Modelo de clasificación del riesgo en intersecciones rurales en T y validación del tiempo de evasión como medición alternativa de la seguridad de tránsito en intersecciones

Risk classification model in rural T-form intersections and time to evasion evaluation as surrogate safety measure

Alejandro Torres Flores*[1], Francisco Torres Avilés*, José Mª Pardillo Mayora**

*Universidad de Santiago de Chile. CHILE
**Universidad Politécnica de Madrid. ESPAÑA

Fecha de recepción: 01/ 06/ 2010
Fecha de aceptación: 31/ 08/ 2010
PAG. 353 - 370

Resumen

En los cruces o intersecciones, los conductores afrontan una multitud de opciones relacionadas con el camino, velocidad, y trayectoria que, en combinación con numerosos movimientos del tránsito, complican la tarea del conductor y aumentan considerablemente la potencial ocurrencia de un accidente. En este artículo se utilizan datos de conflictos de tránsito, tomados en cruces de la Provincia de Talagante en la Región Metropolitana de Santiago, Chile. Usando la metodología de componentes principales, se propone un índice que represente el nivel de riesgo de cada conflicto de tránsito en una intersección, determinándose que los valores más altos del índice están relacionados a un mayor riesgo de que un conflicto termine en accidente, mientras que valores bajos indican un menor riesgo. Posteriormente, se desarrolla un modelo de clasificación del riesgo en intersecciones, aplicando la teoría de árboles de decisión. Por último, se valida el Tiempo de Evasión (TE) como variable cuantificadora de la severidad de los conflictos de tránsito, tomando como base variables ya estudiadas a nivel mundial, como el Tiempo hasta la Colisión (TC) y el Tiempo de Post-invasión (TPI).

Palabras Clave: Análisis de componentes principales, árboles de clasificación, conflicto de tránsito, tiempo a la colisión, tiempo de evasión

Abstract

In the intersections, the drivers confront many options related to the way, speed and trajectory, which ones in combination to many options of traffic movements, they make difficult the task of the conductor and considerably increase the potential occurrence of an accident. In this paper, data of traffic conflicts are used, taken in rural T-form intersections from Talagante district in the Metropolitana region, Chile. Using the principal component analysis, an index are proposed that represents the risk level of each traffic conflict in an intersection, determining that the highest values of the index are related to a greater risk of than a traffic conflict it finished in accident, whereas low values indicate a smaller risk. After that, a model of classification of the risk in intersections is developed, applying the theory of decision trees. Finally, Time to Evasion is validated like quantifying variable of the traffic conflicts, taking as it bases studied variables at world-wide level, like Time to Collision and Post – encroachment Time.

Keywords: Principal component analysis, classification trees, time to collision, time to evasion, traffic conflicts

1. Introduction

One of the main objectives intended to be achieved when designing an intersection is to minimize the events leading to potential accidents, as well as to facilitate the users to understand the displacement through it by employing the highest safety standards. The importance of developing expedite studies leading to asses intersection safety lays in this factor, as well as the prompt evaluation of improvements conducted.

[1] Autor de correspondencia / Corresponding author:
E-mail: alejandro.torres@usach.cl

In order to determine whether an intersection is safe or not, diverse methodologies have been used such as accident prediction models, transit simulation models and techniques not requiring greater data bases to achieve an accident potential risk, as traffic conflict analysis and safety road audits or the so called surrogate measurement, which are the basis of the current research.

The surrogate measurement of road safety might be a useful indicator to quantify safety and help to chose among different alternatives, not needing to conduct expensive accident studies or quite expensive reconstruction projects. An option is to complement such studies with other type of measurements in order to determine the safety of different elements existing in a highway, which is precisely a way to obtain a wide application field from safety surrogate measurements at intersections.

2. Traffic conflict technique

Perkins and Harris (1968) identified the traffic conflict patterns related to some kind of accident. Such technique was considered as accidents potential measurement. Glauz and Migletz (1982), provide the following considerations.

a) Relation with safety: conflicts must be related with accidents.
b) Relation with the place: traffic conflicts may be useful to diagnose problems at a given place or as an efficiency measurement of the improvement conditions at a given point of the road.
c) Soundness: the definition of a conflict must have a minimum variation among different observers registering the same event.
d) Repetition: the definition should have an acceptable variation level at repeated observations conducted by the same observer, at the same place and under identical conditions.
e) Economy: obtained data must be trustworthy and obtained at reasonable time and costs.

