

Research Article

The Role of Journey Purpose in Road Traffic Injuries: A Bayesian Network Approach

Juan Diego Febres,¹ Fatemeh Mohamadi,² Miguel. A. Mariscal ,³ S. Herrera,⁴
and Susana García-Herrero ^{2,3}

¹Department of Chemistry and Exact Sciences, Universidad Técnica Particular de Loja, P.O. Box 11 01 608, Loja, Ecuador

²School of Architecture and Cities, University of Westminster, London, UK

³Departamento de Ingeniería Civil, Universidad de Burgos, Avda. Cantabria s/n, Burgos 09006, Spain

⁴Department of Applied Mathematics & Comp Science, Universidad de Cantabria, Santander, Spain

Correspondence should be addressed to Susana García-Herrero; susanagh@ubu.es

Received 23 May 2019; Accepted 31 July 2019; Published 2 December 2019

Academic Editor: Abdelaziz Bensrhair

Copyright © 2019 Juan Diego Febres et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction. Road traffic injuries are now regarded as the eighth leading cause of death globally. For example, in 2016, 102,362 traffic injuries took place in Spain in which 174,679 drivers suffered injuries. These findings necessitated the development of the current study which focuses on the prime factors that cause this type of injuries. The aim of this study, therefore, is to explore the behavioral factors that entail a higher risk of suffering either a serious or a fatal injury for drivers. **Methods.** The findings are based on information and data provided by “Dirección General de Tráfico” (DGT) in Spain on traffic injuries that occurred in the year 2016. Reviewing a wide range of the literature, the authors identified the most influential variables and created a model using the Bayesian networks. The variables that define the model are grouped into four factors: vehicle factor, road factor, circumstantial factor and human factor. **Results.** The results suggest that the principal variables that determine a higher probability of serious or fatal injuries in traffic injuries are: lack of using appropriate safety accessories, high-speed violations, distractions as well as errors. Finally, the research shows the severity probability based on reason of displacement (“in itinere,” on business, or in leisure).

1. Introduction

Road traffic injuries are one of the main causes of death in the world [1, 2]. Every year 1.24 million people die on the world’s roads and between 20 and 50 million people are injured, making road traffic accidents the eighth cause of death globally [3].

Studies argue that around 10% of road traffic injuries take place when the driver is traveling in the course of work; while a further 18% of injuries take place while a driver is traveling to or from work, i.e., commuting [4]. A bus driver injured in the course of driving for work would be seen as such. A cross Europe, it is estimated that 40% of traffic injuries happen during commute or business journeys [5]. In Spain, in this regard, in 2016 there were 566,235 injuries associated with traveling to/from work in which 64,737 cases were road traffic injuries

occurring during business or work travel, accounting for more than 11.4% of the total [6].

In this study the journey purpose is classified into three groups: “in itinere” refers to commute—journeys or travel from home to work and vice versa, on business refers to when a driver travels for work-related purposes, and in leisure refers to when a driver travels for pleasure.

The variables that influence the occurrence of a traffic injury can be divided and defined in four groups: demographic factor, human factor, vehicle factor, and circumstantial factor. Nevertheless, the focus of the current research is mainly on the human factor. Human factor has been considered the main cause of traffic injuries as highlighted by Mazankova [7].

Sabey and Taylor [8] suggest that the behavior that the driver adopts in front of the steering wheel has become an important factor in the principal cause of promoting traffic

injuries. For that reason, several theories have been developed in which they have explained possible risk behaviors behind the steering wheel. One of these theories is denominated “the zero-risk theory” which discusses the existence of a risk threshold above which the danger is not perceived [9]. This theory considers that the reasons and emotions play an important role in the driver behavior.

If we extrapolate “the zero-risk theory” as suggested by Salminen and Lähdeniemi [10] to the traffic that occurs during the workday, some of these reasons could be argued as the time pressure, work pressure as well as excessive workload and tiredness.

The time saving is argued as a main reason by Summala [9], which can trigger an increase in speed (assuming higher risk) to meet the objective (arrive early). Tiredness behind the wheel is one of the risk factors that has been highlighted by Bener, Yildirim [11]. Driving throughout long periods of time without rest phases makes driving a monotonous task, reducing the ability of driver to drive safely until dangerous limits [12]. Kim and Chung [13] explain the role of job satisfaction in relation to the number of traffic accidents, and Wishart, Somoray [14] suggest strategies should be developed in order to encourage positive work driving safety climate at work. Finally, issues associated with family or work conflicts which in the majority of the cases result in biological imbalances also often trigger a reduction in resting hours as well as drowsiness and subsequently add to risk factors [10, 15–19].

Another relevant variable presented in different campaigns of the Dirección General de Tráfico is the lack of using appropriate safety accessories. Many authors consider the lack of using helmet or seat belt as the main risk factors in work-related injuries [20, 21].

Several authors have taken into account the gender, age, labor sector, and economic remuneration received in order to identify which population groups are most propense to suffering a traffic injury [15, 22].

Regarding gender, studies argue that male gender is more involved in injuries than the female gender. The main reason for such conclusion is that sectors with higher frequency indices are the transport and distribution sectors that are generally run by men. However, these studies highlight that women suffer more work-related traffic injuries during their displacement than men [23].

Studies conducted in relation to age factor demonstrate that young drivers overestimate their driving abilities, using risk maneuvers [25]. The aging process involves the biological and psychological system deterioration, and it is considered that it starts around 45–50 year old. From the point of view of driving, this loss is focused on the sense of sight, slowing down the speed of perception and response to stimuli and the reduction of muscle strength [26, 27].

To this end, we can conclude that speed is one of the most influential behaviors of the driver that causes fatal injuries [28, 29]. Little increases in speed highly increment the risk of an injury and the severity of the injury [30]. An increased speed means a greater kinetic energy; therefore, in the case of an impact, this energy is absorbed by the vehicle, its passengers, and the element against which it interacts, encoring the number and the severity of injuries. A driver traveling at a high

speed lengthens the reaction distance, defined as the distance traveled by the vehicle before the driver reacts to a danger. The pressure of arriving to work on time can cause some reckless and careless manners of drivers such as reaching high speed, which could result in more injury-prone in the roads [15, 21, 31, 32].

