Factors influencing Traffic Accidents in Countries with Fast-growing Economies

Kulothungan, Vimalathithan, Vallabhaneni pratap-naidu, Kuthethur murari pranesh rao, Thimiri gunasagaram venkataramana, Chinmoy Pal

Abstract Road traffic accidents are increasing year on year in low- and middle-income (GDP) countries compared to high-income countries, which are showing a downward trend. These trends could be due to differences in law enforcement, vehicle safety, road standards, and many other factors. A detailed examination is needed of the fatality trends across different categories, such as four-wheeled vehicles, pedestrians, motorized two- and three-wheeled vehicles, bicycles, etc. The objective of this study is to understand the inherent attribute relationships affecting the fatalities for ASEAN+6 countries. This study used data from the World Health Organization's Global Status Report on Road Safety, focusing on 44 selected attributes from the ASEAN+6 countries. The Self-Organizing Map (SOM) technique was applied to map the non-linear relationships between the 44 different attributes. Three characteristic clustering groups were identified for ASEAN+6 countries using the SOM method. Attributes such as enforcement of speed limits, road quality, enforcement of seatbelt laws, drunk driving laws and helmet use can be effective in reducing road fatalities.

Keywords Safety, World Health Organization, Accident Analysis, Cluster Analysis, ASEAN.

I. INTRODUCTION

According to the WHO's Global Status Report 2015 [1], the ninth leading cause of death worldwide is road traffic fatality. Tragically, road traffic fatalities are projected to claim the seventh spot among leading causes of death by 2030. Around 1.2 million people are killed on roads every year, and the numbers of those injured far exceeds the numbers of those killed (approximately 50 million people sustain non-fatal injuries). Road traffic fatality is the leading cause of death among the 15–29-year-old population. Further, road traffic fatalities show an increasing trend for low- and middle-income countries, which account for 90% of road traffic fatalities, despite having 54% of the world's registered vehicles. An economic loss of 5% of GDP is estimated for low- and middle-income countries due to road traffic fatalities. These characteristics are due to the fast urbanization and motorization associated with rapid economic growth (WHO Global Status Report, 2015) [1]. Nearly half of all traffic fatalities are accounted for by pedestrians, by motorized two-/three-wheelers and by cyclists.

A comprehensive dataset is needed to understand the crash characteristics and causes of injury. It is important to estimate the inherent relationship between multiple factors/attributes and to interpret the findings. Attribute relationships can be understood by grouping them based on similar behavior using clustering analysis [2]. Many researchers have explored accident databases using different types of clustering analysis, such as centroidal, density, hierarchical, neural network, etc. To study traffic risk and health risk for different cities in Turkey, k-means and fuzzy c-means clustering was used [3]. Six clustering groups identified in that study and its findings indicate that health risk and traffic risk are high for poorly developed, low-income rural provinces.

Crash patterns in car-to-bicycle collisions in Denmark (2007–2011) were studied through the latent class clustering method [4]. Thirteen different cyclist-to-motorist latent groups were identified, with specific patterns for rural and urban regions, and the influential attributes were found to be speed limit, road surface, type of infrastructure and number of lanes. Uno *et al.* [5] used FARS database 2010 to understand the complicated relationship within variables using self-organizing map (SOM)-based clustering analysis.

V. Kulothungan (e-mail: vimalathithan.kulothungan@rntbci.com), P.-N. Vallabhaneni, M. P. R. Kuthethur and G. V. Thimiri work at Renault Nissan Technology Business Center, Chennai, TN, India. C. Pal works at Nissan Motor Co., 1-1 Aoyama, Morinosato, Atsugi, Kanagawa, Japan.

The data of 16,180 fatally injured drivers were used for clustering analysis and five cluster groups were identified. The SOM clustering method was used to understand the characteristics of Indian road accidents [6]. It found that the influential attributes responsible for accidents were improper overtaking, improper highway parking, and low helmet usage rates. Pal *et al.* [6] used the SOM clustering method to study the WHO data for 176 countries and found the influential attributes to be Gross National Income (GNI), speed, population, gender, and vehicle standards.

The objective of this study is to identify the main influential attributes for fatal accidents on roads in ASEAN+6 countries through SOM-based clustering analysis.

