Do Automated Vehicles Reduce the Risk of Crashes–Dream or Reality?

Árpád Török

Abstract— **In the future, the role of the human factor in the driving processes is expected to decrease continuously. At the same time, based on the global trends, the role of computer-supported decision systems and artificial intelligence (AI)-based control solutions increases in relation to driving processes, which carries a significant safety-enhancing potential. To assess the possible social benefits of automated vehicle systems objectively, it is necessary to analyze the possible negative effects in detail as well. Accordingly, the aim of this article is to present a statistical survey of crashes involving automated vehicles today in order to identify and evaluate the factors that are relevant in the crashes. The analyzed data showed that when the autopilot mode was turned off and the human driver made the control decisions, the severity of crashes on straight roads was greater at $\alpha = 0.1$ significance level than when the vehicle was in autopilot mode and the vehicle system made the control decisions. In addition, if the α significance level is 0.2, then crashes on plain terrain, during the day, or in the speed range of 80-100 km/h are generally less serious for vehicles driven in autopilot mode than for vehicles with autopilot mode turned off. In light of the considerations above, it is also important to emphasize that this paper only investigates crash severity given occurrence but not the probability of occurrence itself.**

Index Terms— **Automated vehicles, crash severity of highly automated vehicles, risk factors, statistical data analysis, collision causes of highly automated vehicles.**

I. Introduction

SEVERAL technical and social benefits are expected if highly automated vehicles with self-driving functions connected into a network spread widely [1], [2]. Safety and Security Research Group of the Department of Automotive Technologies at TU Budapest, in cooperation with the ZalaZONE proving ground, launched a new research project to assess the possible social benefits of highly automated vehicle systems. For this purpose, it is necessary to analyze the possible negative effects in detail, enabling us to mitigate unfavorable processes' negative effects.

Today, in the majority of road vehicle crashes, the human factor plays a significant role. Consequently, if the role of the human factor is diminished, the significant safety risk posed by the human factor will also decrease. At the same time, as future highly automated systems are expected to be more reliable than human drivers, computer-aided and artificial intelligence (AI) based decision-support [3], [4]; and decision-making systems [5] have a high safety enhancing potential [6], [7]. However, the estimated safety-enhancing effect of highly automated vehicle systems might be significantly reduced by the risk related to the artificial intelligence based decision-making processes and the issues of reliability related to the role of communication processes in highly automated systems [8].

Regarding the risk of artificial intelligence based approaches, it has to be mentioned that nowadays, the operation of artificial intelligence based solutions can primarily be described as a black box [9]. As a result, we cannot precisely describe these complex systems' behavior in the entire operational and decision space. One of the consequences of this system characteristic is that we can frequently meet with crashes where all the individual components of the system function perfectly. However, among the infinite possible combinations of a large number of input variables, a case can occur where the entire system's proper functioning was not tested, and the decision made by the system does not meet the safety requirements [10].

Until the final bridging of this set of problems, artificial intelligence-based decision solutions related to highly complex systems may carry potential risk factors in addition to their favorable safety effects.

In addition to the above, compared to classical solutions in the vehicle industry, we must also guarantee the reliability of information flow and data transmission for the safe implementation of highly automated functions. Compared to mechanical systems, highly automated systems can be quite data intensive. In light of this, the reliability of both wired and wireless communication processes will significantly influence the safety of future systems [11].

The objective of this article is to present a statistical survey of crashes involving highly automated vehicles today in order to identify and evaluate the factors that are relevant in the crashes. Accordingly, the paper aims to use various statistical tests to investigate whether crashes in autopilot mode or mode off are more serious under different conditions. The used database consists of 40 crashes, to evaluate the influencing factors of crashes occurred in autopilot mode and when autopilot mode was turned off. According to the reviewed literature, highly automated vehicles cannot be blamed for the crashes under investigation. After all, the person behind the wheel is still responsible for the vehicle's safety.

Several articles discuss the crashes caused by highly automated vehicles. The primary aim of the research conducted by Alrefaie and his colleagues was to evaluate the self-driving

Manuscript received 24 March 2022; revised 28 July 2022 and 28 September 2022; accepted 3 October 2022. This work was supported in part by the Hungarian Academy of Science (HAS) by the Janos BOLYAI Scholarship and in part by the Ministry of Innovation and Technology National Research, Development and Innovation Office (NRDI) within the framework of the Autonomous Systems National Laboratory Program. The Associate Editor for this article was S. A. Birrell.

The author is with the Department of Automotive Technologies, Budapest University of Technology and Economics, 1111 Budapest, Hungary (e-mail: torok.arpad@kjk.bme.hu).

Digital Object Identifier 10.1109/TITS.2022.3212280

function responsible for takeover by the human driver with respect to the psychological characteristics of the driver [12]. They proved that system limits [13] must be defined for SAE[1] 3 level automated vehicle systems. In line with this, scenarios must be determined in which the system cannot be expected to carry out a safe intervention. The experiment proved that based on pulse rate and pupil size, both the time required for the driver to intervene and the efficiency and reliability of the intervention can be reliably forecasted.

According to, as the level of automation is increasing, it is becoming more complex to determine the responsibility for a crash. In such investigations, the person(s) responsible for the crash must be identified, the facts and principles underpinning their responsibility must be specified, and the degree of damage caused by the person responsible for the crash must also be determined [14]. Human error is among the primary causes of road crashes; however, the increasing role of decisions made by machines during the driving process can affect this ratio in the future related to the spreading of highly automated vehicles. Owing to these tendencies, the limits of human responsibility must be explored [15]. In order to be able to determine the degree of the driver's responsibility in crashes involving highly automated vehicles, the degree of automation of the relevant functions that influence the operation of the vehicle must be known. Conditionally automated vehicle systems are categorized as SAE 3 or above.

Let us mention the highway pilot system of the Audi, which is able to drive the vehicle automatically on the condition that the human driver can take over driving at any time in the case of faulty system operation or if the system signals that human intervention is required. The system gives the control to the human driver, i.e., human intervention is required, when in a special situation, the decisions required for safe operation are ambiguous (i.e., not straightforward) for the control system. These traffic situations include cases when an emergency vehicle appears, or roadworks are in progress. Consequently, if an SAE 3 level automated vehicle causes a crash and the system had given the control over to the driver prior to it, the responsibility goes over to the driver even if disturbing conditions hindered appropriate human intervention.

