Methodologies for Aggregating Traffic Conflict Indicators

Karim Ismail
Lecturer, Department of Civil Engineering
Carleton University
1125- 3432 Colonel By Drive
Ottawa, ON, Canada K1S 5B6

Tarek Sayed
Professor and Distinguished Scholar
Department of Civil Engineering
The University of British Columbia
6250 Applied Science Lane
Vancouver, BC, Canada V6T 1Z4

Nicolas Saunier
Assistant Professor
Department of civil, geological and mining engineering
Ecole Polytechnique de Montréal,
C.P. 6079, succ. Centre-Ville
Montréal, QC, Canada H3C 3A7

Word Count

- Text (in Abstract)	=	202
- Text (in Paper)	=	5791
- Tables+Figures: (9 x 250)	=	2225
- Total	=	8215

Methodologies for Aggregating Traffic Conflict Indicators

Abstract

Various objective conflict indicators have been proposed in the literature in order to measure the severity of traffic events. Objective conflict indicators measure various spatial and temporal aspects of proximity on the premise that proximity is a surrogate for severity. It is argued in this paper that these aspects of severity may be partially overlapping and in some cases independent. Two sets of conflict indicators are used in this study. The first set of conflict indicators requires the presence of a collision course common to the interacting road users. The second set of conflict indicators measures severity in terms of mere temporal proximity between road users. The goal of this study is to demonstrate that the integration of the different severity cues, provided by each conflict indicator, can be performed in order to better reflect the true, yet unobservable, severity of traffic events. This study proposes a methodology to aggregate the event-level measurements of conflict indicators into a safety index. First, individual conflict indicator measurements are mapped into severity intervals [0,1]. Second, these severity indices are aggregated to a safety index that includes both individual severities and exposure. The methodology presented in this paper is applied on individual measurements of pedestrian-vehicle conflicts.

BACKGROUND

Traditional road safety analysis has often been undertaken using historical collision records which suffer from quality and availability problems. As well, this is a reactive approach where collisions have to occur in order to properly conduct safety analysis. Moreover, in order to evaluate the effectiveness of safety programs in reducing collisions, adequate before and after periods of observation have to be allowed in order to conduct a statistically valid analysis. Meanwhile, society bears the social burden of road collisions. These limitations motivate the development of surrogate safety measures. A key type of surrogate safety analysis is the traffic conflict technique (TCT) which involves the observation and analysis of traffic conflicts or near misses. The definition of a traffic conflict has evolved since its first proposition by (1). A widely accepted conceptual definition of a traffic conflict is “an observable situation in which two or more road users approach each other in space and time to such an extent that there is a risk of collision if their movements remained unchanged” (2). Traffic conflicts possess important advantages over road collisions for the purpose of road safety analysis. Traffic conflicts are more frequent, can be clearly observed and much less costly, if any, than road collision. Moreover, the observation and analysis of the positions of road users involved in traffic conflicts may provide insight into the failure mechanism that leads to collision.

Despite the well-recognized advantages of the TCTs, they suffer from: the inter- and intra-observer reliability and the high cost required to train field observers and to collect conflict data. Inter-observer reliability concerns the variance across observers regarding the evaluation of a traffic event, *i.e.*, a conflict or not. Intra-observer reliability concerns the inconsistency of a trained observer's assessment of the same event if displayed in different contexts or at different times. These limitations have inhibited a wider application of the technique. Recently, methods to automate the analysis of traffic conflicts have been proposed and shown to be reliable (3) (4) (5) (6). These methods draw on the extensive work in the field of computer vision to analyze data collected by video sensors which provide rich, detailed, inexpensive, and permanent observations of traffic scenes. The final product of video analysis is road user tracks (sequence of positions in space and time). Extracting road user tracks from video sequence enables road safety analysis at a much higher spatial and temporal resolution than current techniques in use. The main advantage of computer vision techniques is the potential to collect microscopic road user data at a degree of automation and accuracy that cannot be feasibly achieved by manual or semi-automated techniques. Microscopic road user data can be used to draw objective inference on their proximity to the risk of collision. The objectiveness and automation of conducting traffic conflict analysis using computer vision techniques empowers the two main challenges of traditional observer-based traffic conflict analysis: cost and subjectivity.

A multitude of conflict indicators have been developed for the purpose of traffic conflict observation. Example of these indicators are objective conflict severity measures, whether deterministic objective conflict indicators, such as Time to Collision (TTC), Post Encroachment Time(PET), Gap Time (GT), and Deceleration to Safety Time (DST), or probabilistic indicators that take into account various chains of events that may lead to collision (3). The various objective severity measures and conflict indicators are hypothesized to be of different and sometimes of independent nature. Each objective severity measure provides a cue for the underlying level of safety. Ultimately, important road safety treatment decisions must be taken based on a singular inference on the underlying level of safety.

The goal of this paper is to develop a new quantitative methodology for the integration of various objective measures of traffic conflicts. The proposed methodology is tested on video data used in a before-and-after safety evaluation of a pedestrian scramble phase (6) .

CONFLICT INDICATORS AS PARTIAL IMAGES

The main advantage of objective conflict indicators over qualitative severity measures is consistency of measurement. In other terms, the advantage is the reliability of objective conflict indicators over other severity measures. Reliability of measurement refers to the invariance of the conflict indicator to all factors extraneous to positional and temporal attributes of road users. For example, if the tracks of conflicting road users are known, their TTC is calculable and in an identical way regardless of the time, the location, and the traffic context of their interaction. However, some of the factors eliminated from consideration in evaluating conflict indicators may in reality be relevant to the true severity of the concerned traffic event. The coupling of subjective assessment of traffic events and conflict indicator measurements has been reported in several studies (8) (9) (10). To corroborate this relevance of subjective severity measurements, in validating the Swedish TCT, it was found that serious traffic conflicts rated as such by subjective human assessment was in stronger correlation with collisions than serious conflicts rated by objective conflict indicators (9). Another study on the correlation between collision and traffic conflicts adopted a combination of TTC and a subjective observer-based severity assessment of traffic conflicts (10). The subjective risk measure was introduced to supplement the intrinsic shortcoming of TTC in comprehensively representing the severity of traffic conflicts.

Based on evidence in the literature, it is plausible that various conflict indicators appear to represent partial images of the true severity of traffic events. Not surprisingly, the trained observer appears to be able to fathom much closer to the true severity of traffic events than conflict indicators based solely on positional data. Unfortunately, the observer however provides this measure at much lower reliability than is sufficient to establish a sound practice of traffic conflict analysis.

