67 (6.36)***
Chanlz	-0.83 (3.04)*
N	812
NagelkerkeR2	0.72
H.R-Right Turn	83%
H.R-Total	87%

*,**,***=Significant at 90%,95%, and 99% levels, respectively

In RTM logistic model 1, a positive sign of rider and serious conflict lane change shows that RTM is likely to accept a shorter gap acceptance. Conversely, a negative sign in the traffic signal and channelisation can be interpreted to mean that RTM is likely to accept a longer gap.

Development of SEM

The dataset applied in the right-turning motorist (RTM) behaviour model for logistic regression is the same dataset implemented in structural equation modelling. The relationship between exogenous and endogenous latent variables can be visualised through the SEM diagram. In addition, SEM's ability to assess both causal impact among these observed and unobserved variables. In this model, endogenous variables were RTM and six exogenous variables (rider, gap, conflict lane change, channelisation, and traffic signal).

The IBM SPSS AMOS 23 computer programming is used to develop the structural equation modelling (SEM). AMOS 23 Arbuckle and Wothke, is a software for analysing, validating, and testing observed data [4]. Meanwhile, SPSS 26 is used to prepare the dataset. Table 3 shows the outcomes result for each variable in the analysis. The detail such as parameter estimate, standard error (S. E), critical ratio (C. R) and level of statistical significance (P) are described in Table 3. All four variables namely gap acceptance (Gap), traffic signal (TSignal) and conflict lane change (CLC) were highly statistical significance at 99%, excluding motorcycle riders got a 95% significance level. Meanwhile, channelisation was found insignificant. Each parameter shows a positive sign, except traffic signal and channelisation get a negative sign. The negative and positive coefficients of a structural equation model are like correlation or regression coefficients McIntosh and Gonzalez-Lima [27]. The positive sign of conflict lane change, gap, and motorcycle rider shows RTM is likely to accept a short gap. Subsequently, the negative sign of the traffic signal shows the RTM is likely to accept a longer gap.

Table 3: Result of SEM for Traffic Behaviour

	Estimate	S.E.	C.R.	P	Label
RTM <--- CLC	.573	.063	9.168	***	par_5
RTM <--- Gap	.080	.003	25.540	***	par_6
RTM <--- Tsignal	-.080	.029	-2.764	***	par_7
RTM <--- Chlzation	-.041	.045	-.926	.354	par_10
RTM <--- Rider	.066	.028	2.394	.017	par_11

*,**,***=Significant at 90%,95%, and 99% levels, respectively

The Chi-square $\chi 2$ was 9.445, meanwhile, $\chi 2$/df statistic index was 3.145, which is less than 5.0, showing a good fit of the model [28]. The root mean square residual (RMR) index has a value of 0.010, the index of (RMR) less than 0.08 mean the index is quite good [29]. Subsequently, the root mean squared error of approximation (RMSEA) is 0.05, which is less than 0.05. A value equal to or less than 0.05 usually shows the good quality of the model and when RMSEA is between 0.07-0.09, the model is in categories of logical estimation. (O. Khassawneh et al., 2022) [30]. The goodness-of-fit index (GFI) and the adjusted goodness-of-fit index (AGFI) had values of 0.996 and 0.973, respectively. The index was close to 1.0, showing a perfect fit for the conflict model [31]. The comparative fit index (CFI), tucker-lewis coefficient (TLI) and normal fit index (NFI) were 0.996, 0.949 and 0.985, respectively. All Incremental Fit Indexes are close to 1.0, representing the best fit of the model [32]. We concluded that the fit of our model is excellent and sufficient to proceed. The summary of the model index value and its requirement is shown in Table 3. Several efforts to construct the final structure of the behaviour model are illustrated in Figure 6.

