



HAL
open science

Continuous Risk Assessment for Automated Driving Systems using Random Forest

Anil Ranjitbhai Patel, Kunjkumar Thummar, Peter Liggesmeyer

► **To cite this version:**

Anil Ranjitbhai Patel, Kunjkumar Thummar, Peter Liggesmeyer. Continuous Risk Assessment for Automated Driving Systems using Random Forest. CARS@EDCC2024 Workshop - Critical Automotive applications: Robustness & Safety, Apr 2024, Leuven, Belgium. hal-04558508

HAL Id: hal-04558508

<https://hal.science/hal-04558508v1>

Submitted on 25 Apr 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Continuous Risk Assessment for Automated Driving Systems using Random Forest

Anil Ranjitbhai Patel
RPTU Kaiserslautern-Landau
Germany
apatel@rptu.de

Kunjkumar Thummar
RPTU Kaiserslautern-Landau
Germany
thummar@rptu.de

Peter Liggesmeyer
RPTU Kaiserslautern-Landau
Germany
peter.liggesmeyer@rptu.de

Abstract—The emergence of Automated Driving Systems (ADS) signifies a remarkable transformation in the realm of vehicular technology, bringing forth new challenges in assessing risks and formulating safety goals.

In this paper, we delve into the evolving nature of risk features in ADS and their prioritization, highlighting the inadequacy of conventional static risk assessment models. Our research has led to the development of a model that utilizes Random Forest (RF) for predicting risk ratings in diverse driving conditions. This model gathers and processes various types of data from On-Board Sensors (OBS), converting them into risk features through mathematical formulas. This process aids in estimating the unknown correlations between the values of risk features and predicting risk ratings as an output.

Utilizing a RF model for runtime risk assessment, we demonstrate how varying triggering conditions, such as unintended acceleration and unintended braking, affect the derived risk ratings. These ratings are then used to guide the implementation of risk-informed actions. They not only influence the modification of existing safety goals but also assist in formulating new, context-specific safety goals. Additionally, the predicted risk ratings can prompt functional modifications if necessary.

Index Terms—Automated Driving System, Risk Assessment, Random Forest, Severity, Controllability

I. INTRODUCTION

The Hazard Analysis and Risk Assessment (HARA) of ISO 26262 [1] plays an important role in ensuring the functional safety of automotive systems. It does so by evaluating hazardous events, their associated severity, exposure, and controllability, to determine an Automotive Safety Integrity Level (ASIL). In similar vein, the Hazard Identification and Risk Evaluation (HIRE) process from ISO 21448 [2] also evaluates factors such as severity, exposure, and controllability, its main focus lies in assessing the hazardous event itself, rather than assigning an ASIL.

A challenge arises from the subjective nature of risk rating assignment, often overlooking the dynamic interaction between vehicle dynamics and the environment. This leads to a lack of emphasis on the dynamic aspects of risk and their prioritization under varying conditions. The HARA and HIRE process, therefore, suffers from inconsistencies, primarily due to the subjective judgments of safety engineers. These judgments, influenced by personal backgrounds, experiences, and preferences, introduce variability in assessing hazard severity, controllability, and exposure [3, 4]. Such inconsistencies in

risk assessment can cause improper allocation of resources for safety goals and hazard mitigation.

This paper introduces a dynamic and responsive risk assessment method tailored for the evolving conditions of the driving environment and advancements in ADS. Central to this approach is the integration of runtime data and predictive analytics, leveraging data from OBS and sophisticated algorithms. This method is designed to be both empirical and adaptable, aiming to enhance the accuracy and relevance of risk evaluations in real-world driving scenarios.

A key aspect of this proposed approach is the use of a RF model for predicting severity and controllability ratings. The choice of this algorithm is strategic, reflecting its capability to handle a multitude of input variables and to model complex interactions effectively. This is coupled with the continuous risk assessment cycle, which facilitates moving from unknown-unsafe to known-safe regions. It does this by integrating multiple interactions between environmental factors, system dynamics, and user behavior, thereby adding nuance to the risk ratings.

