Proceedings of the 35th European Safety and Reliability & the 33rd Society for Risk Analysis Europe Conference
Edited by Eirik Bjorheim Abrahamsen, Terje Aven, Frederic Bouder, Roger Flage, Marja Ylönen

doi: 10.3850/978-981-94-3281-3_ESREL-SRA-E2025-P7154-cd

A POMDP-based approach for obstacle avoidance in autonomous trains

Mohammed Chelouati, Abderraouf Boussif, Julie Beugin, El-Miloudi El Koursi
COSYS-ESTAS, Univ Gustave Eiffel, 20 rue Élisée Reclus, Villeneuve d'Ascq, F-59650, France.
E-mail: {mohammed.chelouati}{abderraouf.boussif}{julie.beugin}{el-miloudi.el-koursi}@univ-eiffel.fr

Autonomous trains must operate in highly dynamic environments where ensuring safety remains a significant challenge. Unlike human operators who can intuitively assess and respond to potential risks, autonomous systems require continuous, real-time evaluation of their surroundings in order to make safe decisions. In this paper, we present an approach for obstacle avoidance and environment monitoring for autonomous trains using Partially Observable Markov Decision Processes (POMDPs). The proposed approach models and assesses the risks while take into account the uncertainties associated with the train status and the various operational and environmental conditions; then outputs the adequate control action to maintain the train in safe state. To evaluate its efficiency, the approach is applied to the anti-collision function of autonomous trains in hazardous scenarios.

Keywords: Autonomous trains, Risk assessment, Partially Observable Markov Decision Process, Safety assurance, anti-collision function

1. Introduction

Autonomous trains are poised to revolutionize the rail industry, offering significant advancements in efficiency, reliability, and operational flexibility (Singh et al., 2021). However, these systems must operate in highly dynamic and unpredictable environments, where ensuring safety remains the main challenge (Chelouati et al., 2022). Unlike human operators who possess intuitive capabilities to assess and respond to hazardous situations in real-time, autonomous systems rely on advanced algorithms to continuously evaluate their surroundings and make informed, safe decisions (Endsley, 2018). This necessitates a robust and adaptive tool for risk evaluation to enable these systems to navigate complex and unpredictable conditions.

Dynamic risk assessment forms the foundation of safety assurance in autonomous systems, allowing them to identify, evaluate, and mitigate risks as operational and environmental conditions evolve (Chelouati et al., 2023; Patel et al., 2024). In railways, autonomous trains face multiple challenges due to their operating environments, including varying track conditions, changing weather, and potential obstacles. In this context, the embedded decision-making processes must balance among between accuracy, computational efficiency, and real-time responsiveness to maintain system safety without compromising performance.

To partially address these challenges, we previously proposed in (Chelouati et al., 2023) a novel risk-based decision-making approach for autonomous trains, using Partially Observable Markov Decision Processes (POMDPs) for continuous monitoring and evaluation of environmental collision risks. The present paper extends our early work by establishing an enhanced POMDP model designed to handle more complex operational driving scenarios with improved efficiency and scalability. Notably, by significantly increasing the state-space in the POMDP model, this approach enhances the risk evaluation, allowing for more precise estimations of potential hazards. Additionally, this increased model complexity is managed to maintain computational efficiency, ensuring its suitability for real-time applications. The approach is illustrated on the obstacle collision avoidance function in autonomous trains.

The rest of this paper is organized as follows: Section 2 presents preliminary concepts of dynamic risk assessment and POMDP. Section 3 details the integration of a POMDP-based risk model into the autonomous driving system of trains. In Section 4, the proposed model is further simulated through its application to the anti-collision function. Finally, concluding remarks and directions for future work are presented in Section 5.

2. Background

In this section, we present the main concepts of the proposed approach, namely dynamic risk assessment (DRA) and POMDPs.

2.1. *Dynamic risk assessment*

DRA addresses the challenges arising from the complexity and unpredictability of operational conditions in autonomous systems (Reich and Trapp, 2020). In contrast to traditional systems, where human operators rely on intuition, experience, and situational awareness to manage operational hazardous situations, autonomous systems employ advanced algorithms to continuously detect, assess, and respond to hazards in real time (Li et al., 2022). This requirement is particularly critical in autonomous train operations, where dynamic elements such as moving obstacles, multiple track conditions, and sudden environmental changes create hazardous scenarios that require constant adaptation. Static risk assessment methods, which rely on predefined (worst case) conditions, fall short in addressing these challenges, underscoring the need for a more flexible and responsive approach (Chelouati et al., 2023). For this purpose, the process of DRA is considered as a robust solution enabling continuous risk evaluation and estimation based on real-time perception information.

