THE BALTIC JOURNAL
OF ROAD AND BRIDGE ENGINEERING

2013

8(2): 124–132

HOW TO COMPARE DIFFERENT NATIONAL DATABASES OF HGV ACCIDENTS TO IDENTIFY ISSUES FOR SAFETY IMPROVEMENTS

Salvatore Cafiso[1], Alessandro Di Graziano[2]✉, Giuseppina Pappalardo[3]

Dept of Civil and Environmental Engineering, School of Engineering, University of Catania, Viale Andrea Doria 6, I-95125 Catania, Italy
E-mails: [1]dcafiso@dica.unict.it; [2]adigraziano@dica.unict.it; [3]giusy.pap@dica.unict.it

Abstract. The objective of this paper is to present a methodological approach and a case study for an international comparison of accident data coming from different national databases. Safety levels and the characteristics of severe crashes involving heavy goods vehicles in different European countries (Italy, France, Germany, Great Britain and Spain) are analysed. Considering that all the countries involved have different inventory structures for the variables reported in their national accident databases, the taxonomy theory was used in order to create a comparable structure for the database used in the analysis. The taxonomy is non-exclusive and the codes are categorical, denoting the absence or presence of a certain feature. Based on the data available in each national database the five European Union databases of accidents involving heavy goods vehicles have been referenced to only one, composed of 11 items (casualty class, injury number and severity, location, light conditions, road conditions, junction, vehicle type, driver age, driver gender, accident type and manoeuvres), which capture common features of heavy goods vehicles accidents. A statistical analysis was carried out in order to highlight significant differences in the proportions of heavy goods vehicles crash categories.

Keywords: accident data, heavy goods vehicle, database, taxonomy, statistical analysis, proportion method.

1. Introduction

The objective of this paper is to compare heavy goods vehicle (HGV) safety levels and characteristics in different European countries. The European Union (EU) was originally composed of 15 countries (EU-15) now extended to 27 (EU-27) including new countries from the East.

In the EU-15, HGV fatal crashes fell from 4988 in 1995 to 3114 in 2006, a fall of more than 30%, although they still represent about 13% of the overall fatalities occurring in road crashes (Broughton *et al.* 2008). Despite the relevance of the phenomenon, few detailed statistics are currently available regarding accidents involving HGVs and even less is known about differences or similarities between different European countries. This lack of comparable data is due to an absence of homogeneity among accident databases at international level. To overcome this problem, in 1993 the Community Road Accident Database (CARE) was created as a useful tool for comparing accidents in EU countries, but, after 15 years of application it has not been able to harmonize the different national accident databases. With particular reference to commercial vehicles in CARE there is a lack of details for a more in-depth analysis.

For this reason, at European level in-depth analyses of accidents involving HGVs are only carried out by specific investigation systems including a high degree of detail but, consequently, with a limited number of available cases. In the European Truck Accident Causation (ETAC) study of 2007 a common database, made up by “only” 600 truck accident reports for seven European countries was used. In all those accidents, the main cause of accident (85.2%) was linked to the human error of one of the road participants (truck driver, car driver, pedestrian etc). Other factors such as weather conditions (4.4%), infrastructure conditions (5.1%) or technical failures of the vehicle (5.3%) played only a minor role. Accidents at intersections (27%) represented the first accident typology followed by accidents in queues (21%).

Accident data analyses highlight immediately some peculiarities characterizing the accident phenomenon. Fig. 1 shows the accident rate (accident/HGV fleet) and the fatality rate (fatality/accident) for each European country. The reference year for all the accident data is the 2006, except for France where the 2005 has been the last available one.

Based on national databases, it is difficult to conduct a more in-depth analysis due to the difference in the variables considered in each national database.

