Journal of Safety Research 92 (2025) 68–80

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Risk of crashes among self-employed truck drivers: Prevalence evaluation using fatigue data and machine learning prediction models

Rodrigo Duarte Soliani [a], Alisson Vinicius Brito Lopes [a], Fábio Santiago [b], Luiz Bueno da Silva [c], Nwabueze Emekwuru [d,*], Ana Carolina Lorena [b]

[a] Federal Institute of Acre, Av. Brazil, 920 - ZIP Code: 69.903-06, Rio Branco/AC, Brazil
[b] Aeronautics Institute of Technology, Praça Marechal Eduardo Gomes, 50 - Zip Code: 12228-900, São José dos Campos/SP, Brazil
[c] Federal University of Paraíba, Via Expressa Padre Zé, s/n – Zip Code: 58051-970, João Pessoa PB Brazil
[d] Engineering Department, Harper Adams University, Newport TF10 8NB, United Kingdom

ARTICLE INFO

Keywords:
Self-employed truck drivers
Fatigue
Occupational conditions
Sleep
Working hours

ABSTRACT

Introduction: Transportation companies have increasingly shifted their workforce from permanent to outsourced roles, a trend that has consequences for self-employed truck drivers. This transition leads to extended working hours, resulting in fatigue and an increased risk of crashes. The present study investigates the factors contributing to fatigue and impairment in truck driving performance while developing a machine learning-based model for predicting the risk of traffic crashes. *Method:* To achieve this, a comprehensive questionnaire was designed, covering various aspects of the participants' sociodemographic characteristics, health, sleep, and working conditions. The questionnaire was administered to 363 self-employed truck drivers operating in the State of São Paulo, Brazil. Approximately 63% of the participants were smokers, while 17.56% reported drinking alcohol more than four times a week, and also admitted to being involved in at least one crash in the last three years. Fifty percent of the respondents reported consuming drugs (such as amphetamines, marijuana, or cocaine). *Results:* The surveyed individuals declared driving for approximately 14.62 h (SD = 1.97) before they felt fatigued, with an average of approximately 5.92 h of sleep in the last 24 h (SD = 0.96). Truck drivers unanimously agreed that waiting times for truck loading/unloading significantly impact the duration of their working day and rest time. The study employed eight machine learning algorithms to estimate the likelihood of truck drivers being involved in crashes, achieving accuracy rates ranging between 78% and 85%. *Conclusions:* These results validated the construction of accurate machine learning-derived models. *Practical Applications*: These findings can inform policies and practices aimed at enhancing the safety and well-being of self-employed truck drivers and the broader public.

1. Introduction

In the transportation industry, there is an ongoing trend of shifting workers from permanent employment to outsourced or self-employed arrangements to reduce overall costs (Messias et al., 2019). By hiring self-employed drivers, transportation companies can lower operating expenses related to fuel, vehicle maintenance, and tires, as well as the responsibilities and costs associated with labor charges (Rocha et al., 2018). However, this model often results in drivers facing more demanding schedules to maintain their earnings, leading to exhaustive working hours.

Evaluating and managing the impact of fatigue on driving safety is a complex challenge (Batson et al., 2023; Caldwell et al., 2019; Sadeghniiat-Haghighi & Yazdi, 2015). This issue is particularly critical in road freight transport, where the main challenge is balancing work hours and rest periods to minimize fatigue-related risks (Wijngaards et al., 2019). In the productivity-driven transportation sector, working time often prioritizes flexibility and timely delivery over safety considerations, sometimes resulting in continuous, around-the-clock operations (De Winter et al., 2024; Hesse, 2016; Soliani, 2022).

Customer demands, supplier relationships, commercial agreements, and new production models drive this work format, yet the occupational conditions of many drivers have deteriorated over the past 15 years (Jaffee & Bensman, 2016). In the road transport sector, supply chains

* Corresponding author.
E-mail address: eemekwuru@harper-adams.ac.uk (N. Emekwuru).

https://doi.org/10.1016/j.jsr.2024.11.002
Received 20 February 2024; Received in revised form 16 June 2024; Accepted 6 November 2024

are pressured to cut costs, forcing carrier owners and drivers to make safety compromises, including working extensive hours, drug usage, speeding, and reduced vehicle maintenance (Belzer & Sedo, 2017; Hege et al., 2015; Lemke et al., 2021; Ryley & Belzer, 2023; Sinagawa et al., 2015). Therefore, researching more effective ways to measure and control fatigue-inducing factors is crucial; especially when this physical health condition is more pronounced due to use of alcohol and drugs (Cardoso et al., 2019; Jakobsen et al., 2023; Rosekind et al., 2024).

To mitigate the dangers associated with driver fatigue, the Brazilian government enacted the Driver's Law (Law 13.103/2015 (Brazil, 2015)). This legislation sets explicit regulations for cargo and passenger vehicle drivers, including a mandatory one-hour meal break, a minimum of 11 h of rest within a 24-hour period, and a 35-hour weekly rest period. Additionally, for long trips, the law requires a 30-minute break after every five and a half hours of continuous driving (Brazil, 2015).

However, Fragoso Junior and Garcia (2019) observed that labor inspections within the Brazilian road freight transport sector failed to adequately address the severity of fatal occupational crash indicators. The study verified this deficiency, which was evident in the insufficient number of inspections conducted, the lack of analysis of work-related crashes, and the absence of inspection of the "working hours" and "rest time" attributes. These factors are considered relevant predisposing factors for occupational crashes in the freight transport sector according to the literature.

Proceeding from this observation, the scientific literature reviewed in this study highlights three main factors influencing truck driver fatigue: sleep (circadian rhythm), working conditions (long driving hours and lack of rest), and health conditions (sleep problems, general health, and lifestyle issues such as alcohol consumption and drug abuse). These factors have guided the design of this research, aiming to address the critical issue of fatigue among truck drivers.

The driver's health, particularly mental health, is significantly impacted by long working hours, interrupted sleep, social isolation, pressure to meet delivery deadlines, stress, low pay, and a lack of medical follow-up (Hege et al., 2019). Professional drivers engaged in long-distance driving, covering over 900 km (500 miles) per day (Zhang et al., 2021), are exposed to numerous mental health risks within the transportation setting. These risks encompass the well-documented consequences of fatigue, as well as additional stressors such as prolonged separation from home and loved ones, heightened time constraints due to Just in Time operations (Makuto et al., 2023).

Further compounding this challenging scenario are procedures intrinsic to the transport activity that extend the working day and reduce sleep time, such as waiting in lines for loading and unloading goods, which have been linked to driver fatigue (Friswell & Williamson, 2019; Mahajan et al., 2019). Drivers often face delays at multiple points in the transport operation, including loading and unloading, which can further extend their working hours (Simonelli et al., 2018).

Self-employed drivers, who are compensated based on productivity (per trip or delivery), rather than time worked, face additional challenges (Girotto et al., 2019). Time spent waiting in queues can diminish their earnings potential, prompting them to extend their driving hours to compensate for the time lost in line (Lemke et al., 2021). Disruptions to sleep patterns can lead to a variety of negative outcomes, including reduced cognitive processing efficiency, slower reaction times, decreased attentional responsiveness, impaired memory, and increased irritability, among other adverse effects on mental and physical health (Bener et al., 2017; Rahman et al., 2023).

This study investigates, analyzes, and compares the factors contributing to the onset of fatigue and diminished truck driving performance through a comprehensive literature review, with a particular focus on self-employed truck drivers. This research provides valuable insights to the broader transportation field, which traditionally does not distinguish between different types of drivers. To assess these impacts, a questionnaire was developed and administered to self-employed truck drivers operating in São Paulo, Brazil's most populous and economically active state, where the majority of goods are transported by truck.

A key contribution of this research is the application of advanced Machine Learning (ML) techniques to the collected data, utilizing eight distinct ML classification algorithms to estimate the risk of traffic crashes for this specific cohort of drivers. This innovative approach offers an accurate assessment of their work conditions, representing a novel contribution to the field of transportation studies.

2. Methodology

The research initially involved conducting a literature review, a structured method enabling the verification of methods used and the reproduction of results obtained by researchers (Seuring et al., 2021). This reduces the risk of bias and subjectivity in the research. The objective of the literature review performed in this study was to locate and examine the most relevant existing studies on the impact of fatigue on the truck driving profession.

To identify the main papers on the subject, searches were conducted in the Scopus, Web of Science, Proquest, PubMed, and Springer Link databases using Boolean operators such as "OR" and "AND." The investigation was limited to works published between 2012 and 2024. This timeframe was specifically chosen to focus on recent studies, aiming to establish a comprehensive understanding of the current state of knowledge in the field of fatigue and safety among truck drivers. Publications had to be written in Portuguese or English, with full-text availability. The search terms used included combinations such as ("Safety" OR "Health" OR "Truck Driver") AND ("Occupational Health and Safety" OR "Fatigue") AND ("Traffic Incidents" OR "Crash-risk" OR "Fatigue").

The research process, as illustrated in Fig. 1, began with consulting five databases and implementing specific search strategies. Boolean operators were used to search for keywords relevant to the study's focus, resulting in the compilation of pertinent records. The results were then filtered to include articles published between 2012 and 2024, available in either Portuguese or English. Subsequently, the gathered records were transferred to the Mendeley reference manager to identify and eliminate duplicates, as well as to verify full-text availability. Finally, the selected records underwent analyses to determine their suitability for inclusion, ensuring alignment with the study's thematic scope.