DRA has been widely applied in transportation sectors. In the automotive sector, DRA has been incorporated into advanced driver assistance systems (ADAS) and autonomous vehicles to manage collision risks and adapt driving strategies in real-time (Reich and Trapp, 2020). In aviation, it is used for air traffic management and onboard collision avoidance systems, enabling aircraft to respond dynamically to changing traffic and weather conditions (Mendes et al., 2022). Similarly, in the maritime sector, DRA supports navigation safety in congested waterways and adverse weather conditions through real-time sensor data and predictive models (Fan et al., 2024). These examples illustrate the versatility of DRA in managing risks across diverse operational contexts, reinforcing its importance for autonomous train operations.

DRA process can be integrated within the functional architecture of autonomous driving systems. The process begins with the perception of the environment, which gathers and processes external data. This is followed by the understanding and prediction stage, where insights into the current state and potential future scenarios are derived. Decision-making then uses this information to evaluate risks and determine appropriate actions. Finally, execution implements these decisions while considering the system's state of health and capabilities, ensuring safe and effective operation in dynamic conditions (Chelouati et al., 2022). This DRA framework is depicted in Figure 1, illustrating its iterative nature and role in enabling adaptive and informed decision-making.

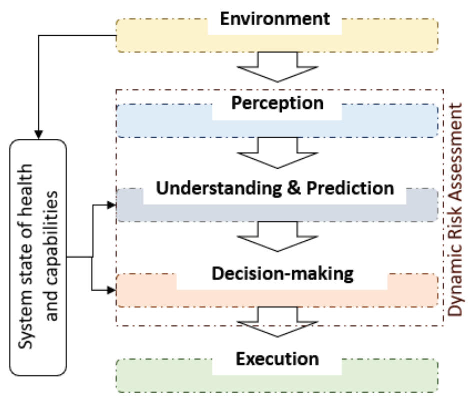

Fig. 1. An illustration of the DRA process for autonomous trains (Chelouati et al., 2023)

The implementation of DRA within a decision-making process relies on the use of robust risk models capable of representing and quantifying uncertainties in the system's operational environment. These models form the foundation for evaluating potential hazards and guiding decision-making processes in real time. Among the various models from the literature, POMDPs stand out as a particularly suitable method for establishing a risk-based decision-making process for autonomous systems. One of the key strength of POMDPs lies in their ability to operate effectively under conditions of partial observability, where the system does not have complete or accurate knowledge of its surroundings. This is the case with autonomous trains which frequently encounter situations where sensor data may be noisy, incomplete, or subject to occlusions, such as when detecting obstacles in complex environments. Additionally,

2.2. *Partially Observable Decision Making Processes (POMDPs)*

POMDPs extend the Markov Decision Processes to model the sequential process of a system under uncertainty. A POMDP is formally defined by a tuple $< S, A, O, T, R, Z, \gamma >$, where S represents the set of possible states, A denotes the set of actions, and O is the set of observations the agent can make. The transition function $T(s, a, s')$ describes the probability of transitioning from state s to state s' after taking action a, while the reward function $R(s, a, s')$ quantifies the benefit of such transitions. The observation function $Z(o \mid s', a)$ models the likelihood of making observation o given the resulting state s' after an action a. Finally, γ is the discount factor that balances immediate and future rewards.

In this framework, the true state of the environment is not directly observable. Instead, a belief $b(s)$, represented as a probability distribution over all possible states, is maintained to estimate the system's current condition. This belief is updated iteratively using Bayes' rule based on the agent's actions and observations, enabling reasoning under uncertainty.

The policies π, which map belief states to actions, define the strategy the agent follows to maximize expected cumulative rewards. The reward structure is carefully designed to align the system's objectives with operational goals, such as minimizing collision risks or optimizing operational efficiency. Additionally, the observation model captures the stochastic nature of perception, which is particularly important for handling noisy sensor data in autonomous systems.

The POMDP model is well-suited for autonomous train operations due to its ability to manage uncertainty and support risk-based decision-making. In this context, the states S represent the conditions of the environment and the system, such as the train's position, obstacle locations, and track status. The actions A correspond to possible maneuvers, including braking, accelerating, or maintaining speed, while the observations O are derived from perception systems, such as LiDAR, cameras, and GPS. The belief state $b(s)$ provides the train with a probabilistic understanding of its surroundings, accounting for sensor inaccuracies and environmental variability. By deriving policies through the POMDP framework, the train can select optimal actions to minimize risks while maintaining operational constraints.