ISSN 1822-427X print / ISSN 1822-4288 online doi:10.3846/bjrbe.2013.16

http://www.bjrde.vgtu.lt

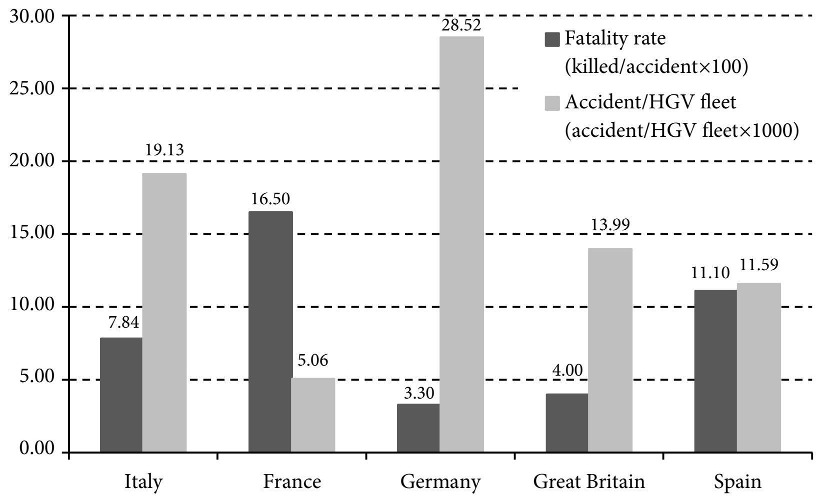

Fig. 1. Fatality and Accident rate in the analyzed countries

For this reason, the aim of the paper is to attempt grouping comparable countries by way of national dataset management and then to compare the countries within a specific group or class.

With this purpose, the paper can be subdivided into two logical parts:

– the taxonomy approach for different national database management;
– statistical analysis of HGV data collected in different EU countries.

Considering all EU-15 or EU-27 countries is time and cost consuming and not useful for the aim of the present work. Therefore, only five representative countries were selected. The transport of goods by road is prevalent in Europe with peaks of about 90% in Spain, Denmark, Greece, Ireland, Italy, Luxemburg and Portugal. In 2006 the fleet of commercial vehicles in the EU-15 countries amounted to about 22 mln vehicles. From these countries five (EU-5: Italy, France, Germany, Great Britain and Spain) were selected representing about 70% of the overall EU HGV fleet.

2. Methodological approaches

As regards road crashes, all national traffic accident databases contain a rich source of information on the different circumstances in which the accidents have occurred: cause of the accident (type of collision, road users, injuries, etc.), traffic conditions (max speed, priority regulations, etc.), environmental conditions (weather, light conditions, time of the accident, etc.), road conditions (road surface, obstacles, etc.), human conditions (fatigue, alcohol, etc.) and geographical conditions (location, physical characteristics, etc.) (Geurts *et al.* 2003). Unfortunately, each national database reports the accidents occurring throughout the country following its own particular choice of dataset.

In the present research, the accident data consists of statistical databases from five European countries (Italy, Germany, Spain, France and the Great Britain).

In Italy, the source statistics for the detection of accidents are provided by ISTAT (National Institute of Statistics). Any injury and/or fatal accident should be reported by the police authorities in the jurisdiction of the crash using the Model *CTT.INC ISTAT*. In Germany, the source statistics for the detection of accidents is provided by the Statistisches Bundesamt (DESTATIS). The accident information is based on a monthly collection of road accidents occurring over their territory collected by the public authorities. The Spanish national accident statistics database is run by the General Directorate of Traffic (DGT) which comes under the Spanish Ministry of the Interior. The database is fed by the police reports for all road accidents where at least one casualty was registered. The French National Road Administration's accident database requires that any accident involving injuries should be reported and coded in a Bullettin d'Analyse d'Accident Corporel de la Circulation (BAAC) by the gendarmerie or police in the jurisdiction of the crash. The STATS19 national database, run by the UK Government, contains comprehensive information about UK road accidents on the public highway which involving human injury or death. The data contains highway, vehicle and human information compiled at the time of accident by the police.

Each national crash database reports accidents occurring throughout the country that meet specific criteria for inclusion and classification. These criteria are different for each country and are not necessarily comparable. Therefore, for a comparison analysis, it was necessary to create a common structure to harmonize the individual differences into one consistent reporting system. For this purpose, the taxonomy approach can be used (Wallace, Ross 2007).

2.1. Taxonomy

The hierarchical structure of the data taxonomy represents a convenient way of classifying data in order to prove it is unique and not redundant (Bryce 2005). Mathematically, a hierarchical taxonomy is a tree structure of classifications for a given set of objects. At the top of this structure there is a single classification, the root node that applies to all the objects. Nodes below this root are more specific classifications that apply to subsets of the total set of classified objects. The reasoning progresses from the general to the more specific. Classifying events using taxonomies designed for that purpose is a common technique in the human sciences (e.g. psychology, sociology, psychiatry) and studies have also been presented for its application to traffic accident analysis (Donnell *et al.* 2010; Elvik 2010; Gstalter, Fastenmeier 2010; Johnson *et al.* 2009; Regan *et al.* 2011; af Wåhlberg 2002). In traffic accident analysis, the point of any taxonomy is simply to help classify the factors that contribute to accidents or injuries and thus establish a starting point to study the causes of accidents. For any taxonomy the categories must strike a balance between incorporating too much and too little. For this reason any taxonomy of traffic accidents is necessarily incomplete as there are always categories which could be included or excluded.