The article screening process involved two reviewers who assessed titles and abstracts to determine the relevance of the articles based on the established inclusion and exclusion criteria. In cases of disagreement, a third reviewer was consulted to reach a consensus. This meticulous internal review process facilitated the careful curation of a final list of articles for comprehensive scrutiny. Following this screening, a full-text reading of the selected articles was conducted to ensure a thorough examination of their content and findings.

Upon a thorough examination of the chosen articles during the review phase, data were diligently extracted and synthesized. This analytical process revealed a recurring theme: issues related to sleep, health, and working conditions consistently emerged as predominant factors influencing the manifestation of fatigue among truck drivers. These factors not only significantly impacted drivers but also contributed substantially to the broader discourse on occupational risks encountered during the demanding workdays of freight transport drivers.

After identifying three key factors influencing fatigue, a purpose-built questionnaire was devised to further investigate the impact of sleep, health, and working conditions on the occupational risks associated with cargo transportation. Moreover, an exploratory analysis was conducted to investigate the intricate relationships between these factors and various variables, including age, years of professional experience, average weekly distance traveled, frequency of traffic crashes, and drug and alcohol consumption. Potential variations in relation to other sociodemographic variables were also examined. Finally, the data gathered from the questionnaires were processed using machine

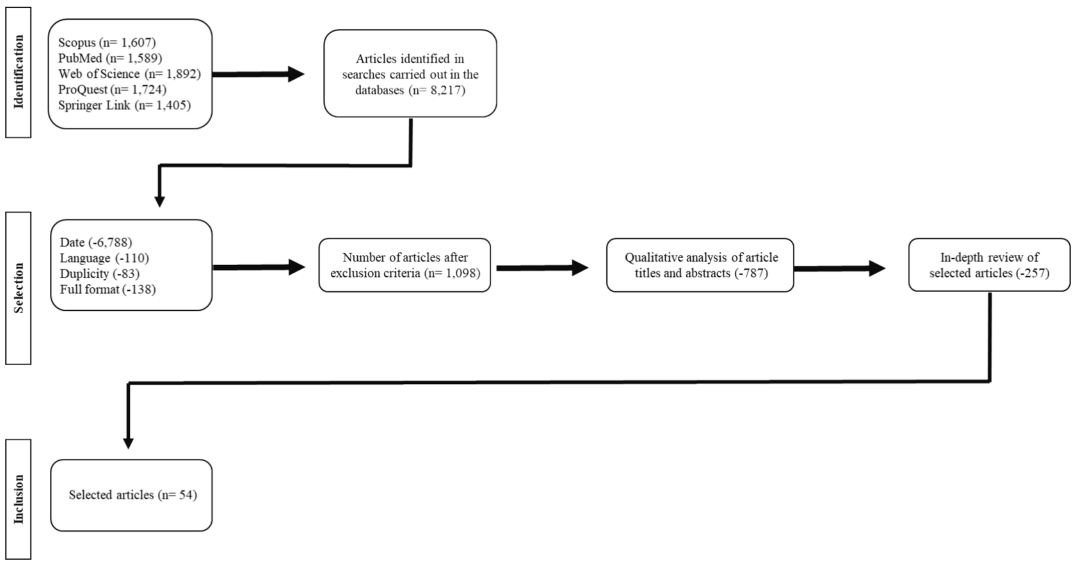

Fig. 1. Methodological approach for literature review.

learning algorithms to determine crash probability.

2.1. Questionnaire

The questionnaire applied consists of 74 questions (71 closed and 3 open), divided into 5 sections, namely:

1. Sociodemographic characteristics including age, marital status, education level, lifestyle factors (such as smoking and alcohol consumption), and working conditions (including traffic crash history, type of shift, daily and weekly driving hours, and years of experience as a full-time driver) (Campos et al., 2020).
2. Sleep: To assess sleepiness levels, the Epworth Sleepiness Scale (ESS) was utilized, which measures sleepiness levels in specific situations, according to a four-point Likert scale (0–3) (Johns, 1991; Filomeno et al., 2019). This is a fast and standardized instrument commonly employed in clinical studies and international research guidelines (Guglielmi et al., 2018).
3. Work: The Job Content Questionnaire (JCQ) is a self-evaluation tool designed to assess the psychosocial aspects of work, including factors such as work control and the psychological demands associated with the job (Useche et al., 2018). The items are structured as a four-point Likert scale, ranging from 'strongly disagree' to 'strongly agree.' This questionnaire focuses on issues related to work demands, the nature of work, decision-making, and social aspects (Karasek et al., 1998).
4. Health: The Fatigue Severity Scale (FSS) is also a self-assessment questionnaire, consisting of nine statements, each one scored from 1 to 7. The Portuguese translation of this questionnaire was validated according to the study by Gouveia et al. (2015). The FSS is applied to assess health aspects and disorders often associated with fatigue (Bener et al., 2017).
5. Visual Analog Scale: similar to other analog scales (pain, depression), it allows participants to indicate on a line the intensity of fatigue they feel, with 0 meaning no fatigue and 10 corresponding to maximum fatigue (Al Salman et al., 2013; Micklewright et al., 2017; Tseng et al., 2010).

2.2. Data collection

The eligible population for this study consisted of self-employed drivers working in food distribution logistics operations in the state of São Paulo. A total of 363 self-employed drivers responded to the questionnaire, representing a 12% increase over the average sample size (n = 320) typically found in the reviewed literature. The questionnaire was administered at a transportation company with the assistance of the administrative team to address any inquiries. It was completed between December 2021 and January 2022, following a researcher's briefing on its objectives and sections.

The choice of the carrier was based on convenience and accessibility, adhering to predefined criteria aligned with the research objectives. Consequently, a medium-sized company with over 10 years of experience in road freight transport was selected. This company operated routes exceeding 300 km and exclusively employed self-employed drivers. All participating drivers held valid driver's licenses, were self-employed, owned their trucks, and were not exclusive service providers for the company.

The data collected through the questionnaire were categorized using Tables 3 and 4, which facilitated a macro analysis of the data. This included obtaining information on the age groups of truck drivers and their levels of education, as shown in Table 3, and identifying the levels of daytime sleepiness stratified by age groups, as displayed in Table 4. Additionally, qualitative data related to truck drivers' fatigue were analyzed using the psychometric Likert scale, enabling the measurement of agreement or disagreement regarding symptoms associated with fatigue.

2.3. Machine learning (ML) techniques

ML techniques have been applied in various contexts to provide information capable of improving decision-making, reducing risk, predicting scenarios, and addressing other applications (Gomes et al., 2023). Fundamentally, these algorithms begin with a database related to the problem under study. Through the ML model implemented by each of them, predictions are made, whether of a regressor or classifier nature. This process enables the application of metrics that quantify the error or probability associated with the projection of interest (Naser & Alavi, 2023).

The present study aims to determine the risk that self-employed truck drivers face of being involved in a crash during their daily work hours. To achieve this, data collected through questionnaires administered to drivers were processed using a collection of ML algorithms. This approach allowed for the determination of the probability of a crash occurring.

Given the classifying nature of the problem studied, that is, based on predictive attributes, the algorithms should decide for one of two distinct classes: the class that suffers or the class that does not suffer. In this way, it was then decided to employ the following collection of classifier algorithms: GaussianNB (GNB), K- Nearest Neighbor (KNN), Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), Logistic Regression (LR), Light Gradient Boosting Machine (LGBM) and the Extreme Gradient Boost known as XGBoost (XGB). These are available in Python language through ML libraries such as Scikit-learn (2022), LightGBM (2022) and DMLC XGBoost (2022).

After selecting the ML algorithms for this work, the next step was to choose the predictive attributes and classes. Thus, the following predictive attributes were selected: alcohol consumption by the truck driver, drug use, discomfort related to the position, perception of excessive working hours, truck driver sleep hours, number of working hours, and the FSS and ESS scales. Two classes were used: suffering or not suffering a crash. Therefore, each of the 363 records available in the database consists of 8 predictive attributes and one of two classes.

Once the ML models were defined, the next task was to choose the hyperparameters and metrics capable of evaluating their performance. Table 1 consolidates each of the algorithms along with the chosen hyperparameters associated with them. Overall, for a comprehensive list of possible hyperparameters readers should refer for instance to Scikit-learn (2022), LightGBM (2022) and DMLC XGBoost (2022).

To quantify the performance of the algorithms, two levels of evaluation were chosen: the first involved obtaining the confusion matrix, as shown in Fig. 2, for each of the algorithms. For this test, a ratio of training and testing data similar to that used in Xu and Goodacre (2018) was considered.

Once the confusion matrix was obtained for each of the algorithms, the associated metrics were then calculated, as shown in Table 2.

At the second level of model evaluation, the objective was to assess their generalization capacity using the K-fold Cross-Validation test, as outlined in Rodriguez et al. (2010). This test aims to evaluate the generalization capacity of a classification algorithm. A total of 400 divisions between the training and test databases were utilized, with 40 cross-validation iterations performed. In each iteration, the database was divided into 10 partitions. From the results of the K-fold test, the probability density distribution for each of the ML models can be obtained, allowing the application of various statistical metrics related to each distribution, as well as comparisons between each model.

3. Results

3.1. Exploratory analysis

A total of n = 363 drivers responded to 74 questions. These questionnaires encompassed extensive socio-demographic inquiries (see Appendix A), along with the FSS, JCQ, and ESS methods, and a visual fatigue scale. All drivers were male and self-employed. The drivers were classified into five age groups with intervals of nine years. Table 3 provides a summary of the socio-demographic data collected, based on the questionnaire presented in Appendix A.

