GIDAS-AIDED QUANTIFICATION OF THE EFFECTIVENESS OF TRAFFIC SAFETY MEASURES IN EU 27

Tobias Niebuhr
Jens-Peter Kreiss
Technische Universität Braunschweig, Institut für Mathematische Stochastik
Germany
Stefanie Achmus
Volkswagen AG
Germany
Paper Number 13-0167

ABSTRACT

The paper describes a sound methodology, which leads to a GIDAS-aided quantification of the effectiveness of traffic safety measures for passenger vehicles for countries in EU 27.
Even based on rather limited accident information from national statistics and under the assumption that comparable accidents in different countries lead to comparable accident outcomes, the described procedure allows defining so-called weighting factors. By weighting each single accident of GIDAS a modified GIDAS database can be established which imitates the accident situation in the region or country of interest to some extend. The main ingredients of the proposal are a proper clustering of European countries according to their accident occurence and a statistical procedure (Iterated Proportional Fitting) which allows the prediction of the common distribution (high dimensional) of accident data of the region or country of interest based on available lower dimensional marginal distributions (even one-dimensional). Since the effectiveness of safety systems reliably can be predicted on the detailed GIDAS accident database our procedure allows a prediction of the effectiveness of such systems for other regions or countries. The method is confirmed by real accident data examples, which shed light on the strengths and weaknesses of the proposed procedure.

INTRODUCTION

A reliable quantification of the effectiveness of driver assistant systems and/or traffic safety infrastructure measures is an important task in accident research. For this task police recorded accident data has successfully been used in the literature, e.g. for the effectiveness of electronic stability control (ESC) in passenger cars. See for example Lie et al. (2004), Kreiss et al. (2005), Sugimoto and Sauer (2005), Lie et al. (2006), Zangmeister et al. (2009) or Kreiss et al. (2011).

Dealing with general traffic safety measures, especially recent driver assistant systems or systems which have not been introduced into the market, quantification is a much more delicate problem. An assessment of many safety systems requires much more detailed accident material as available from national statistics. Beside the availability the frame of survey varies significantly, so comparison between different countries is not straightforward. Therefore it is difficult to transfer developed method from one country to another. For recently introduced safety systems the additional problem arises that vehicles equipped with such systems hardly show up in police recorded accident databases as the market penetration of new systems needs time.

One of the most suitable and internationally accepted accident data sets is the German In-Depth Accident Study (GIDAS). GIDAS contains one of the largest in-depth accident data sets ever collected. The GIDAS teams operate in the two German regions Hanover and Dresden and their surrounding areas. Currently accidents are reported to GIDAS teams by police, if and only if injured participants are to be expected. This consequently leads to a substantial bias in the injury severity. It is well known that within GIDAS accidents with severe injury outcomes are overrepresented.

The GIDAS data contains hundreds of categories which carefully have to be reported for every single accident. A deeper insight into the detailed work of the GIDAS teams can be found in Otte et al. (2003). One advantage of GIDAS accident data is that it contains detailed information on the pre-crash phase and on this basis predictions of the possible benefits are possible even for future driver assistant systems.

Motivated by the wish to take advantage of this enormous mass of detailed accident information for regions and countries beyond these restricted German regions of Hannover and Dresden, a proposal for using GIDAS accident data in order to imitate the accident structure in other European countries, or more specifically within EU27, will be presented in this paper. Of course one main problem is that results obtained from GIDAS

cannot be transferred to either Germany or other countries within EU 27 immediately.

A typical approach for transferring GIDAS to other regions, e.g. the whole of Germany or EU 27, is the method of projection factors. The main objective of this projection method is to obtain country specific weights for the GIDAS data such that the modified accident database imitates the accident occurrence of the region or country of interest. Knowing the effectiveness of a specific safety system on the basis of the original GIDAS data the effectiveness can be transferred to the target region easily. In principle this approach could be carried through for all European countries. The main assumption underlying this approach is that comparable accidents in two different countries result in comparable injury outcomes, at least on the level of accident severity distribution. Comparable accident situations can be interpreted as nearly constant intrinsic accident configurations. As a consequence hidden relations of categories or missing information in different countries are assumed to coincide. However, it should explicitly be mentioned that such an approach transfers intrinsic accident relations of German accident data to other countries within EU 27.

The application of the method to all single European countries is rather time-consuming and difficult to handle in a reasonable way. To avoid an enormous computational expenditure on the one hand and to avoid small sample problems of small countries on the other hand a division of European countries into groups of similar accident occurrence will be favourable. A proposal for such an clustering procedure will be developed and obviously this clustering will lead to an essential reduction of the expenditure on computation. The countries of each resulting cluster have comparable accident occurrences to a sufficient. Thus the effectiveness of driver assistance systems can be quantified cluster by cluster.

For this reasons one of the most important issues is the development of an appropriate clustering procedure resulting in as dissimilar as possible clusters of as similar as possible countries regarding accident occurrence and traffic. In other words, the developed procedure should lead to a heterogeneous classification of homogeneous groups of countries.