While extensive work has been performed on the validity of traffic conflict techniques, most of this work involved a handful of conflict indicators. Surprisingly little work has been done on validating the entire set of conflict indicators proposed in the literature (11). Previous work has been conducted on the validation of TTC against PET in comprehending the severity of traffic events in which the latter conflict indicator was favorable (8). Little, if any, investigation has been conducted on the validity of other conflict indicators found in the literature. As they stand, conflict indicators reflect different and sometimes independent severity aspects. It is however possible to group conflict indicators into two classes. The first class requires the presence of a collision course; the second class measures the mere spatial and temporal proximity of conflicting road users. The first class of conflict indicators, and potentially the more developed, measures the proximity to a collision point. Examples of the first class are TTC and probabilistic representation thereof. The value of TTC at a specific instant, called Time to

Accident, has been extensively used in the development of the Swedish TCT and has been validated for this purpose (9). Most notable of the second class is PET, which represents the observed temporal proximity of the conflicting road users. PET has been adopted in another key study in which it was proven to be a reliable predictor of road collisions if observed over an extended period of time (12). The two conflict indicators, TTC and PET, however do not represent the same collision mechanism. Arguably, they reflect different partially overlapping severity aspects.

TTC represents the proximity of conflicting road users to a potential collision point, while PET represents their proximity to each other. Generally, TTC is more suited to comprehend the severity of traffic events that involve the risk of rear-end collision. PET is of little validity in this case since it is dependent on the speed of the lagging road user as opposed to their relative speed. PET is better suited for representing the severity of crossing events. The two conflict indicators are not necessarily calculable for all events. Moreover, when calculable, they may represent variant severity measurements.

The above discussion leads to the main hypothesis of this paper: “Conflict indicators measure partially overlapping and sometimes independent severity aspects of traffic events.”

METHODOLOGY

A number of systematic approaches have been proposed to combine different road safety cues into composite indices, e.g. (13) (14). A theoretical framework was proposed for the general development of composite road safety indicators (14).. A central component of safety index development is the normalization, weighing, and aggregation of different indicator values. Previous developments focused on the integration of different road safety cues into macroscopic safety indices. The same reasoning and theoretical framework can be adopted at the microscopic level of individual traffic events. As opposed to a single conflict indicator, a set of conflict indicators can be used to measure the severity of traffic events. Different conflict indicators can be integrated in order to obtain a more accurate measure of the severity of traffic events. Two methods are introduced for integrating different conflict indicators and for mapping their composite values into the severity dimension: single-step integration and multi-step integration.

Integration Approach A: Single-step Integration

In this approach an integration function $\Psi(.)$ is constructed to map a set of conflict indicator values into the severity dimension. Let $x_1, x_2, \dots, x_n$ be the individual values of n conflict indicators, then the severity value represented by these conflict indicators is constructed as follows:

$$I = \Psi(x_1, x_2, \dots x_n) \qquad \dots (1)$$

where $I \in [0,1]$ is a dependent variable of which domain is the severity dimension. In subsequent sections of this chapter, I is referred to as severity index. The calibration of this integration approach requires reference severity measurements of a large sample of traffic events. This type of data is currently unavailable. Therefore, this approach has not been implemented in this paper

Integration Approach B: Multi-step Integration

In this approach each conflict indicator value x_i is independently mapped to the severity dimension by an individually defined mapping function $\Psi_i(x_i)$. The last step is to draw a representative value from the set of individual mappings of different conflict indicators. Following are proposals of representative values:

$$I = \begin{cases} \frac{1}{n}\sum_{\mathrm{i}} \Psi_{\mathrm{i}}(x_i) & approach\ \boldsymbol{B1} \\ max_i\big(\Psi_i(x_i)\big) & approach\ \boldsymbol{B2} \\ quantile\big(\Psi_{\mathrm{i}}(x_i)\big) & approach\ \boldsymbol{B3} \end{cases} \quad \ldots (2)$$

The multi-step integration approach can be viewed as a special case of the single-step integration. The interpretation of both is however distinct. The first integration approach (approach A) considers the interdependence of different conflict indicators in representing severity. The second set of approaches (B1, B2, and B3) assumes that every conflict indicator provides a unique and independent severity measure. In multi-step integration, it is necessary to draw a representative value from the individual mappings of conflict indicators. Equation 2 provides sample strategies for drawing representative values from individual mappings of conflict indicators. Selecting the average of individual mappings (approach B1) of conflict indicators is favorable when: 1) comparable validity in representing the severity of a traffic event is assumed for every conflict indicator and 2) differences in severity among conflict indicators are attributable to random road user characteristics. For example, TTC and PET satisfy these conditions since each of them measures independent proximity measures. The adoption of an extreme value of individual conflict indicator mappings, for example the maximum or the minimum value, implies the variability among conflict indicators in comprehending the severity of the concerned traffic event. For example, if it is the case that various conflict indicators each independently tend to underestimate severity, then drawing the maximum of individual mappings is more suited than drawing the average value, as is entailed by approach B2. It is straightforward to show that if severity is overestimated, then selecting a minimum value is potentially a more accurate representation of the true severity.

The integration approach B2 may however lead to erroneous severity measurement in the case when extreme values are induced by tracking errors. Common tracking errors are over-grouping (multiple objects are tracked as one), over-segmentation (one object is tracked as many), and tracking noise. The first two errors have been addressed in the original dataset as outlined in previous work (6). The issue of tracking noise concerns the sudden change in direction of the trajectory of moving objects. While Kalman filtering techniques have been used to mitigate this issue, instances of tracking noise may still exist. A consequence of tracking noise is erroneous TTC values that occur due to vehicle orientation toward a collision with other road users. In order to mitigate this potential for error, an order statistic or quantile value is used as an approximate to estimated extreme value (approach B3). However, in situations when few conflict indicators are used, the use of an order statistic may not be feasible.

Mapping Methods

Two main mappings are proposed in this paper: function mapping and distribution mapping. The mapping development was restricted to four conflict indicators: TTC, PET, DST, and GT. The mappings are also restricted to measuring the severity of pedestrian-vehicle conflicts, as a case study will be presented in this context.

Functional Mapping

In this mapping approach, closed-form functions are established in order to map the value of a conflict indicator (expressed in some unit) into a severity index (unitless). Following are the functional forms of the mappings:

$$I(\mathrm{x}) = e^{-\frac{x}{p1}} \qquad \forall \mathrm{x} = TTC > 0 \qquad \ldots (3)$$

$$I(x) = e^{-p2*\left(p3*x+e^{-p3*x}-1\right)} \qquad \forall \mathrm{x} = PET, GT, or DST > 0 \qquad \ldots (4)$$

where $p1, p2$, and $p3$ are specific mapping parameters that define its shape, $I(.)$ is the mapping function that takes the value of the conflict indicator as an argument and outputs a severity index that ranges from 0 for events with no reasonable exposure to the risk of collision and 1 for all collisions. Note that the proposed mapping is by construction unable to comprehend the variable outcomes of collision events.