Taxonomy has been shown to be highly useful if whether a category finally ends up inside or outside the database meets three different criteria (Ross *et al.* 2004; Stanton, Salmon 2009; Yeraguntla *et al.* 2005):

- the importance of the variable in the analysis of the phenomenon (when it comes to causing accidents and/or the usefulness of the category in accident analysis and prevention);
- the availability of the kind of data needed to code for a variable;
- the balance between the number of variables used and the size of the resulting samples.

With these main guiding principles the accident taxonomy was developed for this study, using the procedure described below, to compare HGV safety levels and characteristics in different European countries (Italy, France, Germany, Great Britain, Spain).

2.2. Procedure

With the aim of harmonizing national databases and thus obtaining useful information for crash analysis and comparison, different items were defined (root node). Then for every node, starting from the different structure of each national database, more specific sub-classifications were defined according to the variables characterizing the datasets. In order to univocally characterize the property matched to each variable the attributes were defined with reference to every sub-category.

Finally, an Identification Data (ID) was applied to each sub-category in order to easily codify information taken from different databases. If there was not enough information to decide on the applicability of the variable it was not used and was marked as "missing". If the evidence for the interpretation of an attribute into a variable was not clear enough it was "not coded".

A taxonomy root structure was carried out for each item using the five European (Italy, Germany, Spain, France and the Great Britain) in order to define the list of attributes.

For example, to identify crashes involving heavy trucks (HGVs) Fig. 2 shows item *E* related to the vehicle type. Other example of hierarchical taxonomy is shown in Fig. 3 for junction/no junction definition (item *D*), highlighting the necessity for a high level of aggregation based on data availability. The variable's names are reported in the national language for a better reference to the original database.

From the data available in each national database eleven tree structures were composed, like those shown in the previous figures, using the taxonomic approach referring to the same number of items (casualty class, injury number and severity, location, light conditions, road conditions, junction, vehicle type, driver age, driver gender, accident type and manoeuvres) which capture common features of road accidents.

Table 1 reports the list of attributes included in the new common database drawn up by referring each national dataset to only one.

3. HGV accident data comparison in EU countries

A simple comparison of the number of accidents referring to different categories doesn't lead to interesting results due to the variability among the various countries in terms of exposure (vehicle fleet, travelled km). Instead, the proportions of the occurrence of different typology of crashes are not influenced by the sample dimension and therefore can be used to compare the characteristics of HGV crashes in the analyzed countries. As each analysis of accident data a simple comparison of data could lead to bias due to the stochastic nature of the phenomenon.

3.1. The Bayes theorem for proportions

The "proportions" method compare proportions of an accident type among different samples (Cafiso *et al.* 2012; Heydecker, Wu 1991; Lyon *et al.* 2007) considering the random characteristics of the phenomenon.

The proportion of a specific collision type for the sample "*i*" is defined μ_i:

$$\mu_i = \frac{x_i}{n_i}, \tag{1}$$

where x_i – the total number of target collisions, during the study period in the sample "*i*"; n_i – the total number of all types of collisions in the sample "*i*" during the same period.

Considering m different samples the mean proportion of the target collision type is given by:

$$\bar{\mu} = \frac{\sum_{i=1}^{m} \mu_i}{m}. \tag{2}$$

VEHICLE TYPE

PASSENGER CAR	1
MOTOR CYCLE	2
BUS AND COACH	3
LIGHT TRUCK	4
HEAVY TRUCK	5
OTHER MOTOR VEHICLE	6

IT

Passenger car	1
Passenger car with trailer	1
Taxi/Private hire car	1
Car rescue or police	1
Urban bus	3
Rural bus	3
Tram	6
Goods vehice (≤ 3.5 t)	4
Goods vehicle (> 3.5 t)	4
Lorry with trailer	5
Articulated vehicle	5
Special vehicle	6
Tractor	6
Agricultural vehicle	6
Pedal cycle	6
M/cycle 50 cc and under	2
M/cycle over 50 cc	2
M/cycle over 50 cc with passenger	2
Other motor vehicle	6
Other non motor vehicle	6
Unknown vehicle	6
Quadricycle	6