For some European countries we have to face the situation that only one dimensional distributions of a small number of accident related factors are available. Think for example of the distribution of accident location (*rural* and *urban*), the distribution of accident type or accident severity. In order to obtain country specific weights (actually cluster specific) we need the common multidimensional distribution concerning the mentioned factors. In case of these higher dimensional distributions not being available for the region of interest we make use of a method called Iterative Proportional Fitting (see Deming and Stephan (1940)). This procedure – on the assumption that main intrinsic accident relations stay nearly constant – allows the transfer of multidimensional distributions from GIDAS to these clusters on the basis of original marginal distributions. Therefore a cluster-wise accident database can be constructed and this eventually leads to a possibility for the evaluation of the effectiveness of safety systems for the whole of the EU27.

Our results show that a reliable assessment of the effectiveness of safety measures can be obtained via the described GIDAS-aided quantification procedure. Relevant real world examples will be presented and the similarity of recorded accident data to projection results from GIDAS will be investigated.

Of course, following the suggested proposal, one could not expect absolutely precise evaluations of the effectiveness of all safety measures for the whole of EU 27. It turns out that based on the obtained clusters of accident-similar countries it would be advantageous to have multidimensional distributions of the considered accident parameters of at least one representative region within the cluster in order to be able to calibrate the procedure of projection of GIDAS accident data onto the cluster. Moreover it would be extremely advantageous to have detailed accident data of at least one representative region within every cluster so the projection could be done from this (new) detailed accident database onto the cluster.

Data

Case studies in all sciences require reliable data. Especially the evaluation of traffic safety measures should be based on very detailed accident material. Otherwise it is nearly impossible to secure an assessment of only the considered safety system. As the frame of survey varies for different European countries, national reports are hard to handle in Europe-wide investigations. Relevant accident categories have to coincide for all countries especially if a clustering procedure shall be installed. While there is no in-depth European data base, the International Road Traffic and Accident Database (IRTAD) will be used. The IRTAD database includes accident and traffic data along with other safety indicators for 29 countries (see www.internationaltransportforum.org). The data base is a collection of national statistics and thus, there is no unique coding for accident

categories. To assure the reliability of later investigations and to avoid different country-specific coding (especially for the injury severity) all following investigations are restricted to fatally injured road users.

Further the distribution of traffic participation for these fatally injured road users of the years 2006 to 2009 is used. In detail the distribution of the traffic participation is limited to the categories *bicycle*, *moped*, *motorcycle*, *passenger car* and *pedestrian*. Additionally the location of accidents divided into the categories *urban* and *rural* will be used. Altogether distributions for Austria (A), Belgium (B), Czech Republic (CZ), Denmark (DK), Finland (FIN), France (F), Germany (D), Hungary (H), Italy (I), Netherlands (NL), Norway (N), Poland (PL), Portugal (P), Slovenia (SLO), Spain (E), Sweden (S) and Switzerland (CH) can be obtained for the ensuing exemplary studies. This is close to EU17.

It should be noted that the method is not limited to the EU17. If there is a wider database for the whole of EU27, the procedure will work as well.

The four years (from 2006 to 2009) will be treated separately and the partial results will be combined afterwards. Using this yearly approach one will obtain the possibility to identify trends and randomly occurring clusters. Furthermore, stability of clusters will be shown if clusters consist of the same countries for several years.

In addition to the IRTAD and GIDAS databases a national accident report of Hungary from 2008 will be used for exemplary applications. The example will use a three-dimensional table containing the injury severity in categories *killed*, *seriously injured*, *slightly injured*; the location in categories *rural*, *urban*; and the type of collision in categories *head on*, *rear-end*, *due to crossing*, *with parked*, *skidding*, *with pedestrians.*

GIDAS-AIDED PROJECTION TO CLUSTERS OF EUROPEAN COUNTRIES

Quantification of the effectiveness of traffic safety measures is of essential interest in accident research and for car manufacturers. While case-specific proposals exist for several regions and countries, a Europe-wide evaluation is much more challenging. On the one hand a single Europe-wide evaluation may not be able to point out local differences of the intrinsic accident structure. On the other hand a partition into too many and too narrow European regions often causes small sample difficulties or the problem that data of sufficient detail is not available for each region. Furthermore, such a big amount of regions results in a lot of expenditure in computations as they have to be carried out for every region separately. Thus, the motivation for a classification of Europe into so-called clusters arises quite naturally. Ideally, every cluster will consist of countries containing equal or at least strongly comparable, say similar intrinsic accident structure.

It should be emphasized that the focus is on the development of a general and flexible evaluation method instead of concrete quantifications of the effectiveness of specific driver assistant systems. However, examples will be given later on.

Step 0: How to process accident data (if necessary)?
Step 1: How to obtain high- dimensional data using the IPF?
Step 2: How to decide distributions to be similar or dissimilar?
Step 3: How to cluster countries using the test for similarity of distributions?
Step 4: How to quantify the effectiveness of driver assistance systems in given clusters?

Figure 1. Working packages.