Table 1
Hyperparameters of ML methods.

ML Models	Hyperparameters
GaussianNB (GNB)	***None***
K-Nearest Neighbor (KNN),	***n_neighbors*** = 5, ***metric*** = "minkowski"
Decision Tree (DT),	***max_depth*** = 3, ***criterion*** = "entropy",
Random Forest (RF)	***n_estimators*** = 150, ***criterion*** = "gini", ***max_depth*** = 4
Support Vector Machine (SVM)	***kernel*** = "rbf", ***C*** = 1
Logistic Regression (LR)	***max_iter*** = 300, ***penalty*** = l2
Light Gradient Boosting Machine (LGBM)	***num_leaves*** = 150, ***objective*** = binary, ***max_depth*** = 7, ***learning_rate*** = 0.05, ***max_bin*** = 200
XGBoost (XGB)	***max_depth*** = 4, ***n_estimators*** = 200, ***random_state*** = 3, ***learning_rate*** = 0.05

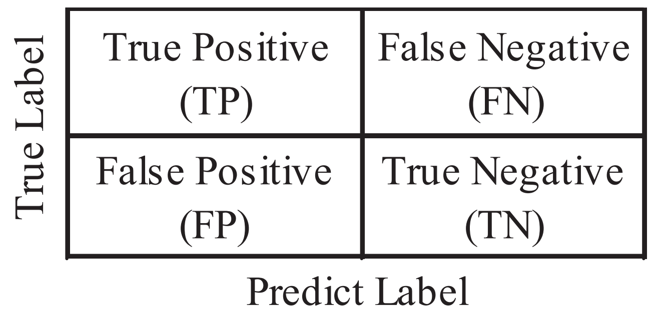

Fig. 2. Confusion Matrix.

Table 2
Predictive performance metrics based on the Confusion Matrix.

Metrics	Expression
Accuracy	(TP + TN)/(TP + FN + FP + TN)
Recall	TP/(TP + FN)
Precision	TP/(TP + FP)

Table 3
Summary of driver's socio-demographics profile (n = 363).

	Age category (in years range)				
	25–34	35–44	45–54	55–64	≥65
Number of drivers	74	102	88	66	33
(% of total number of drivers)	20.38	28.09	24.24	18.20	9.09
Education (%)					
Primary	87.83	0	75	72.72	93.94
Secondary	10.82	100	25	0	0
No education	1.35	0	0	27.28	6.06
Civil status (%)					
Single	5.4	11.68	5.68	0	0
Married	83.79	74.02	75	100	100
Divorced	10.81	14.30	19.32	0	0
Smokes? (%)					
Yes	63.52	60.79	72.73	72.73	72.73
No	36.48	39.21	27.27	27.27	27.27
Takes alcohol? (%)					
No	45.94	8.82	0	0	12.12
2–3 times per week	36.50	84.32	61.36	50	78.78
≥ 4 times per week	17.56	6.86	38.64	50	9.1
Drug use (%)					
Yes	50	71.57	38.64	39.39	24.24
No	50	28.43	61.36	60.61	75.76
Number of traffic crashes (last 3 years) (%)					
None	58.1	71.57	73.86	57.58	100
One	41.90	26.47	25	42.42	0
Two	0	1.96	1.14	0	0

According to Table 3, the majority of drivers (~72.71%) fall within the age range of 25 to 54 years old, with less than 10% aged above 65 years old. Notably, none of the drivers hold a university or college degree. Among the younger drivers (25–34 years), the majority (87.83%) have only completed primary education, and 83.78% are married, while 10.81% are divorced. Within this age group, 89.18% have less than five years of working experience, and over 63% are smokers.

Although 45.94% of the drivers did not consume alcohol, a majority (54.05%) consumed alcohol at least twice a week. Significantly, 17.56%

of drivers reported drinking alcohol more than four times a week, and these drivers reported being involved in at least one crash in the last three years. This timeframe allows for an in-depth analysis, capturing traffic incident patterns and trends relevant for occupational health and safety decision-making and interventions (Roman et al., 2015).

In the 25–34 years age group, out of the 27 drivers who reported drinking alcohol 2–3 times a week, 37% were not involved in any traffic crashes, while 63% experienced at least one traffic incident. Conversely, out of 34 drivers (45.94% of all drivers in this category) who reported not drinking, 33 of them did not use drugs, and none of these non-drinking, non-drug using drivers were involved in any crashes. However, the sole non-drinking driver who used drugs (amphetamines) reported one crash. These results suggest a strong correlation between traffic incidents and alcohol intake in this category. Additionally, half of the drivers in this age group reported drug use, with 33.78% consuming amphetamines and 16.21% combining amphetamines with either cannabis or cocaine.

Drivers participating in this research reported that, during their last experience of fatigue, they were driving for an average of 14.62 h (SD = 1.97). Approximately 43 drivers (12%) reported driving for over 16 h a day, with roughly 10% of all drivers admitting to driving for 18 h. When asked to rate their level of fatigue on a scale from 0 (total absence of fatigue) to 10 (maximum fatigue), none of the drivers rated themselves below five. The average score among all 363 respondents was 7.00 (SD = 1.60). Furthermore, 161 drivers classified their fatigue as equal to or greater than 8, while 20 respondents considered themselves to be at the maximum level of fatigue.

Overall, the Likert chart of the Fatigue Severity Scale (Fig. 3) indicates that truck drivers generally agree with the statements, reflecting the negative impact of fatigue on their lives. The mean fatigue score is 43 with an SD = 8.43, suggesting that self-employed drivers are experiencing fatigue at a level warranting medical attention. Notably, about 78% of drivers feel their motivation decreases when fatigued, with none expressing disagreement and 22% neither agreeing nor disagreeing. Moreover, over 96% of drivers agree they are usually too fatigued to exercise, with 61% mainly or strongly agreeing.

A significant majority of drivers, specifically over 63%, acknowledge they are prone to experiencing fatigue easily, with no disagreement among respondents. However, when asked about the statement "fatigue interferes with my work, family, or social life," drivers' responses varied. Approximately 56% neither agreed nor disagreed, while 26% partially disagreed, indicating uncertainty or divergence in responses. This finding suggests the need for further investigation, potentially involving family member surveys to understand their perceptions of driver fatigue and its impact on daily family and social interactions.

A similar trend is observed regarding the statement "fatigue interferes with my physical functioning," with 43% partially disagreeing and 50% partially agreeing. In summary, the fatigue severity scale reveals not only an alarming score but also negative feelings among drivers regarding their fatigue condition, especially concerning its impact on family and social life.

To further explore sleeping issues concerning drivers, an inquiry was made in the JCQ section regarding whether waiting time for loading/unloading affected their working and resting hours, to which all respondents provided affirmative responses (66 participants left this question blank). Comments from drivers highlighted issues such as poor logistical management by companies, inadequate resting facilities, and long queuing times. Some noted improvements in scheduling loading or unloading time slots by certain companies, resulting in reduced waiting time.

Table 4 presents the results of applying the ESS score to the 363 drivers, categorized by age. Overall, the ESS score remains consistent across all age categories, indicating a similar level of moderate excessive daytime sleepiness regardless of age. Notably, around 9% of drivers scored 16 or higher, signifying severe excessive daytime sleepiness. Among these individuals, approximately 38.70% were involved in at least one traffic crash in the last three years, with all reporting smoking and alcohol consumption, and 83.33% using drugs, particularly amphetamines.

With regards to the 363 drivers, the mean number of sleep hours in the last 24 h averaged 5.92 h (SD = 0.96). Approximately 15.3% of the drivers reported sleeping less than 5 h, with the minimum number of hours reported being 0 (indicating no sleep in the last 24 h prior to taking the survey). Conversely, 23.8% reported sleeping between 7 and 8 h.

Based on this exploratory analysis, significant factors that can impact the risk of crashes were identified. These factors include the duration of sleep among truck drivers, their drug usage, alcohol consumption, as

Table 4
Epworth Sleepiness Scale (ESS) Score.

Age category (years)	Mean	Standard Deviation	Min	Max
25–34	12,92	3,24	8	16
35–44	13,00	2,74	10	17
45–54	12,93	3,06	8	17
55–64	13,06	4,57	8	18
≥65	12,84	4,43	8	16
ESS score (all categories)	12,96	3,41	8	18

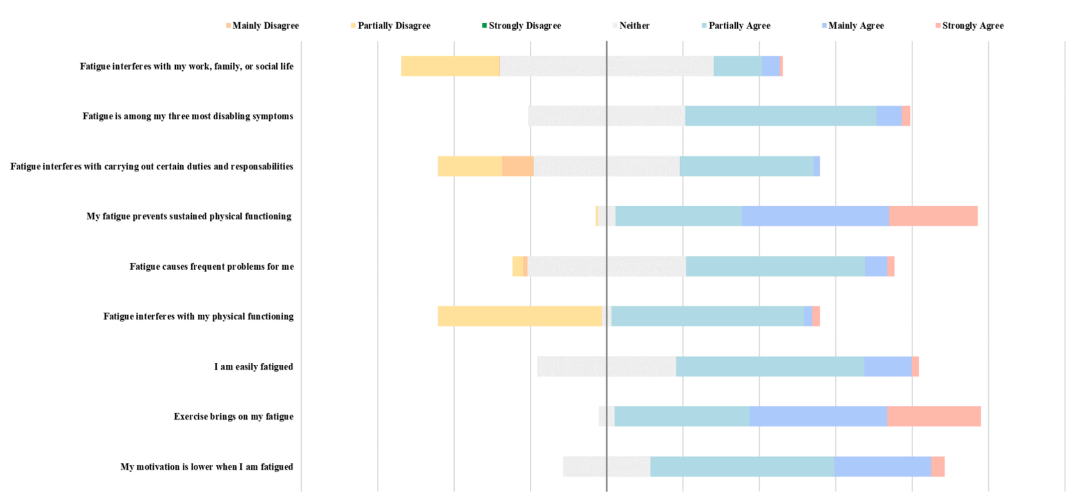

Fig. 3. Fatigue Severity Scale (FSS).

well as indicators obtained from the ESS (Epworth Sleepiness Scale) and FSS (Fatigue Severity Scale) scales. These findings validate the data utilized in this study.