Although POMDPs provide a robust theoretical foundation, their practical implementation in real-time systems presents challenges. Solving POMDPs exactly is computationally infeasible for large state and action spaces. To address this issue, approximate solution methods, such as point-based value iteration, are employed to compute near-optimal policies efficiently. Advances in computational power and algorithmic techniques, including hierarchical POMDPs, further reduce the complexity by breaking decision-making into smaller, manageable sub-problems. Adaptive representations also enhance efficiency by dynamically discretizing the state space, improving both accuracy and computational feasibility.

3. Proposed approach

This section introduces the POMDP model for collision avoidance in autonomous trains.

3.1. *State space*

The state space in the proposed POMDP model is specifically structured to encapsulate the variables influencing collision risk. It is defined as follows, where each variable reflects a key operational or environmental factor:

$$s = (\alpha, \beta, \phi, \psi, \omega) \tag{1}$$

The component α represents the relative distance to an obstacle. It is discretized into four levels: *Far*, *Medium*, *Close*, and *Critical* ($\alpha = \{0; 1; 2; 3\}$ with 0 being Far and 3 being Critical). β denotes the relative speed of an obstacle, categorized as *Moving Away*, *Stable/Approaching Slowly*, or *Approaching Quickly*, to account for how rapidly the train is closing to detected hazards ($\beta = \{0; 1; 2\}$ with 0 being Moving Away and 2 being Approaching Quickly). The obstacle type, denoted by ϕ, is classified into *Stationary/Static Object*, *Dynamic Object* and another *Train* ($\phi = \{0; 1; 2\}$ with 0 being Static Object and 2 being (another) Train). The track condition, represented by ψ, is included to model how external factors such as *Normal* and *Slippery* impact braking efficiency and maneuverability ($\psi = \{0; 1\}$ with 0 being Normal and 1 being Slippery). Finally, ω represents the braking system status, which can be *Normal* or *Degraded*, acknowledging that system health directly influences the train's capabilities ($\omega = \{0; 1\}$ with 0 being Normal and 1 being Degraded). Hence, the (theoretical) state space contains $4 * 3 * 3 * 2 * 2 = 144$ possible states.

3.2. *Actions space*

The action space in the model is intentionally limited to three essential train control actions to balance model simplicity with operational realism. The action space is defined as $A = \{a_1, a_2, a_3\}$, where a_1 is *Maintain Speed*, ensuring that the train continues at its current speed when no immediate action is necessary; a_2, is *Nominal Braking*, which applies service braking to gradually reduce speed in response to emerging risks. Finally, a_3 is *Emergency Braking*, engaging the train's maximum braking capacity to stop the train.

3.3. *Transition function*

The transition function $T(s, a, s') = P(s' \mid s, a)$ models the system's evolution from one state to another based on the chosen action. It is grounded in the kinematic dynamics of train movement, which are crucial for accurately predicting how the system responds to control actions. The evolution of the train's velocity and position is governed by the following dynamic model:

$$\begin{bmatrix} v_T(t+\delta t) \\ x_T(t+\delta t) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \delta t & 1 \end{bmatrix} \begin{bmatrix} v_T(t) \\ x_T(t) \end{bmatrix} + \begin{bmatrix} \delta t \\ \frac{\delta t^2}{2} \end{bmatrix} a^T_{cc}(t) \tag{2}$$

In this formulation, $v_T(t)$ and $x_T(t)$ represent the train's current speed and position, respectively, while δt is the time step. The term $a^T_{cc}(t)$ is the action-dependent acceleration command, defined by:

$$a^T_{cc}(t) = \begin{cases} 0 & \text{if } a = a_1 \text{ (Maintain Speed)} \\ -a_{\text{nominal}} & \text{if } a = a_2 \text{ (Nominal Braking)} \\ -a_{\text{emergency}} & \text{if } a = a_3 \text{ (Emergency Braking)} \end{cases} \tag{3}$$

This dynamic model ensures that the transition function realistically simulates how actions impact the train's behavior over time, accounting for both control inputs and environmental influences.

Moreover, the transition dynamics are designed to capture the dependencies among the state variables. For instance, the evolution of the distance to obstacle is modeled using a simple kinematic relationship, described as:

$$\alpha(t+\delta t) = \alpha(t) - \beta(t).\delta t \tag{4}$$

where the subtraction represents the fact that a higher closing speed leads to a faster decrease in safe distance.