When determining the driver's the responsibility, several factors must be considered that might have affected the intervening ability of the driver significantly. Such factors may include the characteristics of the disturbance, the proper or improper way of using the vehicle, and the length of the time period available for the intervention. If we suppose that the proportion of crashes in which the driver is responsible for the case is to diminish, the proportion of crashes in which the responsibility of manufacturers, maintainers, and operators is expected to grow.

Merriman and her colleagues analyzed five crashes in which highly automated vehicles had been involved. The authors drew general conclusions which can contribute to a better understanding of the reasons for such crashes and thus to the mitigation of the related risks [10]. The study explains what role the human factor plays in such crashes, with a special focus on load, perception speed, characteristic mental models, and trust. Concerning load, in four crashes out of the five analyzed cases, the drivers' mental load was low, owing to the fact that the highly automated system had taken over several tasks. As for the perception speed, it was found that drivers did not watch the road but were engaged with something else. They did not perceive the hazard, thus the information required for the intervention could not be processed or could be processed with significant delay by them. On the question of mental models, it was true for all five analyzed cases that the conditions required for the self-driving mode were not present, but the users still operated the vehicle in autonomous mode.

The enhanced vehicle control model by Monkhouse, Habli and McDermid makes it possible to identify risks connected to highly automated functions, especially risks of human-machine interactions [16]. The enhanced model combines Michon's Hierarchical Control Model and the Motor Industry Software Reliability Association (MISRA) Vehicle Control Model. The study found that improper usage is a high risk for highly automated vehicles, i.e., if the driver uses the self-driving mode in circumstances that do not meet the requirements of the feature. The study also proved that in addition to the above, the decline of the driver's awareness is also a critical risk, which significantly influences the length of time required to take over the control of the vehicle [17].

The aim of the study written by Petrović, Mijailović and Pešić is to evaluate the unique characteristics of crashes involving highly automated vehicles [18]. Their study analyzed the distribution of the crash types and maneuvers of highly automated vehicles and the drivers' mistakes. They examined 300 crashes altogether. The three major findings of the research are summarized as follows. The introduction of highly automated vehicles is expected to reduce the share of broadside and pedestrian accidents, as these cars are able to compensate for the effects caused by drivers who fail to give way. At the same time, the spread of highly automated cars is to increase the share of rear-end crashes due to the short safety distance between cars and unsafe speed choices. One of the reasons for this is that drivers are not used to the vehicle dynamics characteristics (speed choice, braking) of highly automated vehicles.

Dichabeng, Merat and Markkula conducted focus group research to reveal the factors that influence the acceptance of self-driving vehicles. All the 21 participants agreed that the reliability of autonomous systems critically influences trust towards the system and its usage. Additionally, the majority of participants felt that fully automatic systems are less safe due to the impossibility of intervention [19].

Some research focuses on exploring the possible risks of the autopilot mode. Morando and colleagues focus on the developed a model that can investigate the glance pattern observed around driver-initiated, non-critical disengagements of autopilot in naturalistic highway driving [20]. According to their findings, when autopilot was turned changes a lower visual attention was observed in the moajority of the observed cases. Biondi et al. investigated the unexpected safety consequences of semi-automated driving. The research proved the significance of the disadvantageous effects that semi-automatic

[1]Society of Automotive Engineers.

driving causes reduced awareness of traffic hazards [21]. In order to compare the results of this evaluation with other research, the outcomes of other studies dealing with the statistical data of highly automated and autonomous vehicles were reviewed. Many previously conducted analyzes were primarily aimed at comparing the crash rate of autonomous and conventional vehicles. However, the results are relatively widely distributed, mainly depending on the investigated environment and the types of vehicles included in the analysis. For example, the crash rate of Google autonomous vehicles is about 2.19 crashes per million vehicle miles traveled [22], while Favarò et al. estimated 23.9 crashes per million vehicle miles traveled [23]. Xu et al. state that crashes involving connected and autonomous vehicles are less severe than normal crashes [24]. Based on the reviewed research studies, it is advisable to choose the examined sample so that the vehicles participating in the crashes that occurred in self-driving mode and those that occurred under human control are equipped with similar technology. Therefore, this paper aims to examine crashes involving only automated vehicles, applying the principle of ceteris paribus. The cases are primarily distinguished by whether the autopilot mode was engaged or not.

The next section describes the database and the methods used for the analysis. Then the analysis of crashes involving automated vehicles is presented, and conclusions are drawn.

II. Methodology

A. Database

This study is based on an inventory of crashes involving automated vehicles [25]. A primary goal of the analysis was to investigate the crashes as close as possible to the time of the analysis, in contrast to the statistical data sets, which in some cases have a longer processing time, the data came from publicly available, independent sources. In this way, the necessary data set was available and could be inferred and a thorough textual description of the cases was also available. In light of the above, the injury data may differ from the final crash data recorded in the statistical database, as the severity of the injured may have changed over time. At the same time, the above deviations affect both crashes in autopilot mode and when the autopilot was turned off, so expectedly this phenomenon only slightly affects the comparison. In all cases, the data source is indicated in the referenced database created by Török and Mammadli. Regarding the crashes, we determined that at least one participating vehicle had SAE 2 or higher automation level. Tesla vehicles were involved in 70% of the cases, and vehicles from other highly automated vehicle manufacturers were involved in the remaining crashes.

Following the conclusions of the literature review, the database only includes crashes involving highly automated vehicles. This can guarantee that vehicles involved in a crash are equipped with similar technology. Concerning the crashes collected following the introduced selection principles, the available descriptions provided information on whether the vehicle was under human control or whether the autopilot mode was turned on. The analysis as a whole and all its elements were carried out for the 40 crashes collected. As authors of the database collected crashes for which information was available that the vehicle was in autopilot mode or under human control at the time of the collision, the source of the human-driven data could also be the database created by Török and Mammadli. Accordingly, the human-driven cases were also provided by the created database, which contains a total of 40 cases. In line with this, the crashes are primarily distinguished by whether the autopilot mode was turned on or off in the examined vehicle. On the other hand, it must be emphasize that beyond the use of autopilot mode, other important factors can also contribute to crash probability and severity, such as the environmental conditions, the period of the day, or the relationship between the permitted and the driving speed. These factors are represented in the current study by comparing the different subcases. Furthermore, it is also worth noting that the literature review confirms that operation under conditions outside the operational design domain (ODD) of an automated function is often identified as an important crash factor, which is also the responsibility of the driver [9].

Based on the characteristics of registered and analyzed crashes, the critical factors of the operation of highly automated vehicles can be identified. The database comprises 40 crashes from all over the world, which occurred between 2016 and 2021. The following variable were used from the database to evaluate the investigated crashes (Table I).