3.2. Machine learning (ML) techniques

In this section, the results obtained with the ML algorithms will be presented and evaluated after being applied to the problem that is the subject of this research in accordance with the methodology described in section 2.3. Thus, for each of the algorithms listed in Table 1, their confusion matrix was obtained, as shown in Fig. 4 below.

The interpretation of each of the fields in the confusion matrix is shown in Fig. 4. In the present case, the ML algorithms classified each of the records belonging to the test database into one of two possible classes, namely: zero (0) when the driver does not experience a crash and one (1) when the driver experiences a crash.

Once the confusion matrix for each algorithm is obtained, the metrics defined in Table 2 can be calculated, as demonstrated in Table 5.

The values presented by the metrics in Table 5 were evaluated according to the ML libraries in Scikit-Learn (2022) to obtain a broad understanding of the performance of the classification algorithms and their limitations in predicting crashes involving truck drivers. Therefore, the assessments carried out at this stage are primarily related to class (1), that is, the occurrence of crashes.

The analysis of the results began with those provided by the Precision metric of class (1), which showed that most algorithms have hit proportions varying between 74% and 87%, with the Random Forest algorithm standing out in this regard with 91% accuracy, followed by the Naive Bayes and Logistic Regression algorithms.

Continuing with the evaluation of the results, those referring to the Recall metric for class (1) were then considered. This analysis showed that the algorithm with the best performance in this regard was Logistic Regression, with a hit proportion of 89%, followed by the Decision Tree algorithm with 88% of correct answers. The other algorithms presented proportions of correct answers varying between 48% and 64%.

Table 5
Confusion Matrix metrics results.

Algorithm	Accuracy	Class	Precision	Recall
GaussianNB (GNB)	0.75	0	0.71	0.95
		1	0.87	0.48
K-Nearest Neighbor (KNN)	0.77	0	0.75	0.89
		1	0.81	0.60
Decision Tree (DT)	0.82	0	0.90	0.77
		1	0.74	0.88
Random Forest (RF)	0.77	0	0.72	0.96
		1	0.91	0.50
Support Vector Machine (SVM)	0.79	0	0.75	0.95
		1	0.89	0.57
Logistic Regression (LR)	0.81	0	0.78	0.93
		1	0.87	0.89
Light Gradient Boosting Machine (LGBM)	0.76	0	0.80	0.57
		1	0.76	0.57
XGBoost (XGB)	0.80	0	0.78	0.91
		1	0.84	0.64

Finally, the results obtained by the accuracy metric were evaluated, showing that most classifier algorithms presented hit proportions varying between 75% and 80%. The Decision Tree algorithm stood out with 82% accuracy, the highest proportion among all, followed by the Logistic Regression algorithm with 81% accuracy.

Once the analyses on each of the confusion matrix metrics were completed, a decision was made based on these results. The Logistic Regression algorithm, which obtained an accuracy of 81%, Precision in class (1) of 87%, and Recall in class (1) of 89%, stood out among the algorithms used in this test, thus completing the first stage of evaluations.

The second stage of evaluating the algorithms, as provided for in the methodology described in section 2.3, consists of applying the K-fold test. This test aims to evaluate the generalization capacity of the classification algorithms on the data studied. The application of the aforementioned test was considered with k = 10, meaning 10 different partitions of the database. For each of these partitions, 40 iterations

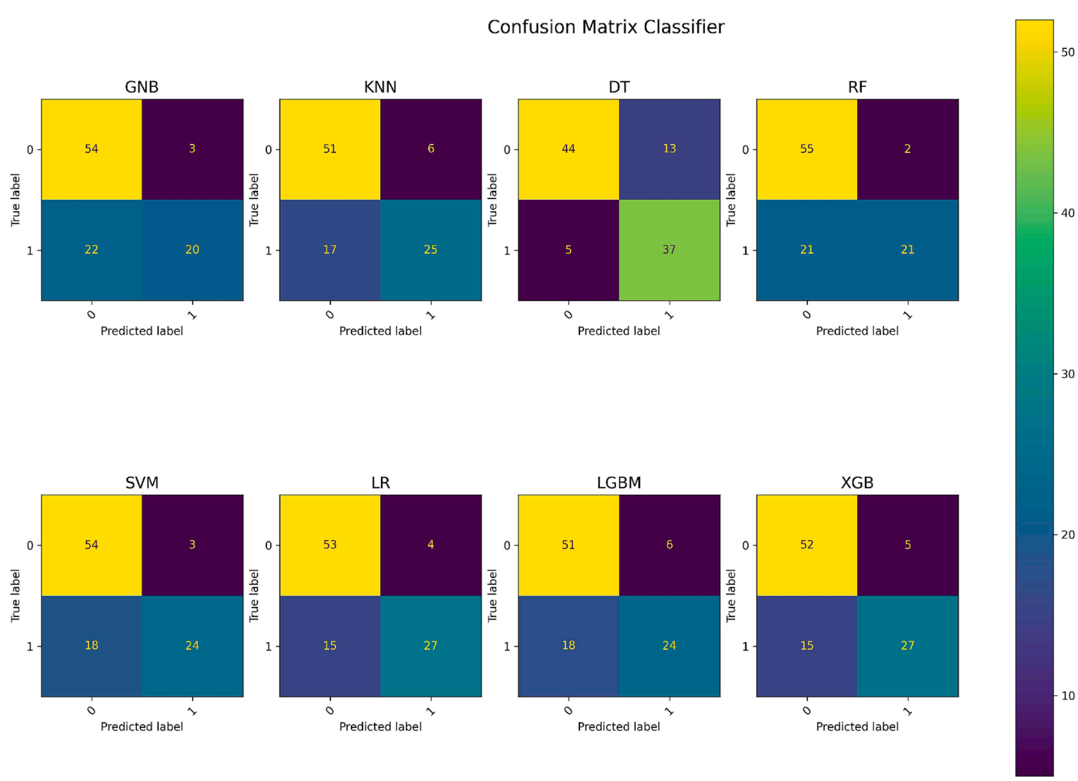

Fig. 4. Confusion matrix to ML models.

were carried out, randomizing the data. In each of the iterations, the accuracy metric was calculated for each of the algorithms, obtaining its distribution, as shown in Fig. 5.

The qualitative analysis of the image in Fig. 5 shows that the ML algorithms in this work have accuracy values greater than 76%, with the SVM algorithm standing out with accuracy values greater than 86%. However, to gain a better understanding of the accuracy distribution of the algorithms and consequently their generalization capacity, classic statistical tools were used, namely: boxplot, dispersion, and position measurements.

The boxplot graph of the data from each of the curves in Fig. 5 can be seen in Fig. 6 below. It shows that among the algorithms used in this work, SVM has the highest accuracy value. For the median, the Random Forest algorithm has the lowest value in this regard. Furthermore, it is observed that the LightGBM, Naive Bayes, and Logistic Regression algorithms generated accuracy distributions with outlier values, as evidenced in their boxplot graph.

Finally, the accuracy distributions obtained by the algorithms were evaluated using position and dispersion metrics, and the results of these evaluations were collected in Table 6 below.

As shown in Table 6, the algorithms have an average accuracy greater than 80%, with the SVM algorithm standing out with an average of 83.58%. In contrast, the Random Forest algorithm presented an average accuracy of 79%, while the others had average accuracy values ranging between 80.60% and 82.36%.

Analyzing the set of algorithms through the combination of the various metrics presented in Table 6, it is concluded that the SVM algorithm is the one that demonstrates the best performance in the k-fold test. It exhibits the highest average accuracy, low standard deviation values, a minimum accuracy value greater than 81%, and a maximum accuracy value of 85.45%, indicating its high generalization capacity.

With the results and evaluations conducted in this section, it is concluded that the ML algorithms used in this work demonstrate high robustness in predicting crashes involving the target audience of this study. This capacity of the models enables proactive decision-making based on a set of metrics, thus preventing crashes, conserving resources, and saving lives.

4. Discussion

Truck drivers face significant challenges regarding fatigue, as evidenced by the average driving time of 14.62 h reported in this study. This duration substantially exceeds the global average of 11 to 13 h (Goel & Vidal, 2014; Lemke et al., 2021), highlighting the prevalence of high levels of fatigue among self-employed drivers in Brazil and its correlation with road crashes.

The ability to respond effectively and manage critical situations is vital for drivers. However, fatigue can significantly affect driver behavior in such situations, potentially leading to crashes. Studies conducted by Narciso and de Mello (2017), Mizuno et al. (2020), Pourabdian et al. (2020), and Zhang et al. (2024) demonstrate how fatigue can reduce truck drivers' ability to perform calculations and increase the frequency of errors. Additionally, fatigue impairs information processing, resulting in longer decision-making times.

