

Peer Review

Review of: "Advancing Autonomous Vehicle Safety: A Combined Fault Tree Analysis and Bayesian Network Approach"

Erwin de Gelder¹

1. Integrated Vehicle Safety, Netherlands Organisation for Applied Scientific Research, Delft, Netherlands

I enjoyed reading the manuscript. It is an important topic, and the writing style contributes to the pleasant read. I have, however, some questions and remarks that may need to be clarified in a revision:

1. Page 2, regarding the literature review: you mention there is a gap on “how failures in individual subsystems contribute to overall system risk and how these risks evolve over time”. Regarding the first part, that is clearly addressed in the current paper. I wonder, however, if there isn't any literature on this topic, potentially in other domains. If so, it would be good to mention it. If not, it might also be good to mention that. Regarding the second part (how risks evolve over time), is that addressed in the current manuscript?
2. Page 2, you mention ASIL B. Why (only) ASIL B?
3. In the same sentence, you mention “ensuring that AV ...for safety compliance”. Given that you only target ASIL B, it seems you are a bit overselling the approach.
4. Page 3, minor point, but you use FIT without first introducing the acronym.
5. Page 3, you mention “dynamic failure probabilities”. I don't know what you mean by this.
6. Page 3, “Chen et al... in Level 2 and 3 AVs”, where AV is short for autonomous vehicle. Please check the terminology, but AV typically refers to level 3 and above.
7. Page 4, it may be trivial, but you don't explain the relationship between C_{min} and $E_1, E_2, \dots, E_n$, which makes it hard to understand the explanation. In addition, it might help to explain also in words what a “minimal cut set” means.
8. Figures 3 to 5, the text is too small.

9. (!) Page 13, “we randomly assign failure rates to basic events”. This concerns me a bit. You draw many conclusions based on the experiment, but if I understand correctly, the failure rates are random, so the conclusions are random too. You even mention the conclusions in the abstract and conclusions section of the manuscript, but if the conclusions are random, it doesn’t seem very valuable to give so much attention to them. If I understood it incorrectly, please explain more clearly how you came to the failure rates and uncertainties in Table 1. And if I understood correctly, it would be good that when mentioning the results in the abstract and conclusion (if you want to do that at all), you clearly indicate how the use of random failure rates impacts the reliability of the conclusions drawn in the study.
10. (!) In general, all events are binary. For example, there is a steering failure or not (MCF3). To my understanding, using the BN, you assign probabilities to these events. However, for many events, it isn’t really binary. For example, for DMF2 (“Delayed Response to Obstacle”), in reality, there will always be some form of delay; the question is how large. And this is not a discrete probability distribution (or even only Bernoulli, which is what you effectively are using, if I understand correctly), but rather a continuous distribution. I am assuming that the mathematics become quite a lot more complicated and that the original FTA (which only uses 0 and 1) isn’t enough. Still, it would be closer to reality. I don’t expect a full solution to this from the authors but would be very happy to learn what your ideas on this are.
11. (!) Related to the previous point, I am missing a “discussion” section. After the experimental results, the conclusions are presented. For me, it would be highly valuable to include a discussion section in which you can present limitations of your work and potential future directions (in more detail than the single sentence on “future work” in the conclusions).
12. Page 17, minor point, [6] and [13] are the same.

Declarations

Potential competing interests: No potential competing interests to declare.