The functional form selected for Equation 3 is adopted from a similar formulation by (15). The functional form presented in Equation 4 is adopted from a generic development of penalty functions with minor modification to yield an indexed value ((16). The function parameters were calibrated based on the severity benchmarks in the literature shown in Table 1.

Those severity thresholds were used to assign a nominal severity rating, *e.g.*, serious or mild, based on conflict indicator measurements. The highest severity level in Table 1, defining severe conflicts was selected to be represented by a severity index value of 0.8. Three more thresholds were selected for lower severity thresholds. Other TTC values in Table 1 were selected from the severity measures found in the Swedish TCT (17) assuming a constant conflicting speed of 20 km/h and assuming that the highest severity level of 30 corresponds to a severity index of 1. The highest severity threshold for GT and PET as well as severity thresholds for DST were reported in (18). The least severe temporal proximity for PET/GT is selected to be the time consumed for a pedestrian to walk corresponding distance of 10.0 m. The spatial proximity threshold is intrinsically defined in the calculation of conflict indicators to demarcate the boundary between exposure events and uninterrupted passages. Exposure events are constituted by any pair of pedestrian and vehicular road users that attain at minimum spacing closer than a spatial proximity threshold and also exhibit at some time convergent movement directions. The spatial proximity threshold was selected to be 10.0 m. Exposure events were selected for further proximity analysis while uninterrupted events were discarded for this purpose. Refer to (6) for further details on the conflict indicator calculations.

Distribution Mapping

The idea behind this mapping is to represent the severity by the relative frequency of a conflict indicator value. Ideally, if a large-scale pool of conflict indicator measurements is available, relative frequency will be closely related to the anomaly in conflict indicator value. According to the severity hierarchy theory as well as empirical evidence in (19) (17) (3), severe events are observed with low frequency. The pool of conflict indicator measurements used to establish the distribution of the four conflict indicators was obtained from the work presented in (6). Instead of using empirical cumulative distributions, which could be expensive to calculate, a Gamma distribution was fit to the conflict indicator observations. To deal with negative values for PET and GT, two sets of distribution parameters were estimated from positive and negative conflict indicator values. For negative PET and GT, their absolute values were used to estimate the set of distribution parameters for negative conflict indicator values.

Aggregation of Severity Measurements

Little statistical work has been conducted on drawing an inference about the level of road safety from the severity distribution of traffic events. The only work found on this subject was in the context of before and after studies (17). The statistical analysis was mainly based on testing the difference in shape between the severity hierarchy before and after the implementation of a safety treatment. However, testing for shape difference is not capable of comprehending the difference in distribution among individual severity levels. Statistical testing for shape difference has to be supplemented with a thorough review of the difference in frequency at each severity level. However, there are no developed models to aid in relating the change in relative frequency at each severity level and the underlying level of safety. In order to circumvent this methodological gap, aggregation of microscopic individual severity measurements should be conducted to produce higher-level measures.

Two main aggregation attributes were adopted in this paper, time and road user. Aggregation over time describes severity of traffic events along the time dimension. All severity measurements are referenced to the moment of analysis. One of the advantages of this aggregation approach is that important temporal patterns can be recognized using this aggregation approach. However, the key advantage of aggregating over time is the simplicity of extrapolating severity measurements outside the time span of observations. A prime example of aggregation over time was adopted in a key study on extreme value model for road collision (12).

The simplicity of adopting time as an aggregation attribute, or as a surrogate for exposure, comes at the expense of lacking insight into road user interactions. The most direct shortcoming of aggregating over time is the inability to represent the variation in severity measurements among traffic events that take place at the same moment. Another critical shortcoming of aggregation over time is the tendency to under-represent severity if traffic events exhibit irregularity over time. For example, if the average severity per moment is selected as a representative value, aggregate severity will be underestimated if the same number of traffic events takes place within shorter time periods. These shortcomings are intrinsic to aggregation over time and can be overcome by adopting different aggregation attribute.

Another aggregation attribute adopted in this study is road users. This aggregation approach provides more insight into road user interactions. For example, it is possible to represent the severity of traffic events irrespective of their temporal regularity. Two main shortcomings remain for aggregation over road users. The first shortcoming is the relative difficulty of extrapolating severity measurements outside the observational time span as compared to aggregation over time. Road user counts, especially pedestrian counts, are expensive to obtain for extended time periods. The second shortcoming is the inability to represent the presence of the multiple interactions in which the same road user may be involved.

In order to address this last shortcoming, aggregation should be conducted along the event dimension. The same pattern emerges; aggregating over events instead of road users provides a more accurate representation of road user interactions. However, this enhancement comes with significantly more expensive extrapolation of severity measurements outside the observational time span. In fact, the authors are not aware of the presence of any temporal conversion factors for the number of traffic events. Aggregation over events was not directly conducted in the case study presented in this chapter because of the significant computational expense. Instead, aggregate measurements were normalized by the number of events.

CASE STUDY

This case study is based on video data collected in 2004 for the evaluation of a pedestrian safety treatment in Oakland, California (20). Using an automated computer vision analysis approach (21) , a total of six hours of video data were analyzed for the before as well as after periods - three hours for each period. The distributions of conflict indicators were obtained for a subset of all video sequences and a total of four hours (6). An additional hour for each observational period was analyzed in this paper.

Empirical Independence of Conflict Indicators

. The correlation between various conflict indicator measurements was conducted in order to investigate the hypothesis that conflict indicators provide different and possibly independent severity measurements. Only TTC, PET, GT, and DST were considered in this analysis.

First, all pairs of conflict indicators that belong to the same traffic event with at least one calculable value were considered. This was conducted for the joint test of correlation as well as the common calculability of conflict indicators. Table 2 shows both the Pearson linear correlation coefficients and Spearman correlation coefficients for different combinations of conflict indicators. The severity interpretation of signed and unsigned values of PET and GT, corresponding to vehicle passage in front of or behind the pedestrian, is not well known. Therefore, the absolute values of PET and GT were also considered in the analysis. Second, testing was conducted for pairs of jointly calculable conflict indicators. For example, pairs of conflict indicators are considered only if both of them report calculable values. This is to separate from the conclusion the effect of whether the two conflict indicators are calculable for the same event. Similarly, Table 3 shows Pearson and Spearman correlation coefficients for different combinations of conflict indicators. Spearman correlation coefficient is slightly more relevant to this context since a linear relationship between the values of conflict indicators may be impacted by the lack of a uniform range definition for conflict indicators, except for the case of pairs of GT and PET.

In general, there is no strong correlation between TTC and any other conflict indicator, except for a 0.67 Spearman correlation with |PET|, when both indicators are mutually calculable. This means that in this video sequence absolute temporal proximity reflects to some extent the existence of a collision course. In addition, there is a strong correlation between PET and GT when both are mutually calculable (0.70 Pearson and

0.87 Spearman correlation coefficients). This is generally expected since the temporal proximity measured by both indicators is to some extent similar. A mild correlation between DST and GT is found for both cases of pairwise calculability. While correlation results are subject to several interpretations, the general conclusion that can be drawn is in support of the hypothesis. It should be noted, however, that the correlation results presented in Tables 2 and 3 are limited to the video data analyzed in this paper and may not be generalized to other data sets.