The development of the procedure briefly is divided in step-wise working packages. The first-of-all working package, namely Step 0, is about the processing of data. This step is optional in terms of leaving out if data of sufficient detail and same form is available for every country. Otherwise the data will have to be processed (e.g. consider two countries with different coding of injury severity). Usually this working package contains the screening of the reliability of data and/or the coding of categories of accident parameters. In the following no further comments on this step will be given because there is no general method for data editing. Step 1 will describe how the Iterative Proportional Fitting Procedure (IPF) can be used to generate (approximated) country-specific multi-dimensional distributions of accident parameters. Probably step 2 is the most innovative and essential working package for the quantification approach in its total. It deals with the development of an approach for testing for similarity of distributions. While mathematical standard tests for equality often fail in application cases, a generalization to testing for similarity will briefly be pointed out. This theoretical result creates a basis for the ensuing clustering procedure, namely step 3. The clustering method will be given as a pseudo-algorithm wherein the test for similarity plays a pivotal role. The clustering procedure will lead to a classification in terms of the countries within each group being as similar as possible and the groups being as dissimilar as possible. Further on, Step 4 is the projection approach. Depending on a given

cluster configuration the projection method can be applied to each cluster. Thus one will obtain cluster-wise evaluations of the effectiveness of the driver assistance system of interest which can be aggregated in a final step if liked. Another possibility for evaluation is to aggregate the received weighting factors to projection factors for the whole of Europe. Based on this projection the effectiveness of driver assistance systems can be determined for the whole of Europe. Each Step of the presented method will be discussed using elementary examples.

Step 1: Iterative proportional fitting

Any clustering requires distributions of the same form for every country. As one could expect, the higher dimensional the distributions are the better the classification is (assuming that we consider essential parameters). However, for some countries only low-dimensional data is available. A review of the Iterative Proportional Fitting Procedure (IPF) will be given to show how higher dimensional data can be computed using low-dimensional statistics. In-depth data bases like e.g. GIDAS will be a helpful tool to generate initialization tables during the IPF algorithm. Using IPF, based on several one-dimensional distributions (e.g. the three distributions of the type of collision, the injury severity and the accident location) one will obtain one higher dimensional objective distribution (e.g. the corresponding three-dimensional distribution). This obejctive distribution is required for each country with respect to ensuing country-wise comparisons.

It should explicitly be pointed out that Step 1 is required if and only if the multidimensional objective distribution is not available for all countries of interest. In this instance the IPF defined by Deming and Stephan (1940) takes country-specific marginal distributions of lower dimension and computes a higher dimensional distribution which fulfills the marginal restrictions. The procedure depends on the initialization table. While computing, the procedure keeps the inner dependence structure of the initialization table as much as possible. Obviously the solution will depend on the used initialization table. Assuming that all countries of the European Union have an equal intrinsic accident structure the GIDAS data base can be used for initialization. As a consequence all inner dependencies of GIDAS which have not been controlled by the used parameters throughout the IPF are assumed for Europe in its total.

The IPF algorithm of Deming and Stephan (1940) operates as follows. For reasons of simplicity the review is restricted to the two-dimensional case. Nevertheless, higher dimensional cases can be treated in an analogous way.
Let

$$(m_{i\#}\,; i = 1, \dots, I) \tag{1}$$

and

$$(m_{\#j}\,; j = 1, \dots, J) \tag{2}$$

be two one-dimensional country-specific marginal distributions (e.g. the accident location (*urban, rural; I=2*) or the injury severity (*fatally, severely, slightly injured; J=3*)). The actual underlying two-dimensional distribution

$$(m_{ij}\,; i = 1, \dots, I\,; j = 1, \dots, J) \tag{3}$$

may be unknown. Further an initialization table is required. With the help of GIDAS one can extract a sample

$$(n_{ij}\,; i = 1, \dots, I\,; j = 1, \dots, J) \tag{4}$$

imitating the true two-dimensional distribution. Of course, the total number of absolute frequencies usually differs. However, the algorithm only relies on the inner relation of the accident configurations. Now the IPF algorithm adjusts the initialization table iteratively by the use of alternating row-wise

$$n_{ij}^{*} = n_{ij}\left(\frac{m_{i\#}}{n_{i\#}}\right) \tag{5}$$

and column-wise fittings

$$n_{ij}^{*} = n_{ij}\left(\frac{m_{\#j}}{n_{\#j}}\right). \tag{6}$$

The convergence of the procedure to a stable limit distribution was discussed by Stephan (1942) and Rüschendorf (1995). It should be emphasized that the IPF limit is an approximated version of the true two-dimensional distribution depending on the assumption that the country-specific and the GIDAS-specific intrinsic accident relations are equal. Thus, sample problems in GIDAS will be transferred implicitly to the approximated distribution. In other words the IPF is an approximating procedure which contains coincidence by the use of data sets like e.g. GIDAS. Thus, exact results cannot be expected. Nevertheless, the IPF is a helpful tool and gives a possibility to handle dimension problems of the data.

Additional results and discussions on the IPF can be found in e.g. Fienberg (1970), Wong (1992) and Norman (1999). Latter ones especially give advice

how to handle the difficulties of occurring zeros in initialization tables.
At this point, it should be mentioned that the IPF is necessary if and only if the required high-dimensional distributions are not available for each country. Using IPF one obtains for every country higher-dimensional distributions of the concerned accident parameters which are the basis for further investigations.

Example 1: An exemplary enumeration is given by applying the IPF to Hungarian data of the year 2008 taken from the official national report (see Hungarian Central Statistical Office (2009)). Using the marginal distributions given in the Tables 1 to 3 a three dimensional case study is computed. The GIDAS data is used for initialization. The resulting IPF limit and thus approximately proposed three-dimensional Hungarian distribution is given in Table 4. A comparison to the actual higher-dimensional Hungarian distribution is possible using Table 5 obtained by the national report.