TABLE I
Analyzed Crash Data

Time of the crash
Global Navigation Satellite System (GNSS) position
Number of fatal injuries of the crash
Number of seriously injured persons related to the crash
Number of slight injuries related to the crash
Number of vehicles involved
Visibility condition
Weather condition
Speed limit
Actual velocity
Autopilot mode
Age of vehicle
Road environment
Curvature
Period of the day
Visibility conditions

Concerning the places of crashes, 54% of examined cases occurred in the United States. When visibility was high, more than 53% of crashes happened during the day. In 48% of cases, the speed was higher than 60 km/h. Autopilot mode was on only in 38% of crashes. 61% of crashes involved cars that were at least two years old. A bit more than one-third of crashes (36%) occurred while the vehicles were driving in a curve. Meanwhile, 43% of cases involved vehicles driving in opposite directions. Furthermore, the primary cause of these crashes was the human fault in 66% of cases, which correlates with the fact that the autopilot mode was on in 38% of cases.

The contents of the autopilot mode database field need to be interpreted. In general, the meaning of autopilot is a system module that controls certain functions of the vehicle independently, without the intervention of the driver [26]. If autopilot mode was turned on during the collision, it is assumed that the machine logic made the related decisions during the pre-crash driving process. If the autopilot mode was switched off, the driver made the decisions related to the driving process. In the first case, it is assumed that the crash happened under the control of a machine. In the second case, it was assumed that the crash was under human control.

B. Normality Test

In order to make the analysis of the crash data well-grounded, various statistical methods were applied. In the first step, it must be checked whether the examined collection of data conforms to the applicability requirements of the statistical tests and models [27].

Normality of the analyzed variables is often a requirement for various statistical tests and models. More precisely speaking, the normality condition is valid for the distribution of sample means taken from the whole population. However, no reliable information on this distribution is available; therefore, the conclusions regarding the population must be drawn from the available sample. For this, it is possible to use the assumption that if the population variable is normally distributed, this must also be true for the distribution of the sample taken from the population. Thus, if the normal distribution of a given variable is proven, the given variable can be regarded as normally distributed.

If a sample does not meet the requirement for normal distribution, the hypothesis should be underpinned by applying methods that allow the analysis of samples with a non-normal distribution: typically, non-parametric tests. It must also be noted that if the sample is cut up into subsamples, each subset must meet the requirement of normal distribution.

As for the evaluation of the results, it must be noted that sample size and the relative element numbers of the population and the sample can significantly influence the results. Both type I and type II errors may occur, thus when in doubt, several types of normality tests are to be run.

Owing to its outstanding performance, the Shapiro–Wilk test (SWT) is widely used to test normal distribution [28].

$$SWT = \frac{\left(\sum_{i=1}^{n} coeff_i \cdot X_{(i)}\right)^2}{\sum_{i=1}^{n} (X_i - \bar{X})^2} \tag{1}$$

where

- X_i : the i^{th} element of the analyzed sample;
- $X_{(i)}$: the i^{th} smallest element of the analyzed ordered sample;
- $coeff_{1..n} = \left(m' \cdot V^{-1}\right) \cdot \left((m' \cdot V^{-1})(V^{-1} \cdot m')\right)^{-0,5}$,
- $m' = (m_1, m_2, \ldots, m_i, \ldots m_n)$, is the vector of the expected values of the ordered sample $X_{(1)}, X_{(2)}, \ldots, X_{(i)}, \ldots, X_{(n)}$.

C. Comparison of Expected Values of Two Samples

During examinations, including serial sampling and during the comparison of experiments comparing various subcases, it is often the task to quantify the relationship of specific samples and the degree of differences between samples in a statistically well-founded manner. The t-test is one of the most widely used methods for comparing the expected values of two samples. However, the basic prerequisite for the application of this statistical test is that the samples to be examined should have a normal distribution. Thus, if the value of the Shapiro–Wilk test is lower than the chosen alpha significance level and thus the null hypothesis must be rejected, another, non-parametric test should be applied, for which the normal distribution of the sample is not a prerequisite. If the normality of the examined independent samples cannot be proven, the Mann–Whitney U test (MWT) should be used to carry out the pairwise comparison of expected values [29].

In essence, for the Mann–Whitney U test the elements of the two analyzed sample sets must be paired up. Thus, all the elements of one sample (X_i) are paired up with all the elements of the other sample set (Y_i). The total number of pairs is $N \times M$, where N is the number of elements in the first set while M is the number of elements in the second set. The number of those pairs must be counted in which the first element is greater than the second one $(X_i > Y_i)$. The number of such pairs is the value of the Mann–Whitney U test. If the expected value of the analyzed sample pair does not differ, the number of element combinations for which $X_i > Y_i$ and $X_i < Y_i$ is almost equal. If one type of element combination significantly outnumbers the other type of element combination, it is highly probable that the expected values of the two samples are different.

$$MWT_1 = R_1 + \frac{N \cdot (N+1)}{2} \tag{2}$$

$$MWT_2 = R_2 + \frac{M \cdot (M+1)}{2} \tag{3}$$

where

- R_1 : sum of ranks for Sample 1;
- R_2 : sum of ranks for Sample 2;
- N : size of Sample 1;
- M : size of Sample 2.

D. Comparison of Expected Values of Several Samples

Multivariable statistical tests are to be applied if there are several samples and several variables. If the samples have a normal distribution, the ANOVA (ANalysis Of VAriance) method should be used [30]. If the normality of the samples cannot be proven, and the samples to be compared are independent of one another, the Kruskal–Wallis Non-parametric test (KWT) should be utilized [31].

For the Kruskal–Wallis Non-parametric test, the independent samples are unified, and all the members of the unified sample are ranked. If there are members with equal values, their rank is determined by calculating the mean of the ranks of that these elements would get if they were not equal. For example, for the sample 2, 3, 3, and 6, the ranking is 1; 2.5; 2.5; 4 because if the two elements did not have equal values, their ranks would be 2 and 3, respectively. The average of 2 and 3 is 2.5.

The test function can be calculated as follows.

$$KWT = (N-1)\cdot \sum_{i=1}^{m} n_i \cdot (\bar{r}_i - \bar{r})^2 \cdot \left(\sum_{i=1}^{m}\sum_{j=1}^{n_i}(r_{ij}-\bar{r})^2\right)^{-1} \quad (4)$$

where

- N: is the total number of members in the unified group;
- m: is the number of analyzed samples;
- n_i: is the number of members in the i^{th} sample;
- r_{ij}: is the rank of the j^{th} element of the i^{th} sample in the ranked unified set;
- $\bar{r}_i$: is the average rank of all members of the i^{th} sample in the ranked unified set;
- $\bar{r}$: the average of ranks r_{ij}.