Results of Different Aggregation Approaches

The average values of different conflict indicators were calculated for various mapping approaches and aggregation approaches. In addition, two bounding percentile values, the 15^{th} and 85^{th}, were obtained to gauge the dispersion of every conflict indicator. Average values and estimated bounds are provided for index values calculated for each traffic events and using two mapping approaches. For the analysis presented in this paper, all function mappings were conducted using parameters inferred from benchmarks found in the literature and presented in Table 1. Mappings were also conducted for average values of the four conflict indicators (integration approach B1, Equation 2). Sample results are presented in Tables 4 and 5 aggregating over road users. Average values of every conflict indicator in their respective units are presented in the second columns, entitled Average Indicator, of Tables 4 and 5. For example, the average of all calculable TTC values for each road user in the before period is shown to be 4.85 sec. The 15^{th} and 85^{th} percentile bounds are provided for each conflict indicator and index in smaller table cells. For example, the 15^{th} percentile value for the distribution of calculable TTC values for all road users in the before period is shown to be $4.85 - 2.93 = 1.92$ sec. The fourth and fifth columns of Tables 4 and 5 show the function mapping of each average conflict indicator and the percentile bounds, respectively. For example, the function mapping of the average TTC value shown in Table 4 can be calculated as: $I(4.85) = e^{-\frac{4.85}{8}} = 0.54$. Similarly, the upper bound for the function mapping of the average TTC can be calculated as follows: $I(4.85 - 2.93) - I(4.85) = e^{-\frac{1.92}{8}} - 0.54 = 0.24$. Using the same steps of calculation, distribution mappings can be conducted by aid of Figure 1. Results of distribution mappings of conflict indicators are shown in columns 6 and 7 of Tables 4 and 5. The average value of individual index values from different conflict indicators is shown in columns 8 and 9 of Tables 4 and 5. For example, the average function mapping of all individual averages of conflict indicators, entitled Individual Aggregation, can be calculated as follows: Individual function Aggregation $= \frac{0.54+0.29+0.28+0.02+0.50+0.47}{6} = 0.35$. The 15^{th} percentile bound can be calculated using elementary error theory as follows: $-\sqrt{\frac{0.14^2+0.23^2+0.23^2+0.02^2+0.25^2+0.32^2}{6}} = -0.22$.

It is noteworthy that Tables 4 and 5 present various aggregations without taking into account the frequency of observations of conflict indicators and indices per road user. Results for other combinations of aggregation approaches including aggregation over time and considering frequency of observation are not provided to economize on space. In general, there was no noticeable effect of taking into account frequency of observation on the variance of conflict indicators and indices from before and after periods.

The following observations are noted from the analysis of results of different aggregation approaches:

1. There is a significant dispersion in all conflict indicators and indices values. It is difficult to provide explanation for this observation except that the severity hierarchy was investigated into adequate depth and that a wide variation of severity levels was observed.
2. There was no evidence of a measurable difference in average values between before and after conditions.
3. There was no significant difference in results with and without using frequency for calculating average values. This indicates that there was a general balance for the number of conflict indicator observations per frame and per road user.
4. Function mapping tends to consistently yield results lower than distribution mapping. A direct explanation of this observation, as also exhibited in Figures 1, is that if compared with a larger pool of observations, the distribution mapping may yield fewer abnormality values. In other words, the limited reference observations collected in this study created a bias toward overestimating severity if the distribution mapping is used.

Aggregation results in Tables, 4 and 5 mainly concern the average severity of all exposure traffic events. However, change in average severity between before and after periods cannot represent the change in exposure between the same periods. For example, Figure 2 shows the distributions of the severity index mapped using function mapping shown B1. The distributions exhibit a clear reduction in frequency of observation of traffic events at almost all severity levels. This safety improvement was not evident in Tables 4 and 5 mainly because averaging conflict indicators and indices measurements implicitly discards the effect of variant exposure.

Figures 3 and 4 further demonstrate the distinct safety information obtained when normalizing various severity measurements by number of exposure events. In Figures 3 and 4, the distributions of various conflict indicators are shown after normalizing their frequencies by the total number of exposure events. The magnitude and sign of the difference in distributions between before and after periods is mixed. Some indicators, such as |GT| exhibit stable severity for every instance of road user exposure in before and after conditions. PET exhibits different trends for positive and negative values, with positive PET exhibiting increase in severity after the treatment. Other indicators such as DST and TTC exhibit increase in severity per instance of road user exposure after the safety treatment. The distinct information contained in severity measures normalized by number of exposure events can be misinterpreted as all-encompassing safety cue. A more comprehensive severity index can be constructed by including the following aspects:

1. Severity of each exposure event.
2. Observed number of exposure events, and
3. Maximum number of possible exposure events,

A simple mechanism to combine the first and second aspects is the summation of all severity indices measurements, $\sum_{\text{event}} \text{severity indices}$. In order to further incorporate the third aspect, the previous summation can be divided by the number of maximum possible exposure E_{max}. This is to account for the safety differential between situations where the same summation of severities originates from different levels of traffic volume. This normalized safety measure $\bar{I}$ can be constructed as follows:

$$\bar{I} = \frac{\sum_{event} severity\ indices}{E_{max}} \qquad \ldots(5)$$

Theoretically, the maximum number of possible exposure events is the product of two conflicting traffic streams. In the context of pedestrian safety, $\mathrm{E_{max}}$ is the product of the number of pedestrians and the number of vehicles present during the observational period. Another plausible surrogate for $\mathrm{E_{max}}$ the total pedestrian and vehicle volumes during the observational period. Results of $\bar{I}$ calculation using different $\mathrm{E_{max}}$ estimates and different aggregation approaches are presented in Tables 5-9.

It is important to note the difference between use of $\mathrm{E_{max}}$ proposed in Equation 5 and its use as a surrogate for the total number of exposure events. The construction of the normalized safety measure presented in Equation 5 sets clear boundary between the estimation of maximum possible exposure and the accurate observation of exposure represented by the number of exposure events. Putting the two quantities in perspective, or dividing them as is shown in Equation 5, represents the distinct safety benefit of reducing actual exposure.