II. METHODS

Data

This study used the Global Status Report on road safety compiled by WHO [1]. The report contains data from 180 countries worldwide, covering 6.97 billion people. The report on road safety was developed based on the inclusive collection of documents shared by participating countries. Attribute selection was based on experience and prior research findings, as discussed in the Introduction, above. Attributes with a frequency of occurrence below 5% were removed, resulting in a final selection of 44 attributes for this study (see Appendix, Fig. A1). Out of 44 attributes, five attributes are associated with accidental deaths, five attributes are associated with vehicle, 20 attributes are associated with legal regulations and enforcement, seven attributes are associated with safety standards, five attributes are associated with policies, and two attributes are associated with income of the country. Road quality attribute information was derived from the World Economic Forum's Global Competitiveness Report: 2014-15 [7].

Approach

Clustering is a process of organizing objects into groups based on their similarities. This process helps to divide data into useful groups, which helps us to understand the inherent relationships among the attributes. In this study, a SOM-based clustering method was selected. SOM utilizes neural networks logic and decreases data dimensionality for better visualization. Commercial software package Viscovery was used in this study to construct the two-dimensional maps.

III. RESULTS

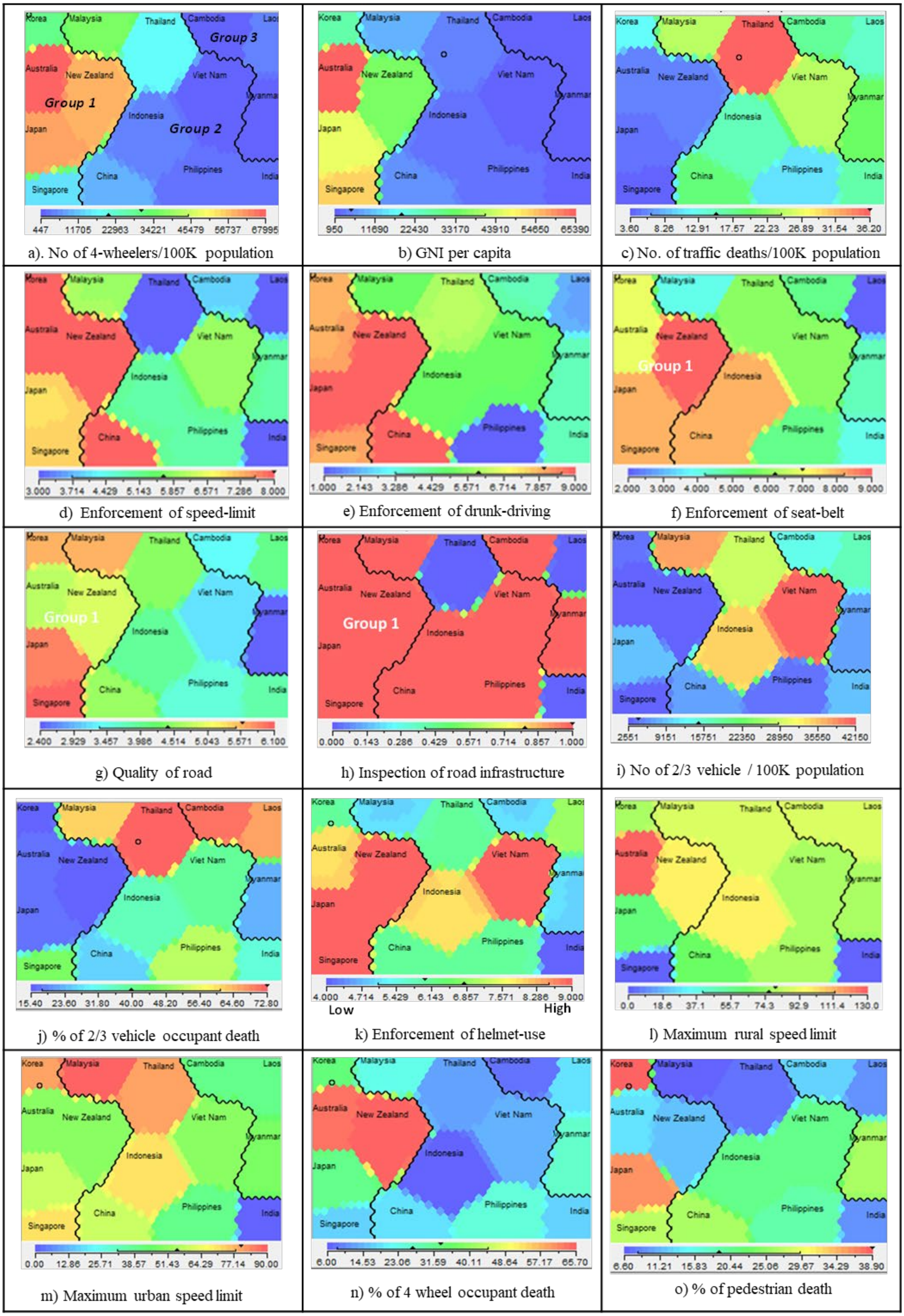
This section discusses the results of SOM cluster analysis for ASEAN+6 countries.