Step 2: Testing for similarity of distributions

The most usual way to decide whether two distributions are equal is the χ^2-test. However, in application cases using accident data the test rejects the hypothesis of equal distributions in nearly every situation. This is due to noisy data and coding problems along with several other reasons. If distributions (countries) shall be clustered, they do not necessarily have to be equal though in detail equality would represent the optimum. In the context of clustering procedures it is mainly sufficient to have the possibility to decide whether or not distributions are similar. This directly motivates the task to find a method indicating similar distributions.

Table 1.
Hungarian marginal distribution for the type of collision in 2008.

type of collision					
head-on	rear-end	due to crossing	with parked	skidding etc.	with pedes-trians
3533	3886	7564	4048	2958	3539

Table 2.
Hungarian marginal distribution for the injury severity in 2008.

injury severity		
fatal	severe	slight
962	6001	11892

Table 3.
Hungarian marginal distribution for the location of the collision in 2008.

accident location	
urban	rural
15927	9601

Table 4.
Resulting IPF-limit for the three-dimensional Hungarian distribution of the year 2008 using GIDAS as initialization table.

	urban			rural		
	fatal	severe	slight	fatal	severe	slight
head-on	22.4	362.6	1053.3	162.1	724.9	1207.9
rear-end	10.7	149.7	1369.0	59.2	450.5	1846.9
crossing	64.6	1339.0	4445.4	52.5	420.2	1242.3
parked	54.4	676.3	2545.1	54.6	211.0	506.6
skidding	37.7	212.8	325.3	166.7	869.4	1346.2
pedestrian	204.6	1471.5	1626.9	60.9	81.4	83.7

Table 5.
Actual three-dimensional Hungarian distribution of the year 2008 obtained of the national report.

	urban			rural		
	fatal	severe	slight	fatal	severe	slight
head-on	33	387	971	181	723	1238
rear-end	27	393	1710	65	428	1263
crossing	82	1420	4714	57	359	932
parked	64	548	1458	118	587	1273
skidding	22	313	595	52	556	1420
pedestrian	160	1138	1892	83	140	126

While the little difference between equality and similarity is clear in interpretation, from a mathematical point of view it is of essential effect. The testing problem, whether distributions are equal, is well-studied (see Hartung (2005) for a review and detailed examples). Especially the χ^2-test and the Kolmogorov-Smirnov-test are well understood. However, only less satisfying results dealing with the similarity of distributions exist.

One recent analysis on similarity of distributions is given in the work of Niebuhr et al. (2011). Their main results are shortly reviewed below. For reasons of simplicity technical details are neglected in here.

Transferring the idea of the standard χ^2-test to the more general non-central χ^2-pattern their approach introduces a correction term representing additional tolerance to the standard test hypothesis of strict identity of two distributions. From a less statistical point of view one could interpret this approach as a modification of the test of equality. The modification itself, represented by a so-called term of tolerance (see equation (11)), weakens the claim of equal distributions to at least similar ones.
In consequence also the test statistic has to be changed in comparison to the standard test, see (9). The additional term represents tolerated deviation. Vividly, this tolerance is interpreted as an overlaying jittering to the original distribution. From another point of view, instead of comparing a sample to one specific distribution, the sample is compared to all possible distributions close to the specific distribution or in other words to a class of similar distributions.

The mathematical construction of the correction term includes both small overall deviations and larger deviation in few cells of the distribution. The resulting test for similarity is reviewed below. It should explicitly be emphasized that the developed test statistic is consistent to the one of the standard χ^2-test for equality if the additional tolerance level is set to zero. Thus, similarity without tolerance clearly becomes equality. The test is as follows. Let

$$(n_{ij}\,; i = 1, \dots, I\,; j = 1, \dots, J) \tag{7}$$

be the assumed original distribution (e.g. yearly national accident distributions of Hungary, see Table 5, containing the type of collision, the accident location and the injury severity) and let

$$(n^*_{ij}\,; i = 1, \dots, I\,; j = 1, \dots, J) \tag{8}$$

denote the assumed sample of (7) (e.g. a GIDAS-aided imitation of the national statistic or a computed IPF approximation, see Table 4). Further, n may describe the total number of cases, which is the sum of the n_{ij}. As already mentioned, the test statistic χ^2 from the test for equality plays an important role in this generalization. It is defined via

$$\chi^2 = \sum_{i,j} \frac{\left(n^*_{ij} - n_{ij}\right)^2}{n_{ij}}. \tag{9}$$

The null hypothesis H_0 and alternative H_1 are given as expected:

H_0: The distributions are similar.
H_1: The distributions are not similar.

The accompanying test decision is as follows. The null hypothesis H_0 will be rejected if

$$\frac{\chi^2 - \lambda - IJ}{\sqrt{\frac{2IJ + 4\lambda}{2IJ}}} - IJ \;>\; \chi^2_{IJ,1-\alpha}, \tag{10}$$

where

$$\lambda = u^2_{0.5+\beta} \sum_{i,j} n_{ij} \left(1 - \frac{n_{ij}}{n}\right) \tag{11}$$

and u_δ describes the δ-percentile of the standard normal distribution. The additional tolerance is introduced by β which implicitly scales λ and thus, at least has a decreasing effect on the test statistic value. If β is set to zero, the test will simplify to standard χ^2 -test.