The study focuses on the prioritization of various risk features in diverse conditions and their impact on runtime risk ratings. It particularly emphasizes different triggering conditions encountered during operation, illustrating how this method aligns safety measures with actual on-road conditions, thus ensuring a more objective assessment of risks.

II. RELATED WORK

The field of risk assessment for ADS encompasses a diverse array of methodologies, ranging from qualitative approaches that rely on subjective judgments [5] to quantitative methods grounded in numerical data and statistical analysis [6]. Dynamic models stand out for their adaptability to runtime data and changing conditions [7], while simulation and scenario-based methods [8] offer risk foresight through virtual modeling. Additionally, data-driven and machine learning approaches [9], particularly using the RF model, have gained traction for their ability to analyze vast data sets, with studies exploring its application in both accident data analysis [10], ongoing risk monitoring [11], and impact on accident frequency [12]. These varied perspectives provide a comprehensive view of risk assessment in ADS, laying a foundation for the development of continuous risk assessment

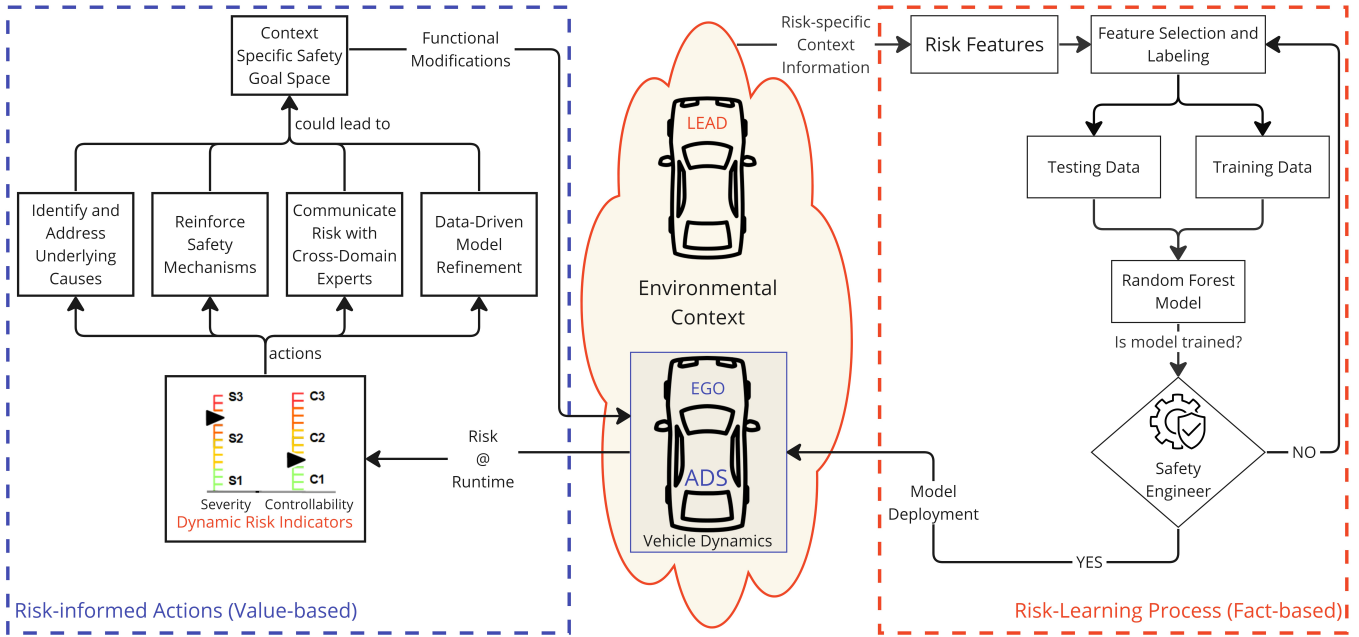


Fig. 1. Schematic Representation of the Continuous Risk Assessment Cycle

cycles that are responsive, adaptable, and grounded in a blend of empirical data and expert insight.