SOM Analysis (ASEAN+6 economic zones)

The selected data (44 variables of ASEAN+6 countries) were used as input to the SOM algorithm for clustering with 100x100 neuron cells two-dimensional plane. Each cell (data point) represents one neuron in each map, as shown by thick black boundary lines in Fig. 1. ASEAN+6 countries were classified into three main groups from the results of the SOM analysis, and their separation is shown in Fig. 1(a)–(o). The countries' locations demonstrate the characteristics of those in each three clusters and the influential parameter appears to be Gross National Income (GNI).

(i) Cluster of High GNI per Capita Group (Singapore, Japan, Australia, and New Zealand, Korea).
(ii) Cluster of Medium GNI per Capita Group (Malaysia, Indonesia, China, Thailand, Vietnam, Philippines, India).
(iii) Low GNI per Capita Group (Cambodia, Laos, Myanmar).

The SOM map color varies from blue (low value) to red (high value). For example, with the quality of road level between 1 and 7, the lowest grade is for Myanmar, and the highest grade is for Singapore (refer to the scale at the bottom of Fig. 1(g)). The variation of intermediate colour from deep blue to thick red represents the values in between them.



a). No of 4-wheelers/100K population
b) GNI per capita
c) No. of traffic deaths/100K population
d) Enforcement of speed-limit
e) Enforcement of drunk-driving
f) Enforcement of seat-belt
g) Quality of road
h) Inspection of road infrastructure
i) No of 2/3 vehicle / 100K population
j) % of 2/3 vehicle occupant death
k) Enforcement of helmet-use
l) Maximum rural speed limit
m) Maximum urban speed limit
n) % of 4 wheel occupant death
o) % of pedestrian death

Fig. 1. Three main groups were identified from the results of SOM analysis.

IV. DISCUSSION

Total Number of Deaths

The SOM map for total traffic deaths per 100,000 population was compared with those for GNI per capita and the total number of registered vehicles. This analysis attempted to understand where higher numbers of deaths occur and to identify relationships between these variables. While the set of SOM maps provides qualitative insights, it does not offer quantitative measurements. To identify all influential attributes (independent variables) related to the output variable (number of deaths in this study), careful observation of the same regions across all SOM maps is necessary.

(a) Fig. 1(c), (d), (h) suggest that low levels of speed-limit enforcement and road infrastructure inspection may contribute to the high number of deaths per 100,000 population in Thailand among ASEAN+6 countries.

(b) Fig. 1(c), (g) indicate that a low score for road quality could be associated with the moderately higher number of deaths per 100,000 population in Vietnam.

(c) Fig. 1(a), (c), (f) suggest that the low level of enforcement of seatbelt laws might be a contributing factor to the increased number of deaths per 100,000 population in Malaysia, which has a relatively high number of registered four-wheel vehicles.

(d) Fig. 1(c), (h), (g) suggest that the low number of deaths per 100,000 population in highly populated Singapore among ASEAN+6 countries could be attributed to the very high quality of roads, strict enforcement of seatbelt and helmet use, and thorough inspection of road infrastructure.

(e) Fig. 1(a), (c), (e) suggest that stricter enforcement of laws related to drunk driving in Japan and New Zealand has contributed to their lower numbers of deaths per 100,000 population.