In 1987 at Lund, Sweden, Professor Christer Hydén introduced a concept named Time to Collision, as part of the Swedish Traffic Conflicts Technique, which was mainly characterized by three aspects.

i) Defined objective measurement to evaluate severity conflicts: Time to Collision (TC) Controversial Speed (CS).

ii) Only severe conflicts were considered.

iii) Information is collected in a subjective way by means of observers.

The researchers Almqvist and Hydén (1994) believe this technique is the most direct method to evaluate an intersection with high accident risks. Gettman and Head (2003) defined traffic conflicts as observable situations that may indicate the relative safety different intersection designs may have. Such conflict events take place between two vehicles running on a collision trajectory, but not due to an evasive action. These conflicts take place at a particular place and space known as conflict point.

3. Road safety alternative or surrogate measurement

Surrogate safety measurements are time measurement, which are registered while the traffic conflict is in progress. Such measurements evaluate how close the conflict between vehicles facing an intersection may end up in a collision and the severity level the collision might have.

The most accepted surrogate measurement of conflict severity is Time to Collision (TC), even though some others (post–encroachment time, PET; or deceleration rate, DR) have been proposed as measurement for different conflict situations (Gettman and Head, 2003). The problem is to define a series of secondary measurements having a reasonable connection with the assessment of improvements developed to road safety. Besides, it is necessary to register, as easily as possible, such measurements in field studies.

When reviewing bibliography and the main studies conducted, it is clearly noticeable that authors consider TC as one of the most important measurement of conflict severity (Perkins and Harris, 1968; Hayward, 1972; Allen et al., 1978; Glauz and Migletz, 1982; Hydén, 1987; Brown and Cooper, 1990; Van der Horst, 1990; Gettman and Head, 2003; Archer, 2005), which severe conflict threshold is fixed at TC<1.5 seconds.

For a conflict point, TC measurement is defined as the difference between the final encounter time for a turning vehicle and the estimated time to reach the conflict by another vehicle displacing in a straight line, which has the right of way; provided that the latter continued the same trajectory and kept the same speed before starting to restrain itself in order to avoid collision.

For a conflict point, the measurement of post-encroachment time (PET) is defined as the difference of time between the exit of conflict point in the encounter zone and the arrival to such conflict point by a vehicle displacing in straight line, which has the right of way. Documentary review establishes that closest to zero value the PET was the closest probability of collision was (Allen et al., 1978).

Graph in Figure 1, indicates time-space allocation of conflict point concepts, TC and PET.

The current study will introduce the measurement Time to Evasion (TE), defined as time expressed in seconds, from the moment an evasive movement is made (breaking) until the front end bumper reaches the defined point of conflict. This new concept is associated to the driver's perception when approaching the intersection, which provides time to make an evasive movement in case an unexpected event occurs. Calculation is defined by the quotient between Breaking Distance (BD) and Approaching Speed (AS). This new surrogate safety measurement (TE) will be correlated with measurement already studied, such as TC and PET, thus becoming 3 surrogate measurements used by this study.

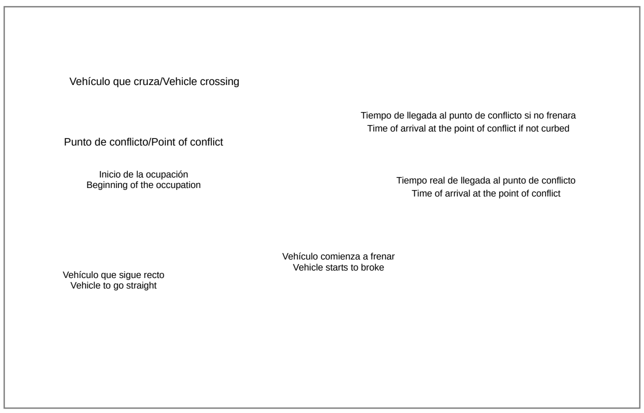

Figura 1. Ubicación tiempo-espacio de un punto de conflicto, tiempo hasta la colisión y tiempo posterior a la invasión
Figure 1. Ubication of time-space conflict point, time before colision and time after invasion

At the same time, Breaking Distance (BD) is defined as the length expressed in meters, from the point the vehicle break is pushed (which starts an evasive movement) up to the conflict point defined at the intersection studied. The Approaching Speed (AS) is measured 150 meters before the conflict point selected.