III. CONTINUOUS RISK ASSESSMENT CYCLE

The continuous risk assessment cycle is a methodical approach that splits into two essential yet interlinked modules: the “Risk-Learning Process” and the “Risk-Informed Actions” as shown in Fig. 1.

On its right side, the cycle embraces a fact-based methodology, rooted firmly in empirical data collected from runtime environments. This phase begins with the gathering of specific risk-related context through OBS, establishing a solid, data-driven foundation for assessing risk. A key element of this segment is the incorporation of machine learning, particularly the use of a RF model to analyze these risk features. Importantly, this process is supervised by a safety engineer who brings their expertise to the fore, fine-tuning and enhancing the model. This step ensures that the development of the model is not only reliant on data but also enriched by professional insight and judgment. Once the model has been approved by the safety engineer, it is deployed on the EGO vehicle model to generate runtime risk ratings, thereby encapsulating the data-centric nature of this phase.

In contrast, the cycle’s left side is anchored in a value-based approach, focusing on how predicted risk ratings are practically applied in decision-making processes. This segment involves using runtime risk ratings as a facilitator for actions such as identifying root causes, strengthening safety mechanisms, and fostering communication across various domains. Safety experts harness the model’s outputs to craft specific safety goals that blend empirical data with expert judgment. Such critical reviews help pinpoint areas where the ADS can

be enhanced, striking a balance between its technological prowess and safety requirements. The cycle’s fluidity is further emphasized by its ongoing refinement process, where both the machine learning model and safety tactics are continuously polished based on feedback from real-world operations.

The continuous risk assessment cycle smoothly combines a fact-based learning methodology with a value-based action. This balanced integration ensures that the framework is not only firmly anchored in empirical data but is also adjusted to the value-driven safety imperatives that are essential for ADS. Having outlined the overarching cycle of the continuous risk assessment, we now delve into the detailed exploration of its core modules, to understand their individual functionalities and synergistic interactions within the system.

A. Risk-Learning Process

In our study, we adopt a methodology for runtime risk assessment inspired by [13, 14], focusing on three fundamental aspects: data acquisition, feature selection and labeling, and model training and deployment. The initial phase of our research concentrates on data acquisition from simulated environments. This phase is particularly tailored to Highway Lane Following scenarios with Adaptive Cruise Control using Matlab/Simulink [15]. Here, the emphasis is placed on the nature of the data collected and its preprocessing. We utilize OBS data, transforming it into risk features through mathematical equations, as illustrated in Fig. 2.

Subsequently, our next steps include detailed feature extraction, precise labeling based on predetermined risk thresholds, and a strict selection process to ensure the relevance and accuracy of the features. In the context of severity classification, our study delineates four categories: S0 (no injuries), S1