Motorised two-/three-wheeler Deaths

Apart from Singapore, countries with lower GNI per capita, such as Malaysia, Thailand, Cambodia, and Laos, generally exhibit a higher percentage of motorised two-/three-wheeler deaths compared to Group-1 countries with high GNI per capita. These countries often have a larger number of motorised two-/three-wheelers in densely populated urban and rural areas. A closer observation reveals the following:

(a) Fig. 1(j), (k): due to the strict enforcement of helmet use for two-wheelers in Vietnam and Indonesia, the percentage of two-wheeler deaths is significantly lower. A reverse trend is observed in neighbouring countries like Thailand, Malaysia, Cambodia, and Laos.

(b) Fig. 1(j), (m): the higher maximum urban speed limit also contributes to the higher percentage of motorised two-/three-wheeler deaths in Thailand and Malaysia. Notably, the maximum urban speed limit in Singapore is higher compared to Japan, Australia, and New Zealand, which cannot be fully explained by the 44 attributes chosen in this study and warrants further investigation.

Four-Wheeler Deaths

By examining the SOM map showing the percentage of four-wheel vehicle occupant deaths (Fig. 1(n)) in Australia and New Zealand, along with other SOM maps such as maximum rural speed limit (Fig. 1(l)), GNI per capita (Fig. 1(b)), number of four-wheelers per 100,000 population (Fig. 1(a)), and enforcement of seatbelt laws (Fig. 1(f)), it can be concluded that the high maximum rural speed limits and the enforcement of seatbelt laws significantly influence the higher death rates of four-wheeler occupants in these two countries. This is particularly notable among the high GNI per capita countries in Group-1.

Pedestrian Deaths

A detailed examination of the percentage distribution of pedestrian deaths in Group-1 (Fig. 1(o)) reveals that pedestrian fatalities are higher in South Korea and Japan compared to Australia and New Zealand. Pedestrian accidents, which are more common in urban areas, occur at lower urban speeds in Australia and New Zealand compared to South Korea and Japan. Conversely, the maximum rural speed limit is higher in Australia and New Zealand than in South Korea and Japan (refer to Fig. 1(l)), leading to more fatal accidents involving four-wheeler vehicles. The higher population density in South Korea and Japan compared to Australia and New Zealand could be a contributing factor to the increased pedestrian deaths in these countries. However, since Singapore has a higher population density than both South Korea and Japan, further investigation is needed to understand this discrepancy.

V. CONCLUSION

Advanced statistical clustering analysis can be highly beneficial in developing future vehicle and road safety strategies for any country. This study aimed to investigate the relationships between fatalities, policies, vehicle populations, and other factors using self-organising map (SOM)-based clustering analysis, utilising data from the WHO 2015 report for ASEAN+6 countries.

Countries were classified into three groups based on the SOM clustering analysis. A closer examination revealed that this classification was primarily influenced by Gross National Income (GNI) per capita.

Measures such as speed-limit enforcement, road infrastructure inspection, road quality improvements, and the enforcement of seatbelt, drunk driving, and helmet laws appear to be effective in reducing traffic fatalities.

VI. ACKNOWLEDGEMENTS

The authors would like to acknowledge the work the World Health Organization, for its data collection and provision and for its continuous efforts to reduce road traffic accidents across the globe.

VII. REFERENCES

[1] World Health Organization accident fatalities report (2015) Global Status Report on Road Safety. http://www.who.int/violence_injury_prevention/road_ safety _status/2015/en/ (accessed 25 August 2016).

[2] Brian, S., Everitt, Sabine L., Morven, L., Daniel, S. (2011) Cluster Analysis. Wiley Series in Probability and Statistics; 5th edition.

[3] Ahmet, T., Ahmet, A. (2015) Clustering Analysis of Traffic Accident Risk in Turkey. *Iran Journal of Public Health*, **44**(3): pp. 425–426.

[4] Sigal, K., Carlo, G. P. (2012) Cyclist-Motorist Crash Patterns in Denmark: A Latent Class Clustering Approach. *Traffic Injury Prevention*, **14**(7): pp. 725–733.