CONCLUSIONS

This paper presented a hypothesis that conflict indicators represent partially overlapping severity aspects. A number of approaches have been proposed in order to map into the severity dimension and to integrate conflict indicator measurement into a severity index. In addition, aggregation of conflict indicator and severity index measurements was advocated. A number of aggregation approaches have been proposed. For this purpose, three approaches were developed: aggregations over time, over road users, and over exposure events. The order of the three approaches reflects the accuracy of exposure measurement. However, the data required for projecting such aggregation measures outside the period of observation is proportional to their accuracy. With progress in road user tracking technologies, surrogates of exposure will be gradually abandoned in favor of more accurate measures of exposure. Part of the analysis presented in this paper dealt with average conflict indicators and severity indices per road user or time frame. A case study presented in this paper regarding a safety treatment of pedestrian scramble. The findings were consistently in favor of the effectiveness of this treatment in reducing the potential for conflicts between pedestrians and vehicles.

An important distinction was made in this paper between maximum possible exposure and actual exposure. The two quantities reflect the effectiveness of a safety treatment in limiting road user exposure to collision risk. The two quantities were augmented with the summation of all severity indices obtained from each traffic event (total severity) to produce a novel safety measure. The proposed safety measure is based on normalizing the summation of all severity indices by the maximum possible exposure. There is a well-recognized shortcoming of the naïve division of total severity by exposure. It may be the case that, similar to collision frequency, total severity independent of the underlying safety level is non-linearly related to maximum exposure. In this case, for reasons extraneous to safety, the mere increase in traffic volume would unreasonably lead to reduction in the safety measure. This non-linearity should to be further investigated.

REFERENCES

1. *Traffic Conflict Characteristics: Accident Potential at Intersections.* **Perkins, R. and Harris, J.I.** 1968, Highway Research Record, Vol. 225, pp. 35-43.

2. **Amundsen, F. and Hydén, C.** Oslo, Norway : Institute of Transport Economics, 1977. First Workshop on Traffic Conflicts.

3. *A Probabilistic Framework for Automated Analysis of Exposure to Road Collisions.* **Saunier, N. and Sayed, T.** 2008, Transportation Research Record: Journal of the Transportation Research Board, Vol. 2083, pp. 96-104.

4. *Large Scale Automated Analysis of Vehicle Interactions and Collisions.* **Saunier, N., Sayed, T. and Ismail, K.** 2010, Transportation Research Record: Journal of the Transportation Research Board. In press.

5. *Automated Analysis of Pedestrian-Vehicle Conflicts Using Video Data.* **Ismail, K., et al.** Washington, DC : s.n., 2009, Transportation Research Record: Journal of the Transportation Research Board.

6. *Automated Analysis Of Pedestrian-vehicle Conflicts: A Context For Before-and-after Studies.* **Ismail, K., Sayed, T. and Saunier, N.** 2010, Transportation Research Record: Journal of the Transportation Research Board. In press.

7. **van der Horst, R.** *A Time-based Analysis of Road User Behaviour in Normal and Critical Encounters.* Delft : Delft University of Technology, 1990.

8. **Grayson, G., et al.** *The Malmö Study: a Calibration of Traffic Conflict Techniques.* Institute of Road Safety Research. Leidschendam, the Netherlands : s.n., 1984. Technical Report. R-84-12.

9. **Svensson, Å.** *Vidareutveckling och Validering av den Svenska Konflikttekniken (in Swedish). Further Development and Validation of the Swedish Traffic Conflicts Technique.* Department of Traffic Planning and Engineering, Lund University. 1992.

10. *Traffic Conflict Standards for Intersections.* **Sayed, T. and Zein, S.** 4, 1999, Transportation Planning and Technology, Vol. 22, pp. 309-323.

11. *Road User Behaviour Indictators in Automated Video Analysis Systems.* **Laureshyn, A. and Svensson, A.** Journal of Intelligent Transportation Systems Technology, Planning, and Operations. Submitted. Available in A. Laureshyn's Ph.D. thesis..

12. *The Extreme Value Theory Approach to Safety Estimation.* **Songchitruksa, Praprut and Tarko, Andrew.** 4, 2006, Accident Analysis and Prevention, Vol. 38, pp. 811-822.

13. **Al-Haji, G.** *Towards a Road Safety Development Index.* Linköpings University. 2005. PhD Thesis.

14. *Uncertainty Assessment of the Road Safety Index.* **Hermans, Elke, Van den Bossche, Filip and Wets, Geert.** 7, 2009, Reliability Engineering & System Safety, Vol. 94, pp. 1220-1228. Special Issue on Sensitivity Analysis.

15. *Traffic Accident Prediction using 3D Model Based Vehicle Tracking.* **Hu, W., et al.** 3, 2004, IEEE Transactions on Vehicular Technology, Vol. 53, pp. 677–694.

16. *Optimization of a Design Code for Wind-turbine Rotor Blades in Fatigue.* **Ronold, Knut O. and Christensen, Carl J.** 8, 2001, Engineering Structures, Vol. 23, pp. 993-1004.

17. **Svensson, Å.** *A Method for Analysing the Traffic Process in a Safety Perspective.* Department of Traffic Planning and Engineering, Lund University. 1998. PhD Thesis.

18. *The Development of an Automatic Method of Safety Monitoring at Pelican Crossings.* **Malkhamah, S., Miles, T. and Montgomery, F.** 5, 2005, Accident Analysis and Prevention, Vol. 37, pp. 938-946.

19. **Archer, J.** *Methods for the Assessment and Prediction of Traffic Safety at Urban Intersections and their Application in Micro-simulation Modelling.* Royal Institute of Technology. 2004. PhD Thesis.

20. *Pedestrian Scramble Signal in Chinatown Neighborhood of Oakland, California: an Evaluation.* **Bechtel, A, MacLeod, K and Ragland, D.** 2004, Transportatoin Research Record: Journal of the Transportation Research Board, Vol. 1878, pp. 19-26.

21. *A feature-based tracking algorithm for vehicles in intersections.* **Saunier, N. and Sayed, T.** s.l. : IEEE, 2006.

Table 1: Severity benchmark values for constructing mapping functions

Conflict Indicator	Severity Level	Severity Index	TTC (sec)	PET/GT (sec)	DST (m/s^2)
Severity Thresholds	Most severe	0.8	1.6	3	1
	Severe	0.6	5	-	2
	Medium	0.4	8	-	4
	Minimal	0.2	11	8.5	6

Table2 Correlation coefficients for only pairs of commonly calculable conflict indicators

Conflict Indicator	TTC	PET	DST	GT	|PET|	|GT|
TTC	1	-0.07 (0.37)	-0.30 (0.09)	-0.09 (0.07)	0.42 (0.14)	0.14 (0.08)
PET	0.28 (0.32)	1	0.49 (0.29)	0.70 (0.10)	0.25 (0.20)	0.06 (0.26)
DST	-0.57 (0.09)	0.46 (0.34)	1	0.22 (0.09)	-0.04 (0.13)	-0.08 (0.03)
GT	-0.10 (0.07)	0.87 (0.05)	0.59 (0.12)	1	0.30 (0.20)	0.58 (0.28)
|PET|	0.67 (0.06)	0.35 (0.36)	0.01 (0.14)	0.56 (0.21)	1	0.43 (0.15)
|GT|	0.23 (0.05)	0.40 (0.30)	-0.01 (0.07)	0.50 (0.11)	0.70 (0.07)	1

Upper echelon contains pair-wise Pearson linear correlation coefficients. Lower echelon (shaded) contains Spearman ρ rank correlation coefficient. Values in parentheses are the standard deviation of the correlation coefficients calculated for all pairs within a sample of all ½ hours of video data (12 samples).