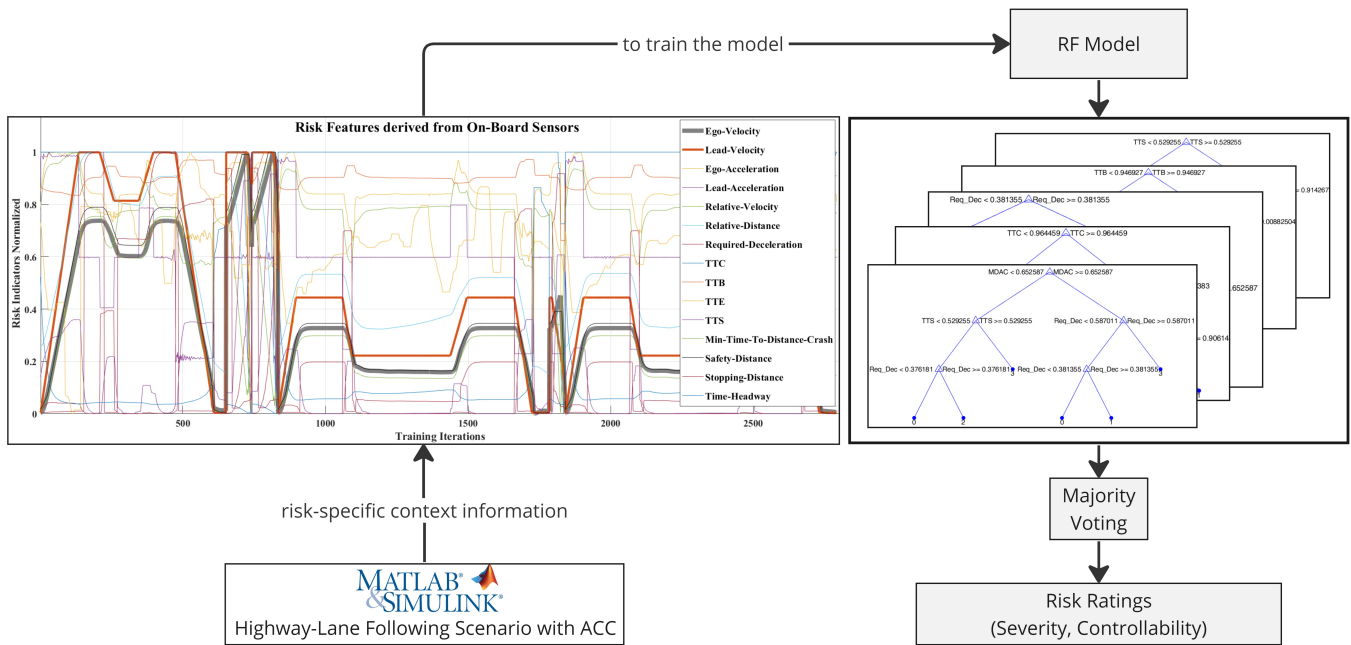


Fig. 2. Implementation of the Risk-Learning Process: Focusing on RF Model Development and Deployment

(light and moderate injuries), S2 (severe and life-threatening injuries), and S3 (fatal injuries). In terms of controllability, the levels range from C0 (controllable) to C3 (uncontrollable), underscoring their system-dependent nature.

Our scenario creation is grounded on three pillars: Failure Conditions, Design Parameters, and Non-Design Parameters. We analyze Failure Conditions to evaluate the system's response to unexpected issues, such as unintended acceleration and unintended braking. Design Parameters, which fall under the manufacturer's control like vehicle acceleration/braking efficiency, and Non-Design Parameters, such as weather and traffic conditions (other traffic participants), are rigorously simulated to generate comprehensive data representing dynamic driving conditions. In the final phase of model training and deployment, our focus shifts to developing the RF model. This phase encompasses critical processes such as model training and hyperparameter tuning, and model deployment on ADS.

In our implementation using the MATLAB/Simulink model (as shown in Fig. 2), we observed the significant impact of lead vehicle behavior on risk assessment in various driving scenarios. When the lead vehicle decelerated rapidly, our RF model highlighted a high severity (S3) with low controllability (C3). This situation was particularly critical because the sudden decrease in speed of the lead vehicle increased the likelihood of a collision, and the ADS had limited time to respond effectively. Conversely, when the lead vehicle maintained a steady speed and distance, the model indicated a moderate severity (S2) with higher controllability (C1). In this scenario, the consistent behavior of the lead vehicle, both in terms of speed and distance, allowed the ADS more room for corrective actions. This consistency in the lead vehicle's

behavior provides a safer and more controlled environment for the ADS to operate, reducing the risk of a collision.