4. Methodological strategy

It is based on multivariate methods of classic statistics, which are used in order to analyze the information available. In this research each experimental unit is defined by traffic conflicts.

Such analysis considers:

a) Information sampling of original variables (TC, TE, PET, BD and AS) for each conflict identified at intersections, including a correlation analysis of conflicts.

b) Summary of available information by means of a new variable or index, using the methodology known as Principal Components (Peña, 2002). Besides, it will lead to a development of a quantitative risk index for each intersection comprised by this study, using a multivariate standardization of the original variables.

c) Eventually, it is quite important to study the proposed index behavior in order to determine possible groups among data, which will represent the risk levels.

Proposed classification rules are created using the methodology presented by Hothorn et al. (2006), which general procedure is based on the construction of a regression/classification tree by successively measuring the association between the response variable and the covariate partitions by means of multiple permutation tests (Strasser and Weber, 1999), until no significant association is detected, consequently the recursive process ends. So as to reach this objective three phases were developed.

1. An independent regression tree is designed for each intersection, which response variable shall correspond to the quantitative risk index obtained from the analysis of main components. This regression theory is intended to relate a set of original variables (BD, AS, TC, PET and TE) with a response variable that in this case, corresponds to the intersection risk index.

2. The assessment of each tree model will be done by employing the adjusted models for each intersection and, the most suitable of them will be selected using as discrepancy measurement the average resulting from the difference between observed and predicted square values or the estimated mean square error (MSE).

3. Once the regression tree model is selected with a minimum MSE, and by using a quantitative index, a classification rule is proposed. Such rule is based on a qualitative classification tree, which indicates the presence or absence of risk thus providing a thorough response to evaluate when an observed intersection conflict is considered as an "accident high risk" or "accident low risk".

Calculations and analysis in the current research will be carried out by using the "R" statistic software and considering a 5% significance level.

5. Information collection and analysis

Information Features

The intersections selected for information collection at field sites, are located in a rural zone in Talagante District, Santiago Metropolitan Region, Chile. Such intersections are G-78 highway intersecting G-374 (I1), G-30 highway intersecting G-374 (I2) and G-40 highway intersecting G-46 (I3), which are shown in Figure 2.

Figura 2. Intersecciones seleccionadas para el estudio: a) ruta G-78 con G-374; b) ruta G-30 con G-374; c) ruta G-40 con G-46 Fuente: Google Earth

Figure 2. Selected intersections for the study: a) road G-78 and G-374; b) road G-30 and G-374; c) road G-40 and G-46 Source: Google Earth

The three crossroads are shaped up by the encounter of two 2-ways roads and there is a STOP signal regulating the right of way. Measurements were conducted during labor periods and working hours (from 10 am to 5 pm), during three days for each intersection, from which 635, 463 and 591 measured conflicts were obtained for intersections I1, I2 and I3, respectively. Such intersections have similar features, so it is expected that no external factors would make a difference in results.

Association Analysis

In Table 1, the upper triangular matrix shows the linear association level existing among the involved variables in this study, while lower triangular matrix shows p-values associated with the correlation hypothesis equal cero. Notice that the association level among variables is high, confirmed by p-values lower than 0.05 for variables related to measure times in the 3 proposed intersections.

Tabla 1. Coeficiente de correlación y p-valores correspondiente a las intersecciones 1, 2 y 3
Table 1. Correlation coefficient and p-values corresponding to intersections 1, 2 and 3