The result obtained with the Fatigue Severity Scale (FSS) indicates that truck drivers interviewed tend to agree with statements suggesting a negative impact of fatigue on their lives. The average fatigue score obtained was 43, suggesting a high level of fatigue that requires medical attention. Approximately 78% of drivers report that their motivation decreases when they are fatigued, while more than 96% agree that they are generally too tired to exercise. This finding is consistent with studies by Crizzle et al. (2024), Passey et al. (2014), Shaw et al. (2023) and Thiese et al. (2018), which show that heart disease and lack of physical exercise are among the most prominent health issues affecting drivers.

The studies by Casey et al. (2024), Mahajan et al. (2019), and Wise et al. (2018) support the challenges reported by drivers regarding waiting and queuing for loading and unloading. These challenges, including poor logistical management, inadequate resting facilities, and long queuing times, emphasize the need for proactive measures to improve freight management efficiency and alleviate driver fatigue.

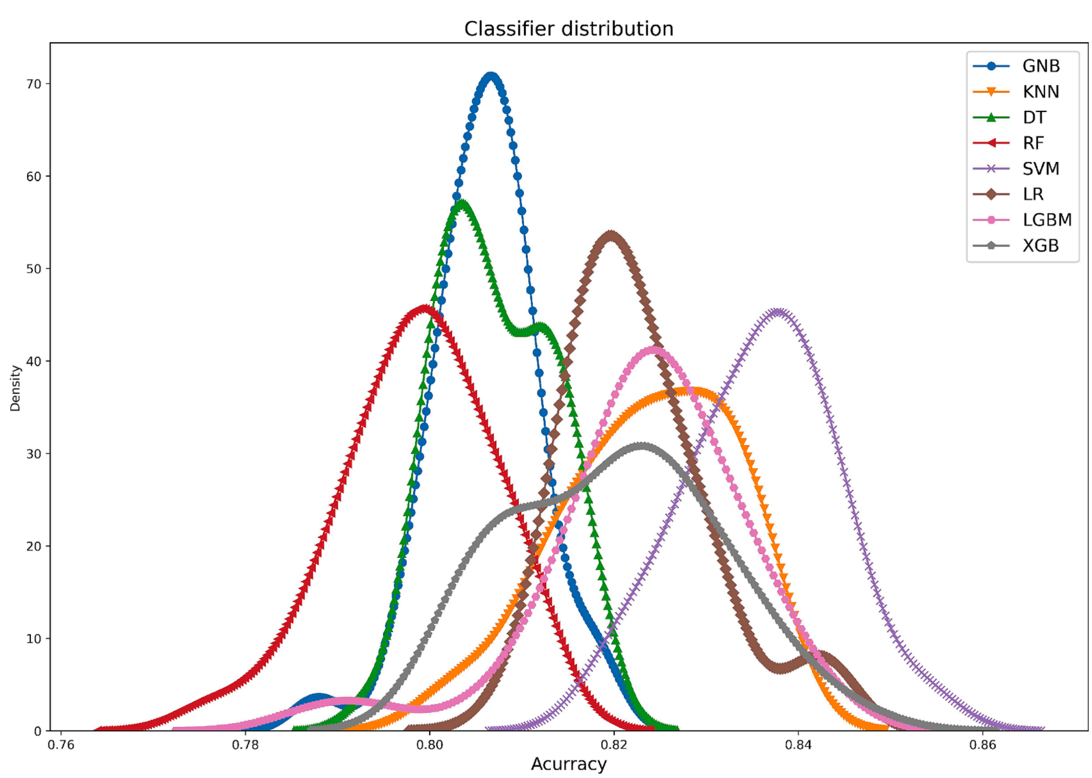

Fig. 5. Distributions of predictive accuracies of the classifiers in the K-fold Cross-Validation procedure.

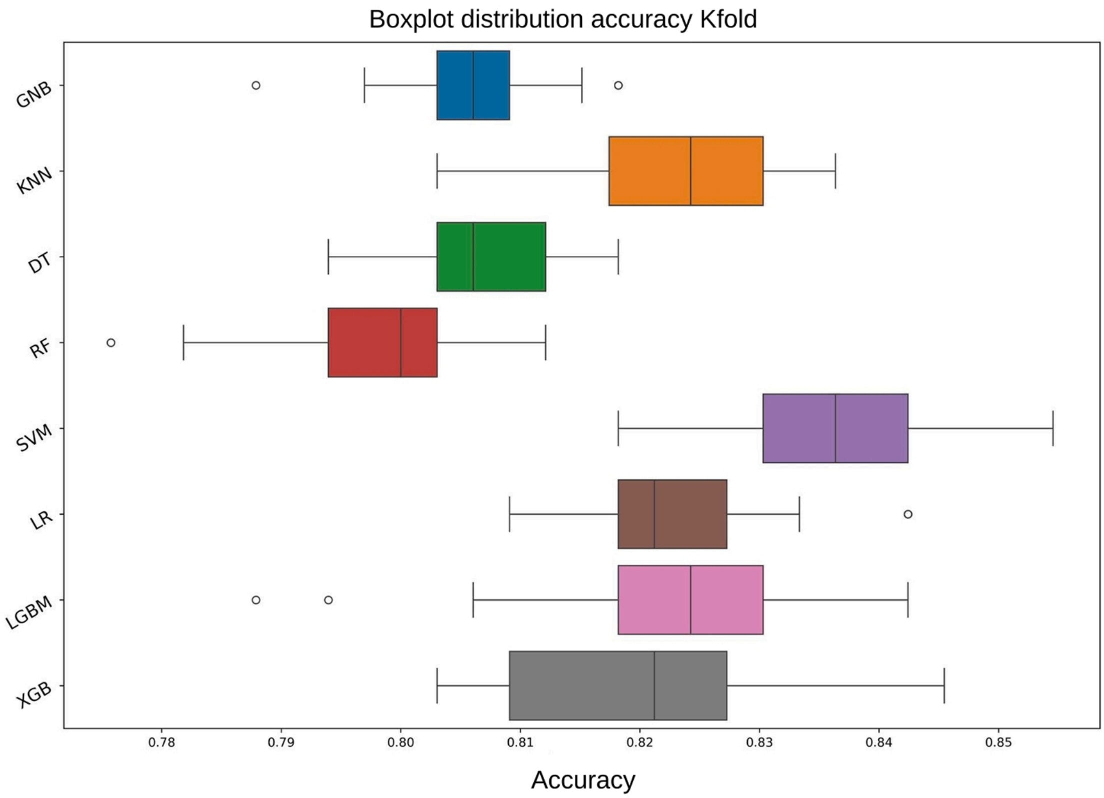

Fig. 6. Boxplot distribution of accuracies obtained by K-fold Cross-Validation procedure.

Table 6
K-fold Cross-Validation accuracy results.

	Classifiers							
Metrics	**GNB**	**KNN**	**DT**	**RF**	**SVM**	**LR**	**LGBM**	**XGB**
mean	0.8060	0.8236	0.8071	0.7979	0.8358	0.8228	0.8231	0.8189
std	0.0058	0.0091	0.0060	0.0102	0.0082	0.0079	0.0108	0.0122
min	0.7879	0.8030	0.7939	0.7758	0.8182	0.8091	0.7879	0.7939
25%	0.8030	0.8174	0.8030	0.7902	0.8303	0.8182	0.8182	0.8091
50%	0.8061	0.8242	0.8061	0.7985	0.8364	0.8212	0.8242	0.8212
75%	0.8091	0.8303	0.8121	0.8061	0.8424	0.8273	0.8303	0.8273
max	0.8182	0.8364	0.8182	0.8121	0.8545	0.8424	0.8424	0.8424

Friswell and Williamson (2019) suggest that by reducing inefficiencies in freight management and enhancing communication between drivers, carriers, and customers, the industry can address these challenges and ensure drivers have sufficient rest before driving. Additionally, incentivizing efficient freight management practices may encourage transport customers to prioritize reducing waiting times and queuing, thus minimizing the negative impact of disrupted sleep patterns on drivers' mental and physical health.

Among the 363 drivers surveyed, the average sleep duration in the past 24 h was 5.92 h. Notably, around 15.3% of drivers reported sleeping less than 5 h, and some even admitted to having had no sleep at all. Conversely, 23.8% reported sleeping between 7 and 8 h. These results are consistent with the research of Mahajan et al. (2019), who identified a significant link between sleep duration and accident rates among drivers, emphasizing that those with less than 7 h of sleep faced heightened risks. Particularly concerning was their discovery that individuals sleeping less than 4 h per day had a staggering 4.8 times higher chance of experiencing drowsiness while driving. These findings align with the observations of Hege et al. (2015), indicating that truck drivers typically average less than 6 h of sleep per day, highlighting a worrisome lack of improvement in drivers' overall rest duration.

Studies conducted by researchers such as Filomeno et al. (2019), Garbarino et al. (2018), Jakobsen et al. (2023), Léger and Stepnowsky (2020), Mello et al. (2013), Soliani and Da Silva (2023), Ward et al. (2013), and Zhang and Chan (2014) have provided substantial evidence in the literature highlighting excessive daytime sleepiness as a significant factor in road incidents worldwide. Fatigue contributes to approximately 20% to 30% of all traffic crashes globally (Nazari et al., 2017), resulting in an estimated annual cost of around R$ 50 billion to Brazilian society (Ipea, 2020), mainly due to expenses related to lost productivity among victims and hospital costs.