Additionally, our simulations explored scenarios involving unintended braking. On dry roads, if the lead vehicle initiated braking, the model predicted a moderate severity (S2) with moderate controllability (C2), factoring in the ADS braking efficiency and the existing distance to the lead vehicle. However, in wet road conditions, the same action by the lead vehicle led to a high severity (S3) rating with low controllability (C3). The wet surface increased the potential for skidding, thereby reducing the ADS ability to brake safely and effectively.

The adaptability of the RF model was particularly evident in its feature prioritization, which varied depending on the lead vehicle's behavior. This dynamic adjustment of risk assessments based on runtime actions of the lead vehicle underscores the model's capability to offer nuanced risk predictions. By focusing on the immediate behavior of the lead vehicle, the model provided targeted and relevant risk assessments, crucial for ensuring the safety of the ADS in various operational scenarios. This approach marks a significant stride in enhancing the precision and context-awareness of risk assessments in ADS.

B. Risk-Informed Actions

In this subsection, we propose applying severity and controllability risk ratings to guide critical actions for transitioning from static safety goals to context-specific safety goals. These four key actions contribute to this transition as follows:

1. Identifying and Analyzing Underlying Causes: The risk ratings provided by the predictive RF model, offer insights into the severity and controllability of potential hazards. By analyzing these ratings, underlying causes of risks can be iden-

tified. For instance, a high severity rating in highway scenarios might indicate issues with the vehicle's ACC under certain conditions. This analysis enables targeted investigations into specific system behaviors or environmental interactions that contribute to elevated risk.

2. Reinforcing Safety Mechanisms: Based on the understanding gained from analyzing risk ratings, safety mechanisms within the ADS can be reinforced. If a particular scenario consistently results in high-risk ratings, it indicates a need for strengthening the corresponding safety protocols or system responses. For example, a scenario with low controllability might necessitate enhancements in the vehicle's emergency braking system or evasive maneuver algorithms.

3. Communicating Risk with Cross-domain Experts: The quantified risk ratings serve as a valuable communication tool with cross-domain experts. By presenting clear, data-driven assessments of risk, discussions with safety engineers, software developers, and other stakeholders become more focused and productive. This collaboration is essential for a comprehensive safety strategy, ensuring that all aspects of the ADS, from software algorithms to hardware reliability, are scrutinized and optimized.

4. Refinement of the Data-Driven Model: Finally, the iterative nature of the process allows for continuous refinement of the predictive model. The feedback from real-world performance and expert analyses informs subsequent iterations of model training. This cyclic improvement ensures that the model stays relevant and accurate in evolving driving contexts and technological advancements.

In the context of developing safe ADS, by following these four actions to create a context-specific safety goals are not only enhancing safety but also facilitate functional modifications, such as improving sensor and actuator performance/accuracy. This process enables the augmentation of recognition and decision algorithms, which is pivotal for ADS development. It improves testability, ensuring rigorous testing of the system. It increases the safety margin for future ADS development. This iterative process continues until the previously unknown and unsafe regions are transformed into known and safe ones. This methodology ensures that ADS remains responsive to current challenges and adaptable to the evolving driving environments and triggering conditions.

IV. CONCLUSION

This study has carefully explored the development and implementation of a continuous risk assessment cycle for ADS. It excels in providing a dynamic framework that intelligently prioritizes various risk features in line with the evolving scenarios of driving.

Nevertheless, it's crucial to recognize certain constraints. Notably, while the utilization of the RF model is advantageous for handling diverse input variables and intricate interactions, it may not entirely grasp the uncertainty and non-linearity found in extremely rare or unseen situations. Moreover, the dependence on sensor-based model predictions, though efficient for runtime data processing, is subject to the inherent limitations

of sensor precision. This can lead to data gaps or noise, which in turn, might affect the reliability of the predictions.

Despite these challenges, the model stands out for its ability to adjust to the shifting significance of risk features across different operational conditions. However, there is still room for improvement in its verification processes and broader applicability. Rigorous testing is vital to bolster the safety case for unforeseen scenarios and to confirm that the model effectively fills the gap left by traditional risk assessment methods.