	Intersección/Intersection	DF	VA	TE	TC	TPI
	I1:G78					
	G374					
	I2:G30		-0.063	0.740	0,734	0.596
	G374		0.002	0.768	0.733	0.631
DF	I3:G40 G46		0.047	0.689	0.619	0.480
	I1:G78					
	G374					
	I2:G30	0.111		-0.668	-0,665	-0,574
	G374	0.974		-0.583	-0.616	-0.485
VA	I3:G40 G46	0.251		-0.635	-0.628	-0.527
	I1:G78					
	G374					
	I2:G30	<0.001	<0.001		0.996	0.812
	G374	<0.001	<0.001		0.989	0.727
TE	I3:G40 G46	<0.001	<0.001		0.878	0.663
	I1:G78					
	G374					
	I2:G30	<0.001	<0.001	<0.001		0.811
	G374	<0.001	<0.001	<0.001		0.736
TC	I3:G40 G46	<0.001	<0.001	<0.001		0.789
	I1:G78					
	G374					
	I2:G30	<0.001	<0.001	<0.001	<0.001	
	G374	<0.001	<0.001	<0.001	<0.001	
TPI	I3:G40 G46	<0.001	<0.001	<0.001	<0.001	

The existing negative values in the upper triangular matrix shown in the table, are consistent with field observations, since it is expected that at higher AS shorter the associated time would be for the vehicle to reach the intersection. It is important to mention that between BF and AS values are close to cero (p-values>0.05), which means that BD would not be linearly correlated with the vehicle speed before executing an evasive movement. This conclusion makes sense, since this calculation used Pearson coefficient correlation, which only detects linear association.

Quantitative intersection risk classification model

Principal components are defined as new variables summarizing the information contained in the original ones. Such new variables are not correlated and they are placed in decreasing order in accordance to importance level. Only the two first components are indicated, since the accumulated variability percentage expressed by both dimensions exceeds the 70% in each intersection.

For the analysis of main components, employed to build up the risk quantitative index, an estimated correlation matrix will be used for each intersection. The principal components and their variables are shown in Table 2.

Tabla 2. Componentes principales correspondientes a las intersecciones en estudio

Table 2. Main components corresponding to intersections under study

	Intersección/Intersection G78 G374		Intersección/Intersection 2 G30 G374		Intersección/Intersection 3 G40 G46	
-	Comp. 1	Comp. 2	Comp. 1	Comp. 2	Comp. 1	Comp. 2
DF	0.382	0.675	0.406	0.620	0.361	0.710
VA	-0.353	0.737	-0.322	0.784	-0.361	0.701
TE	0.510	0.015	0.515	0.017	0.510	0.048
TC	0.509	0.012	0.515	-0.031	0.519	-0.008
TPI	0.459	-0.026	0.449	0.016	0.458	-0.053
Varianza/Variance	3.740	0.938	3.595	1.003	3.425	1.051
% Variabilidad Explicada/Variability Explained	74.80	18.75	71,91	20.06	68.50	21.02

In addition to make sure that selected components accumulate an explained variability higher than 60%, it was demanded that components had a logical interpretation related to the intersection risk. Consequently the construction of index is consistently justified.

Since the first principal component, which coefficients are the ones associated with the linear combination of the original variable set, to be comprised by the index, expresses the threshold of demanded accumulated percentage value (> 65%) for the three intersections. Therefore, it can be observed that these new variables properly summarize the information from original variables associated with surrogate safety measurement, then, the first components are potential candidates to represent a risk index.

Considering the first components for each intersection, the proposed indexes for every involved intersection are given by formulas (1) – (3).

Índice de I1/Index for I1

$$I_1 = -0.382DF + 0.353VA - 0.510TE - 0.509TC - 0.459TPI \quad (1)$$

Índice de I2/Index for I2

$$I_2 = -0.406DF + 0.322VA - 0.515TE - 0.509TC - 0.449TPI \quad (2)$$

Índice de I3/Index for I3

$$I_3 = -0.361DF + 0.360VA - 0.510TE - 0.519TC - 0.458TPI \quad (3)$$

It is possible to conclude that the interpretation of the first principal component is associated with an intersection risk index, since the coefficients of these components are negative for the variables. If the latter had smaller values, it would reflect a greater intersection risk, such as variables BD, TC, PET and TE. On the other side the coefficient of SA variable is positive, so when yielding greater values, a greater intersection risk is also obtained. Summarizing, lower index values indicate a lower intersection risk and higher values represent a higher risk.