FR

Pedal cycle	6
Moped	6
Quadricycle	6
M/cycle 50 cc and under	2
M/cycle over 50 cc	2
Side-car	2
Passenger car	1
Passenger car + caravan	1
Passenger car + trailer	1
Goods vehicle (≤ 3.5 t)	4
Goods vehicle (≤ 3.5 t) + caravan	4
Goods vehicle (≤ 3.5 t) + trailer	4
Goods vehicle (3.5–7.5 t)	5
Goods vehicle (> 7.5 t)	5
Lorry with trailer	5
Tractor	5
Tractor + semi-trailer	5
Bus	3
Tram	6
Special vehicle	6
Agicultural vehicle	6
Other motor vehicle	6

D

Passenger car	1
Pedal cycle	6
Motorcycle	2
Bus	3
Goods vehicle (< 3.5 t)	4
Goods vehicle (> 3.5 t)	5
Semi-Trailers	5
Tractor	5
Truck with special design	4
Moped	6
M/cycle under 50 cc	6
M/cycle 50 cc	2
M/cycle over 50 cc	2
Motor scooter	2
Rural Bus	3
Lorry with trailer	5
Other motor vehicle	6
Tram	6
Tram	6
Other and unknown vehicle	6

GB

Pedal cycle	6
M/cycle 50 cc and uder	2
M/cycle over 50 cc and up to 125 cc	2
M/cycle over 125 cc and up to 500 cc	2
Motocycle over 500 cc	2
Taxi/Privale hire car	1
Car	1
Minibus (8–16 passenger seats)	3
Bus or coach (17 or more passenger seats)	3
Other motor vehicle	6
Other non motor vehicle	6
Ridden horse	6
Agricultural vehicle	6
Goods vehicle 3.5 t mgw and under	4
Goods vehicle over 3.5 t mgw and under 7.5 t mgw	5
Goods cehicle 7.5 t mgw and over	5

SP

Pedal cycle	6
M/cycle 50 cc and over	2
Disabled vehicle	1
M/cycle over 50 cc	2
Passenger car up to 9 seats	1
Passenger car without trailers	1
Passenger car with trailers	1
Ambulance car	6
Agicultural vehicle	6
Tractor without trailers	6
Tractor with trailers	6
Goods vehicle (≤ 3.5 t) without trailers	4
Goods vehicle (≤ 3.5 t) with trailers	4
Van	4
Goods vehicle (> 3.5 t) without trailers	5
Goods vehicle (> 3.5 t) with trailers	5
Tankers without trailer	5
Tankers with trailer	5
Articulated vehicle	5
Bus	3
School bus	3
Other bus	3
Train	6
Other vehicle	6
Unknown vehicle	6

Fig. 2. Example of taxonomy per vehicle type (item *E*)

JUNCTION / NO JUNCTION

JUNCTION	1
NON JUNCTION	2

IT

Junction	1
Roundabout	1
Signalized junction	1
Un-signalized junction	1
Rail crossing	1
Tangent	2
Curve	2
Slope	2
Narrow	2
Lit tunnel	2
Until tunnel	2

FR

Not at junction	2
X junction	1
T junction	1
Y junction	1
Multiple junction	1
Roundabout	1
Square	1
Rail crossing	1
Other junction	1

D

Junction	1
Other junction	1
Ramp	1
Slope	2
Curve	2

GB

Not ar or within 20 metres of junction	2
Roundabout	1
Mini roundabout	1
T or staggered junction	1
Slip road	1
Crossroads	1
Multiple junction	1
Using private drive or entrance	1
Other junction	1

SP

Tangent	2
Smooth curve	2
Unmarked sharp curve	2
Marked sharp curve without posted speed signal	2
Marked sharp curve with posted speed signal	2
T o Y junction	1
X o + junction	1
Entrance ramp	1
Exit ramp	1
Roundabout	1
Other junction	1

Fig. 3. Example of taxonomy per junction / no junction (item *D*)