Furthermore, the obstacle type directly influences how the relative speed evolves. Depending on the current relative speed, the model defines the following effects:

$$\beta(t+\delta t) = \begin{cases} \beta(t) & \text{if } \phi(t) = 1 \\ \beta(t) + \Delta v_{obs} + \epsilon & \text{if } \phi(t) = 2 \\ \beta(t) + \Delta v_{obs} & \text{if } \phi(t) = 3 \end{cases} \tag{5}$$

where $\epsilon \sim \mathcal{N}(0, \sigma^2)$ is the Gaussian additional noise and $_{obs}$ is the speed difference (between the instant t and the instant $t + \delta t$) of the obstacle (in the second case, the obstacle is an object, while in the third case, the obstacle is another train). Finally, we consider that the performance of the braking system is sensitive to environmental conditions as follows:

$$\omega(t+\delta t) = \begin{cases} \omega(t) & if\psi = 1 \\ \omega(t) + Degradation & if\psi = 2 \end{cases} \tag{6}$$

3.4. *Observation function*

Given that the train operates in an environment with incomplete and noisy information, the observation function is critical for modeling uncertainty. The observation function is defined as $Z(o \mid s, a) = P(o \mid s, a)$, representing the probability of observing o given the current state s and action a. To account for sensor inaccuracies, the observation is modeled as a multivariate Gaussian distribution centered around the true state with covariance Σ:

$$Z(o \mid s, a) = \mathcal{N}(o; s, \Sigma) \tag{7}$$

Moreover, the observation vector is defined as

$$o = (\alpha_0, \beta_0, \phi_0, \psi_0, \omega_0) \tag{8}$$

Each component is subject to observation error, captured by the misperception probability ϵ_i:

$$P(o_i \mid s_i) = \begin{cases} 1-\epsilon_i & \text{if } o_i = s_i \\ \epsilon_i & \text{if } o_i \neq s_i \end{cases} \tag{9}$$

This formulation ensures that the model realistically reflects perception uncertainty, allowing the system to make informed and robust decisions in dynamic and uncertain environments.

3.5. *Reward function*

The reward function $R(s, a)$ is designed to penalize unsafe actions and highlight risk-reducing behaviors. It is structured as follows:

$$R(s,a) = \begin{cases} 0 & \text{if } R(s) = 1 \text{ and } a \in \{a_1\} \\ -5 & \text{if } R(s) = 2 \text{ and } a \notin \{a_3\} \\ -20 & \text{if } R(s) = 3 \text{ and } a \notin \{a_2\} \\ -100 & \text{if } R(s) = 4 \text{ and } a \neq a_3 \end{cases} \tag{10}$$

This structure ensures that high-risk situations are met with appropriate responses, heavily penalizing inaction or inadequate reactions when risks escalate.

3.6. *Policy and risk categorization*

The optimal policy π^* maps belief states to the most effective action, guiding the train's behavior in minimizing risk. Risk levels are categorized using the weighted scoring function:

$$\rho(s) = 3\alpha + 2\beta + 2\phi + 1\psi + 2\omega \tag{11}$$

Risk categories are defined as:

$$r(s) = \begin{cases} 1 & \text{if } \rho(s) \leq 7 \quad \text{(Low Risk)} \\ 2 & \text{if } 8 \leq \rho(s) \leq 11 \quad \text{(Moderate Risk)} \\ 3 & \text{if } 12 \leq \rho(s) \leq 15 \quad \text{(High Risk)} \\ 4 & \text{if } \rho(s) \geq 16 \quad \text{(Critical Risk)} \end{cases} \tag{12}$$

This method ensures every state is systematically evaluated, guiding effective, risk-based decision-making.

It is important to note that the coefficients in Equation 11 can be determined either by expert judgment or by using historical data from the system. In this study, we established the coefficients based on our assessment of the importance of each state variable in influencing collision risk. For example, we assigned a higher weight to the distance variable because it plays a critical role in determining safety, while other variables received lower weights. This method was chosen to ensure that the risk score accurately reflects the system's behavior as understood by the authors.

4. Simulations and results

The simulation design was constructed to exhaustively explore many combinations of the relevant state variables, resulting in a series of episodes of the state space. Each episode involved an initial condition and was allowed to evolve for a fixed number of steps, ensuring that the POMDP model was tested under diverse scenarios. The plots (Figure 2and3) illustrate how the risk classification, belief updates, and chosen braking actions evolve over time, particularly during obstacle encounters.

