Can we trust Microscopic Traffic Simulation Tools for Automated Vehicles Evaluation? A Scenario-Based Survey to Assess Confidence Levels of Simulation Tools Output.

H. Blache[a*], P-A. Laharotte[a], N-E. El Faouzi[a]

[a] Univ. Gustave Eiffel, Univ. Lyon, ENTPE, LICIT-ECO7 UMR T9401,F-69675, Lyon (France)
E-mail: hugues.blache@univ-eiffel.fr, pierre-antoine.laharotte@univ-eiffel.fr, nour-eddin.elfaouzi@univ-eiffel.fr
* Corresponding author

Extended abstract submitted for presentation at the Conference in Emerging Technologies in Transportation Systems (TRC-30)
September 02-03, 2024, Crete, Greece

Avril 26, 2024

Keywords: Analytical Hierarchy Process, Micro-Simulation Tools, Trust, Scenario-Based, Survey, Connected and Automated Vehicles

INTRODUCTION

Connected and Automated Vehicles (CAV) must undergo a series of rigorous tests before being authorised to operate on open roads. However, due to the high number of possible scenarios, conducting each test becomes economically unfeasible. To address this issue, scenario-based approaches [1] are being implemented to significantly reduce the number of tests required for CAV. These tests are divided between costly but realistic field trials versus easier-to-implement and cost-effective simulations. However, simulation tools provide a simplified view of reality and focus on a subset of attributes during the calibration stage. Therefore, assessing the reliability of the simulation tool per scenario is required, before admitting their outputs as acceptable. Based on expert knowledge, this paper aims to evaluate the degree of confidence that can be placed in simulation tools using Multi-Criteria Decision-Making methods (MCDM). Our paper copes with the problem of assigning labels (i.e. Confidence Levels) to any traffic scenario according to the perceived trust in the capacity of microscopic simulation tools to mimic the real world. We conducted a survey to feed a multi-criteria decision-making tool, namely the Analytic Hierarchy Process (AHP) [2], using in some studies to determine the complexity of a scenario [3].

Methodology

Developed by Prof. Saaty [2], the AHP is a branch of MCDM that unfolds in several stages.

1. **Step 1 - Identification of Respondent Type:** The participants we targeted in this study are individuals belonging to the traffic community with various experience levels when using microscopic simulation tools (*e.g* universities, R&D institutes, ...)
2. **Step 2 - Determination of Factors Under Study:** We modeled any scenario as a bag of words resulting from an ontology. Then, the factors under study directly result from the attributes, i.e., words featuring scenarios. The goal is to assign weights to each word and layer (i.e. group of words organised by topic) to achieve an overall score on the scenarios, depicted by the indices *Fij* in Figure 1.
3. **Step 3 - Value Assignment to Each Factor:** We conduct pairwise comparisons to evaluate the weights of each factor. This method involves comparing elements at the same level with shared parents using a judgement scale defined in Table 1.

Table 1 – *Judgement scale*

Importance value	*Linguistic Scale*	*Definition*
1	Equal Importance	The two attributes are equally important.
3	Moderate Importance	The first attribute assigns some importance to the objective compared to the second attribute
5	Strongly Importance	The first attribute assigns significant importance to the objective compared to the second attribute
7	Extremely Importance	The first attribute assigns extreme importance to the objective compared to the second attribute.

The weight W_i of attribute *i* is defined as:

$$W_i = \frac{V_i}{\sum_{j=1}^{n} V_j} \quad (1)$$

where V_i is the geometric mean of the values associated with criterion *i*, defined as the geometric mean of the values in the row of criterion *i*.

$$V_i = \left(\prod_{j=1}^{n} m_{ij}\right)^{1/n} \quad (2)$$

where m_{ij} is the judgement value between criterion *i* and criterion *j* given by the expert according to the judgement of Table I.

4. **Step 4 - Assigning Confidence Values to Each Scenario:** The final score results from the sum of weights assigned to factors selected by a scenario, expressed as:

$$W_{sf} = \sum_{i=1}^{n} W_i \quad (3)$$

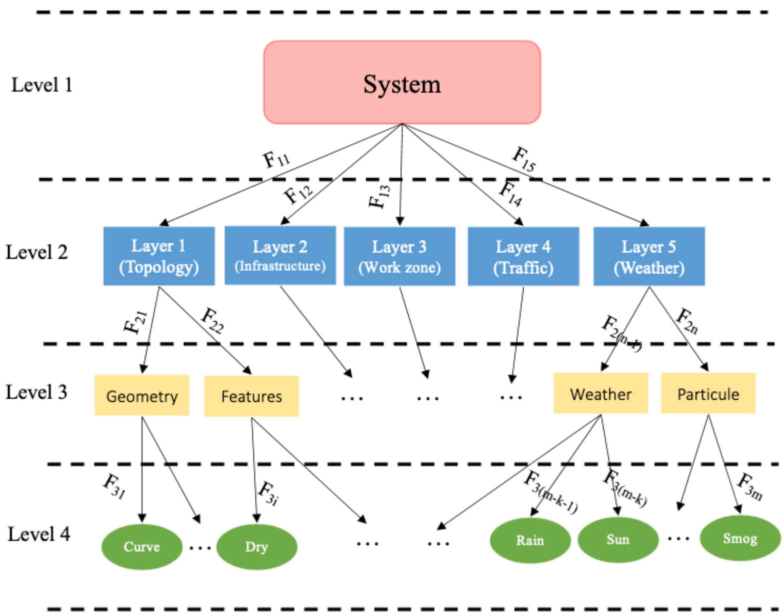

Figure 1 – *Hierarchy of ontology components*

Results

Functional Scenario Representation

Table 2 displays an example of a scenario produced by our ontology. The ontology generates scenarios with 5 layers representing: the road topology (layer 1), the road infrastructure (layer 2), the temporary modifications, like roadworks (layer 3), the static and dynamic elements (layer 4) and weather conditions (layer 5).