Table 3 Summary results for different aggregation strategies for <u>before</u> conditions. Representative statistics are drawn only from calculable values of each indicator or index for each <u>road user</u>

Conflict Indicator	Average Indicator			Individual Index Value						Individual Aggregation	
				Function			*Distribution*			*Function*	*Distribution*
TTC (sec)	4.85	*-2.93*	*2.54*	0.54	*-0.14*	*0.24*	0.35	*-0.20*	*0.44*	0.35 *(-0.22 : 0.41)*	0.49 *(-0.28 : 0.39)*
PET+ (sec)	7.52	*-4.36*	*4.28*	0.29	*-0.23*	*0.49*	0.46	*-0.29*	*0.43*		
PET- (sec)	-6.63	*-3.50*	*3.47*	0.28	*-0.23*	*0.48*	0.47	*-0.27*	*0.38*		
DST (m/s^2)	0.29	*-0.52*	*0.67*	0.02	*-0.02*	*0.14*	0.37	*-0.37*	*0.39*		
GT+ (sec)	5.50	*-4.24*	*2.50*	0.50	*-0.25*	*0.45*	0.66	*-0.25*	*0.32*		
GT- (sec)	-5.16	*-2.77*	*4.03*	0.47	*-0.32*	*0.49*	0.63	*-0.28*	*0.35*		
|PET| (sec)	7.06	*-3.90*	*3.99*	0.33	*-0.25*	*0.45*	0.50	*-0.30*	*0.39*	-	-
|GT| (sec)	5.35	*-4.16*	*2.62*	0.52	*-0.26*	*0.44*	0.68	*-0.26*	*0.30*	-	-
Index (function)	0.34	*-0.23*	*0.20*	-			-			-0.006 *-0.32* *0.46*	-
Index (distribution)	0.51	*-0.25*	*0.23*	-			-			-	0.01 *-0.38* *0.45*

Values in italic are the 15th percentile value minus the mean and the 85th percentile value minus the mean.

Table 4 Summary results for different aggregation strategies for <u>after</u> conditions. Representative statistics are drawn only from calculable values of each indicator or index at each <u>road user</u>

<table>
<tr><th rowspan="2">Conflict Indicator</th><th rowspan="2" colspan="2">Average Indicator</th><th colspan="4">Individual Index Value</th><th colspan="4">Individual Aggregation</th></tr>
<tr><th colspan="2">Function</th><th colspan="2">Distribution</th><th colspan="2">Function</th><th colspan="2">Distribution</th></tr>
<tr><td rowspan="2">TTC (sec)</td><td rowspan="2">4.14</td><td>-2.99</td><td rowspan="2">0.59</td><td>-0.14</td><td rowspan="2">0.44</td><td>-0.23</td><td rowspan="12" colspan="2">0.37
(-0.26 : 0.42)</td><td rowspan="12" colspan="2">0.53
(-0.34 : 0.38)</td></tr>
<tr><td>2.30</td><td>0.27</td><td>0.47</td></tr>
<tr><td rowspan="2">PET+ (sec)</td><td rowspan="2">7.54</td><td>-4.17</td><td rowspan="2">0.28</td><td>-0.21</td><td rowspan="2">0.45</td><td>-0.27</td></tr>
<tr><td>3.77</td><td>0.47</td><td>0.42</td></tr>
<tr><td rowspan="2">PET- (sec)</td><td rowspan="2">-5.81</td><td>-3.61</td><td rowspan="2">0.38</td><td>-0.31</td><td rowspan="2">0.56</td><td>-0.32</td></tr>
<tr><td>4.11</td><td>0.54</td><td>0.40</td></tr>
<tr><td rowspan="2">DST (m/s^2)</td><td rowspan="2">0.37</td><td>-0.44</td><td rowspan="2">0.04</td><td>-0.04</td><td rowspan="2">0.44</td><td>-0.44</td></tr>
<tr><td>0.54</td><td>0.12</td><td>0.30</td></tr>
<tr><td rowspan="2">GT+ (sec)</td><td rowspan="2">5.34</td><td>-4.19</td><td rowspan="2">0.52</td><td>-0.32</td><td rowspan="2">0.68</td><td>-0.33</td></tr>
<tr><td>3.43</td><td>0.44</td><td>0.30</td></tr>
<tr><td rowspan="2">GT- (sec)</td><td rowspan="2">-5.39</td><td>-4.14</td><td rowspan="2">0.44</td><td>-0.37</td><td rowspan="2">0.61</td><td>-0.38</td></tr>
<tr><td>4.63</td><td>0.54</td><td>0.38</td></tr>
<tr><td rowspan="2">|PET| (sec)</td><td rowspan="2">7.09</td><td>-4.33</td><td rowspan="2">0.33</td><td>-0.24</td><td rowspan="2">0.50</td><td>-0.29</td><td rowspan="2" colspan="2">-</td><td rowspan="2" colspan="2">-</td></tr>
<tr><td>3.83</td><td>0.50</td><td>0.42</td></tr>
<tr><td rowspan="2">|GT| (sec)</td><td rowspan="2">5.36</td><td>-4.38</td><td rowspan="2">0.52</td><td>-0.34</td><td rowspan="2">0.68</td><td>-0.35</td><td rowspan="2" colspan="2">-</td><td rowspan="2" colspan="2">-</td></tr>
<tr><td>3.69</td><td>0.45</td><td>0.31</td></tr>
<tr><td rowspan="2">Index (function)</td><td rowspan="2">0.36</td><td>-0.26</td><td rowspan="2" colspan="2">-</td><td rowspan="2" colspan="2">-</td><td colspan="2">-0.017</td><td rowspan="2" colspan="2">-</td></tr>
<tr><td>0.24</td><td>-0.37</td><td>0.49</td></tr>
<tr><td rowspan="2">Index (distribution)</td><td rowspan="2">0.51</td><td>-0.27</td><td rowspan="2" colspan="2">-</td><td rowspan="2" colspan="2">-</td><td rowspan="2" colspan="2">-</td><td colspan="2">-0.01</td></tr>
<tr><td>0.26</td><td>-0.43</td><td>0.47</td></tr>
</table>

Values in italic are the 15th percentile value minus the mean and the 85th percentile value minus the mean.