If no significant result is arrived at, there is no pair in the compared samples that would have a significant difference in the expected value. However, if the test proves to be significant, there is a sample whose expected value is greater than that of another sample.

E. Pairwise Comparison of the Expected Values of Multiple Samples

If the analyzed samples are independent of one another, in reality, m independent statistical tests are carried out. The type I error, i.e., the probability of the mistaken acceptance of the alternative hypothesis (H_1) in m tests is marked by p_i, where $i = 1 \ldots m$.

In this case, for the whole test series, the probability of mistakenly accepting the H_1 hypothesis at least in one test can be calculated as the difference between the whole event space and the probability of the error-free case.

$$\begin{aligned} P &= 1 - \prod_{i=1}^{m}(1-p_i) \\ &= \sum_{i=1}^{m} p_i - \sum_{i_1=1}^{m-1}\sum_{i_2=i1+1}^{m} p_{i_1}\cdot p_{i_2} \\ &\quad - \sum_{i_1=1}^{m-2}\sum_{i_2=i1+1}^{m-1}\sum_{i_3=i_2+1}^{m} p_{i_1}\cdot p_{i_2}\cdot p_{i_3} - \ldots - \sum_{i_1=1}^{m-(m-2)} \\ &\quad \ldots \sum_{i_{m-2}=i_{m-3}+1}^{m-1}\ \sum_{i_{m-1}=i_{m-2}+1}^{m} p_{i_1}\cdot\ldots\cdot p_{i_{m-1}} - \prod_{i=1}^{m} p_i \end{aligned} \quad (5)$$

According to the above, if the number of compared experiments grows, the risk of mistakenly accepting the phenomena and effects described in hypothesis H_1 grows exponentially.

In the case of multiple hypothesis tests, the risk of type I error, i.e., mistakenly identifying a non-existent effect, is called family-wise error rate (FWER).

In order to correct the hypothesis test, the confidence interval (p) is modified. In the present analysis, the Bonferroni correction was applied [32], [33].

$$p = p * \frac{m\,(m+1)}{2} \quad (6)$$

TABLE II

CORRELATION BETWEEN CRASH SEVERITY INDEX (B) AND OTHER CRASH CHARACTERISTICS. OUTLIERS ARE GIVEN IN BOLD

Auto-pilot	*Re-lief*	*Time of day*	*Visi-bility*	*Wea-ther*	*Speed limit*	*Speed*	*Age*	*Curve*
On	*-0.04*	***0.49***	*-0.37*	*0.17*	*0.36*	***0.48***	*0.30*	*0.15*
Off	*-0.04*	*0.14*	*-0.09*	*0.01*	*-0.03*	*-0.11*	*-0.08*	*0.06*

where

- m: is the number of samples examined with the Kruskal–Wallis test.

F. Preliminary Analysis of Expected Effects of Crash Factors on Severity

The next step is the preliminary analysis of correlations between crash severity and various factors. Thus, based on the analysis of rank correlations between the cells of the database and the newly introduced crash severity index (B), the expected relevance of each factor's effect on the severity of crashes is to be estimated.

The crash severity index (B) is calculated for each crash based on the number of injuries (S_i) represented with the applied severity weight (b_i). The severity index is derived from the number of injured people involved in each crash, not the number of injuries sustained. Given that many studies examine the crash risk based on the crash rate, it is reasonable to investigate the relationship between the severity index and the crash rate. Since the severity index does not contain vehicle-miles-traveled data, it is primarily suitable for evaluating the severity component of the risk, so it does not carry information about the probability of crash occurrence. The severity weights are identified based on the costs of crashes of different severity [34], [35], adjusted to round values (fatal: serious: slight – 100 : 10 : 1).

$$B = \sum b_i \cdot S_i \quad (7)$$

The introduced approach to express severity index is often applied in selecting dangerous crash locations; hence its ability to discriminate is not further investigated [36], [37].

Correlation (R) is calculated by the following formula (8), shown at the bottom of the next page, where $rank(X_i)$ and $rank(Y_i)$ are the ranks of the members of samples X and Y, respectively, and $\overline{rank(X_i)}$ and $\overline{rank(Y_i)}$ are the means of the ranks of members in the corresponding samples.

The results of the correlation analysis between the crash severity index (B) and the relevant cells of the database are given in Table II.

Although far-reaching conclusions cannot be drawn from the comparison of rank correlation parameters, based on the observed correlations, hypotheses can be set up. Some of the correlation coefficients are characterized by outliers in autopilot mode (see bold data in Table II).

Consequently, there is a much stronger correlation between two of the factors, i.e., the time of the day and the driving speed [38], [39], and the crash severity index in autopilot mode than if the autopilot mode was turned off. Consequently, the crash data of automated vehicles suggests that, extra attention should be paid to the time of the day and the driving speed when determining critical test cases in the development phase.

The above results suggest that a comprehensive statistical analysis is due to explore the effect size of individual crash factors. In the next step, I will examine whether the effect of each factor is greater in crashes when the autopilot mode is on or off. For this, the two types of crashes (autopilot on vs. off) are compared for each factor. In the first step of the analysis, the averages of crash severity indices for each factor are compared. As the samples are not distributed normally, this comparison is solely applied for the setting up of hypotheses. To test the hypotheses, problem-specific, non-parametric tests are applied: for the pairwise comparison of independent samples, the Mann–Whitney test, while for group-wise comparison of independent samples, the Kruskal–Wallis test are used. In the next step, a pairwise comparison of the crash groups that occurred in autopilot mode and when the autopilot was turned off is performed within each factor. Based on these results, the relevant factor categories are compared, while the cases of the autopilot mode being turned on and off are distinguished.

To reveal whether it is relevant to rank crash factors proven to be critical in pairwise comparisons, the averages of each sample are analyzed and compared. It must be emphasized that the samples are not normally distributed; thus, the sample mean cannot characterize the expected value of the sample. Therefore, the comparison below only serves to facilitate setting up sensible hypotheses, which are further analyzed with other statistical methods. Based on the sample means, it can be assumed which crash factors have an exceptionally high crash severity index value in autopilot mode on or off.