Figure 2 (figure in top) illustrates the evolution of key state variables throughout the simulation. This composite figure plots the trajectories of the five main variables: distance (α), relative speed (β), obstacle type (ϕ), track condition (ψ), and

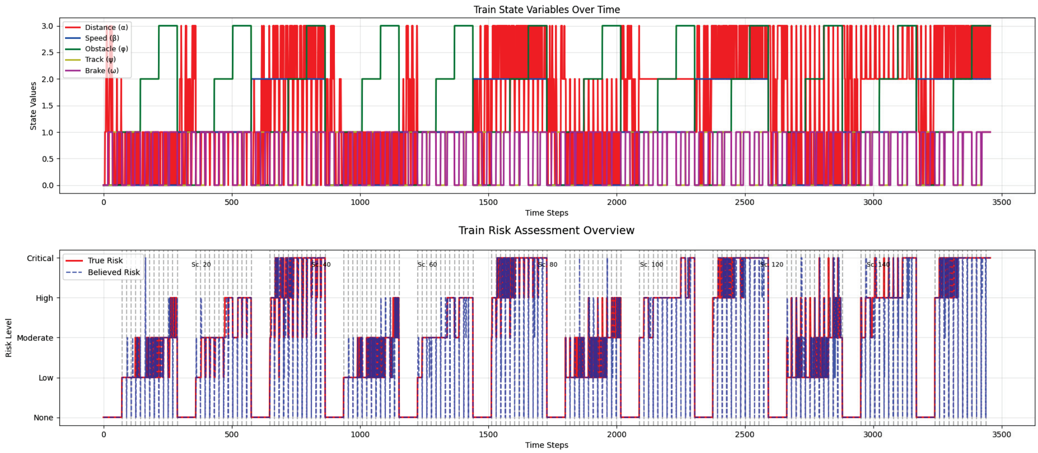

Fig. 2. The evolution of states and risk levels over time

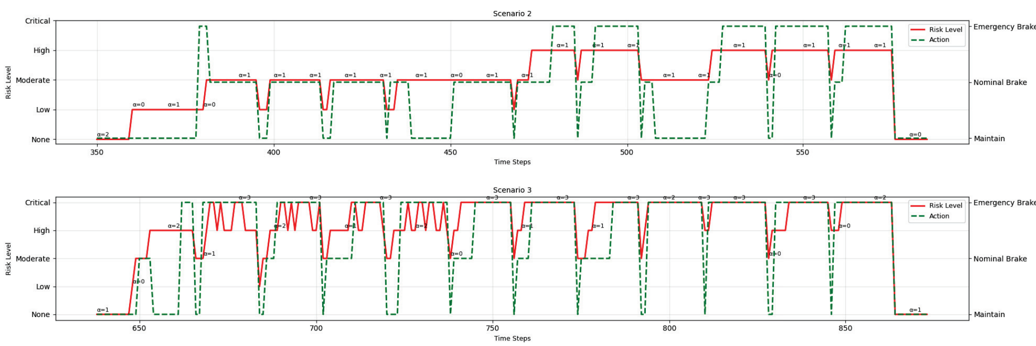

Fig. 3. The actions taken over time for two scenarios of the simulation

brake status (ω). Each variable is plotted with a different color and line style, making it possible to observe how these factors interact over time. The evolution of these state variables under the influence of the braking actions and environmental dynamics confirms that the model reacts coherently to sudden changes. This comprehensive view of the state evolution supports the conclusion that the model is robust and adaptive across a wide range of scenarios.

In Figure 2 (figure in bottom), the overall mission performance is presented over the entire simulation duration. The horizontal axis represents the simulation time, while the vertical axis shows the risk level on a scale from 0 (no risk) to 4 (critical risk). The solid red curve depicts the true risk, which is calculated directly from the current state of the system, and the dashed blue curve represents the risk estimated from the belief state maintained by the POMDP model. The mission history is segmented into portions corresponding to different initial configurations. To avoid visual clutter, only every 20th segment boundary is labeled. This figure illustrates how the model tracks the evolution of risk over time and highlights the consistency between the true and estimated risk levels despite sensor noise and uncertainties.

Figure 3 zooms on two specific obstacle encounter scenarios. Here, the red curve again represents the actual risk, which shows a moderate increase as the obstacle is encountered. In parallel, the green dashed line indicates the braking actions taken by the system, where the value 0 corresponds to maintaining speed, 1 to nominal braking, and 2 to emergency braking. This detailed view demonstrates that once the risk surpasses a certain threshold, the system initiates a braking response. The figure clearly shows that the model avoids unnecessary braking when the risk remains low, thereby validating the decision-making process under a moderate risk scenario.