To conclude, the key point of this study is to establish a probabilistic model based on Bayesian networks. Such analysis was conducted in order to predict the risk of suffering an injury in function of displacement reasons: whether “in itinere”, on business, or in leisure trips and others. The model narrows down its focus on four groups of factors including demographic factors, vehicle factors, circumstantial factors, and human factors. Thus, the model determines those drivers’ behaviors that entail a greater risk of suffering an injury. Therefore, research directly focusing on a systematic relationship between the journey purpose and harmfulness of drivers while taking into account these four groups of factors in road traffic injuries in a Spanish context remains limited in the field. To this end, the justification behind conducting this research was to address this gap in the field and aims to add to the existing knowledge as well as the literature around the topic.

2. Materials and Methods

2.1. Data Base Acquisition. The data base used to develop this study has been provided by the Dirección General de Tráfico (DGT). Institution on charge to register the traffic injuries in Spain.

In Spain, when a traffic accident occurs, the agents of the authority in charge from the surveillance and control of traffic, within the scope of their respective competences, send the information related to traffic accidents to the National Registry of Victims of Traffic Accidents. This information includes the information concerning the traffic accidents with victims, and through the form, this information is included in the annex of the official document BOE-A-2014-12411 [33]. The micro-data set used in this study has three tables: general table, vehicles table, and drivers table, which gather information about the traffic injuries that happened in 2016. In that year, 102,362 injuries took place in which 172,971 drivers were implicated [34]. This research specifically focuses on those drivers whom harmfulness is known, and at the same time the study focuses on the type of their known displacement. The degree of severity of such drivers has been defined as: fatal (FI), seriously injured (SI), lightly injured (LI), and unhurt (U). These drivers are registered by traffic police as drivers who were taking a journey either to go to work from home or vice versa to home from work. They also could be registered as drivers who were driving for work purposes or driving was their job. Finally, they could be registered as traveling for leisure and pleasure purposes. Taking this harmfulness of the driver and cause of displacement aspects into account, the final dataset includes a total of 66,253 drivers.

To this end, the sampling technique employed in this study is a systematic sampling method. The authors have excluded the data for traffic accidents in 2016 in which the purpose of

the journey and the driver harmfulness were not reported by (DGT). Utilizing data from the sample population collected by (DGT) and employing a Bayesian network, the current study focuses on four relevant variables and discusses results in which the study highlights the importance of relationship between drivers' behaviors in road traffic injuries with the level of drivers' harmfulness.

2.2. Study Variables. The variables that contribute to the occurrence of a traffic injury and result in driver harmfulness can be assembled into four groups: demographic factors, vehicle factors, circumstantial factors, and human factors.

Each of these factors in turn includes a series of variables, with their corresponding states.

- (i) Demographic factors: combination of the gender and the age of the driver.
- (ii) Vehicle factors: type of vehicles.
- (iii) Circumstantial factors: type of trips or reasons for displacements, type of roads or zones and distance or kilometers of travel.
- (iv) Human factors: the behavioral factors, or modifiable factors by the driver. These could include wearing a seat belt, wearing a helmet, the speed violation as well as distraction and errors made by the driver.
- (v) Study variable: driver harmfulness represents driver injury severity.

This study focuses mainly on the human factor, being considered as the principal cause of traffic injuries (between 70% and 90%) [35].

2.3. Bayesian Network. In order to characterize the dependences between the different factors and the target variable, the probabilistic graphical models (PGMs) have been considered. Several studies have previously employed Bayesian network in their analysis of traffic accidents to express certain relationships between the different factors [36–39]. These models are based on a graph in which each node represents a variable or factor and each link between variables represents a dependence between them. These dependences/independences let us to factorize the joint probability distribution (JPD), which is the second element of these models, dramatically reducing the number of parameters of our model and, as a result, simplifying the learning and inference processes. In addition, the graph obtained is a visual and easily interpretable tool to illustrate the factors affecting our target variable. In particular, in our study, we have considered the discrete Bayesian networks [40] in which the graph of the model is a directed acyclic graph (DAG). The link's direction introduces two additional concepts in the nodes of our model, parents and children, depending on whether the arrow departs or points to the node, respectively. As a result, the JPD can be expressed mathematically as

$$p(x_1, x_2, \dots, x_n) = \prod_{i=1}^n p(x_i | \pi_i), \quad (1)$$

where π_i corresponds to the parents of x_i , being the BN the model defined by both the DAG and the corresponding JPD in Equation (1).

Once the Bayesian network has been defined, the probability of any node or set of nodes given any information on the state of the others variables (evidence) can be efficiently obtained by using both the factorization and the DAG (inference), letting us to analyze the impact of each of the variables in the injury severity grade suffered by the driver. As an example, we could have some evidence about the motive of the displacement, the age, and the gender with which we can determine the probability of a serious injury in the accident by means of the expression:

$$p(\text{harmfulness} | \text{Type of Trip} \cap \text{Age} \cap \text{Gender}) = \frac{p(\text{harmfulness intersect Type of Trip} \cap \text{Age} \cap \text{Gender})}{p(\text{Type of Trip} \cap \text{Age} \cap \text{Gender})}. \quad (2)$$

Moreover, from the definition of the Bayesian network, a natural classifier for the injury severity can be obtained defining a threshold for the probability above/below of which serious/no injury is assigned. To evaluate this classifier, the receiver operating characteristic (ROC) curve was considered. This technique was introduced in the clinical investigation by two radiologists which allow us to represent the true positives (sensitivity) based on false positives (specificity) [41]. The area enclosed under the curve (AUC) allows to evaluate the model. This area can take values between 0 (perfect predictor of the contrary state) and 1 (perfect predictor), corresponding the 0.5 value to a random prediction (unreliable model).