Citation: Fajaruddin Mustakim, Mohammad Nazir Ahmad, Rabiah Abdul Kadir, Azlan Abdul Aziz, Othman Che Puan (2023) Traffic Behaviour Analysis using Logistic Regression Method (LRM) and Structural Equation Modelling (SEM). Journal of Artificial Intelligence & Cloud Computing. SRC/JAICC-130.
DOI: doi.org/10.47363/JAICC/2023(2)121

Table 3: Goodness-of-fit indexes

Indexes Values	Values	Requirement
Chi-Square	9.45	Significance > 0.05
χ^2/df	3.15	Between 2-5 (Tabachnick and Fidell, 2007)
Goodness of Fit Index (GFI)	0.99	close to 1, (Hair et al., 2010)
Adjusted Goodness of Fit Index (AGFI)	0.97	close to 1, (Hair et al., 2010)
Comparative Fit Index (CFI)	0.99	> 0.90, (Benitez et al., 2018)
Tucker-Lewis Coefficient Index (TLI)	0.95	> 0.90, (Benitez et al., 2018)
Normal Fit Index (NFI)	0.99	> 0.90, (Benitez et al., 2018)
Relative Fit Index (RFI)	0.93	> 0.90, (Benitez et al., 2018)
Incremental Fix Index (IFI)	0.99	> 0.90, (Benitez et al., 2018)
Root Mean square Residual (RMR)	0.01	< 0.08 (Benitez et al., 2020)
Root Mean Square Error of Approximation (RMSEA)	0.05	< 0.08 (O. Khassawneh et al., 2022)

Root mean squared error of approximation calculation formula is defined by Equation 2.

$$RMSEA = \sqrt{\frac{(\chi^2 - df)}{df(N-1)}}$$

Where N number of observations (812), df the degrees of freedom (3) and Chi-square $\chi 2$ of the model (9.45).

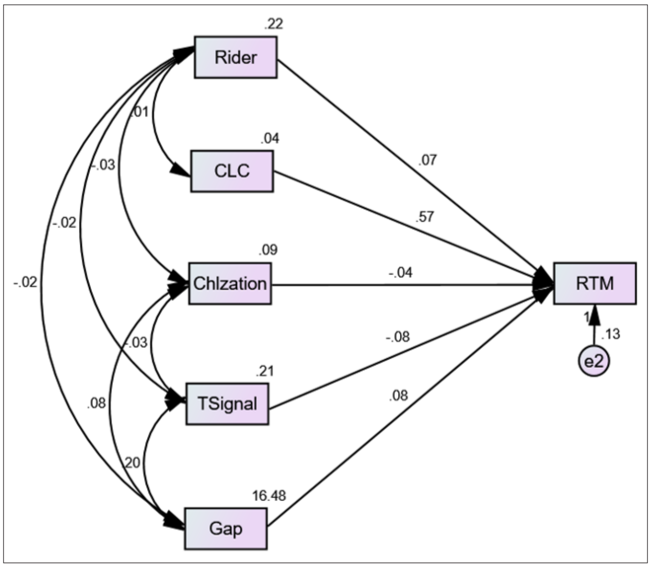

Figure 6: Structural Equation Modelling Network (Multiple Regression Path)

Discussion

- Using SEM is quite challenging at the early stage, this is because the researcher needs to have basic knowledge regarding statistical analysis, model fit index, variable network, connecting between endogenous and exogenous variables, observe and unobserved (latent) quantitative variables. It became more interesting after regularly practising the development of the structure model with the right guideline. Especially on the function of modification indices (MI) to determine the excellent result on model index fitness. By increasing covariance between the independent variables and modification based on the MI.
- Overall traffic volume at five UI represents stable fluctuation from early morning until midday except in the afternoon (17:00-18:00) at UI20 was 4,804 veh/hr, which is the highest traffic volume.
- Subsequently, right turning volume (RTV) from a minor road onto a major road have an average volume between 34 veh/hr to 92 veh/hr from morning until afternoon except for UI10 remark the peak turning volume was 244 veh/hr during (17:00-18:00).
- Meanwhile, the left turning volume (LTV) from a minor road onto a major road has almost the same turning pattern as the right turning volume (RTV) from a minor road onto a major road except for UI 10 experiences two-time peak hours during midday (12:00-13:00 = 115 veh/hr) and afternoon (17:00-18:00 = 184 veh/hr).
- The Right turning volume (RTV) from the major road as shown in Figure 7 illustrates the most dynamic fluctuation, especially UI 8 and UI 20. Meanwhile, UI10, UI9 and UI 2 represent moderate fluctuation in traffic volume flow.
- Although UI 2 achieved the highest traffic volume during (17:00-18:00= 4804 veh/hr), however, in right turning volume from a minor road onto a major UI2 received the lowest (17:00-18:00= 34veh/hr). Conversely, UI 10 recorded the lowest traffic volume in the afternoon (17:00-18:00= 2218 veh/hr), yet in right turning volume (RTV) from a minor road onto a major road stated the highest turning volume (17:00-18:00= 244 veh/hr).