Truck drivers operate in stressful conditions that promote unhealthy lifestyles, making them more susceptible to health issues. Prolonged driving periods, inadequate sleep, and disruptions in circadian rhythms lead to diminished driving performance (Bener et al., 2017; Philip et al., 2019; Rahman et al., 2023), with inadequate sleep significantly affecting physical and cognitive functioning. Tasks requiring complex mental processes and sustained vigilance are particularly affected by sleep deprivation, resulting in decreased cognitive efficiency (Heaton et al., 2021), slower reaction times, and pronounced cognitive distortions (Cardoso et al., 2019).

The tests conducted here demonstrated that the algorithms are highly robust in predicting crashes. This enables proactive decisions to prevent accidents based on various metrics, ultimately saving lives and resources. Furthermore, it can alert authorities to the high risk faced by truck drivers due to long and exhausting working hours, thereby sparking discussions and potentially leading to legislation to protect these workers.

5. Conclusion

Comparatively, few studies investigate the occupational conditions specific to truck drivers, focusing primarily on the prevalence of diseases and the health effects of their work. This study identifies three main factors contributing to truck driver fatigue: sleep issues (circadian rhythm disruptions), work-related factors (long driving periods, inadequate rest), and health factors (sleep disorders, overall health status, lifestyle habits).

Self-employed truck drivers face strenuous workloads leading to significant physical and mental fatigue, adversely affecting their health and quality of life. Unhealthy behaviors such as sedentary lifestyles, poor eating habits, and substance use (alcohol and drugs) are common. Night shifts further disrupt biological rhythms, leading to long-term health impacts and affecting social and family life. These lifestyle choices increase fatigue, raising the risk of incidents and compromising road safety.

Among surveyed drivers, most (72.71%) were aged 25 to 54 years. Of those aged 25 to 34 years, over 63% were smokers, and 17.46% drank alcohol four or more times per week. Notably, many had experienced crashes in the past three years. Among 27 drivers who drank alcohol 2–3 times a week, 63% had at least one traffic incident, highlighting a strong link between alcohol consumption and crashes. About half of the drivers used drugs, with 33.78% consuming amphetamines and 16.21% using amphetamines, marijuana, or cocaine.

Drivers reported driving for about 14.62 h the last time they felt fatigued. On a fatigue scale from 0 (no fatigue) to 10 (maximum fatigue), the average score was 7.00 (SD = 1.60). Around 78% reported lower motivation when tired, and more than 59% identified fatigue as one of the three most disabling symptoms, corroborated by 95% who agreed that fatigue prevents continuous physical functioning. The average sleep in the last 24 h was 5.92 h (SD = 0.96). The Epworth Sleepiness Scale (ESS) scores indicated similar moderate excessive daytime sleepiness across all age groups. About 9% scored 16 or greater, indicating severe excessive daytime sleepiness, with 38.70% involved in crashes in the last three years. All these drivers smoked and drank alcohol, while 83.33% used drugs, mainly amphetamines. Predictive models using machine learning techniques achieved about 80% accuracy in predicting crash risks, aiding in understanding risk factors.

Drivers reported that the waiting time for loading goods directly affected their working and resting hours, emphasizing the critical need for proactive measures to improve freight management efficiency and alleviate driver fatigue. The study suggests truck drivers are working under extreme conditions, necessitating policies to enhance their working conditions. Disease prevention policies and specific health promotion measures are crucial. The findings can inform public policies to improve working conditions, regulate working hours, prevent drug and alcohol use, and reduce traffic crashes. Future studies should conduct surveys with a representative national sample and deepen discussions on the impact of lifestyle, health, and work satisfaction on truck drivers.

CRediT authorship contribution statement

Rodrigo Duarte Soliani: Conceptualization, Investigation, Methodology, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing. **Alisson Vinicius Brito Lopes:** Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. **Fábio Santiago:** Data curation, Formal analysis, Methodology, Software, Validation, Visualization, Writing – review & editing. **Luiz Bueno da Silva:** Conceptualization, Supervision, Writing – original draft, Writing – review & editing. **Nwabueze Emekwuru:** Supervision, Writing – review & editing. **Ana Carolina Lorena:** Software, Supervision, Validation, Visualization, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A

Questionnaire

Sociodemographic characteristics

Driver's Profile
Variables
Age (years)
18–24
25–34
35–44
45–54
55–64
>65
Sex
Female
Male
Schooling
Illiterate
Elementary School
High School
College
Graduation
Marital Status
Single

(*continued on next page*)

(*continued*)

Driver's Profile
Stable Union
Married
Divorced
Widower
Employment
Hired
Self-employed
Length of Work (years of driving experience)
<5
5–15
15–30
30–45
>45
Daily hours of work
Up to 8 h
Up to 11 h
> 11 h
Hours at work
Between 8 AM to 10 PM
From 6 or 8 AM and from 10 PM and 2 AM
Part of the journey is between 2 and 6 AM
Weekly Kilometrage (average)
<500 km
500–700 km
700–900 km
900–1.100 km
>1.100 km
Traffic crashes
0
1
2 or more

Job Content questionnaire (JCQ)

Job Content Questionnaire (JCQ)

Dimensions	Questions	Strongly agree	Agree	Disagree	Strongly disagree
Control over work	At work you can learn things				
	Your job is repetitive				
	Your job demands a high level of qualification				
	Your job requires you to be creative				
	You are in charge of different tasks				
	at work you can develop special abilities				
	You can take decisions				
	You have a lot to tell about your job				
	You have little freedom in how to perform tasks				
Macro power of decision	Is it possible for your ideas to be considered in the elaboration of the policies adopted in the company? (Hiring, salary level, dismissal, purchase of equipment, etc.)				
	Part of your job is to supervise others				
Psychological demands	You have to work fast				
	Your job requires you to work a lot				
	There is too much to be done				
	The time you are given is enough				
	You are free of conflicting demands from others?				
	Your job is developed at a frantic pace				
	You are frequently interrupted so you have to finish tasks later				
	Your job demands long hours of concentration				
	Does waiting for other people's work often slow down the pace of your work?				
Physical demands	Your job demands fast movements				
	Your job demands physical effort				
	You need to keep head and arms in difficult positions for long hours				
	You need to keep your body in difficult physical positions				
	Along the day you are asked to move or lift				
Social support from supervisor	(_) Don't have a supervisor				
	Your supervisor is concerned about the employees' well being				
	Your supervisor pays attention to what you say				
	Your supervisor is hostile at times				
	Your supervisor helps you				
	Your supervisor succeeds in team work management				
Social support from workmates	(_) Work by myself				
	Your team is competent				
	Your team is concerned with the job developed				
	Your team may be hostile at times				
	Your team is friendly				

(*continued on next page*)

(*continued*)

Job Content Questionnaire (JCQ)					
Dimensions	Questions	Strongly agree	Agree	Disagree	Strongly disagree
	Your team is collaborative				
	Your team encourage each other				
Work situation instability	Your stability at work is				
	In 5-year-time your stability will be				
	Your promotion in the career is feasible				
	Last year were you unemployed or in temporary job?				
	Some people lose permanently their dream job. What is the possibility of losing your job in the next 2 years?				
Macro power of decision	(_) Work by myself				
	How many people are there in your team?				
	Do you have any influence over the decisions of your team?				
	Are the decisions democratic?				
	Are you in a union of some sort?				
	Can your union influence the company's policies?				
	Do you have a voice inside your union?				
Are the waiting lines for loading /unloading affecting your work or your free time? Make comments.					

Epworth sleepiness Scale (ESS)

Epworth Sleepiness Scale (ESS)	Chance to nap			
Situation	0	1	2	3
1) Sitting and reading				
2) Watching TV				
3) Sitting in a public place				
4) As a passenger in a car riding for over 1 h				
5) Lying down in the afternoon when I get the chance				
6) Sitting talking to someone				
7) Calmly sitting after lunch, no alcohol consumption				
8) In the car, when the car stops due to heavy traffic				

0 – No chance to nap
1- Little chance to nap
2- Moderate chance to nap
3- High chance to nap

How many hours have you slept in the last 24 h? Make comments on the quality of your sleep.

Fatigue Severity Scale (FSS)

Fatigue Severity Scale (FSS)	Don't Agree Agree						
Questions	1	2	3	4	5	6	7
1. I'm less motivated when I'm tired							
2. Physical exercise gives me fatigue							
3. I'm easily tired							
4. Fatigue interferes in my performance							
5. Fatigue often causes me problems							
6. My fatigue prevents me from working long hours							
7. Fatigue interferes in my performance in certain activities							
8. Fatigue is one of the 3 incapacitating symptoms I have							
9. Fatigue interferes in my job, in my family life and in my social life							

How long have you been working the last time you felt fatigue? Make comments on your work journey

Visual Analog Scale

Visual Analog Scale										
From 0 to 10, what is your level of fatigue right now?										
Light			Moderate					Intense		
0	1	2	3	4	5	6	7	8	9	10

REFERENCES

Al Salman, A., St Clair Gibson, A., & Micklewright, D. (2013). The Rating of Fatigue Scale (ROF). Medicine and Science in Sports and Exercise, 45(5, 1), 282-282.

Batson, A., Newnam, S., & Koppel, S. (2023). A preliminary study on the barriers and facilitators to improving the health, safety, and well-being of aging heavy vehicle drivers. *Journal of Safety Research, 86*, 262–273. https://doi.org/10.1016/j.jsr.2023.07.005

Belzer, M. H., & Sedo, S. A. (2017). Why do long distance truck drivers work extremely long hours? *The Economic and Labour Relations Review, 29*(1), 59–79. https://doi.org/10.1177/1035304617728440

Bener, A., Yildirim, E., Özkan, T., & Lajunen, T. (2017). Driver sleepiness, fatigue, careless behavior and risk of motor vehicle crash and injury: Population based case and control study. *Journal of Traffic and Transportation Engineering, 4*(5), 496–502. https://doi.org/10.1016/j.jtte.2017.07.005

Brazil. (2015). Lei N° 13.103, de 2 de março de 2015. Presidency of the Republic, General Secretariat, Brasilia.