3. Results and Discussion

3.1. Theory Model. The proposed model can be appreciated in Figure 1.

Below is the list of the factors and the interactive variables with their definitions that contributed to the development of our model (please see Figure 2. Theory model).

The factor vehicle refers to the type of vehicle variable that has been discretized in six groups: cars, bikes, motorcycles, buses and coaches, trucks and others.

The demographic factor included two types of variables: age and gender. The variable "gender" remains the same as mentioned in the questionnaire. However, the variable "age" has been grouped into four groups: less than 18, 18 to 24, 25 to 60, and over 60.

The human factor has been grouped in five types of variables: seat belt, helmet, speed, distraction, and error. The variables seat belt and helmet indicate if the driver was using such safety accessories in the moment of the accident. The speed variable has the same four states as shown in the questionnaire; the first state is "none" and indicates that the driver was driving in the correct speed, the second group indicates if the speed was inadequate, the third state shows when the driver was driving over the limit speed allowed, and the fourth state indicates if driver was driving the vehicle too slow—below the standards. Finally, the group of the variables' errors and

distraction indicates that the driver did not make any error or distraction. On the other hand, the state “yes” indicates the contrary. All the errors and distractions included in the analysis are shown in the section of comments in Table 1.

The circumstantial factor has been grouped into two types of variables: zone and type of trip. The variable type of road or zone remains in the same four types as in the questionnaire filled by the police (road, crossing, street, and highway). The variable type of trip shows the cause of displacement in three groups including “in itinere,” on business and in leisure. Another variable taken into consideration is the distance that the driver undertakes. The variable shows the same three groups with the answers that drivers answered the transport policemen. The distances are categorized as: local (less than 50 km), medium (between 50 km and 200 km), and long distance (more than 200 km).

Lastly, the objective variable, is the object of the study, injury severity, has been created considering the severity of the driver’s injury. This variable has two values: firstly “light” if the driver was slightly injured, and secondly kill serious injuries “KSI” if the driver was either fatally injured or seriously injured. It would be worth mentioning here that this study focuses merely on injury severity for the driver himself or herself.

3.2. Validation. A k -fold cross-validation approach, with $k = 10$, was considered to evaluate the model. This method divides the data into 10 folds including the 10% of the sample (i.e., ~6625 data for each fold). For each fold the other 90% of the sample (~59628 data) is used as training data to predict the sample included in the corresponding fold, used as test data. This procedure was performed ten times in order to ensure that all data with no exceptions were calculated since it has been part of the training and testing analysis. The area under the curve indicates the ability to determine the probability of suffering whether it was a major, a fatal, a minor, or an unharmed injury. In this case, the AUC is in a range of (0.767–0.801).

3.3. Initial Probabilities of Serious Injury in a Traffic Injury. A sensitivity analysis of each of the variables carried out in this study is to determine the initial probability of death or major injury for the drivers versus minor injury in each of its states. A sensitivity analysis of each of the variables is also carried out to determine the initial probability of death or serious injury (KSI risk) for the drivers versus slight injury in each of its states. The results are shown in Table 1.

After carrying out the sensitivity analyses, showing the initial probabilities, we can argue that the most influential variables are respectively as follows: the type of vehicle, distance, age, seat belt, and, finally, speed.

It is important to take into account the interrelation that may exist between the different variables. Therefore, the Bayesian network presents their strong point for their ability to extract knowledge through the search of the joint probabilities of all the variables among themselves.

The factors that contribute to the severity of accidents are related to each other and do not act on merely. As a result, the accident occurs with complex interactions between road user

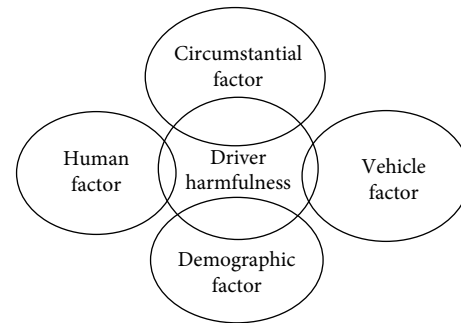


FIGURE 1: Principal risk factors.

behavior, vehicle factors, road geometric characteristics, and environmental factors [42].

This is important; however, the result of the analysis conducted in this study employing the Bayesian network (presented in Figure 1) gives us information on how all these variables are interrelated with each other (Figure 3).

3.4. Probability of Serious Traffic Accident Based on the Cause of Displacement and the Behavior of the Driver. To analyze the influence of displacement reason in the harmfulness in the accident, a sensitivity analysis has been done to establish driver’s harmfulness probability in function of two evidences (see Table 2). The first evidence, in all analysis carried out, is always the type of trip, and the second evidence is in relation to one of these human behavior variables: seat belt, helmet, speed, distraction, and errors.

Not wearing safety accessories including seat belt and helmet results in serious accidents. That reaches levels of 19.9% and 13.0%, respectively. Focusing on the type of displacement, not wearing a seat belt, “in itinere,” on business, and leisure, the figures are shown as 19.1%, 17.4%, and 19.3%, respectively. A cross-tabulation test illustrated in Table 1 examined the relationship between speed variables and type of trip. The test was statistically significant and illustrated 0.202, the worst figure, suggesting that there is a highly significant relationship between “exceeding speed” and “travel for leisure purpose” for serious injuries. Specifically, if we keep our focus on the type of trip, exceeding speed on pleasure trips is the factor that mostly determines the probability of suffering a serious and/or fatal accident. These probabilities account for 17.5% in “in itinere” trips, 15.9% in on business trips, and 20.2% in leisure trips.

The possible distractions made by the drivers are using mobile phones, focusing on GPS devices, being distracted by radio and music, smoking while driving, and some other types of distractions. On the other hand, the errors made by the drivers are related to making mistakes in terms of not paying enough attention to traffic signs, to other vehicles, to the pedestrians, and so on. The results show the probability of death or major injury in an accident based on these variables and the reasons for displacement.