TABLE III

COMPARISON OF SAMPLE MEANS. THE RANK OF THE CRASH SEVERITY INDEX IS SHOWN BY THE BACKGROUND COLOUR (GREEN – LOW; RED – HIGH)

crash circumstances	*autopilot mode off*	*autopilot mode on*
mountanous relief	*66.5*	*123*
flat terrain	*121.13*	*96.9*
day	*103.4*	*82.8*
night	*119.5*	*140.7*
good weather	*119.17*	*83.89*
bad weather (rain/snow/fog)	*122.3*	*120.67*
at a speed of 80 km/h or lower	*123.82*	*96.33*
at a speed between 80 and 100 km/h	*152*	*103.67*
at a speed over 100 km/h	*100.5*	–
straight road	*118*	*93.4*
in curve	*125.3*	*109*

III. RESULTS AND DISCUSSION

A. Comparison of Crash Group Averages

In the case of crashes that happened in autopilot mode, the exceptionally high value of the crash severity index indicates that in industrial development processes, it is advisable to allocate considerable resources to diminishing the unfavorable effects of the given factor. This can lead to a significant increase in the safety of highly automated systems. In the case of crashes that happened in human-driving mode, however, the exceptionally high value of the crash severity index might help to identify how and to what extent the highly automated systems can make transportation safer.

It must be noted that the database does not contain any crashes in which the driving speed was over 100 km/h, and the vehicle was in autopilot mode. Therefore, the present analysis cannot determine whether, in this speed range, there is a difference in the severity of crashes in the two driving modes.

Concerning crashes in autopilot mode, the highest values in the crash severity index compared to the crashes involving human-driven vehicles in the same factor are found if the crash happened at night or if it took place in mountainous terrain. Therefore, the data show that in these circumstances, the safety risk of automated vehicles is higher than that of human-driven vehicles.

I must note here that the vehicles are not responsible for the crashes in these cases either, but primarily the inappropriate control decisions made by the driver in autopilot mode. The wrong control decision refers either to a failure to resume control by the driver when requested or the activation of autopilot mode at an inappropriate time.

On the other hand, the crash severity index is much higher for vehicles with autopilot mode turned off than for vehicles with autopilot mode turned on, on plain terrain, during the day, when the travel speed is between 80-100 km/h, or on straight roads. It is expected that for these factors, more detailed statistical analyses can reveal how highly automated systems can enhance traffic safety today.

For other factors, the crash severity index for cases where the autopilot mode is off is higher, though not significantly. Developments aiming to reduce the unfavorable effects of these factors have further safety-enhancing potential.

As the above data prove, the average crash severity index for crashes that occurred in the speed range 80–100 km/h or flat terrain is much higher if the involved vehicle was driven in autopilot mode turned off than if the involved vehicle was driven in autopilot mode. Therefore, for these factors, tests matching the characteristics of the samples may reveal a significant difference in the severity of the crashes.

Thus, the members of the non-normally distributed, independent samples were compared pairwise by the Mann–Whitney test. The results are presented in Table IV.

$$R = \frac{\sum \left(rank(X_i) - \overline{rank(X_i)}\right) \cdot \left(rank(Y_i) - \overline{rank(Y_i)}\right)}{\sqrt{\sum \left(rank(X_i) - \overline{rank(X_i)}\right)^2} \cdot \sqrt{\sum \left(rank(Y_i) - \overline{rank(Y_i)}\right)^2}} \tag{8}$$

TABLE IV
RESULTS OF THE MANN–WHITNEY TEST

crash circumstances	***crashes if the autopilot mod is turned off are more severe than the ones in autopilot mode***
mountanous relief	*1.00*
flat terrain	*0.18*
day	*0.17*
night	*0.62*
good weather	*0.28*
bad weather (rain/snow/fog)	*0.28*
at a speed of 80 km/h or lower	*0.37*
at a speed between 80 and 100 km/h	*0.14*
at a speed over 100 km/h	*–*
straight road	*0.08*
in curve	*0.60*

As the above data prove, the p-value of the Mann–Whitney test for crashes in straight road sections approaches the $\alpha = 0.05$ significance limit, but it does not reach it. Thus, the probability of a type I error (i.e. rejecting an accurate null hypothesis) is the lowest in this case.

As mentioned earlier, the significance level refers to the tolerable probability of type I error, in other words, the probability of rejecting a true null hypothesis. Accordingly, in some cases, if the probability of rejecting the correct null hypothesis exceeds the originally determined acceptance level, but is close to it, a higher significance level is also investigated, which is already higher than the calculated probability of type I error. This concept tries to show that if we can improve some of the limiting factors affecting the representativeness of the current study (e.g., the small sample number would be larger), then the hypothesis would probably be acceptable.

Accordingly, the data of the examined database, therefore, supports that the probability that crashes with vehicles driven in autopilot mode turned off is more severe than crashes with vehicles driven in autopilot mode if the road is straight.

At significance level $\alpha = 0.2$, crashes involving vehicles driven in autopilot mode are less severe than those in autopilot mode turned off in the following circumstances: in plain terrain, during the day, or at a speed between 80 and 100 km/h.

If a human drives a vehicle, the severity of crashes in plain terrain, during the day, or on a straight road might be due to the risk compensation phenomenon [40] Chen et al., 2017), which does not modify the risk level of machine-driving in an unfavorable way. Contrarily, the increase in severity in the 80–100 km/h speed range in the human-driven category may be due to the fact that automated driving is more effective; the time to react and act for a machine is shorter for a machine than for a human [41], [42].

As the above pairwise comparison suggests, the examined database shows that it is only crashes occurring in straight roads where the crashes involving vehicles driven in autopilot mode turned off are more severe than those involving vehicles driven in autopilot mode in the same circumstances (significance level $\alpha = 10\%$). Consequently, in all other circumstances, the development of automated driving systems has good safety-enhancing potential, safety risks can further be diminished.

After the pairwise comparison, the expected values of the crash severity index for each factor were compared in groups. The Kruskal–Wallis test was applied to the non-normally distributed independent samples. The error rate for the family-wise error arising due to multiply hypothesis testing was corrected by the Bonferroni method.

B. Group-Wise Comparison of the Effects of Critical Crash Factors Regarding Vehicles Driven in Autopilot Mode

The next step of the analysis explored the crash factors with a high safety-enhancing potential for autopilot system, in order to identify factors that are connected to significantly more severe crashes. For a family-wise comparison, the Kruskal–Wallis test was applied with Bonferroni correction. The test proved that the null hypothesis could not be rejected as $p > \alpha$ (0.006). It is supposed that the average rank of all groups is equal. In other words, the differences between the means of all groups are not big enough to be statistically significant. If the result is not significant, it does not prove that H_0 can be accepted. It only proves that H_0 cannot be rejected. Therefore, concerning the present analysis, it is true that the probability that the value of the crash severity index is higher for a randomly chosen crash from the given group than this value for a randomly selected crash from any other group is the same for each group. Consequently, there is no significant difference between the means of ranks for any pairs.