Careful analysis gives the fact that any two distributions are decided to be similar if the tolerance β is generously increased. More detailed, there is a specific tolerance value where the test decision switches from rejection to non-rejection. This specific and also unique value may be denoted as the *p*-value of β in the following. Depending on a fixed level of significance α, one is able to obtain the *p*-value β^* by solving the implicit equation

$$\chi^2 - u^2_{0.5+\beta} \sum_{i,j} n_{ij} \left(1 - \frac{n_{ij}}{n}\right) - \chi^2_{\alpha,IJ} = 0. \tag{12}$$

However, this *p*-value of β represents a kind of dissimilarity measure which will be of essential importance for the ensuing clustering procedure.

Example 2: Besides using the *p*-value as a measure of dissimilarity (see Example 3 for its application in the clustering method), the test for similarity of distributions is a helpful tool in the evaluation of the quality of projections. Returning to the situation of Example 1, the *p*-value gives advice about the quality of the IPF approximation of the Hungarian

statistic. A direct application using Table 4 as the sample and Table 5 as the original distribution leads to

$$\beta^* = 0.133 \tag{13}$$

what corresponds to an required additional tolerance of 13.3 %. It has to be individually decided if this value is acceptable or not. However, empirical investigations of the IPF done by the authors let them conclude that tolerance levels up to 15 % seem appropriate.

Step 3: Clustering procedure

The test for similarity of distributions and especially the corresponding p-value β^* allows evaluating the degree of similarity of country-specific accident data. Using β^* as a measure of dissimilarity one is able to identify the two countries with the most similar distributions out of a pool of countries. Based on this thought a clustering procedure is presented in the following. The procedure can be described as an Iterative Nearest-Neighbor-Centroid-Procedure. In terms of a greedy heuristic the most similar distributions are clustered (nearest-neighbor) and their absolute frequencies are summarized. Therefore, the distribution of the resulting cluster is a weighted interpolation of the two used original distributions (centroid). Finally iterative application leads to one single cluster containing all countries. Further investigations on nearest-neighbor clustering procedures can be found in Wong and Lane (1983).

It should explicitly be pointed out that the test presented in Step 2 compares a given sample with an actual distribution. The clustering procedure instead measures the dissimilarity of samples of two countries in each step. Formally, the statistic of country A is seen as a sample of the distribution of country B and vice versa. Afterwards the mean of both computations is obtained as an approximated p-value. Thus an underlying mutual distribution is assumed that is located between the distributions of the two country-specific samples. The step-wise clustering procedure is as follows:

Step 0:	Start with the distributions of m countries/clusters. Stop if m is equal to 1.
Step 1:	Compute the p-values β^*_{AB} for all possible pairs of countries/clusters A and B.
Step 2:	Compute the mean of β^*_{AB} and β^*_{BA} for all pairs of countries/clusters A and B.
Step 3:	Note the minimum mean of p-values and the corresponding countries/countries.
Step 4:	Aggregate the raw data of the two countries with minimal dissimilarity.
Step 5:	Go back to Step 0 with m replaced by m-1.

Figure 2: Clustering algorithm.

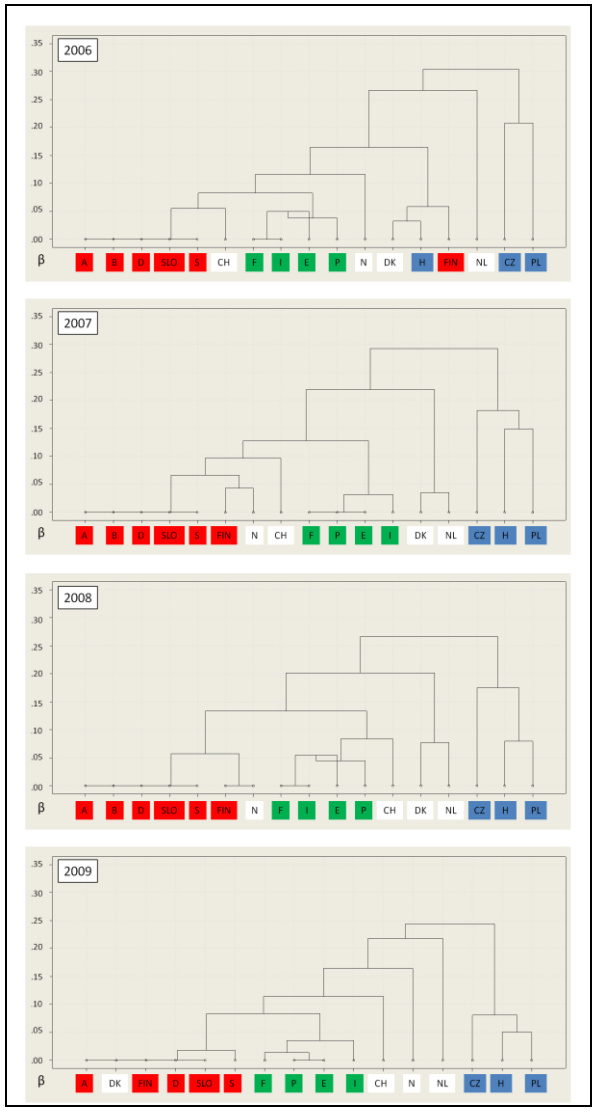

Figure 3. Development of the EU17 clusters from 2006 to 2009 for fatal injuries in dependence on the traffic participation. (Remark: There is no data available for Belgium in 2009).