Distractions and errors, provoke very similar probabilities of suffering a serious and/or a fatal accident. This fact was also verified in several studies as mentioned by Cordazzo, Scialfa [43]. However, if we analyze the probability of distraction and

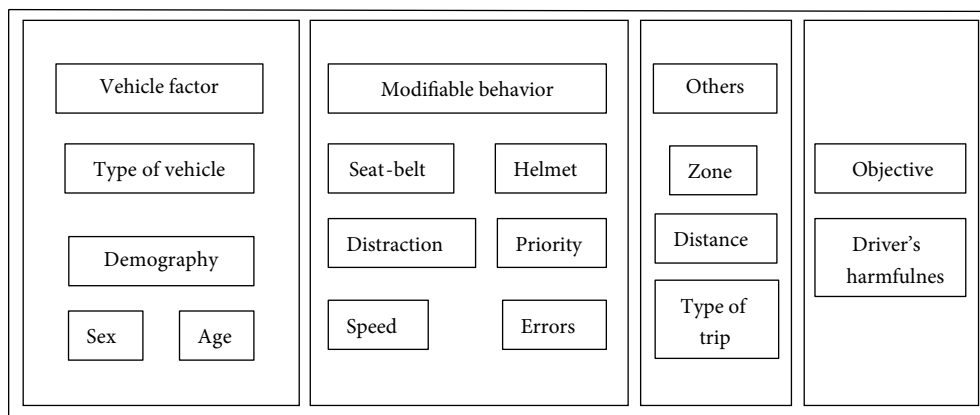


FIGURE 2: Theory model.

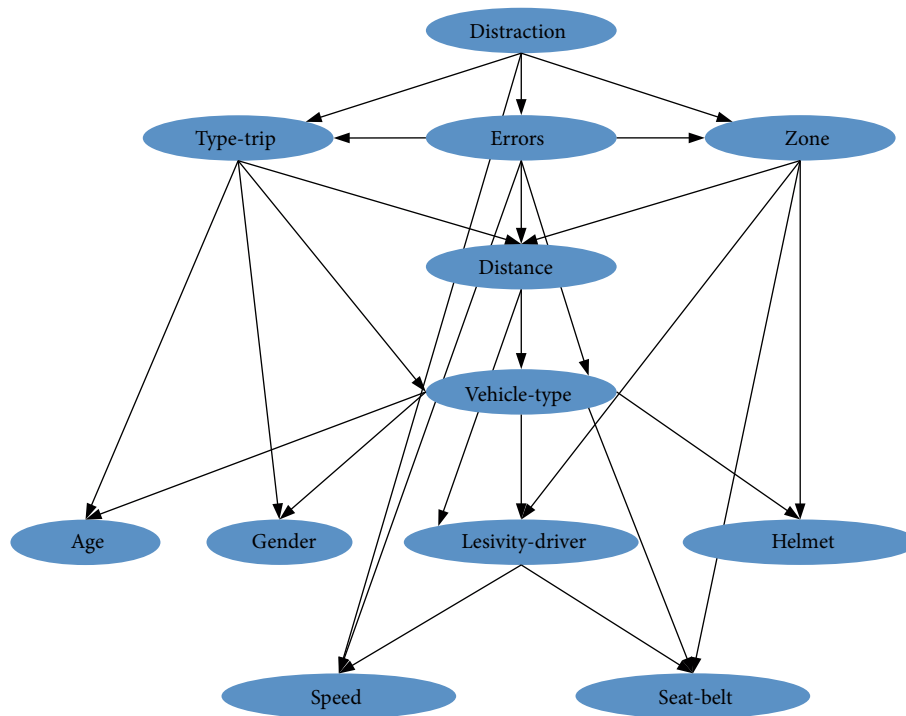


FIGURE 3: Bayesian network. Processed by the authors.

error depending on the type of trip, the results suggest that the highest chances of suffering a serious and/or a fatal accident occur in leisure trips (7.5% and 6.7%). This might be due to the fact that drivers on leisure trips may drive on the routes that are less familiar with and they are likely to be more distracted by other factors such as talking to their fellow travelers.

Table 3 shows the relative probability of having a distraction or errors in function of the purpose of displacement.

Depending on the type of trips taken by drivers, different probabilities of getting distracted or making mistakes during the trip are shown in Table 3. As the table illustrates, the probability to make a mistake due to the errors is higher than

mistakes due to distractions. This is quite evident in leisure trips, reaching a probability of 42.4% of making an error. Likewise, the probability to make a mistake due to distraction is bigger in leisure trips reaching a probability of 17.8%. This suggests that the reasons highlighted earlier resonate well with these findings.

On the other hand, business trip is less probable to result in high-risk accidents. This is in keeping with the work of de Oliveira, Petroianu [44] who explain how recklessness while driving a motorcycle could be argued as the main cause of traffic accidents. In their study only 7% of displacement with motorcycles was for work. Our analysis shows similar results as motorcycles are used on business with a probability of 8.2%.

TABLE 1: Number of cases and a priori probability of the severity of the injury based on behavior variables. Processed by the authors.