C. Group-Wise Comparison of the Effects of Critical Crash Factors Regarding Vehicles Driven in Autopilot Mode Turned off

As preliminary tests proved, if the significance level $\alpha = 0.05$, there is no factor category in which crashes involving automated vehicles in autopilot mode are less severe than in autopilot mode turned off. In order to identify critical factors for human driving, those factors were selected that exhibit a statistically relevant difference, for which the significance level is higher than $\alpha = 0.05$, but it is still in the "acceptable" range due to the low number of elements.

Based on the present database, those categories were identified for which the null hypothesis H_0 can be rejected if the significance level $\alpha = 0.2$. In other words, if the significance level α equals 0.2, crashes involving automated vehicles in autopilot mode are less severe that crashes involving vehicles driven in autopilot mode turned off in the following circumstances:

- plain terrain;
- during the day;
- in the speed range 80–100 km/h;
- in a straight road.

In these circumstances, if significance level $\alpha = 0.2$, vehicles driven in autopilot mode turned off took part in more severe crashes than vehicles driven in autopilot mode. Consequently, in these cases, the higher level of safety of automated systems

TABLE V
RESULTS OF PAIRWISE COMPARISONS WITH KRUSKAL–WALLIS TEST (NEARLY SIGNIFICANT P-VALUES ARE BOLDED)

Sample pair	*KWT*	*Critical value*	*p*
plain terrain – day	*5.0101*	*6.2385*	*0.03*
plain terrain – 80–100 km/h	*2.0356*	*6.2385*	*0.15*
plain terrain – straight road	*0.1429*	*6.2385*	*0.71*
day – 80–100 km/h	*5.7691*	*6.2385*	***0.02***
day – straight road	*5.1759*	*6.2385*	***0.02***
80–100 km/h – straight road	*1.3241*	*6.2385*	*0.25*

can be detected even at the present level of technological development.

The next step of the analysis explored the factors that can be associated with crashes that are significantly more severe than crashes occurring in other circumstances. For a group-wise comparison, the Kruskal–Wallis test (KWT) was applied with Bonferroni correction. The test proved that as p (0.02995) $< \alpha$, the null hypothesis should be rejected. It is supposed that the average rank of a given group is different from that of other groups. In other words, the differences between the means of all groups are big enough to be statistically significant. Therefore, concerning the present analysis, for certain groups, it is true that the probability is higher for the given group that the value of the crash severity index is higher for a randomly chosen crash from the given group than this value for a randomly selected crash from any other group. Consequently, there is a significant difference between the means of ranks for certain pairs.

Table V shows the results of the Kruskal–Wallis test, together with the critical values and the calculated p-values for sample pairs. The significance level determined with the Bonferroni correction is $\alpha = 0.0125$. Based on the above, the family-wise comparison cannot provide a ranking of factors. If the significance level is moderately raised ($\alpha = 0.025$), it is proven that the expected mean rank of the severity of crashes is lower for crashes during the day with vehicles driven in autopilot mode turned off than for crashes happening in the 80–100 km/h range or for crashes occurring in straight roads.

IV. CONCLUSION

This article presents the most important factors of safety analysis for automated vehicles. Based on the evaluation results and the reviewed literature, considering the legal, social, and technical conditions, in most crashes with automated vehicles, the driver is generally responsible for the safety of driving. Therefore, among other things, my investigation sought answers to whether automated vehicle crashes are more severe in autopilot mode or with autopilot mode turned off.

As the first step of the analysis, the outcomes of other research studies were compared briefly with the presented concept. The research carried out by Morando and colleagues, and Biondi et al. showed the potential risks of the autopilot systems, which, in line with my suggestions, supports the need for further development [19]. According to their findings, during the application of autopilot mode, lower visual attention was observed in most cases. Biondi et al. investigated the unexpected safety consequences of semi-automated driving. The research proved the significance of the disadvantageous effects that semi-automatic driving causes reduced awareness of traffic hazards [20]. It was also revealed that the results of the different analysis are relatively widely distributed, mainly depending on the investigated environment and the types of the investigated highly automated and autonomous vehicles (from 2 to 24 crashes per million vehicle miles) [22], [23], [24]. Based on the reviewed research studies, it is advisable to choose the examined sample so that the vehicles participating in the crashes that occurred in autopilot mode and those that occurred when autopilot mode was turned off control are equipped with similar technology. Therefore, this paper examined crashes involving only highly automated vehicles, applying the principle of ceteris paribus. The crashes are primarily distinguished by whether the autopilot mode was engaged or not. Furthermore, in addition to the autopilot mode, other relevant factors can also influence crash probability and severity, such as the environmental conditions, the period of the day, or the relationship between the permitted and the driving speed [43]. These factors were represented in the current study by comparing the different subcases. However, in future investigations, it is advisable to consider these factors as separate regression parameters when estimating the crash risk.

As the first step of the analysis, it was determined whether the data to be analyzed are normally distributed. This hypothesis had to be rejected, so for further analysis, statistical tests that do not require a normally distributed sample were applied. The comparison of mean values of the data correlations justified further analysis of the database.

The Mann– Whitney test was applied to analyze the data. The p-value for the Mann–Whitney test approached the $\alpha = 0.05$ significance level for crashes on straight roads; thus, the possibility of a type I error for the rejection of an accurate H_0 is the lowest for this factor. Furthermore, the data evaluation showed that the severity of crashes on straight roads, if autopilot mode was turned off is higher than if the vehicle was in autopilot mode at a significance level $\alpha = 0.1$. In sum, crashes in the analyzed database were significantly less severe on straight roads if the vehicle was in autopilot mode than in the case of human driving.

Moreover, if significance level α equals 0.2, crashes that occurred in plain terrain, during the day, or in the speed range of 80–100 km/h also tend to be less severe for vehicles driven in autopilot mode than for vehicles driven in autopilot mode turned off.

The increase in the severity of crashes by human-driven vehicles in plain terrain, during the day, or on straight roads might result from the risk compensation phenomenon that characterizes human drivers, while it has no effect on the risk levels of automated driving systems. In contrast, the fact that the severity of crashes occurring in the 80–100 km/h speed range is higher for human-driven vehicles might be due to the effectiveness of automated driving systems, which exhibit a shorter reaction time than humans.