A visualization of the clustering procedure can be done by a so-called dendrogram (see the plots in Figure 3). The dendrogram has to be read bottom-up. Starting with all countries the development of the clustering is visualized step by step. Each dissimilarity value is represented by the height of the horizontal lines and can be interpreted as the clustering level.

Because the p-value β^* gives no indication about a potential test decision, one has to fix the maximum level of tolerance. This corresponds to a level of significance of the test. The level can individually be chosen and it can easily be integrated into the dendrogram plot as a horizontal line. All clusters below this horizontal line are accepted to be similar; all clusters above the level are rejected. Imagine deleting the graph above this horizontal

line. All still connected countries are accepted to be similar and to define clusters.

As already discussed in the introduction, there is no advantage of too many or too few clusters. To obtain a satisfying number of clusters, a more useful approach is to fix the accepted level of tolerance afterwards. This alternative proposal is used in Example 3 to determine Figure 4 in dependence of Figure 3.

Example 3: An exemplary clustering is given. To assure the reliability of this example, the investigation is restricted to fatally injured road users. The distributions of the traffic use of these fatally injured road users for the years 2006 to 2009 were extracted of the IRTAD data base. In detail the categories *bicycle*, *moped*, *motorcycle*, *passenger car* and *pedestrian are used for traffic use*. Using the aforementioned clustering procedure the development of the similarities (2006-2009) can be seen in Figure 3. All computations were implemented using the free software *R*.

The year-wise separated investigation allows to identify stable clusters and trends in the data. Figure 3 shows two very stable clusters (red and green) and a group of countries coming closer to each other (blue). White countries cannot be classified to one specific cluster. One reason may be the small total numbers of accidents and especially fatalities of these countries. In these countries, small fluctuation in the data causes significant changes in the distributions.
It should be emphasized that this example doesn't use an a priori fixed level of tolerance. The alternative proposal is used to determine a satisfying number of clusters.

This four year analysis of fatal traffic participation ends up in Figure 4. It seems that the localization of the countries is essential for the distribution of the traffic use of fatally injured people.

Figure 5 is developed using exactly the same approach based on the distribution of the accident location (*urban* and *rural*) of the fatalities instead of the distribution of road use. Again the IRTAD database was used. It should be mentioned that these two-point-distributions of the location are very sensitive. However, certain clusters can be indicated (see Figure 5).

Both examples give the impression that the country localization is essential for several accident categories. Further, especially small countries with few accident samples fluctuate a lot and cannot be classified to one specific cluster.

It should be mentioned that at this point of research it is not clear if aggregation is possible and if yes how to aggregate several results like Figure 4 and Figure 5 to one final solution. One would expect that equal clusters would be kept but differences in the clusters are hard to handle and at this point of research there are no specific results.

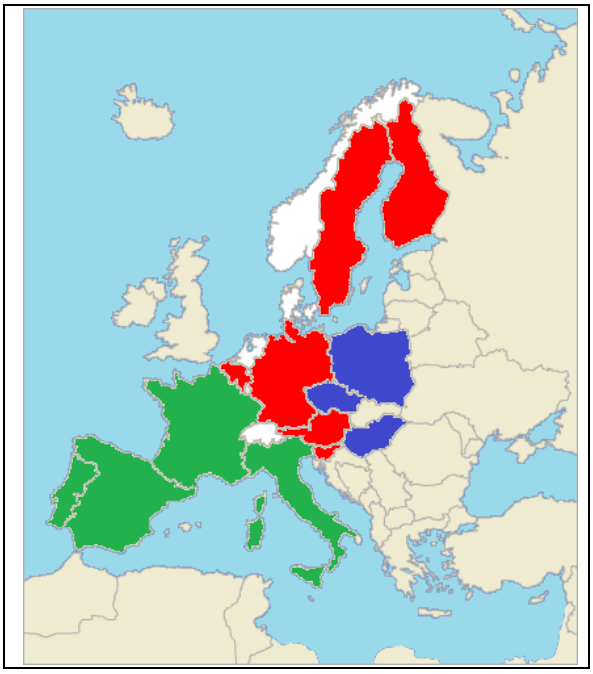

Figure 4. EU17 clusters for fatalities in dependence on the traffic participation as consequence of Figure 3. (white: non-allocated countries).