Caldwell, J. A., Caldwell, L., Thompson, L. A., & Lieberman, H. R. (2019). Fatigue and its management in the workplace. *Neuroscience & Biobehavioral Reviews, 96*, 272–289. https://doi.org/10.1016/j.neubiorev.2018.10.024

Campos, C. I., Pitombo, C. S., Delhome, P., & Quintanilha, J. A. (2020). Comparative analysis of data reduction techniques for questionnaire validation using self-reported driver behaviors. *Journal of Safety Research, 73*, 133–142. https://doi.org/10.1016/j.jsr.2020.02.004

Cardoso, M., Fulton, F., Callaghan, J. P., Johnson, M., & Albert, W. J. (2019). A pre/post evaluation of fatigue, stress and vigilance amongst commercially licensed truck drivers performing a prolonged driving task. *International Journal of Occupational Safety and Ergonomics, 25*(3), 344–354. https://doi.org/10.1080/10803548.2018.1491666

Crizzle, A. M., Wawzonek, P. A., & Bigelow, P. L. (2024). Health Comparisons between Truck Drivers and the General Population Using the Canadian Community Health Survey. *Journal of Occupational and Environmental Medicine, 66*(1), 20–27. https://doi.org/10.1097/jom.0000000000002988

De Winter, J., Driessen, T., Dodou, D., & Cannoo, A. (2024). Exploring the challenges faced by Dutch truck drivers in the era of technological advancement. *Frontiers in Public Health, 12*. https://doi.org/10.3389/fpubh.2024.1352979

DMLC XGBoost. (2022). XGBoost Documentation. Available: https://xgboost.readthedocs.io/en/stable/. Accessed: 20 Dec 2022.

Filomeno, R., Ikeda, A., Maruyama, K., Wada, H., & Tanigawa, T. (2019). Excessive daytime sleepiness and alcohol consumption among commercial drivers. *Occupational Medicine, 69*(6), 406–411. https://doi.org/10.1093/occmed/kqz091

Fragoso Junior, A., & Garcia, E. G. (2019). Transporte rodoviário de carga: Acidentes de trabalho fatais e fiscalização trabalhista. *Revista Brasileira De Saúde Ocupacional, 44*. https://doi.org/10.1590/2317-6369000018317

Friswell, R., & Williamson, A. (2019). Management of heavy truck driver queuing and waiting for loading and unloading at road transport customers' depots. *Safety Science, 120*, 194–205. https://doi.org/10.1016/j.ssci.2019.06.039

Garbarino, S., Guglielmi, O., Sannita, W., Magnavita, N., & Lanteri, P. (2018). Sleep and mental health in truck drivers: Descriptive review of the current evidence and proposal of strategies for primary prevention. *International Journal of Environmental Research and Public Health, 15*(9), 1852. https://doi.org/10.3390/ijerph15091852

Girotto, E., Bortoletto, M. S. S., González, A. D., Mesas, A. E., Peixe, T. S., Guidoni, C. M., & de Andrade, S. M. (2019). Working conditions and sleepiness while driving among truck drivers. *Traffic Injury Prevention, 20*(5), 504–509. https://doi.org/10.1080/15389588.2019.1609670

Goel, A., & Vidal, T. (2014). Hours of service regulations in road freight transport: An optimization-based international assessment. *Transportation Science, 48*(3), 391–412. https://doi.org/10.1287/trsc.2013.0477

Gomes, A. C., De Lima, F. B., Junior, Soliani, R. D., De Souza Oliveira, P. R., De Oliveira, D. A., Siqueira, R. M., Da Silva Nora, L. a. R., & De Macêdo, J. J. S. (2023). Logistics management in e-commerce: challenges and opportunities. GeSec, 14(5), 7252–7272. https://doi.org/10.7769/gesec.v14i5.2119.

Gouveia, V. V., Oliveira, G. F., Mendes, L. A. C., Souza, L. E. C., Cavalcanti, T. M., & Melo, R. L. P. (2015). Fatigue assessment scale: Adaptation for health professionals. *Revista Psicologia Organizações e Trabalho, 15*, 3. https://doi.org/10.17652/rpot/2015.3.594.

Guglielmi, O., Magnavita, N., & Garbarino, S. (2018). Sleep quality, obstructive sleep apnea, and psychological distress in truck drivers: A cross-sectional study. *Social Psychiatry and Psychiatric Epidemiology, 53*, 531–536. https://doi.org/10.1007/s00127-017-1474-x

Heaton, K., Mumbower, R., & Childs, G. (2021). Sleep and safety decision-making among truck drivers. *Workplace Health & Safety, 69*(3), 134–140. https://doi.org/10.1177/2165079920950255

Hege, A., Lemke, M. K., Apostolopoulos, Y., & Sönmez, S. (2019). The impact of work organization, job stress, and sleep on the health behaviors and outcomes of US long-haul truck drivers. *Health Education & Behavior: The Official Publication of the Society for Public Health Education, 46*(4), 626–636. https://doi.org/10.1177/1090198119826232

Hege, A., Perko, M., Johnson, A., Yu, C. H., Sönmez, S., & Apostolopoulos, Y. (2015). Surveying the impact of work hours and schedules on commercial motor vehicle driver sleep. *Safety and Health at Work, 6*(2), 104–113. https://doi.org/10.1016/j.shaw.2015.02.001

Hesse, M. (2016). The City as a Terminal: The Urban Context of Logistics and Freight Transport. *Routledge*. https://doi.org/10.4324/9781315614748

Ipea. (2020). Custos dos Acidentes de Trânsito no Brasil: Estimativa Simplificada com Base na Atualização das Pesquisas do Ipea Sobre Custos de Acidentes nos Aglomerados Urbanos e Rodovias. Institute for Applied Economic Research (Ipea). Available: https://www.ipea.gov.br/atlasviolencia/arquivos/artigos/7018-td2565.pdf.

Jaffee, D., & Bensman, D. (2016). Draying and Picking: Precarious Work and Labor Action in the Logistics Sector. *WorkingUSA, 19*(1), 57–79. https://doi.org/10.1163/17434580-01901006

Jakobsen, M. D., Seeberg, K. G. V., Møller, M., Kines, P., Jørgensen, P., Malchow-Møller, L., Andersen, A. B., & Andersen, L. L. (2023). Influence of occupational risk factors for road traffic crashes among professional drivers: Systematic review. *Transport Reviews, 43*(3), 533–563. https://doi.org/10.1080/01441647.2022.2132314

Johns, M. W. (1991). A New Method for Measuring Daytime Sleepiness: The Epworth Sleepiness Scale. *Sleep, 14*(6), 540–545. https://doi.org/10.1093/sleep/14.6.540

Karasek, R., Brisson, C., Kawakami, N., Houtman, I., Bongers, P., & Amick, B. (1998). The Job Content Questionnaire (JCQ): An instrument for internationally comparative assessments of psychosocial job characteristics. *Journal of Occupational Health Psychology, 3*(4), 322–355. https://doi.org/10.1037//1076-8998.3.4.322

Léger, D., & Stepnowsky, C. (2020). The economic and societal burden of excessive daytime sleepiness in patients with obstructive sleep apnea. *Sleep Medicine Reviews, 51*(101275), Article 101275. https://doi.org/10.1016/j.smrv.2020.101275

Lemke, M. K., Hege, A., Apostolopoulos, Y., & Sönmez, S. (2021). Hours-of-service compliance and safety outcomes among long-haul truck drivers. *Transportation Research Part F: Traffic Psychology and Behaviour, 76*, 297–308. https://doi.org/10.1016/j.trf.2020.11.017

LightGBM.. (2022). LightGBM Classifier. *Available: https://lightgbm.readthedocs.io/en/latest/index.html. Acessed, 19*. Dec 2022.

Mahajan, K., Velaga, N. R., Kumar, A., Choudhary, A., & Choudhary, P. (2019). Effects of driver work-rest patterns, lifestyle and payment incentives on long-haul truck driver sleepiness. Transportation Research. *Part F, Traffic Psychology and Behaviour, 60*, 366–382. https://doi.org/10.1016/j.trf.2018.10.028

Makuto, N., Kristman, V., Bigelow, P., & Bédard, M. (2023). Factors associated with depressive symptoms in long-haul truck drivers. *Transportation Research Interdisciplinary Perspectives, 21*, Article 100851. https://doi.org/10.1016/j.trip.2023.100851

Mello, M. T., Narciso, F. V., Tufik, S., Paiva, T., Spence, D. W., BaHammam, A. S., Verster, J. C., & Pandi-Perumal, S. R. (2013). Sleep disorders as a cause of motor vehicle collisions. *International Journal of Preventive Medicine, 4*(3), 246.

Messias, J. C., Cavieres-Higuera, H., Silva, R. A., de Facundos., G. N., & Lessa, R. T. (2019). Being a truck driver in Brazil: From implicit self-stereotypes to system justification. *Estudos de Psicologia (Campinas), 36*. https://doi.org/10.1590/1982-0275201936e180139

Micklewright, D., Gibson A., S. C., Gladwell, V., & Al Salman, A. (2017). Development and validity of the rating-of-fatigue scale. *Sports Medicine (Auckland, N.Z.), 47*(11), 2375–2393. https://doi.org/10.1007/s40279-017-0711-5

Mizuno, K., Ojiro, D., Tanaka, T., Minusa, S., Kuriyama, H., Yamano, E., Kuratsune, H., & Watanabe, Y. (2020). Relationship between truck driver fatigue and rear-end collision risk. *PloS One, 15*(9), e0238738.