The above results were arrived at by analyzing a sample of a relatively low element number (40), which cannot be regarded as representative. However, the methodology presented here can be readily applied to larger, more comprehensive databases. As a limitation of the introduced methodology, it has to be noted that the introduced severity index does not contain any value related to probability; thus, it is not possible to comment on how much self-driving vehicles might reduce the crash probability. However, we can use the severity index in the future to estimate the aggregate severity level of a fleet of vehicles for a given penetration level of vehicles with a specific SAE level. Furthermore, I there are other factors that contribute dramatically to the overall severity (such as the impact of the crash partner or location). This paper did not examine these factors due to the limitations of the current study. Still, in the future continuation of the research, we must examine the mentioned factors in detail when estimating severity. Besides, this, further analysis should be performed with a more extensive data set. The identified assumptions and limitations so far have helped to avoid false conclusions. Nevertheless, it is recommended to expand the dataset in the next phase of related research. Thus, this allows the new findings to be compared with the results of the baseline analysis.

Finally, it must be noted that even those factors that have not proven to be significantly more serious for vehicles driven in autopilot mode turned off than for vehicles driven in autopilot mode, can also be potentially important for the safety developments of vehicle industry.

REFERENCES

[1] A. Soteropoulos, M. Mitteregger, M. Berger, and J. Zwirchmayr, "Automated drivability: Toward an assessment of the spatial deployment of level 4 automated vehicles," *Transp. Res. A, Policy Pract.*, vol. 136, pp. 64–84, Jun. 2020.

[2] Z. Zhou, X. Dong, Z. Li, K. Yu, C. Ding, and Y. Yang, "Spatio-temporal feature encoding for traffic accident detection in VANET environment," *IEEE Trans. Intell. Transp. Syst.*, early access, Feb. 15, 2022, doi: 10.1109/TITS.2022.3147826.

[3] D.-J. Lin, M.-Y. Chen, H.-S. Chiang, and P. K. Sharma, "Intelligent traffic accident prediction model for Internet of Vehicles with deep learning approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 3, pp. 2340–2349, Mar. 2022.

[4] N. S. Hadjidimitriou, M. Lippi, M. Dell'Amico, and A. Skiera, "Machine learning for severity classification of accidents involving powered two wheelers," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 10, pp. 4308–4317, Oct. 2020.

[5] Y. K. Ki and D. Y. Lee, "A traffic accident recording and reporting model at intersections," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 2, pp. 188–194, Jun. 2007.

[6] S. S. Ahmed, S. S. Pantangi, U. Eker, G. Fountas, S. E. Still, and P. C. Anastasopoulos, "Analysis of safety benefits and security concerns from the use of autonomous vehicles: A grouped random parameters bivariate probit approach with heterogeneity in means," *Analytic Methods Accident Res.*, vol. 28, Dec. 2020, Art. no. 100134.

[7] A. Y. Bin-Nun and I. Binamira, "A framework for the impact of highly automated vehicles with limited operational design domains," *Transp. Res. A, Policy Pract.*, vol. 139, pp. 174–188, Sep. 2020.

[8] M. Zöldy, Z. Szalay, and V. Tihanyi, "Challenges in homologation process of vehicles with artificial intelligence," *Transport*, vol. 35, no. 4, pp. 447–453, Nov. 2020.

[9] D. Castelvecchi, "Can we open the black box of AI?" *Nature*, vol. 538, no. 7623, pp. 20–23, Oct. 2016.

[10] S. E. Merriman, K. L. Plant, K. M. A. Revell, and N. A. Stanton, "What can we learn from automated vehicle collisions? A deductive thematic analysis of five automated vehicle collisions," *Saf. Sci.*, vol. 141, Sep. 2021, Art. no. 105320, doi: 10.1016/j.ssci.2021.105320.

[11] A. Sarker, H. Shen, M. Rahman, M. Chowdhury, K. Dey, F. Li, Y. Wang, and H. Narman, "A review of sensing and communication, human factors, and controller aspects for information-aware connected and automated vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 1, pp. 7–29, Mar. 2019.

[12] M. T. Alrefaie, S. Summerskill, and T. W. Jackon, "In a heart beat: Using driver's physiological changes to determine the quality of a takeover in highly automated vehicles," *Accident Anal. Prevention*, vol. 131, pp. 180–190, Oct. 2019, doi: 10.1016/j.aap.2019.06.011.

[13] (2018). *Society of Automotive Engineers (SAE) J3016 Taxonomy and Definitions for Terms Related to Driving Automation Systems for on-Road Motor Vehicles*. [Online]. Available: https://www.sae.org/standards/content/j3016_201806/

[14] T. Mackie, "Proving liability for highly and fully automated vehicle accidents in Australia," *Comput. Law Secur. Rev.*, vol. 34, no. 6, pp. 1314–1332, Dec. 2018, doi: 10.1016/J.CLSR.2018.09.002.

[15] J. R. Perello-March, C. G. Burns, R. Woodman, M. T. Elliott, and S. A. Birrell, "Driver state monitoring: Manipulating reliability expectations in simulated automated driving scenarios," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 6, pp. 5187–5197, Jun. 2022.

[16] H. E. Monkhouse, I. Habli, and J. McDermid, "An enhanced vehicle control model for assessing highly automated driving safety," *Rel. Eng. Syst. Safety*, vol. 202, Oct. 2020, Art. no. 107061, doi: 10.1016/j.ress.2020.107061.

[17] A. Ulahannan, S. Thompson, P. Jennings, and S. Birrell, "Using glance behaviour to inform the design of adaptive HMI for partially automated vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 5, pp. 4877–4892, May 2022.

[18] D. Petrović, R. Mijailović, and D. Pešić, "Traffic accidents with autonomous vehicles: Type of collisions, manoeuvres and errors of conventional vehicles' drivers," *Transp. Res. Proc.*, vol. 45, Jan. 2020, pp. 161–168, doi: 10.1016/j.trpro.2020.03.003.

[19] P. Dichabeng, N. Merat, and G. Markkula, "Factors that influence the acceptance of future shared automated vehicles—A focus group study with united kingdom drivers," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 82, pp. 121–140, Oct. 2021, doi: 10.1016/j.trf.2021.08.009.

[20] A. Morando, P. Gershon, B. Mehler, and B. Reimer, "A model for naturalistic glance behavior around Tesla autopilot disengagements," *Accident Anal. Prevention*, vol. 161, Oct. 2021, Art. no. 106348.

[21] F. N. Biondi, M. Lohani, R. Hopman, S. Mills, J. M. Cooper, and D. L. Strayer, "80 MPH and out-of-the-loop: Effects of real-world semi-automated driving on driver workload and arousal," in *Proc. Hum. Factors Ergonom. Soc. Annu. Meeting*, vol. 62, no. 1. Los Angeles, CA, USA: Sage, Sep. 2018, pp. 1878–1882.