Variables	No. cases	KSI risk	Comments
<i>Type of vehicle</i>			
Car	47928	0.038	Car, van, all-terrain
Bike	3203	0.146	Bikes
Motorcycle	9617	0.213	Moped, motorcycle <125 cc, motorcycle >125 cc
Bus	939	0.025	Minibus (up to 17 passengers), bus, articulated bus
Truck	3994	0.047	Rigid truck, truck, articulated, articulated vehicle
Others	572	0.120	
<i>Zone</i>			
Road	39235	0.090	Road, secondary road
Crossing	701	0.102	Crossing through roads
Urban area	26166	0.039	Street
Highway	116	0.083	Highway/motorway
<i>Type of trip</i>			
“In itinere”	13406	0.065	“in itinere,” student to student center, transporter minors to school
On bussiness	7893	0.058	Like drive any bus, taxi, transport goods, others
Leisure	28855	0.073	Leisure and entertainment, particular sport activities
Others	16099	0.075	
<i>Distance</i>			
Local	42138	0.050	Less than 50 km
Medium	8269	0.082	Between 50 and 200 km
Long	3206	0.077	More than 200 km
Unknown	12640	0.127	
<i>Age</i>			
<18	904	0.181	
18–24	8124	0.073	
25–60	49027	0.069	
>60	7920	0.062	
Unknown	278	0.074	
<i>Gender</i>			
Men	49109	0.076	
Women	17082	0.052	
Unknown	62	0.087	
<i>Seat belt</i>			
Yes	44626	0.038	Seat belt fastened
No	6584	0.199	Seat belt not fastened
Unknown	15043	0.106	
<i>Helmet</i>			
Yes	11123	0.204	Wearing a helmet or it was supposedly expelled
No	964	0.130	Not wearing a helmet
Unknown	54166	0.041	
Variables	No. cases	K/SI risk	Comments in English
<i>Speed</i>			
No infraction	44338	0.048	Adequate speed
Inadequate	6196	0.120	Inadequate speed for road conditions
Exceeding	693	0.190	Exceeding the established speed
Slow	49	0.068	Slow march / hindering circulation
Unknown	14977	0.107	
<i>Distraction</i>			
No	27474	0.070	No factor is appreciated
Yes	5205	0.069	Use of mobile phone, use of hand-free devices, use of GPS devices, radio or music on, watching DVD, or video device, wearing headphones, smoking, simultaneous driving activities (eating, drinking, finding objects...), interacting with other occupants, distracted by a previous accident, looking at the environment (landscape, advertising, signs...), being lost in thought or absent minded, sleep, fatigue, sudden illness, indisposition.

TABLE 1: Continued.

Variables	No. cases	KSI risk	Comments
Unknown	33574	0.070	
<i>Errors</i>			
No	29773	0.069	There are no errors
Yes	20038	0.060	Failing to see a road sign, failing to see a vehicle/pedestrian/obstacle, not understanding a road sign or confusing it, hesitation or delay in making a decision, incorrect execution of a maneuver or inadequate maneuver, forgetting to signalise (with the vehicle indicators or lights...)
Unknown	16438	0.084	

3.5. Probability of Serious Traffic Accident Based on the Cause of Displacement and the Type of Vehicle, Zone and Distance. To analyze the probability of KSI risk, a sensitivity analysis has been conducted to establish the probability in function of two evidences (see Table 4). In our study, based on our findings, we argue that where a car driver experiences an injury, in 3.8% of cases, the injury is whether serious or mortal. The risk is somewhat higher for truck drivers, at 4.7%. The other two vehicular modes with elevated KSI risk are, respectively, cycles, at 14.6%, and motorcycles, at 21.3%. However, focusing on the type of trips, it is important to emphasize that 14.8% of the serious and/or fatal accidents occur in displacement for pleasure when the vehicle used is a bicycle. Concerning motorcycles, the “in itinere,” on business, and leisure displacements, the figures show higher probability of suffering a serious and/or a fatal accident demonstrating 20.1%, 21.6%, and 20.2%, respectively.

In general, the risks of suffering a serious and/or a fatal accident for road users are less harmful when they travel on urban areas as Olszewski, Szagala [45] mentioned in their study. Table 4 confirms the figure 3.9% of having a serious accident on street, especially on business displacements, which reaches 2.7%. In contrast to this, the analysis shown in Table 4 confirms that the highest probability of suffering a serious and/or a fatal accident occurs on the motorways with 10% in leisure journeys.

The last variable in Table 4 is the distance. It can be noted that among the local displacements, medium and long range, are mid-range displacements that cause more risk to drivers (8.2%). Within these medium-range trips, pleasure trips are the most dangerous trips, reaching the figure 10.3%, in comparison with local displacement for business travels that shows the figure 4.2%.

Leisure trips, as presented in the table, encompass the higher risk of suffering a serious and/or a fatal accident in trips in comparison with the others. These results are consistent with the findings by Bellos et al. (2019). In their article, Bellos et al. explain that the risk of suffering accidents in general is increased with the tourists who drive during holiday periods, those who are obviously doing leisure trips. The article highlights that this may be due to the increase in vehicles during the tourist season and also because tourists do not know the city nor its traffic regulations or signage [46].

The data presented in Table 4 demonstrates an elevated severity risk for road users, those who are involved in leisure-related “in itinere” journeys. As the table indicates, the severity risk for road user leisure travelers is at 9.5%, indicating a higher frequency than the other trip purposes.

3.6. Probability of a Serious Traffic Accident Based on the Reason for Displacement, Age and Gender. A sensitivity analysis examined the age and gender of the driver in relation to the cause for displacement. The analysis presents the probability of having a serious or fatal accident. Table 5 illustrates these analyses.

As in previous studies, the results confirm how men are more likely to have more serious accidents on the road [47]. On the other hand, focusing on the reason for the displacement, the test revealed that there is a big difference in KSI risk in function of the age and gender in relation to the cause of displacement. First, as the figures show man drivers over 25 year old have the highest probabilities of having a serious accident in leisure trips (8.1% in leisure in comparison with 5.5% for business). Looking at the table, however, it becomes apparent that, young drivers (less than 18 year old), regardless of their gender whether a woman or man, reach the highest probabilities on business trips (20.1% for men and 18.7% for women). According to Korpinen and Paakkonen [48], younger people tend to have more accidents while on their mobile phones (distraction, in our study).

4. Conclusions

In the year 2016, in Spain, 177,356 vehicles and 172,972 drivers were involved in traffic accidents resulting in 102,362 traffic injuries. The focus of this study was, therefore, on drivers who were injured in traffic accidents, and the focus shifted on their harmfulness in relation to the type of trips they were undertaken. Therefore, the dataset for this study includes a total number of 66,253 drivers.