The coefficients accompanying each one of these original variables represent the importance they have on the response variables. In the obtained indexes, the TE and TC variables are the most important ones, which suggest a greater importance of these measurements. Table 3 shows the contribution on percentage terms. By using standardized variables, the importance provided by marginal variability can be controlled. Concluded this analysis, the use of a second component is discarded, since it expresses a lower variability than the first one.

Tabla 3. Contribuciones de las variables originales en los índices de riesgo propuestos
Table 3. Contribution of original variables on proposed risk indexes

	Contribuciones por índice de riesgo/Index of risk contributions (%)		
-	Intersección/Intersection G78 G374	Intersección/Intersection 2 G30 G374	Intersección/Intersection 3 G40 G46
DF	17.26	18.40	16.35
VA	15.95	14.59	16.30
TE	23.05	23.33	23.10
TC	23.00	23.33	23.51
TPI	20.74	20.34	20.74

The distribution of estimated values for each index is shown in Table 4. The most relevant fact is the possibility of proposing a clear cut off point, since the 5% of vehicles in the sample, yield values higher than 2.529 for I1; 3.073 for I2 and 3.248 for I3; which is a clear evidence that an index close to 3.0 could be a reference value to distinguish whether a vehicle is under imminent risk of accident. Furthermore, the 5% of vehicles in the sample yielding such values, matches the TC<1.5 sec, which is the threshold to be considered as severe, according to searched bibliography, as well as from experiences by Allen et al. (1978), Hydén (1987), Van der Horst (1990) and Archer (2005).

Once the index is determined for each intersection, a tree was fitted for each one of them, which is considered as a training stage, where three models were fitted and which response corresponds to the index value computed from principal component analysis. The tree will proceed to use the original variables as classification factors of the response or index (Hothorn et al., 2006). After elaborating the three tree models, the other two intersections were considered as independent samples and were fitted with the different models obtained from the training stage (validation stage). Finally the mean square errors associated with each predicted index, for three models in the training stage, are calculated. Table 5 indicates the “predictive” square errors, which may be considered as low, provided that the proposed index takes values in the range of (-5.0, 5.0). Although the information associated to I1 provides errors quite similar to I2, the model and data from the latter intersection will be used in the process of building a decision tree of qualitative risk.

Tabla 4. Medidas de distribución de los índices de riesgo propuestos
Table 4. Distribution measurements of proposed risk indexes

Intersección/Intersection	Mínimo/Minimum	1er. Cuartil/1st Quartile	Mediana/Median	Media	3er. Cuartil/3rd Quartile	Percentil/Percentile 95	Máximo/Maximum
G78 G374	-7,919	-1,524	0,557	-1,1e-15	1,455	2,529	3,771
G30 G374	-5,715	-1,281	-0,354	8,1e-16	1,601	3,073	4,569
G40 G46	-5,663	-1,263	-0,384	-2,5e-15	1,673	3,248	4,580

Tabla 5. Error cuadrático medio del índice de riesgo, según la base de datos ajustada
Table 5. Average quadratic error of risk index, as per adjusted data base

Error cuadrático medio/Mean square error			
Observado/Observed	Predicción/Prediction		
Intersección/Intersection	G78 G374	G30 G374	G40 G46
G78 G374	0.040	0.134	0.340
G30 G374	0.137	0.046	0.290
G40 G46	0.148	0.163	0.089

The classification tree selected to classify the index risk is shown in Figure 3. Notice the importance represented by TC variable in the classification of risk index. Other variables that facilitate the determination of quantitative risk level are TE and PET. The effect "the greater index value, the greater risk at intersection" can be initially observed in I2 tree (Figure 3) provided that when TC $\leq$ 3.25 [sec], TC $\leq$ 2.12 [sec] and TC $\leq$ 1.47 [sec], the index reaches values higher than 3, which correspond to 5.2% of total conflicts in the training sample. If TC > 1.47 [sec], the index values show a relatively high risk, because it reaches values between 2.0 and 3.0. In this stage, TE and PET are potential variables to classify indexes that do not express a "very high risk".