Table 1. List of attributes included in the common data set

Data typology	Item	Variable	Definitions	ID
CRASH DATA	A Road type	Motorway	Public roads with dual carriageways and at least two lanes each way. Entrance and exit signposted	A1
		Rural road	Public roads with single or dual carriageways but no motorway restrictions	A2
		Urban street	Public roads in urban area	A3
		Other	Other roads	A4
	B Light conditions	Daylight	Daylight condition	B1
		Darkness	Darkness without artificial light, darkness with artificial light unlit, darkness with artificial light lit	B2
	C Roadway surface	Dry	Dry pavement condition	C1
		Wet	Wet pavement condition	C2
		Ice	Icy pavement condition	C3
		Snow	Snowy pavement condition	C4
		Other	Other pavement condition	C5
	D Junction/ no junction	Junction	Intersection police officer, intersection traffic lights and traffic signs, intersection priority to right, roundabout	D1
		No junction	Straight road, right curve, left curve, flat road and slope	D2
	E Vehicle type	Passenger car	Motor vehicle with three or four wheels. Used to transport only or mainly people	E1
		Motor cycle	Motor vehicle with two or three wheels, with engine size of more than 50 cylinders.	E2
		Bus and coach	Motor vehicle with at least four wheels, used for transportation of people	E3
		Light truck (< 3.5 t)	Used only for the transport of goods	E4
		Heavy truck (> 3.5 t)		E5
		Other motor vehicle	Other motor vehicles	E6
	F Injury severity	Killed	Any person who was killed outright or who died within 30 days as result of the accident	F1
		Injured	Any person, who was not killed, but sustained one or more serious or slight injuries as a result of the accident	F2
	G Accident type	Accident between vehicle and pedestrian	Accidents involving one or several vehicles and pedestrians irrespective of whether the pedestrian was involved in the first or a later phase of the accident and of whether the pedestrian was injured or killed on or off the road	G1
		Single vehicle accidents	Accidents involving no collision with other users, even though they may be involved or accident caused by collision with obstructions or animals on the road	G2
		Rear-end collisions	Accident caused by a rear-end collision with another vehicle using the same lane of a carriageway and moving in the same direction or temporarily stopping due to the traffic conditions	G3
		Front side and side-swipe collisions	Accident caused by a collision with another vehicle moving in a lateral direction due to leaving or entry from/to another lane, road, or premises	G4
		Head on collisions	Accident caused by a head-on collision with another vehicle using the same lane of a carriageway and moving in the opposite direction or temporarily stopping due to traffic conditions	G5
		Other collisions	Other type of accident	G6
	H Manoeuvres	Reversing		H1
		Slowing or stopping		H2
		Turning left/right		H3

Continued Table 1

Data typology	Item	Variable	Definitions	ID
CASUALTY DATA	K Casualty class	Driver	Person driving or riding any motorized vehicle or pedal cycle	K1
		Passenger	Person on or in a vehicle who is not the driver	K2
		Pedestrian	Person on foot	K3
	I Driver age	Young	Age of road user 0–17	I1
		Normal	Age of road user 18–60	I2
		Elderly	Age of road user over 60	I3
	J Driver gender	Male		J1
		Female		J2

The premise of the "proportions" method is that if the true proportion of sample i is μ_i, then the probability of observing x_i target accidents with n_i total accident is given by the Binomial distribution:

$$f(x_i / n_i, \mu_i) = \binom{n_i}{x_i} \mu_i^{x_i} (1-\mu_i)^{n_i - x_i}, \; 0 < x_i < n_i. \quad (3)$$

Moreover, the parameter μ_i will vary between similar sites and is assumed to follow the Beta distribution, defined as:

$$g(\mu / \alpha, \beta) = \frac{\mu^{\alpha-1}(1-\mu)^{\beta-1}}{B(\alpha,\beta)}, \; 0 \le \mu \le 1, \quad (4)$$

where $B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$, with $\Gamma(.)$ = gamma function.