[5] Uno, H., Kageyama, Y., Yamaguchi, A., Okabe, T. (2013) Method Development of Multi-Dimensional Accident Analysis Using Self Organizing Map. *SAE International Journal of Transport Safety*, **1**(1): pp. 64–75, https://doi.org/10.4271/2013-01-0758.

[6] Pal, C., Hirayama, S., *et al*. (2017) An insight of World Health Organization (WHO) accident database by cluster analysis with Self Organizing Map (SOM). AAAM 61st Annual Scientific Conference. *Traffic Injury Prevention*, **17**(No.S1): pp. 175–218.

[7] The Global Competitiveness Report (2014-2015) Full Data Edition. World Economic Forum. Available at: http://www3.weforum.org/docs/WEF_GlobalCompetitivenessReport_2014-15.pdf (accessed 14 September 2016).

VIII. APPENDIX

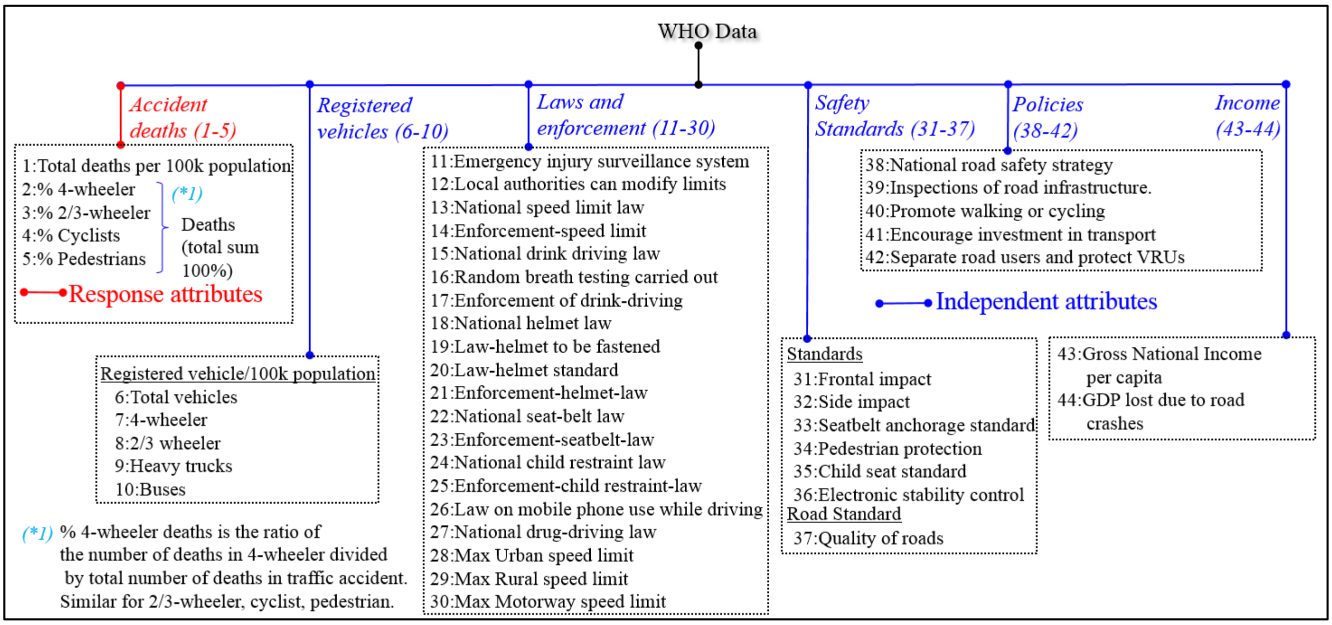


Fig. A1. Attributes/variables.

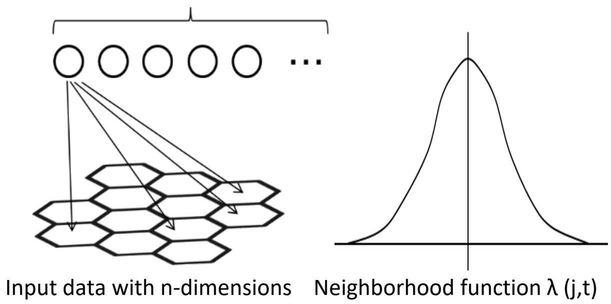


Fig. A2. Self-Organising Map.