Notice that only two obstacle encounter zones are zoomed in Figure 3, despite the existence of multiple obstacle segments in the full simulation. This decision was made to maintain clarity in the visualization. By selecting three representative zones that capture a range of responses, from moderate risk in Figure 2 (top) to highly dynamic scenarios in Figure 2 (bottom), we provide focused insights into the model's behavior without overwhelming the viewer with redundant details from every single obstacle encounter.

At the end of each simulation, three key metrics are derived by analyzing all the recorded steps. First, the fraction of time spent with obstacles is determined by counting the total number of steps during which an obstacle is present and dividing by the overall number of steps in the simulation, resulting in approximately 75.0%. Second, the risk assessment accuracy is obtained by comparing the true and estimated risks at each step, yielding about 96.3% of steps where both risks match. Finally, the average reward is found by summing the rewards over all steps and dividing by the total number of steps, leading to a value of about -3.19. These measures provide a concise overview of how frequently obstacles arise, how accurately the model assesses risk, and how the system's actions balance safety and performance throughout the simulation.

5. Conclusion and future work

The development of autonomous train systems that are able to function effectively and safely in dynamic and uncertain environments demands robust decision-making structures. In this work, we proposed a POMDP-based safety assurance model for real-time collision risk management of autonomous trains. By an enlarged state space and probabilistic structure, the model proposed here captures more operational scenarios, allowing precise risk computation while maintaining computational feasibility.

Simulation results demonstrate the effectiveness of this model in evaluating and responding to various risk scenarios created by environmental uncertainties and dynamic obstacle motion patterns. The framework ensures adaptive control actions, including acceleration, speed keeping, nominal braking, and emergency braking, that minimize the risk of collision under various operation scenarios. Importantly, the model achieves an optimal balance between computational complexity and decision accuracy, rendering it feasible for real-time high-risk scenario implementation.

Future work should focus on extending the model to incorporate additional operational factors, such as energy efficiency and obstacle's movement predictions, and on validating the approach in multi-train environments. Enhancements in sensor fusion and online adaptation techniques may further improve the system's reliability and responsiveness, paving the way for safer and more robust autonomous railway operations in real-world settings.

References

Chelouati, M., A. Boussif, J. Beugin, and E.-M. El Koursi (2022). A framework for risk-awareness and dynamic risk assessment for autonomous trains. In *Proc. 32nd Eur. Saf. Rel. Conf.(ESREL)*, pp. 2128–2135.

Chelouati, M., A. Boussif, J. Beugin, and E.-M. El-Koursi (2023). A risk-based decision-making process for autonomous trains using pomdp: Case of the anti-collision function. *IEEE Access 12*, 5630–5647.

Endsley, M. R. (2018). Automation and situation awareness. In *Automation and human performance*, pp. 163–181. CRC Press.

Fan, H., H. Jia, X. He, and J. Lyu (2024). Navigating uncertainty: A dynamic bayesian networkbased risk assessment framework for maritime trade routes. *Reliability Engineering & System Safety 250*, 110311.

Li, G., Y. Yang, S. Li, X. Qu, N. Lyu, and S. E. Li (2022). Decision making of autonomous vehicles in lane change scenarios: Deep reinforcement learning approaches with risk awareness. *Transportation research part C: emerging technologies 134*, 103452.

Mendes, N., J. G. V. Vieira, and A. P. Mano (2022). Risk management in aviation maintenance: A systematic literature review. *Safety science 153*, 105810.

Patel, A. R., K. Thummar, and P. Liggesmeyer (2024). Dynamic risk assessment: Leveraging ensemble learning for context-specific risk features. In *2024 IEEE Intelligent Vehicles Symposium (IV)*, pp. 1735–1742. IEEE.

Reich, J. and M. Trapp (2020). Sinadra: towards a framework for assurable situation-aware dynamic risk assessment of autonomous vehicles. In *2020 16th European dependable computing conference (EDCC)*, pp. 47–50. IEEE.

Singh, P., M. A. Dulebenets, J. Pasha, E. D. S. Gonzalez, Y.-Y. Lau, and R. Kampmann (2021). Deployment of autonomous trains in rail transportation: Current trends and existing challenges. *IEEE access 9*, 91427–91461.