[22] E. R. Teoh and D. G. Kidd, "Rage against the machine? Google's self-driving cars versus human drivers," *J. Saf. Res.*, vol. 63, pp. 57–60, Dec. 2017.

[23] F. Favarò, S. Eurich, and N. Nader, "Autonomous vehicles' disengagements: Trends, triggers, and regulatory limitations," *Accident Anal. Prevention*, vol. 110, pp. 136–148, Dec. 2018.

[24] C. Xu, Z. Ding, C. Wang, and Z. Li, "Statistical analysis of the patterns and characteristics of connected and autonomous vehicle involved crashes," *J. Saf. Res.*, vol. 71, pp. 41–47, Dec. 2019.

[25] A. Török and F. Mammadli, "Accidents of highly automated vehicles," *Mendeley Data*, Jan. 2022. [Online]. Available: https://data.mendeley.com/research-data/?search=mammadli%20t%C3%B6r%C3%B6k

[26] V. Butakov and P. Ioannou, "Driving autopilot with personalization feature for improved safety and comfort," in *Proc. IEEE 18th Int. Conf. Intell. Transp. Syst.*, Sep. 2015, pp. 387–393.

[27] T. Móri. (2011). *Statisztikai HipotÉzisvizsgálat*. [Online]. Available: https://edu.interkonyv.hu/book/1067-statisztikai_hipotezisvizsgalat

[28] N. Takácová, "Normalitásvizsgálati módszerek egy dimenzióban," M.S. thesis, Eötvös Loránd Tudományegyetem, Budapest, Hungary, 2015.

[29] H. B. Mann and D. R. Whitney, "On a test of whether one of two random variables is stochastically larger than the other," *Ann. Math. Statist.*, vol. 18, no. 1, pp. 50–60, 1947.

[30] C. P. Doncaster and A. J. H. Davey, *Analysis of Variance and Covariance: How to Choose and Construct Models for the Life Sciences*. Cambridge, U.K.: Cambridge Univ. Press, 2007. Accessed: Nov. 4, 2021. [Online]. Available: https://eprints.soton.ac.uk/49504/

[31] W. H. Kruskal and W. A. Wallis, "Use of ranks in one-criterion variance analysis," *J. Amer. Stat. Assoc.*, vol. 47, no. 260, p. 583, 1952, doi: 10.2307/2280779.

[32] C. E. Bonferroni, "Teoria statistica delle classi e calcolo delle probabilita," *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze*, vol. 8, pp. 3–62, May 1936. Accessed: Nov. 12, 2021. [Online]. Available: https://www.scienceopen.com/document?vid=06182bb9-afa9-4e09-9d1b-fe199febbe81

[33] O. J. Dunn, "Multiple comparisons among means," *J. Amer. Stat. Assoc.*, vol. 56, no. 293, pp. 52–64, Apr. 2012, doi: 10.1080/01621459.1961.10482090.

[34] A. Montella, "A comparative analysis of hotspot identification methods," *Accident Anal. Prevention*, vol. 42, no. 2, pp. 571–581, Mar. 2010.

[35] G. Alkan et al., "Predictive modeling of maximum injury severity and potential economic cost in a car accident based on the general estimates system data," *Comput. Statist.*, vol. 36, no. 3, pp. 1561–1575, Sep. 2021.

[36] K. Geurts, G. Wets, T. Brijs, K. Vanhoof, and D. Karlis, "Ranking and selecting dangerous crash locations: Correcting for the number of passengers and Bayesian ranking plots," *J. Saf. Res.*, vol. 37, no. 1, pp. 83–91, Jan. 2006.

[37] R. Elvik, "State-of-the-art approaches to road accident black spot management and safety analysis of road networks," Oslo Transportøkonomisk Inst., Oslo, Norway, Tech. Rep., Sep. 2007.

[38] E. Paschalidis et al., "Deriving metrics of driving comfort for autonomous vehicles: A dynamic latent variable model of speed choice," *Analytic Methods Accident Res.*, vol. 28, Dec. 2020, Art. no. 100133.

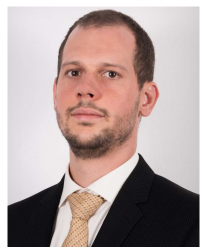
[39] G. Zhai, K. Xie, D. Yang, and H. Yang, "Assessing the safety effectiveness of citywide speed limit reduction: A causal inference approach integrating propensity score matching and spatial difference-in-differences," *Transp. Res. A, Policy Pract.*, vol. 157, pp. 94–106, Mar. 2022.

[40] B. Yu, S. Bao, Y. Chen, and D. J. LeBlanc, "Effects of an integrated collision warning system on risk compensation behavior: An examination under naturalistic driving conditions," *Accident Anal. Prevention*, vol. 163, Dec. 2021, Art. no. 106450.

[41] M. Khashayarfard and H. Nassiri, "Studying the simultaneous effect of autonomous vehicles and distracted driving on safety at unsignalized intersections," *J. Adv. Transp.*, vol. 2021, pp. 1–16, Jun. 2021.

[42] K. Pammer, C. Gauld, A. McKerral, and C. Reeves, "They have to be better than human drivers! Motorcyclists' and cyclists' perceptions of autonomous vehicles," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 78, pp. 246–258, Apr. 2021.

[43] NHTSA. (2019). *U.S. Department of Transportation Releases Policy on Automated Vehicle Development*. Accessed: Oct. 11, 2021. [Online]. Available: https://www.transportation.gov/briefing-room/us-department-transportation-releases-policy-automated-vehicle-development



Árpád Török graduated the M.Sc. degree in transport engineer from TU Budapest in 2006. He received the Ph.D. degree from TU Budapest in 2010. His Ph.D. thesis focused on the conditions of sustainable urban transport and the toolkit of development. He has been with the Department of Automotive Technologies since 2018, where he is currently leading the Safety and Security Research Group. His research interests include automotive safety and security co-engineering and co-analysis techniques.

He worked as a Guest Researcher with INRETS, France, in 2006. In 2008, he was a Guest Researcher with FH Erfurt. From 2013 to 2018, he was the Deputy Head at the Research Centre for Road Safety, KTI. Since 2018, he has been a Leader at the Automotive Safety and Security Research Group, TU Budapest. He has been a member of the Association for Transport Sciences since 2006 and the Hungarian Scientific Academy since 2010.