Table 5 Summary results for before and after index values normalized by the total number of tracked road users. Indices representing an event are the maximum and average of all mapped conflict indicators

Selection	Agg. Type	Freq.	Distribution		Function	
			Before	*After*	*Before*	*After*
Maximum	Time	*Without Freq.*	2.69	1.15	3.37	1.54
		With Freq.	47.41	19.95	56.89	24.05
	Road User	*Without Freq.*	0.10	0.06	0.13	0.08
		With Freq.	0.41	0.16	0.54	0.22
Average	Time	*Without Freq.*	1.58	0.72	2.34	1.12
		With Freq.	23.59	10.29	35.47	15.25
	Road User	*Without Freq.*	0.07	0.05	0.11	0.07
		With Freq.	0.25	0.11	0.39	0.17

Table 6 Summary results for before and after index values normalized by the product of the numbers of pedestrians and vehicles in millions. Indices for every event are the maximum and average of all mapped conflict indicators

Selection	Agg. Type	Freq.	Distribution		Function	
			Before	*After*	*Before*	*After*
Maximum	Time	*Without Freq.*	191	99.4	239	132
		With Freq.	3360	1710	4000	2100
	Road User	*Without Freq.*	7.30	5.25	9.75	7.11
		With Freq.	29.7	14.2	38.7	19.2
Average	Time	*Without Freq.*	111	62.5	165	96.7
		With Freq.	1670	883	2510	1300
	Road User	*Without Freq.*	5.33	4.32	7.85	6.16
		With Freq.	18.2	9.82	27.8	14.7

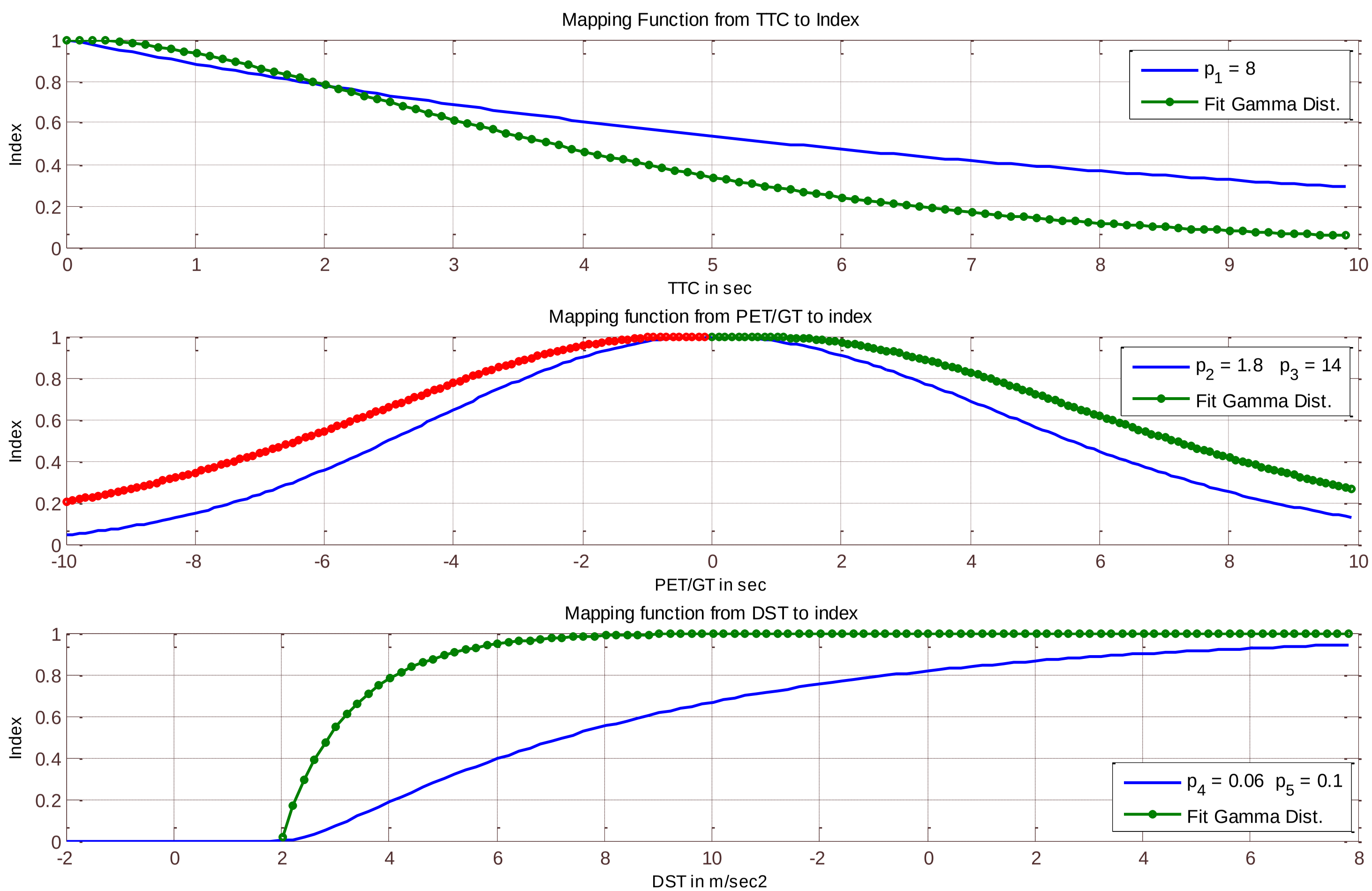

Figure 1: A depiction of two mappings from conflict indicators to severity index. Shown also are the parameters for function mapping 1. Mapping parameters (p1 to p5) are shown in the legend that were collected from benchmarks in the literature. For example p1 = 8.0.