The parameters α and β of the Beta distribution can be estimated from the sample mean and the variance of a reference population using the following equations:

$$\alpha = \frac{\bar{\mu}^2 - \bar{\mu}^3 - s^2 \bar{\mu}}{s^2}, \quad (5)$$

$$\beta = \frac{\alpha}{\bar{\mu}} - \alpha, \quad (6)$$

where s^2 – the variance given by

$$s^2 = \frac{1}{m-1}\left[\sum_{i=1}^{m}\left(\frac{x_i^2 - x_i}{n_i^2 - n_i}\right) - \frac{1}{m}\left(\sum_{i=1}^{m}\frac{x_i}{n_i}\right)^2\right], \; n \ge 2.$$

Using Bayes theorem, the prior Beta distribution is combined with sample "i" specific accident data (n_i, x_i) to derive the adjusted posterior distribution that is again a Beta distribution:

$$g(\mu / \alpha_i', \beta_i') = \frac{\mu^{\alpha_i'-1}(1-\mu)^{\beta_i'-1}}{B(\alpha_i', \beta_i')}, \quad (7)$$

where α_i' and β_i', posterior parameters defined as:

$$\alpha_i' = \alpha + x_i, \; \beta_i' = \beta + n_i - x_i.$$

For the posterior distribution the mean value and variance for each site "i" can be calculated with the following equations:

$$E(\mu_i) = \frac{\alpha_i'}{\alpha_i' + \beta_i'}, \quad (8)$$

$$Var(\mu_i) = \frac{\alpha_i' \beta_i'}{(\alpha_i' + \beta_i')^2 (\alpha_i' + \beta_i' + 1)}. \quad (9)$$

Defining $\bar{\mu}_m$ and μ_{mi}, respectively the median of the prior and posterior distributions the probability $P(\mu_i > \bar{\mu}_m)$ is given by:

$$P(\mu_i > \bar{\mu}_m) = 1 - \int_0^{\bar{\mu}_m} g(\mu, \alpha_i', \beta_i') d\mu. \quad (10)$$

Based on the large sample dimension a probability of 99% can be assumed as acceptable for considering the difference significant.

If $\bar{\mu}_m$ is assumed as reference value of proportion for the accident type to be screened, the Potential for Safety (PfS) can be defined as the difference between the median in the sample "i" μ_{mi} and the reference value of the proportion, $\bar{\mu}_m$:

$$\text{PfS} = \mu_{mi} - \bar{\mu}_m. \quad (11)$$

Basing on the definition of PfS, the value of the potential reduction of accident number Δx_i can be calculated as the product of PfS and the observed number of accident x_i:

$$\Delta x_i = \text{PfS} x_i. \quad (12)$$

A positive value of Δx_i represents the potential reduction in the number of crashes, of the analyzed category, due to the abnormal proportion in the sample "i" with respect to the reference population.

3.2. Study results

Accidents type showing significantly higher proportions (μ_{mi}) in relation to the reference value $\bar{\mu}$ are the best candidate for improvement interventions. In this sense, may have significant potential reductions in the number of accidents at junctions (Table 2) for Great Britain and Spain (respectively 124 and 1282 accidents per year).

As average 40.1% of HGV crashes at intersections can be expected, though SP has a particularly high percentage of crashes with a potential reduction of 1282 crashes per year.

Specifically PfS and $\frac{\Delta x_i}{\text{year}}$ values show for each country the crash type involving HGV with the higher potential reduction in the number of accidents: in Germany it is expected to reduce 267 pedestrian and 353 single accidents per year, in Italy 37 rear and 40 side crashes and in Spain 353 single accidents of HGV. Specifically if the reference proportion $\overline{\mu}_m$ is considered as "normal" the appropriate 4E safety strategies (Engineering, Education, Enforcement, Emergencies) could be particularly effective on accident types that have significant higher proportions than the expected one. It is generally assumed that new technologies can improve safety. In particular, a great deal of attention has been paid to the effects of driver assistance systems on driver performance (Cafiso, Di Graziano 2012; Lin *et al.* 2008). However, challenges still remain in quantifying the benefits of these systems in terms of their impact on reliability, profitability and safety (Cafiso *et al.* 2013).

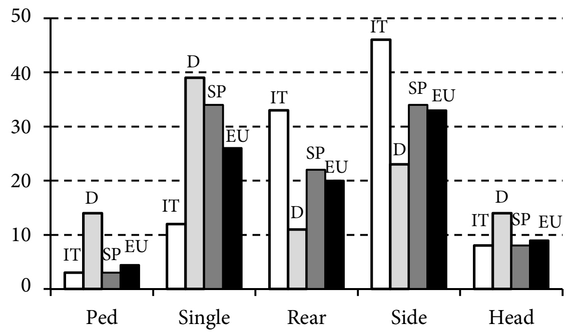

Fig. 4. Proportions for crash type (for all accidents involving HGV)