Conclusion

- The study found a significant relationship between (RTM) as the dependent variable and four independent variables (gap, motorcycle rider, conflict lane change and traffic signal). Two variables that encourage RTM to accept a short gap are RTM motorcycle riders and conflict lane change meanwhile, traffic signals are likely to accept a longer gap. This result describes that motorcycle rider is the vulnerable type of motorist to make a right turn from a minor road at UI. Meanwhile to reduce conflict lane change by introducing artificial intelligent traffic signals on the high-risk unsignalized intersection.

- The resulting outcome from both scientific methods (SEM and LRM) revealed similarity independent variables such as gap, traffic signal and conflict lane change, acquiring a significance level at 99%. Only the RTM motorcycle achieved a 95% confidence level. The same -/+ sign of each variable is given in both methods.
- Despite the independent variable of channelisation insignificance in structural equation modelling, these parameters have statistical significance at 90% in the logistic regression model. All ten goodness-of-fit indices support the analysis having more accuracy in the SEM. SEM can present a visualisation modelling network. Meanwhile, LRM has the advantage of assist explaining the result and early preparation of selection variables in the model before execution of the SEM. Thus, a combination of both scientific and statistical might complement each other and create essential understanding in our research work.

References

1. Joreskog KG, Sorbom D (1988) PRELIS: A Program for Multivariate Data Screening and Data Summarization. A Preprocessor for LISREL - SSI, Inc, Chicago https://www.scirp.org/(S(lz5mqp453edsnp55rrgjct55))/reference/ReferencesPapers.aspx?ReferenceID=884372.
2. Joreskog KG, Sorbom D (1989) LISREL 7: A Guide to the Program and Applications. SPSS, Inc, Chicago https://www.scirp.org/(S(i43dyn45teexjx455qlt3d2q))/reference/ReferencesPapers.aspx?ReferenceID=1624759.
3. Joreskog K G, Sorbom D (1993) LISREL 8: Structural Equation Modelling with the SIMPLIS Command Language. International Inc, Chicago https://www.scirp.org/(S(351jmbntvnsjt1aadkozje))/reference/referencespapers.aspx?referenceid=1458224.
4. Arbuckle J L (2013) Amos 22 User Guide: Amos Development Corporation https://www.scirp.org/%28S%28czeh2tfqyw2orz553k1w0r45%29%29/reference/referencespapers.aspx?referenceid=3043546.
5. Hair J F, Black W C, Babin B J, Anderson R E, Tatham R L (2006) Multivariate data analysis (6th ed). Upper Saddle River, NJ: Prentice-Hall https://www.scirp.org/(S(lz5mqp453edsnp55rrgjct55))/reference/ReferencesPapers.aspx?ReferenceID=1747215.
6. Bentler P M (1995). EQS structural equations program manual. Encino, CA: Multivariate Software https://www.scirp.org/(S(351jmbntvnsjt1aadkozje))/reference/referencespapers.aspx?referenceid=213476.
7. Shaun Mc Quitty, Marco Wolf (2013) Structural Equation Modeling: A Practical Introduction. Journal of African Business 14: 58-69.
8. Fajaruddin M, Fujita M (2009) Crash Prediction Model for Rural Roadway. http://library.jsce.or.jp/jsce/open/00039/201111_no44/pdf/150.pdf.
9. Fajaruddin M, M Fujita (2011) Development of accident predictive model for rural roadway. An International Journal of Science Engineering and Technology 58: 126-131.
10. Fajaruddin M, Motohiro Fujita, Wisinee W (2013) Analysis of fatal, serious accidents and dangerous vehicle movements at access points on Malaysian rural federal roads. Journal of Japan Society of Civil Engineers 69: 286-299.
11. Fajaruddin M, Ismail A R, Erwan S, Zulkifli S, Shamsul K A, et al. (2014) Gap Acceptance Behavior Model for Non-signalized T-Intersections on Malaysia Rural Roadway. Australian Journal of Basic and Applied Sciences 8: 91-97.
12. Fajaruddin M, A A Aziz, M Fujita, Wisinee W, M N Ahmad, et al. (2021) An Analysis of Right-Turning Vehicles and Gap Acceptance Behaviour Models on Malaysian Rural Roads. International Journal of Road Safety 2: 45-53.