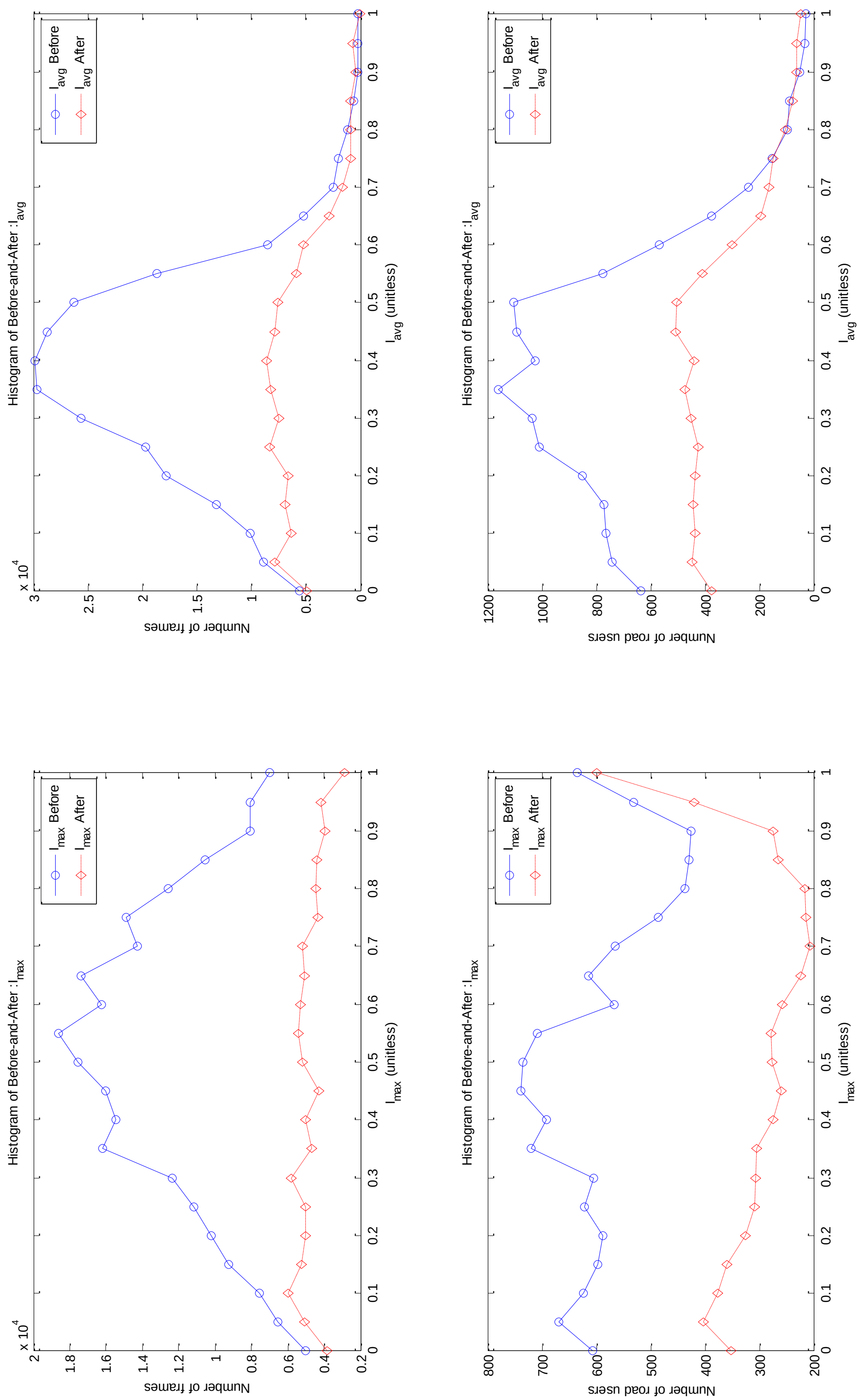

Figure 2 Severity index distributions for before and after conditions. Function mapping was used. Maximum indices were selected for every frame (upper row) and road user (bottom row).

1

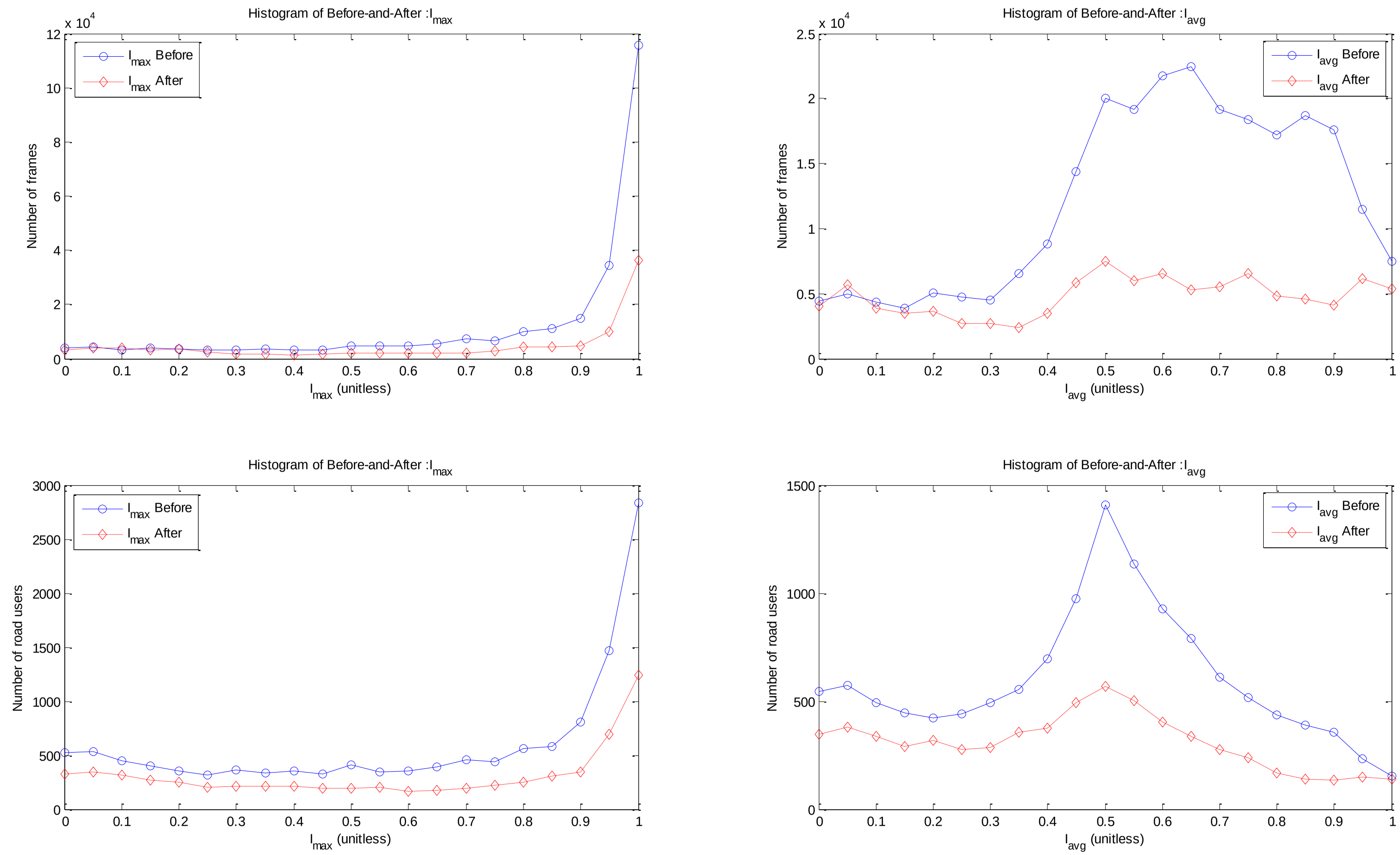

Figure 3 Severity index distributions for before and after conditions. Function mapping was used. Average indices were selected for every frame (upper row) and road user (bottom row).