According to the dataset, out of the 66,253 initial drivers involved in a traffic accident, only 4,542 were seriously injured (6.8%). According to the analysis carried out in this study, the high probability to suffer a serious injury in leisure purpose was (7.3%), “in itinere” (6.5%), and on business (5.8%). Based on these results, it can be argued that there is in general a greater probability of having accidents in leisure trips. These data resemble with the article published by Mitchell, Bambach [49], where the authors conclude that factors such as alcohol, speed and fatigue are less likely to be involved in accidents when they are associated with business issues.

The main risk factors involved in road traffic injuries were, respectively, driving a motorcycle (21.3%), not wearing a seat belt (19.9%), exceeding speed limit (19.0%), drivers under 18 year old (18.1%), not wearing a helmet (13.0%), while crossing

TABLE 2: Probabilities of suffering a serious accident, depending on the type of displacement and the different variables of the human factor. Processed by the authors.

Variable	KSI risk	“In itinere”	On business	Leisure	Others
<i>Seat-belt</i>					
Yes	0.038	0.036	0.042	0.036	0.042
No	0.199	0.191	0.174	0.193	0.230
Unknown	0.106	0.100	0.060	0.114	0.113
<i>Helmet</i>					
Yes	0.204	0.199	0.215	0.191	0.237
No	0.130	0.133	0.130	0.128	0.134
Unknown	0.041	0.039	0.045	0.039	0.045
<i>Speed</i>					
No infraction	0.048	0.045	0.038	0.053	0.048
Inadequate	0.120	0.112	0.096	0.130	0.121
Exceeding	0.190	0.175	0.159	0.202	0.194
Slow	0.068	0.060	0.058	0.071	0.075
Unknown	0.107	0.102	0.094	0.101	0.129
<i>Distraction</i>					
No	0.070	0.064	0.060	0.078	0.066
Yes	0.069	0.061	0.057	0.075	0.066
Unknown	0.070	0.067	0.057	0.068	0.083
<i>Errors</i>					
No	0.069	0.063	0.059	0.076	0.065
Yes	0.060	0.052	0.051	0.067	0.056
Unknown	0.084	0.087	0.061	0.075	0.109

TABLE 3: Probabilities of distracting or making mistakes during driving depending on the type of displacement made by the drivers. Processed by the authors.

Variable	Distraction		Errors	
	No	Yes	No	Yes
“in itinere”	86.6%	13.4%	62.6%	37.4%
On business	87.3%	12.7%	62.9%	37.1%
Leisure	82.2%	17.8%	57.6%	42.4%
Others	83.6%	16.4%	60.2%	39.8%

road (10.2%), driving a medium distance (8.2%), being distracted (6.9%), and finally making a mistake (6.0%).

The findings of the current study, according to the type of vehicle, suggest that motorcycles account for a probability of 21.3% of suffering a serious or a fatal. These results are consistent with the results of the study conducted by de Oliveira, Petroianu [44], and these authors argue that recklessness of motorcyclists while driving is the main cause of traffic accidents. Also, the authors emphasize that besides motorcycles, bicycle cyclists have a high probability of suffering a serious accident, reaching (14.0%) and coinciding with the result of our study (14.6%).

Regarding the sex and age of the driver, the masculine gender is the sex with greater probability to suffer a serious injury and/or a fatal one. Young man drivers (<18) are in

TABLE 4: Probabilities of suffering a serious accident, depending on the type of displacement and distraction or errors made by the drivers. Processed by the authors.

Variable	KSI risk medium	“In itinere”	On business	Leisure	Others
<i>Type of vehicle</i>					
Car	0.038	0.037	0.042	0.036	0.042
Bike	0.146	0.139	0.139	0.148	0.136
Motorcycle	0.213	0.201	0.216	0.202	0.244
Bus	0.025	0.024	0.025	0.022	0.037
Truck	0.047	0.041	0.050	0.034	0.039
Others	0.120	0.123	0.118	0.116	0.123
<i>Zone</i>					
Road	0.090	0.081	0.079	0.095	0.095
Crossing	0.102	0.094	0.081	0.097	0.125
Urban area	0.039	0.039	0.027	0.039	0.044
Motorway	0.083	0.09	0.041	0.100	0.072
<i>Distance</i>					
Local	0.050	0.052	0.042	0.050	0.051
Medium	0.082	0.064	0.060	0.103	0.067
Long	0.077	0.068	0.066	0.089	0.073
Unknown	0.127	0.137	0.075	0.142	0.132

TABLE 5: Probabilities of suffering a serious accident, based on the reason of displacement, age and gender. Processed by the authors.

Age	Gender	KSI risk—cause of displacement			
		“In itinere”	On business	Leisure	Others
<18	Man	0.183	0.201	0.176	0.204
	Woman	0.176	0.187	0.17	0.186
18–25	Man	0.079	0.106	0.071	0.093
	Woman	0.052	0.107	0.047	0.059
25–60	Man	0.071	0.055	0.081	0.081
	Woman	0.049	0.055	0.052	0.055
>60	Man	0.063	0.054	0.067	0.071
	Woman	0.046	0.052	0.047	0.051
Unknown	Man	0.061	0.067	0.1	0.067
	Woman	0.047	0.062	0.063	0.051

particular affected on business trips with a (20.1%) of probability. This is while young women drivers represent the (18.7%) of probability. Another important factor concerning the gender and age is that the lower risk of suffering serious and/or a fatal accident occurs in the age range above 60 year old.

Another important factor that is considered in the current study is the behavior of the driver. The main risk factor associated with driver’s behavior in relation to displacement is not wearing a seat belt, in the case of “in itinere” displacement representing (19.1%) and for business trips (17.4%). Exceeding the speed limit is another factor associated with driver’s behavior in the case of displacement reaching (20.2%) in leisure trips. In relation to the driver’s behavior, it is further observed that the probability of having a serious and/or a fatal accident due to making mistakes or being distracted is not so high. The result

shows an average of (5.7%) for mistakes and (6.4%) for distractions. Nevertheless, the probability of committing an error or distraction during driving is high, which reaches an average of (38.99%) for mistakes and (14.6%) for distractions.