Narciso, F. V., & de Mello, M. T. (2017). Safety and health of professional drivers who drive on Brazilian highways. *Revista de Saúde Pública, 51*. https://doi.org/10.1590/s1518-8787.2017051006761

Naser, M. Z., & Alavi, A. H. (2023). Error metrics and performance fitness indicators for artificial intelligence and machine learning in engineering and sciences. *Architecture, Structures and Construction, 3*(4), 499–517. https://doi.org/10.1007/s44150-021-00015-8

Nazari, S. S. H., Moradi, A., & Rahmani, K. (2017). A systematic review of the effect of various interventions on reducing fatigue and sleepiness while driving. *Chinese Journal of Traumatology, 20*(5), 249–258. https://doi.org/10.1016/j.cjtee.2017.03.005

Passey, D. G., Robbins, R., Hegmann, K. T., Ott, U., Thiese, M., Garg, A., Kinney, A., Murtaugh, A., & M.. (2014). Long haul truck drivers' views on the barriers and facilitators to healthy eating and physical activity: A qualitative study. *International Journal of Workplace Health Management, 7*(2), 121–135. https://doi.org/10.1108/IJWHM-08-2013-0031

Philip, P., Taillard, J., & Micoulaud-Franchi, J. A. (2019). Sleep restriction, sleep hygiene, and driving safety. *Sleep Medicine Clinics, 14*(4), 407–412. https://doi.org/10.1016/j.jsmc.2019.07.002

Pourabdian, S., Lotfi, S., Yazdanirad, S., Golshiri, P., & Hassanzadeh, A. (2020). Evaluation of the effect of fatigue on the coping behavior of international truck drivers. BMC. *Psychology, 8*(1). https://doi.org/10.1186/s40359-020-00440-2

Rahman, M. A., Das, S., & Sun, X. (2023). Understanding the drowsy driving crash patterns from correspondence regression analysis. *Journal of Safety Research, 84*, 167–181. https://doi.org/10.1016/j.jsr.2022.10.017

Rocha, F. P., Fischer, F. M., & Moreno, C. R. C. (2018). Organization of truck drivers' work: Need for an intersectoral policy. *Revista Brasileira de Medicina do Trabalho, 16*, 2. https://doi.org/10.5327/Z1679443520180238

Rodriguez, J. D., Perez, A., & Lozano, J. A. (2010). Sensitivity analysis of k-fold cross validation in prediction error estimation. *IEEE transactions on pattern analysis and machine intelligence, 32*(3), 569–575. https://doi.org/10.1109/tpami.2009.187

Roman, G. D., Poulter, D., Barker, E., McKenna, F. P., & Rowe, R. (2015). Novice drivers' individual trajectories of driver behavior over the first three years of driving. *Accident; Analysis and Prevention, 82*, 61–69. https://doi.org/10.1016/j.aap.2015.05.012

Rosekind, M. R., Michael, J. P., Dorey-Stein, Z. L., & Watson, N. F. (2024). Awake at the Wheel: How auto technology innovations present ongoing sleep challenges and new safety opportunities. *Sleep, 47*(2). https://doi.org/10.1093/sleep/zsad316

Ryley, W. T., & Belzer, M. H. (2023). Compensation and crash incidence: Evidence from the National Survey of Driver Wages. *The œEconomic and Labour Relations Review, 34* (1), 118–139. https://doi.org/10.1017/elr.2022.13

Sadeghniiat-Haghighi, K., & Yazdi, Z. (2015). Fatigue management in the workplace. *Industrial Psychiatry Journal, 24*(1), 12. https://doi.org/10.4103/0972-6748.160915

Scikit-learn. (2022). Scikit-learn: Machine Learning in Python. Available: https://scikit-learn.org/stable/. Acessed: 6 Jan 2023.

Seuring, S., Yawar, S. A., Land, A., Khalid, R. U., & Sauer, P. C. (2021). The application of theory in literature reviews – illustrated with examples from supply chain management. *International Journal of Operations & Production Management, 41*(1), 1–20. https://doi.org/10.1108/IJOPM-04-2020-0247

Shaw, K. A., Zello, G. A., & Crizzle, A. M. (2023). Body mass index and waist circumference correlates with lifestyle and health in Long-Haul truck drivers. *Journal of Occupational and Environmental Medicine, 65*(12), 1051–1057. https://doi.org/10.1097/jom.0000000000002968

Simonelli, G., Bellone, G., Golombek, D., Pérez Chada, D., Glozier, N., Capaldi, V. F., Vigo, D. E., & Kryger, M. H. (2018). Hours of service regulations for professional drivers in continental Latin America. *Sleep Health, 4*(5), 472–475. https://doi.org/10.1016/j.sleh.2018.07.009

Sinagawa, D. M., Carvalho, H. B. D., Andreuccetti, G., Prado, N. V. D., Oliveira, K. C. B. G. D., Yonamine, M., Muñoz, D. R., Gjerde, H., & Leyton, V. (2015). Association between travel length and drug use among Brazilian truck drivers. *Traffic Injury Prevention, 16*(1), 5–9. https://doi.org/10.1080/15389588.2014.906589

Soliani, R. D. (2022). The concept of sustainability in the Brazilian road freight transportation sector. In: Sustainable Development Dimensions and Urban Agglomeration. IntechOpen. https://doi.org/10.5772/intechopen.97280.

Soliani, R. D., & Da Silva, L. B. (2023). Segurança e Saúde Ocupacional dos Motoristas de Caminhão Autônomos: Uma Revisão da Literatura. *Revista Transporte Y Territorio, 28*, 200–213. https://doi.org/10.34096/rtt.i28.10953.

Thiese, M. S., Hanowski, R. J., Moffitt, G., Kales, S. N., Porter, R. J., Ronna, B., Hartenbaum, N., & Hegmann, K. T. (2018). A retrospective analysis of cardiometabolic health in a large cohort of truck drivers compared to the American working population. *American Journal of Industrial Medicine, 61*(2), 103–110. https://doi.org/10.1002/ajim.22795

Tseng, B. Y., Gajewski, B. J., & Kluding, P. M. (2010). Reliability, responsiveness, and validity of the Visual Analog Fatigue Scale to measure exertion fatigue in people with chronic stroke: A preliminary study. *Stroke Research and Treatment, 2010*, 1–7. https://doi.org/10.4061/2010/412964

Useche, S. A., Cendales, B., Montoro, L., & Esteban, C. (2018). Work stress and health problems of professional drivers: A hazardous formula for their safety outcomes. *PeerJ, 6*(e6249), e6249.

Ward, K. L., Hillman, D. R., James, A., Bremner, A. P., Simpson, L., Cooper, M. N., Palmer, L. J., Fedson, A. C., & Mukherjee, S. (2013). Excessive daytime sleepiness increases the risk of motor vehicle crash in obstructive sleep apnea. *Journal of Clinical Sleep Medicine: JCSM: Official Publication of the American Academy of Sleep Medicine, 09*(10), 1013–1021. https://doi.org/10.5664/jcsm.3072

Wijngaards, I., Hendriks, M., & Burger, M. J. (2019). Steering towards happiness: An experience sampling study on the determinants of happiness of truck drivers. *Transportation Research Part A: Policy and Practice, 128*, 131–148. https://doi.org/10.1016/j.tra.2019.07.017

Wise, J. M., Heaton, K., & Patrician, P. (2018). Fatigue in Long-Haul Truck Drivers: A concept analysis. *Workplace Health & Safety, 67*(2), 68–77. https://doi.org/10.1177/2165079918800509

Xu, Y., & Goodacre, R. (2018). On Splitting Training and Validation Set: A Comparative Study of Cross-Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Supervised Learning. *Journal of Analysis and Testing, 3*, 249–262. https://doi.org/10.1007/s41664-018-0068-2

Zhang, C., Kotz, A., Kelly, K., & Rippelmeyer, L. (2021). Development of heavy-duty vehicle representative driving cycles via decision tree regression. *Transportation Research. Part D, Transport and Environment, 95*(102843), Article 102843. https://doi.org/10.1016/j.trd.2021.102843

Zhang, C., Ma, Y., Chen, S., Zhang, J., & Xing, G. (2024). Exploring the occupational fatigue risk of short-haul truck drivers: Effects of sleep pattern, driving task, and time-on-task on driving behavior and eye-motion metrics. Transportation Research. *Part F, Traffic Psychology and Behaviour, 100*, 37–56. https://doi.org/10.1016/j.trf.2023.11.012

Zhang, T., & Chan, A. H. S. (2014). Sleepiness and the risk of road accidents for professional drivers: A systematic review and meta-analysis of retrospective studies. *Safety Science, 70*, 180–188. https://doi.org/10.1016/j.ssci.2014.05.022

Rodrigo Duarte Soliani is with the Federal Institute of Acre, in Brazil.

Alisson Vinicius Brito Lopes is affiliated with the Federal Institute of Acre, in Brazil.

Fábio Santiago is with the Aeronautics Institute of Technology, in Brazil.

Luiz Bueno da Silva is with the Federal University of Paraíba, in Brazil.

Nwabueze Emekwuru is with the Engineering Department at Harper Adams University, in the United Kingdom.

Ana Carolina Lorena is with the Aeronautics Institute of Technology, in Brazil.