Another example of the use of proportion method is reported with reference to accident type. In this case, data from the Great Britain and France were not available as disaggregated variables even if they are part of the national data source, thus the analysis is presented only for the other three countries (Fig. 4). Fig. 4 shows as front/sideswipe (33.2%) and single crashes (26.5%) have the higher

Table 2. Potential for Safety (PfS) and potential reduction in the number of crashes (Δx_i) for different category

Sample	x_i	n_i	μ_{mi}, %	$\overline{\mu}_m$, %	$P(\mu_i > \overline{\mu}_m)$, %	PfS, %	$\frac{\Delta x_i}{\text{year}}$
Crash at intersection (at least one HGV involved)							
IT	3998	10 523	38.0	40.1	0.00		
F	808	4730	17.1		0.00		
GB	4839	11 336	42.7		100	2.6	124
SP	4573	6707	682		100	28.0	1.282
Crash between HGV and pedestrian							
IT	320	10 511	3.0	4.5	0.00		
D	2838	20 383	13.9		100	9.4	267
SP	119	4691	2.5		0.00		
Single HGV crash							
IT	1239	10 511	11.8	26.5	0.00		
D	7932	20 383	38.9		100	12.5	353
SP	1587	4691	33.8		100	7.4	9
Rear end crash (at least one HGV involved)							
IT	3330	10 511	31.7	20.1	100	11.5	37
D	2202	20 383	10.8		0.00		
SP	1030	4691	22.0		100	1.8	2
Front and sideswipe collision (at least one HGV involved)							
IT	4816	10 511	45.8	33.2	100	12.6	40
D	4586	20 383	22.5		0.00		
SP	1568	4691	33.4		62	0.2	0
Head on collision (at least one HGV involved)							
IT	806	10 511	7.7	9.6	0.00		
D	2825	20 383	13.9		100	4.3	121
SP	387	4691	8.2		0.00		

frequencies in terms of accident type. With reference to the EU median values, Germany is characterized by significant high proportions of single and pedestrian accidents, Italy of rear and side crashes and Spain of single accidents of HGV.

4. Conclusions

In the present research, five European countries (Italy, France, Germany, Great Britain and Spain) were considered representing about 70% of the overall EU HGV fleet. Due to an absence of homogeneity in national accident databases, taxonomy was used to create a common structure to harmonize the individual differences into one consistent reporting system. The five EU databases were referenced to only one structure composed of 11 items (casualty class, injury number and severity, location, light conditions, road conditions, junction, vehicle type, driver age, driver gender, accident type and manoeuvres) which captures common features of HGV accidents.

Referring to this new common source it was possible to carry out comparable analyses of accidents involving HGVs using the proportion method to avoid the influence of exposure factors. At European level, as average, 40.1% of HGV crashes at intersections can be expected; while front/sideswipe (33.2%) and single crashes (26.5%) have the higher frequencies in terms of accident type. With reference to the EU median values, Spain has a particularly high percentage of crashes at intersections with a potential reduction of 1282 crashes per year; Germany is characterized by significant high proportions of single and pedestrian accidents with a potential reduction of 353 and 267 accident/year respectively; Italy is characterized by significant high proportions of rear and side crashes HGV with a potential reduction of 37 and 40 accident/year respectively.

Due to the limited availability of data only few comparisons were performed but the structure of the data defined using the taxonomy has identified eleven items that can be used as a reference for future studies and the proposed methodology can be used to compare crash proportions avoiding statistical bias.

References

Broughton, J.; Lawton, B.; Walter, L.; Hoeglinger, S.; Angermann, A.; Weiss, V.; Yannis, G.; Evgenikos, P.; Bos, N.; Reurings, M. 2008. *Heavy Goods Vehicles and Buses. European Road Safety Observatory Documentation. Traffic Safety Basic Facts 2008* [cited 15 January, 2011]. Available from Internet: http://www.dacotaproject.eu/Links/erso/safetynet/fixed/WP1/2008/BFS2008_SN-INTRAS-1-3-HGVs.pdf.

Bryce, T. 2005. *The Benefits of a Data Taxonomy* [cited 13 July, 2009]. Available from Internet: http://it.toolbox.com/blogs/irm-blog/the-benefit-of-a-data-taxonomy-4916.

Cafiso, S.; Di Graziano, A.; Pappalardo, G. 2013. Using the Delphi Method to Evaluate Opinions of Public Transport Managers on Bus Safety, *Safety Science* 57: 254–263. http://dx.doi.org/10.1016/j.ssci.2013.03.001.