Finally, a sensibility analysis was conducted in order to identify the probability of serious accidents as to what extent they determine the cause of displacement, the zone as well as distance. The higher probabilities of suffering a serious and/or a fatal accident according to the zone are in leisure trips in motorway (10%), on business and “in itinere” in crossing areas (8.1% and 9.4% respectively). This is while displacements caused by driving long distance reach (6.8%) “in itinere” trips and (6.6%) on business trips, on the other hand, in the case of leisure trips the high probability occurrence in medium distance reaching (10.3%).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

We would like to thank the Dirección General de Tráfico (DGT) for providing us with the necessary data for conducting this study (SPIP2015-1852). The MICRODATOS_ACCIDENTES_2016, MICRODATOS_CONDUCTORES_2016, and MICRODATOS_VEHICULOS_2016 used to support the findings of this study were supplied by the Dirección General de Tráfico (DGT) and so cannot be made freely available. Requests for access to these data should be made to DGT, C/ Josefa Valcarcel, 28 y 44, 28071 Madrid, España. This work is part of the research project “Modelización mediante técnicas de machine learning de la influencia de las distracciones del conductor en la seguridad vial. Diseño de un sistema integrado: simulador de conducción, eye tracker y dispositivo de distracción. Ref. BU300P18” supported by funds from FEDER (Fondo Europeo de Desarrollo Regional - Junta de Castilla y León).

References

- [1] S. Yasmin, N. Eluru, A. R. Pinjari, and R. Tay, “Examining driver injury severity in two vehicle crashes—a copula based approach,” *Accident Analysis & Prevention*, vol. 66, pp. 120–135, 2014.
- [2] P. Hamalainen, “The effect of globalization on occupational accidents,” *Safety Science*, vol. 47, no. 6, pp. 733–742, 2009.
- [3] OMS, “Lesiones causadas por el tránsito,” 2016. <http://www.who.int/mediacentre/factsheets/fs358/es/>.
- [4] B. Charbotel, J. L. Martin, and M. Chiron, “Work-related versus non-work-related road accidents, developments in the last decade in France,” *Accident Analysis & Prevention*, vol. 42, no. 2, pp. 604–611, 2010.
- [5] Eurogip, “Understanding occupational risks in Europe,” Eurogip, 2013.
- [6] OECT, “Informe de accidentes laborales de tráfico 2016,” 2016, <http://www.oect.es/Observatorio/5%20Estudios%20tecnicos/Riesgos%20especificos/Estudios%20sobre%20Accidentes%20de%20trabajo%20relacionados%20con%20el%20trafico/Ficheros/Trafico%202016.pdf>.
- [7] M. J. Mazankova, “Increased risk of accidents due to human behavior,” in *Applied Mechanics, Behavior of Materials, and Engineering Systems*, Springer, Berlin. Boukharouba, G. Pluinage, and K. Azouaoui, Ed., pp. 551–560, 2017.
- [8] B. E. Sabej and H. Taylor, “The known risks we run: the highway,” in *Societal Risk Assessment: How Safe is Safe Enough?*, C. Schwing and W. A. Albers, Eds., pp. 43–70, Springer, Boston, MA USA, 1980.
- [9] H. Summala, “Risk control is not risk adjustment: the zero-risk theory of driver behaviour and its implications,” *Ergonomics*, vol. 31, no. 4, pp. 491–506, 1988.
- [10] S. Salminen and E. Lähdeniemi, “Risk factors in work-related traffic,” *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 5, no. 1, pp. 77–86, 2002.
- [11] A. Bener, E. Yildirim, T. Özkan, and T. Lajunen, “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*, vol. 4, no. 5, pp. 496–502, 2017.
- [12] M. Dozza, “What factors influence drivers’ response time for evasive maneuvers in real traffic?,” *Accident Analysis & Prevention*, vol. 58, pp. 299–308, 2013.
- [13] S. J. Kim and E. K. Chung, “The effect of organizational justice as perceived by occupational drivers on traffic accidents: Mediating effects of job satisfaction,” *Journal of Safety Research*, vol. 68, 2019.
- [14] D. Wishart, K. Somoray, and A. Evenhuis, “Thrill and adventure seeking in risky driving at work: the moderating role of safety climate,” *Journal of Safety Research*, vol. 63, pp. 83–89, 2017.
- [15] R. Mitchell, T. Driscoll, and S. Healey, “Work-related road fatalities in Australia,” *Accident Analysis & Prevention*, vol. 36, no. 5, pp. 851–860, 2004.
- [16] S. Boufous and A. Williamson, “Work-related traffic crashes: a record linkage study,” *Accident Analysis & Prevention*, vol. 38, no. 1, pp. 14–21, 2006.
- [17] M. Chiron, M. Bernard, S. Lafont, and E. Lagarde, “Tiring job and work related injury road crashes in the GAZEL cohort,” *Accident Analysis & Prevention*, vol. 40, no. 3, pp. 1096–1104, 2008.
- [18] L. Di Milia, M. H. Smolensky, G. Costa, H. D. Howarth, M. M. Ohayon, and P. Philip, “Demographic factors, fatigue, and driving accidents: an examination of the published literature,” *Accident Analysis & Prevention*, vol. 43, no. 2, pp. 516–532, 2011.
- [19] M. Pylkkönen, M. Sihvola, Hk Hyvärinen, S. Puttonen, C. Hublin, and M. Sallinen, “Sleepiness, sleep, and use of sleepiness countermeasures in shift-working long-haul truck drivers,” *Accident Analysis & Prevention*, vol. 80, pp. 201–210, 2015.
- [20] R. O. Phillips and S. F. Meyer, “Kartlegging av arbeidsrelaterte trafikulykker. Analisar dødsulykker av i Norge fra 2005, til 2010, TØI relasjon 1188/2012,” 2012.
- [21] T.-O. Nævestad, R. O. Phillips, and B. Elvebakk, “Traffic accidents triggered by drivers at work—a survey and analysis of contributing factors,” *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 34, pp. 94–107, 2015.
- [22] J. L. Machado-León, J. de Oña, R. de Oña, L. Eboli, and Gabriella Mazzulla, “Socio-economic and driving experience factors