Table 2 – *Functional scenario example*

Layer	Attribu
Layer 5	'sunny' with 'no disturbance'
Layer 4	Presence of 'car' in 'free-flow' traffic condition
Layer 3	'No Temperorary'
Layer 2	Signs
Layer 1	'NoRoundabout' in "traffic lane-simple-junction' with 'straight ' and 'plane' geometry

Functional scenario scores and Gaussian mixture

We have standardised the scores obtained from the AHP to encompass it between 0 and 1. 0 matches low confidence, while 1 represents high confidence level. Visually, as shown in Figure 2, we observe a possible Gaussian mixture of distributions in the scores. This could be interpreted as clusters of confidence levels for scenarios. Within these clusters, there are groups where we have higher confidence compared to others.

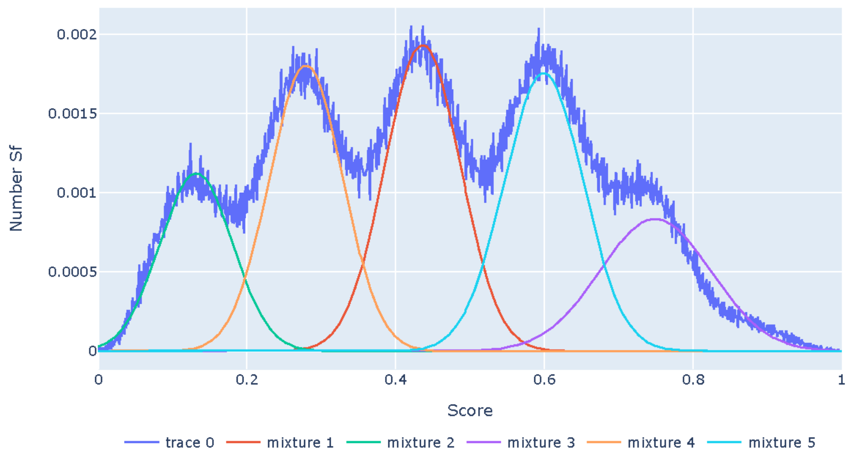

Figure 2 – *Gaussian Mixtures*

Word Cluster

We then attempted to identify, for each cluster, the types of words present in these groups. The results show expected consistencies, with confident groups characterised by words like "sunny weather in free flow condition." In contrast, scenarios involving "construction work" and "degraded weather" are associated with lower confidence. The intermediate groups represent cases where mixed conditions can degrade confidence in the scenario.

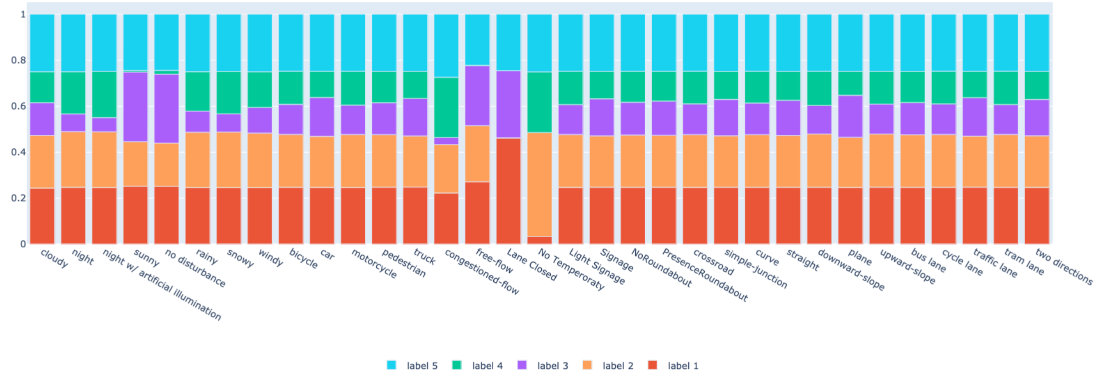

Figure 3 – *Word distribution in each cluster*

Discussion

In this article, we emphasise the confidence we can have in the output indicators of simulation tools. We conducted an extensive survey among domain-specific experts to apply an AHP for identifying confidence levels across different types of scenarios. Our preliminary results show promising and consistent distinctions between scenarios with higher confidence and those with lower confidence, conditioned by the terms used.

However, this approach has certain limitations, such as the independent consideration of layer combinations. It is highly likely that there are conditional probabilities associated with attribute combinations (*e.g* the combination of snowy weather and urban environment). The next step will be to account for these conditions.

References

[1] S. Riedmaier, T. Ponn, D. Ludwig, B. Schick, and F. Diermeyer, "Survey on Scenario-Based Safety Assessment of Automated Vehicles," IEEE Access, vol. 8, pp. 87 456–87 477, 2020, conference Name: IEEE Access.
[2] Saaty, Thomas L. "The analytic hierarchy process (AHP)." *The Journal of the Operational Research Society* 41.11 (1980): 1073-1076.
[3] Xia, Q., Duan, J., Gao, F., Chen, T., & Yang, C. (2017). *Automatic generation method of test scenario for ADAS based on complexity* (No. 2017-01-1992). SAE Technical Paper.