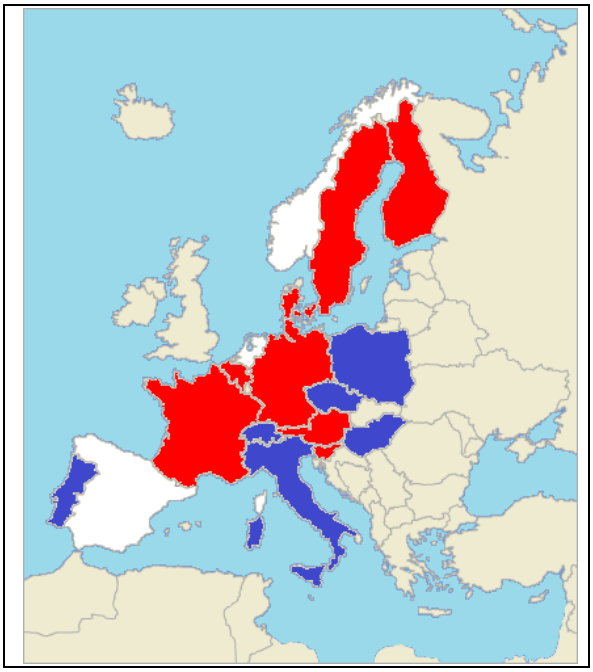

Figure 5. EU17 clusters for fatal injured participants in dependence on the accident location. (white: non-allocated countries).

Step 4: Method of projection

A common way to evaluate driver assistance systems within Germany is to imitate the distribution of interest on the basis of the GIDAS data set. Thus, one has the original distribution (e.g. a two-dimensional distribution containing information on the accident location (*rural*, *urban*) and the injury severity (*fatal*, *severe*, *slight*)) and one distribution mimicked by GIDAS. Obviously, the mimicked one will differ from the original distribution. To solve this differing so-called projection factors can be computed in a cell-by-cell device by dividing the actual number of cases through the number of cases in the mimicked GIDAS table (in the hypothetical two-dimensional example six factors have to be computed, i.e. [*urban,rural] x [fatal,severe,slight]*). Further, using these projection factors the GIDAS-based imitation is weighted in terms of weighting each involved person by the projection factor instead of the weight one. The result is a weighted GIDAS data set that fulfills the correct original distributions. The evaluation of a driver assistance system can be done by using the weighted GIDAS version.

Justification of this procedure is given if inner relations of GIDAS are assumed to represent relations in other regions as well. This is based on the fact that weighting regarding certain parameters (e.g. injury severity, accident location, road use) implicitly influences further parameters (e.g. type of accident). Obviously, positive correlated variables are influenced in the right direction. Nevertheless, misjudgment may occur in few cases and has to be taken into account. However, this straightforward approach is reasonable for empirical investigations.

Furthermore smoothing effects within Europe-wide investigations may hush up local characteristics. The wish to avoid such possibly conflicting regional effects of the weighting a posteriori leads to another motivation for an appropriate clustering procedure.

As a consequence, the projection idea can be applied to each cluster identified in Step 3. While the clustering procedure separates groups with similar distributions one assumes similar intrinsic relations in each group, and thus the projection can directly be applied. The evaluation of a driver assistance system can be evaluated in each cluster separately and possibly be combined to an all-European result afterwards. Another possibility to obtain Europe-wide results would be the aggregation of cluster-specific projection factors to all-European projection factors. Thus evaluations could base on a GIDAS data set adapted to Europe in its total.

It should explicitly be pointed out that the worth of this approach extensively depends on type and scope of the original distribution. Obviously, higher-dimensional distributions result in much better projections than lower-dimensional ones. Thus, the IPF procedure computing higher-dimensional distributions is of far use for this approach. The combination of the technical IPF, the in-depth GIDAS data base and the cell-by-cell projection creates a powerful tool for the evaluation of traffic safety measures, especially for the whole of Europe. Some countries may not offer such detailed data than others do. However, the IPF can be applied and thus the quality of the projection may increase. The only theoretical limit for this approach is given by the dimension of the GIDAS data set and the availability of national statistics.

<u>Example 4:</u> While the presented procedure works for any dimension of data, for reasons of simplicity this example is restricted to one single dimension. Nevertheless, the proposal can be directly transferred to higher dimensional data. The situation of Example 1 is shortly reviewed. The Hungarian distribution of the injury severity is given in Table 2. It is divided into three categories, namely *fatally*, *severely* and *slightly injured* traffic users, with frequencies

$$(962, 6001, 11892) = (5.1\%, 31.8\%, 63.1\%). \tag{14}$$

The corresponding GIDAS distribution is given by

$$(775, 8590, 24349) = (2.3\%, 25.5\%, 72.2\%). \tag{15}$$

Obviously, the Hungarian national statistic differs from the GIDAS one. In consequence, the GIDAS sample has to be weighted. Every case (person) of the data set is valued depending on its injury severity category. Thus, every *fatal* case in GIDAS obtains the weighting factor

$$962/\,775 = 1.24, \tag{16}$$

every *severe* case obtains the weighting factor

$$6001/8590 = 0.70 \tag{17}$$

and every *slightly injured person* the weighting factor

$$11892/24349 = 0.49. \tag{18}$$

Equations (16)-(18) represent the cell-by-cell projection factors. Again it is commented that in this example the weighting is restricted to only one accident property with three categories, namely the injury severity. However, projections containing

any number of properties with any number of categories would follow an analogous approach.

Assuming a hypothetical driver assistance system which is able to prevent 25% of all urban fatalities an exemplary evaluation is computed. Using the weighted GIDAS version as described before a two-dimensional table is build, see Table 6. The effect of the hypothetical system is restricted to one single cell in Table 6. It causes a reduction of 25% of all urban fatalities, thus

$$341.1 * 0.25 = 85.3 \qquad (19)$$

fatalities will be avoided. The total number of fatalities decreases from 962 to at least 876.7 that is a percentage decreasing of

$$1 - 876.7/962 = 0.089, \qquad (20)$$

respectively 8.9%. Note that the described projection approach can directly be transferred to other driver assistance systems and thus one has the possibility to quantify the effectiveness of systems in an appropriate way.