- affecting drivers' perceptions of traffic crash risk," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 37, pp. 41–51, 2016.
- [23] MAPFRE, "Evolutionary analysis of road occupational accidents (2003–2012)," 2015.
- [24] D. M. DeJoy, "An examination of gender differences in traffic accident risk perception," *Accident Analysis & Prevention*, vol. 24, no. 3, pp. 237–246, 1992.
- [25] L. Åbele, S. Haustein, M. Møller, and L. M. Martinussen, "Consistency between subjectively and objectively measured hazard perception skills among young male drivers," *Accident Analysis & Prevention*, vol. 118, pp. 214–220, 2018.
- [26] Scherrer, "Accurate physiology of work notions of ergonomics," 1992.
- [27] Holliday, "Understanding Ageing," Cambridge University Press, 1995.
- [28] Ginebra, "Speed Control: United Nations Road Safety Manual for Decision Making and Professionals," 2008.
- [29] J. Q. Wu and H. Xu, "Driver behavior analysis for right-turn drivers at signalized intersections using SHRP 2 naturalistic driving study data," *Journal of Safety Research*, vol. 63, p. 177, 2017.
- [30] L. Aarts and I. van Schagen, "Driving speed and the risk of road crashes: a review," *Accident Analysis & Prevention*, vol. 38, no. 2, pp. 215–224, 2006.
- [31] S. Newman, B. Watson, and W. Murray, "Factors predicting intentions to speed in a work and personal vehicle," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 7, pp. 287–300, 2004.
- [32] C. Chen, G. Zhang, Z. Tian, S. M. Bogus, and Y. Yang, "Hierarchical Bayesian random intercept model-based cross-level interaction decomposition for truck driver injury severity investigations," *Accident Analysis and Prevention*, vol. 85, pp. 186–198, 2015.
- [33] DGT, "Traffic Accident Form with Victims Anuario Estadístico de Accidentes Statistical Accident Yearbook," 2014, https://www.boe.es/diario_boe/txt.php?id=BOE-A-2014-12411.
- [34] DGT, "Anuario Estadístico de Accidentes," 2016.
- [35] R. L. Blanco, "Perception of Risk and Safety in the Driving of Vehicles," *Carreteras*, pp. 48–56, 2013.
- [36] X. Zou and W. L. Yue, "A bayesian network approach to causation analysis of road accidents using Netica," *Journal of Advanced Transportation*, vol. 2017, pp. 1–18, 2017.
- [37] X. Ma, Y. Xing, and J. Lu, "Causation analysis of hazardous material road transportation accidents by bayesian network using genie," *Journal of Advanced Transportation*, vol. 2018, pp. 1–12, 2018.
- [38] T. Tang, S. Zhu, Y. Guo, X. Zhou, and Y. Cao, "Evaluating the safety risk of rural roadsides using a Bayesian network method," *International Journal of Environmental Research and Public Health*, vol. 16, no. 7, p. 1166, 2019.
- [39] J. De Oña, R. O. Mujalli, and F. J. Calvo, "Analysis of traffic accident injury severity on Spanish rural highways using Bayesian networks," *Accident Analysis and Prevention*, vol. 43, no. 1, pp. 402–411, 2011.
- [40] E. Castillo, J. M. Gutiérrez, and A. S. Hadi, *Expert Systems and Probabilistic Network Models*, Springer, Verlag, 1997.
- [41] J. A. Hanley and B. J. McNeil, "The meaning and use of the area under a receiver operating characteristic (ROC) curve," *Radiology*, vol. 143, no. 1, pp. 29–36, 1982.
- [42] C. Xu, Jie Bao, C. Wang, and P. Liu, "Association rule analysis of factors contributing to extraordinarily severe traffic crashes in China," *Journal of Safety Research*, vol. 67, pp. 65–75, 2018.
- [43] S. T. D. Cordazzo, C. T. Scialfa, K. Bubric, and R. J. Ross, "The driver behaviour questionnaire: a North American analysis," *Journal of Safety Research*, vol. 50, pp. 99–107, 2014.
- [44] A. L. de Oliveira, Andy Petroianu, D. M. V. Gonçalves, G. A. Pereira, and L. R. Alberti, "Characteristics of motorcyclists involved in accidents between motorcycles and automobiles," *Revista da Associação Médica Brasileira*, vol. 61, no. 1, pp. 61–64, 2015.
- [45] P. Olszewski, P. Szagała, D. Rabczenko, and A. Zielińska, "Investigating safety of vulnerable road users in selected EU countries," *Journal of Safety Research*, vol. 68, pp. 49–57, 2019.
- [46] V. Bellos, A. Ziakopoulos, and G. Yannis, "Investigation of the effect of tourism on road crashes," *Journal of Transportation Safety & Security*, pp. 1–18, 2019.
- [47] S. Regev, J. J. Rolison, and S. Moutari, "Crash risk by driver age, gender, and time of day using a new exposure methodology," *Journal of Safety Research*, vol. 66, pp. 131–140, 2018.
- [48] L. Korpinen and R. Paakkonen, "Accidents and close call situations connected to the use of mobile phones," *Accident Analysis and Prevention*, vol. 45, pp. 75–82, 2012.
- [49] R. J. Mitchell, M. R. Bambach, and R. Friswell, "Work and non-work-related vehicle crashes: the contribution of risky driving practices," *Safety Science*, vol. 68, pp. 65–72, 2014.