Cafiso, S.; Di Graziano, A. 2012. Evaluation of the Effectiveness of ADAS in Reducing Multi-Vehicle Collisions, *International Journal of Heavy Vehicle Systems* 119(2): 188–206. http://dx.doi.org/10.1504/IJHVS.2012.046834.

Cafiso, S.; La Cava, G.; Pappalardo, G. 2012. A Comparative Analysis of Powered Two Wheelers Crash Severity among Different Urban Areas, *Procedia – Social and Behavioral Sciences* 53: 891–900. http://dx.doi.org/10.1016/j.sbspro.2012.09.938.

Donnell, E. T.; Porter, R. J.; Shankar, V. N. 2010. A Framework for Estimating the Safety Effects of Roadway Lighting at Intersections, *Safety Science* 48(10): 1436–1444. http://dx.doi.org/10.1016/j.ssci.2010.06.008.

Elvik, R. 2010. Why Some Road Safety Problems are More Difficult to Solve than Others, *Accident Analysis and Prevention* 42(10): 1089–1096. http://dx.doi.org/10.1016/j.aap.2009.12.020.

Gstalter, H.; Fastenmeier, W. 2010. Reliability of Drivers in Urban Intersections, *Accident Analysis and Prevention* 42(1): 225–234. http://dx.doi.org/10.1016/j.aap.2009.07.021.

Geurts, K.; Wets, G.; Brijs, T.; Vanhoof, K. 2003. Profiling of High-Frequency Accident Locations by Use of Association Rules, *Transportation Research Record* 1840: 123–130. http://dx.doi.org/10.3141/1840-14.

Heydecker, B. J.; Wu, J. 1991. Using the Information in Road Accident Record, in *Proc. of the 19th PTRC Summer Annual Meeting*. London, United Kingdom.

Johnson, C.; Lewis, J.; Thew, R. 2009. Assessing and Influencing Driver Attitudes in the United Kingdom, *Transportation Research Record* 2138: 46–53. http://dx.doi.org/10.3141/2138-08.

Lin, T.-W.; Hwang, S.-L.; Su, J.-M.; Chen, W.-H. 2008. The Effects of In-Vehicle Tasks and Time-Gap Selection While Reclaiming Control from Adaptive Cruise Control (ACC) with Bus Simulator, *Accident Analysis and Prevention* 40(3): 1164–1170. http://dx.doi.org/10.1016/j.aap.2007.12.009.

Lyon, C.; Gotts, B.; Wong, W.; Persaud, B. 2007. Comparison of Alternate Methods for Identifying Sites with a High Propensity for Specific Accident Type, *Transportation Research Record* 2019: 212–218.

Regan, M. A.; Hallett, C.; Gordon, C. P. 2011. Driver Distraction and Driver Inattention: Definition, Relationship and Taxonomy, *Accident Analysis and Prevention* 43(5): 1771–1781. http://dx.doi.org/10.1016/j.aap.2011.04.008.

Ross, A. J.; Wallace, B.; Davies, J. B. 2004. Technical Note: Measurement Issues in Taxonomic Reliability, *Safety Science* 42(8): 771–778. http://dx.doi.org/10.1016/j.ssci.2003.10.004.

Stanton, N. A.; Salmon, P. M. 2009. Human Error Taxonomies Applied to Driving: a Generic Driver Error Taxonomy and Its Implications for Intelligent Transport Systems, *Safety Science* 47(2): 227–237. http://dx.doi.org/10.1016/j.ssci.2008.03.006.

af Wåhlberg, A. E. 2002. Characteristics of Low Speed Accidents with Buses in Public Transport, *Accident Analysis and Prevention* 34(5): 637–647. http://dx.doi.org/10.1016/S0001-4575(01)00063-X.

Wallace, B.; Ross, A. 2007. Effective Taxonomies in Organisational Safety, in *Decision Making in Complex Environments*. Ed. by Cook, M.; Noues, J.; Masakowski, Y. Chapter 33: 361–372. ISBN 9780754649502.

Yeraguntla, A.; Bhat, C. R. 2005. A Classification Taxonomy and Empirical Analysis of Work Arrangements, *Transportation Research Record* 1926: 233–241. http://dx.doi.org/10.3141/1926-27.

Received 23 February 2011; accepted 21 June 2011