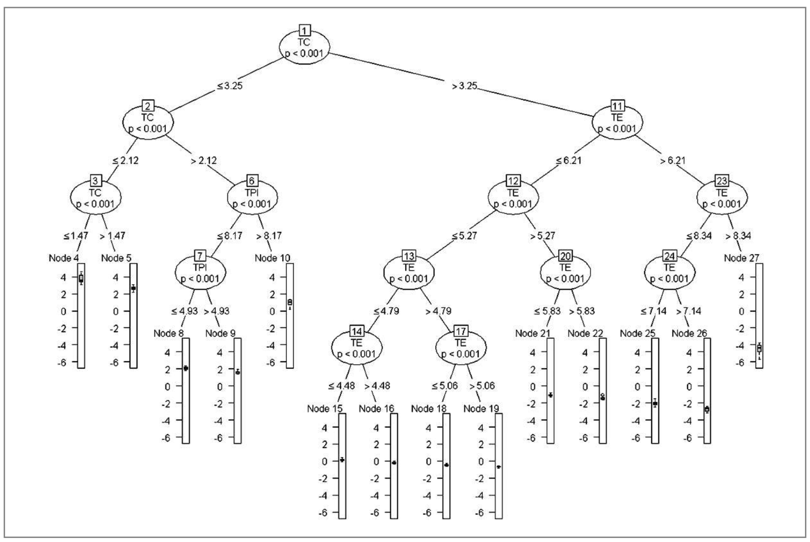

Figura 3. Árbol de clasificación índice cuantitativo del riesgo Intersección I2
Figure 3. Classification tree of risk quantitative index for intersection I2

Classification model of qualitative risk at intersections

In this point the risk at intersections is classified in a thorough and simpler manner. From 14 out of resulting groups in the quantitative model (Figure 3), a new response variable associated only to two groups was built. The criterion employed was to categorize the previous index, by considering diverse results obtained during the current study and available bibliography. The categorization was done based on the classification obtained from the highest risk group and the methodology by Hothorn et al. (2006) was applied to construct a tree, in this case called Classification Tree, because its response is of a qualitative nature. Therefore, Index values ≥ 3 y TC ≤ 1.47 [sec], are useful reference values to classify an intersection event as a high accident risk. This categorization is supported by previous researches, because it states that a high TC threshold, 1.5 sec, indicates a high collision risk.

"High accident risk" categories were established (HR), corresponding to an Index ≥ 3 and TC ≤ 1.47 [sec], and "Low accident risk" (LR), corresponding to an Index < 3 and TC > 1.47 [sec], providing as a result the adjustment of a qualitative classification tree with a binary response (Figure 4).

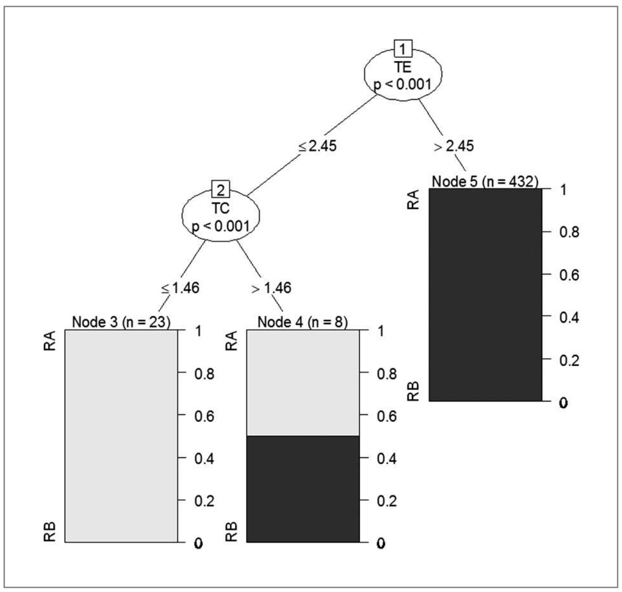

Figura 4. Árbol de clasificación cualitativo del riesgo para la intersección I2

Figure 4. Classification tree of risk qualitative risk for Intersection I2

The resulting binary tree shows that only TC and TE variables are comprised in the model. If values are TE > 2.45 [sec], the accident risk is low. In contrast, if TE≤ 2.45 [sec] and TC ≤ 1.46 [sec], the accident risk is high. For TE≤ 2.45 [sec] and TC > 1.46 [sec], there is a group which is not directly associated to classifications HR and LR and represents an uncertainty group, since the 50% of them would represent a higher accident risk. A percentage of these observations are subject to misclassifications.