Table 6.
Two-dimensional table from GIDAS with Hungarian weights containing accident location and injury severity.

	fatal	severe	slight	total
urban	341.1	3613.9	8693.0	12648.2
rural	620.6	2387.1	3199.0	6206.8
total	962	6001	11892	18855

CONCLUSION & DISCUSSION

A method of the evaluation of the effectiveness of driver assistance systems was presented. Using a step-wise approach it was pointed out how data can be processed to mutual form by the use of the IPF. Though case studies result in very satisfying results, the quality and sensitivity has to be quantified. Further, based on a developed testing procedure for the similarity of distributions, a powerful clustering algorithm was described. The intrinsic relations of higher-dimensional distributions within the obtained clusters result in similar characteristics. Nevertheless, it is not clear in detail how to combine several partial solutions to an overall result.

The method should be expanded to higher-dimensional distributions and all countries within EU 27. The probably largest difficulty will be the common database. However, the combination of the IPF, the GIDAS data base and the projection procedure creates a trustworthy method that seems appropriate for European-wide investigations.

REFERENCES

Bishop, Y.M.M. and Fienberg, S.E. (1969): Incomplete Two-Dimensional Contingency Tables. Biometrics, Vol. 25, No. 1, pp. 119-128.

Deming, W.E. and Stephan, F.F. (1940): On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals are Known. The Annals of Mathematical Statistics, Vol. 11, No. 4 (Dec., 1940), pp. 427-444.

Fienberg, S.E. (1970): An Iterative Procedure for estimation in Contingency Tables. The Annals of Mathematical Statistics, Vol. 41, No. 3, pp. 907-917.

Hartung, J. (2005): Statistik: Lehr- und Handbuch der angewandten Statistik (14nd Ed.). Oldenbourg.

Hungarian Central Statistical Office (2009): Traffic accidents 2008. ISSN 0319 2824, the National Accident Preventive Commettee.

Kreiss, J.-P., Schüler, L. and Langwieder, K. (2005): The Effectiveness of Primary Safety Features in Passenger Cars in Germany, Proc. 19th Int. Tech. Conf. Enhanced Safety Vehicle, Washington D.C. (USA), Paper number 05-0145.

Kreiss, J.-P., Stanzel, M. and Zobel, R. (2011): On the Use of Real-World Accident Data for Assessing the Effectiveness of Automotive Safety Features – Methodology, Timeline and Reliability, Proc. 22nd Int. Tech. Conf. Enhanced Safety Vehicle, Washington D.C. (USA), Paper number 11-0054.

Lie, A., Tingvall, C., Krafft, M. and Kullgren, A. (2004): The Effectiveness of ESP (Electronic Stability Program) in Reducing Real Life Accidents, Traffic Injury Prevention, Vol. 5, No. 1, pp. 37-41.

Lie, A., Tingvall, C., Krafft, M. and Kullgren, A. (2006): The Effectiveness of ESC (Electronic Stability Control) in Reducing Real Life Crashes and Injuries, Traffic Injury Prevention, Vol. 7, No. 1, pp. 38-43.

Niebuhr, T., Achmus, S. and Kreiss, J.-P (2011): Testing for Similarity of Distributions. Technical Report, TU Braunschweig.

Norman, P. (1999): Putting Iterative Proportional Fitting on a Researcher's Desk. Working Paper, University of Leeds.

Otte, D., Krettek, C., Brunner, H. and Zwipp, H. (2003): Scientific Approach and Methodology of a New In-Depth-Investigation Study in Germany so called GIDAS, Proc. 18th Int. Tech. Conf. ESV, Nagoya (Japan), Paper number 03-0161.

Rüschendorf, L. (1995): Convergence of the Iterative Proportional Fitting Procedure. The Annals of Mathematical Statistics, Vol. 23, No. 4, pp. 1160-1174.

Stephan, F.F. (1942): An Iterative Method of Adjusting Sample Frequency Tables when Expected Marginal Totals are known. The Annals of Mathematical Statistics, Vol. 13, No. 2, pp. 166-178.

Sugimoto, Y. and Sauer, C. (2005): Effectiveness Estimation Method for Advanced Driver Assistance System and its Application to Collision Mitigation Brake System, Proc. 19th Int. Tech. Conf. Enhanced Safety Vehicle, Washington D.C. (USA), Paper number 05-0148.

Wong, D.W.S. (1992): The Reliability of Using the Iterative Proportional Fitting Procedure. The Professional Geographer, 44: 3, 340 - 348.

Wong, M.A. and Lane, T. (1983): A k-th Nearest Neighbour Clustering Procedure, Journal of the Royal Statistical Society, Series B, Vol. 45, No. 3, pp. 362-368.

Zangmeister, T., Kreiss, J.-P., Page, Y. and Cuny, S. (2009): Evaluation of the safety benefits of passive and/or on-board active safety Applications with mass accident data-bases, Proc. 21st Int. Tech. Conf. Enhanced Safety Vehicle, Stuttgart (Germany), Paper number 09-0222.