Further research is essential to enhance the model's effectiveness. Expanding the integration of diverse data sources, including driver behavior, and refining the algorithmic approach, will address existing challenges and improve the model's relevance and efficacy in risk assessment.

REFERENCES

- [1] ISO 26262-1:2018, "Road vehicles - Functional safety - Part 1: Vocabulary," 2018.
- [2] ISO 21448:2022, "Road vehicles — Safety of the intended functionality," International Organization for Standardization, 2022.
- [3] T. Aven, "On how to define, understand and describe risk," in *Reliability Engineering & System Safety*, vol. 95, no. 6, pp. 623-631, 2010.
- [4] S. Khastgir, S. Birrell, G. Dhadyalla, H. Sivencrona, and P. Jennings, "Towards increased reliability by objectification of Hazard Analysis and Risk Assessment (HARA) of automated automotive systems," *Safety Science*, vol. 99, pp. 166-177, 2017.
- [5] E. De Gelder, H. Elrofai, A. Khabbaz Saberi, J. P. Paardekooper, O. Op Den Camp, and B. De Schutter, "Risk quantification for automated driving systems in real-world driving scenarios," *IEEE Access*, vol. 9, pp. 168953-168970, 2021.
- [6] F. Warg, M. Skoglund, A. Thorsén, R. Johansson, M. Brännström, M. Gyllenhammar, and M. Sanfridson, "The quantitative risk norm - a proposed tailoring of HARA for ADS," in *2020 50th Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshops (DSN-W)*, pp. 86-93, IEEE, 2020.
- [7] B. Wu, Y. Yan, D. Ni, and L. Li, "A longitudinal car-following risk assessment model based on risk field theory for autonomous vehicles," *International Journal of Transportation Science and Technology*, vol. 10, no. 1, pp. 60-68, 2021.
- [8] S. Hallerbach, Y. Xia, U. Eberle, and F. Koester, "Simulation-based identification of critical scenarios for cooperative and automated vehicles," *SAE International Journal of Connected and Automated Vehicles*, vol. 1, no. 2018-01-1066, pp. 93-106, 2018.
- [9] A. R. Patel and P. Liggesmeyer, "Machine learning-based dynamic risk assessment for autonomous vehicles," in *2021 International Symposium on Computer Science and Intelligent Controls (ISCSIC)*, pp. 73-77, IEEE, 2021.
- [10] A. J-P. Tixier, M. R. Hallowell, B. Rajagopalan, and D. Bowman, "Application of machine learning to construction injury prediction," *Automation in Construction*, vol. 69, pp. 102-114, 2016.
- [11] Y. Zhou, S. Li, C. Zhou, and H. Luo, "Intelligent approach based on random forest for safety risk prediction of deep foundation pit in subway stations," *Journal of Computing in Civil Engineering*, vol. 33, no. 1, pp. 05018004, 2019.
- [12] C. Yang, E. Trudel, and Y. Liu, "Machine learning-based methods for analyzing grade crossing safety," *Cluster Computing*, vol. 20, pp. 1625-1635, 2017.
- [13] A. R. Patel and P. Liggesmeyer, "LADRI: LeArning-Based Dynamic Risk Indicator in Automated Driving System," *2023 IEEE International Test Conference (ITC)*, Anaheim, CA, USA, 2023.
- [14] L. Westhofen, C. Neurohr, T. Koopmann, M. Butz, B. Schütt, F. Utesch, B. Neurohr, C. Gutenkunst, and E. Böde, "Criticality metrics for automated driving: A review and suitability analysis of the state of the art," *Archives of Computational Methods in Engineering*, vol. 30, no. 1, pp. 1-35, 2023.
- [15] The MathWorks Inc., "MATLAB version: 9.13.0 (R2022b)," Natick, Massachusetts: The MathWorks Inc., 2022. [Online]. Available: <https://www.mathworks.com>