Technically, from total studied conflicts under HR and LR classifications, a missclassification rate was computed for each intersection. Table 6 shows the estimated missclassification percentages when classifying data at each intersection, reaching a 2.5% at the most for I1, from the total observations registered in each one crossroads in study, which is a reasonable value for this binary response tree.

For a better understanding of selected groups for the categorization of qualitative classification model, the graphs of TC dispersion v/s Risk Index and TE v/s Risk Index for I2 are shown in Figure 5. It shows two risk groups generated by cut off points of the qualitative classification tree, taking into account that there are a 0.87% of misclassified observations (Table 6).

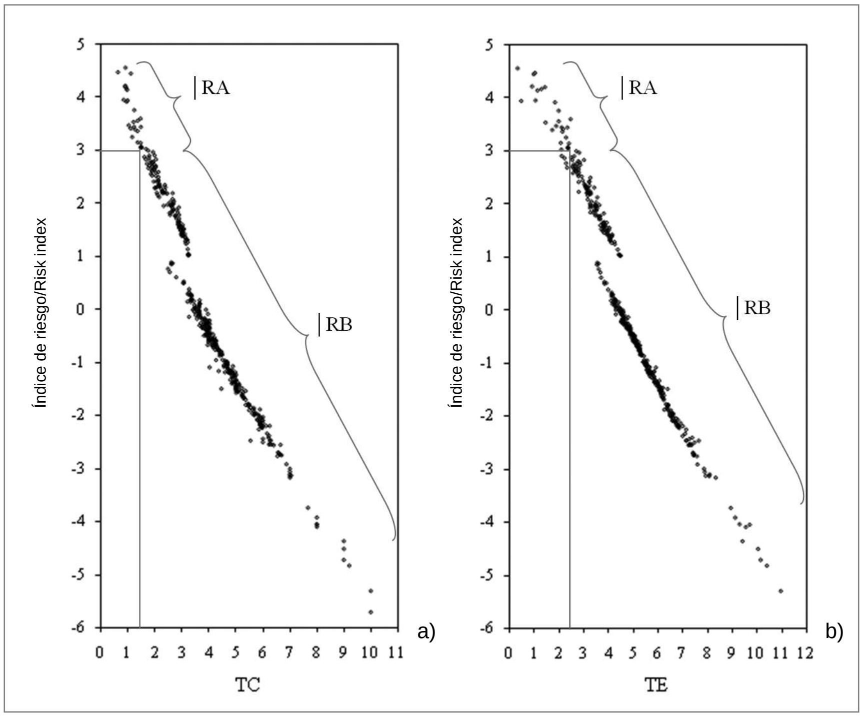

Figura 5. Gráficos de dispersión Índice de Riesgo versus a) Tiempo hasta la Colisión (TC) , b) Tiempo de Evasión (TE)

Figure 5. Graphs for dispersion risk index versus a) Time to Collision (TC), b) Time to Evasion (TE)

Tabla 6. Porcentaje de clasificación incorrecta de los datos de cada intersección con el árbol de respuesta binaria
Table 6. Percentage of data base incorrect classification for each intersection with binary response tree

	Base Predicha/Predicted Base		
Intersección/Intersection	G78 G374	G30 G374	G40 G46
Total observaciones/Total observation	635	463	591
N° mal clasificadas/Nº misclassified	16 (RB: 16; RA: 0)	4 (RB: 4; RA: 0)	13 (RB: 2; RA: 11)
Mal clasificadas/Misclassified (%)	2,5	0,87	2,2

6. Discussion

The proposed methodology for the construction of Intersection Risk Index becomes a useful tool to summarize the set of measurement obtained by employing the Traffic Conflict Technique. Such methodology facilitated the diagnosis of road safety for each intersection. Its greater benefit is the interpretation of a new indicator variable, since this study permitted the association of the first principal component with intersection risks, where each surrogate safety measurement in the index has a representative importance in relation to its intersection risk importance.

It is concluded that, for each intersections risk index, the most influent original variables are time measurement, especially TC and TE. It is reflected by the coefficients or weights associated to those variables included in each intersection index.