13. F Mustakim, A A Aziz, M Fujita, M N Ahmad (2021) Motorcycle and Passenger Cars Behavior at Unsignalized Intersection in Malaysia Rural Roadway. Turkish Online Journal of Qualitative Inquire 12: 2249-2264.
14. F Mustakim, M N Ahmad, A A Aziz, S Jamian, A Mahmud, et al. (2023) Sustainable Hydrodynamic of Artificial Neural Networks and Logistic Regression Model to Lane Change Serious Conflict at Unsignalized Intersection on Malaysia's Federal Route. General Clinic Surgery Medicine 1: 82-89.
15. Tabachnick B G, Fidell L S (2007) Using Multivariate Statistics (5th ed). New York: Allyn and Bacon https://www.scirp.org/(S(czeh2tfqyw2orz553k1w0r45))/reference/ReferencesPapers.aspx?ReferenceID=1887285.
16. Steiger J H (1990) Structural model evaluation and modification: An interval estimation approach. Multivariate Behavioral Research 25: 173-180.
17. Byrne B M (1998) Structural Equation Modeling with LISREL, PRELIS and SIMPLIS: Basic Concepts, Applications and Programming. Mahwah, New Jersey: Lawrence Erlbaum Associates https://www.scirp.org/(S(351jmbntvnsjt1aadkposzje))/reference/ReferencesPapers.aspx?ReferenceID=148465.
18. Mc Quitty S (2004) Statistical power and structural equation models in business research. Journal of Business Research 57: 175-183.
19. MacCallum R C, Browne M W, Sugawara H M (1996) Power Analysis and Determination of Sample Size for Covariance Structure Modeling, Psychological Methods 1: 130-149.
20. Miles J, Shevlin, M (1998) Effects of sample size, model specification and factor loadings on the GFI in confirmatory factor analysis. Personality and Individual Differences 25: 85-90.
21. Hu L T, Bentler P M (1999) Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives. Structural Equation Modeling 6: 1-55.
22. Bentler P M (1990) Comparative Fit Indexes in Structural Models. Psychological Bulletin 107: 238-246.
23. Mousa Alavi, Denis C Visentin, Deependra K Thapa, Glenn E Hunt, Roger Watson, et al (2020) Chi-square for model fit in confirmatory factor analysis. Journal Advance Nursey 76: 2209-2211.
24. Shaun McQuitty, Marco Wolf (2013) Structural Equation Modeling: A Practical Introduction. Journal of African Business, 14: 58-69.
25. Loehlin, J C (2004) Latent Variable Models: An Introduction to Factor, Path, and Structural Equation Analysis (Fourth Edition ed.). London: Lawrence Erlbaum Associates, Publishers Mahwah, New Jersey https://www.isical.ac.in/~arnabc/discrete/loehlin.pdf.
26. Timothy Teo, Tsai L T, Yang C C (2013). Applying Structural Equation Modeling (SEM) in Educational Research. An Introduction. In M. S. Khine (Ed.), Application of Structural Equation Modelling in Educational Research and Practice. Sense Publishers 3-21.
27. McIntosh A R, Gonzalez Lima F (1994) Structural equation modeling and its application to network analysis in functional brain imaging. Hum Brain Mapp 2: 2-22.
28. Schumacker R E, Lomax R G (2004) A Beginner's Guide to Structural Equation Modeling. London: Lawrence Erlbaum Associates Inc 1-304.
29. Benitez J, Henseler J, Castillo A, Schuberth F (2020) How

to perform and report an impactful analysis using partial least squares: Guidelines for confirmatory and explanatory IS research. Inf Manag 57: 103168.

30. Osama Khassawneh, Tamara Mohammad, Rabeb Ben-Abdallah (2022) The Impact of Leadership on Boosting Employee Creativity: The Role of Knowledge Sharing as a Mediator. Administrative Sciences 12: 175.
31. Hair J F, Black W C, Babin B J, Anderson R E (2010) Multivariate data analysis (7th ed). Englewood Cliffs: Prentice Hall https://www.drnishikantjha.com/papersCollection/Multivariate%20Data%20Analysis.pdf.
32. Benitez J, Ray G, Henseler J (2018) Impact of information technology infrastructure flexibility on mergers and acquisitions. MIS Q 42: 25-43.

Copyright: ©2023 Fajaruddin Mustakim, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.