From TE importance among risk indexes, it was possible to prove that this is an useful surrogate safety measurement to diagnose an intersection risk, thus being a significant feature that introduces a driver’s safety perception when approaching a cross road. So it was possible to determine that higher risk index values are related to higher risk probabilities for a conflict to become an accident, meanwhile lower values indicate a lower risk.

The classification model allowed finding the existing relation between surrogate safety measurements (BD, AS, TE, TC and PET) and the constructed risk index, generating different risk groups from decision criteria associated with original variables.

The resulting model (Figure 3) shows that only time measurements TE, TC and PET influence the classification index for new data, which is explained because these variables implicitly gather the information related to BD and AS of a vehicle during a conflict. Therefore the creation model algorithm eliminates BD and AS variables, considering them as redundant information. From this model it was possible to identify different risk groups generated. In the first node, those observations presenting positive risk indexes are placed at the left and the negative ones to the right, in relation to higher and lower risk groups, respectively. Consequently, the initial partition generates an adequate classification in terms of simplifying the identification of two greater risk groups.

From the two main branches new partitions were generated, that finally end up into 14 classification groups. This huge amount of groups makes the association with a specific risk level very difficult, since it obstructs a practical interpretation for each one of them. However, its importance lays in the fact that this analysis allowed the identification of the "highest risk group", among data and conditions to be met by conflicts in order to become part of such group, Index $\geq$ 3 y TC $\leq$ 1.47 [sec], being this latter variable the one that allows the best classification of conflicts with the highest risk indexes. The percentage of cases considered by this group reaches the 5.2% of all conflicts.

The qualitative classification model built up from a binary response index makes the classification more practical against the quantitative classification tree, because it allows the association of each conflict with only two specific risk groups. In this way, the group of most serious conflicts was identified as those associated with a "High Accident Risk", and the group that represents a low risk was identified as "Low Accident Risk". The first binary category (HR) corresponds to the group of highest risk in the quantitative model and, the second binary category (LR) gathers the thirteen remaining groups.

Such models facilitate the classification of a conflict into high or low accident risk, since they only require TE and TC measurements for an adequate categorization. So it is possible to verify the importance of TC and TE variables, since from the analysis beginning they showed a high contribution to the index construction and later they became part of quantitative and qualitative classification trees, then setting the qualitative classification tree aside for BD, AS and PET.

Among qualitative model variants, TC is the most studied surrogate safety measurement, being described as the main severity conflicts measurement. Previous researches indicate that for a TC minor than 1.5 seconds there is a high risk of accident probability, which is a similar value to the one indicated for both classification trees developed by the current study, classifying conflicts as high accident risk those whose values are TC < 1.47 [sec] for quantitative model and TC < 1.46 [sec] for qualitative model. Consequently, it is possible to conclude that the algorithm employed in the creation of classification models yields a correct risk categorization for studied intersections.

As far as TE variable is concerned, which is part of qualitative model, the studies conducted for previous degree thesis are focused in its measurement and finding its relation with parameters associated to safety. However, its effectiveness as a surrogate safety measurement is not assessed. The current research had as main objective the validation of TE as a quantifying variant of traffic conflicts, obtaining as result the importance of such variant in the conducted analysis. In first instance this is observed by the almost linear correlation between TE and TC (0.878 a 0.996), which importance has been worldwide demonstrated, and due to a high contribution to the intersection risk index (26.2% average) in regards to other employed variants. TE validation was completed with the qualitative classification model, where its importance laid in becoming part of such tree and obtaining a 2.45 [sec] critical value, associated with the accident risk.

It can be concluded from a qualitative classification model (Figure 4), that it is possible to determine each group, either “high risk” or “lower risk”, only under TC $\leq$ 1.46 [sec] or TE $\leq$ 2.45 [sec] conditions, for a high risk conflict. On the contrary, TC > 1.46 [sec] or TE > 2.45 [sec] for a conflict with lowers accident risks. This is explained because the group that meets the requirements TE $\leq$ 2.45 [sec] and TC > 1.46 [sec] cannot be associated with a specific risk level (HR or LR), since it is simultaneously made of high and low risk conflicts, meaning that this group is quite an error in the model. Definitively, the error associated to the qualitative model, a 2.5% at the most, is the error made by indicating that a conflict with TE $\leq$ 2.45 [sec] is equivalent to condition TC $\leq$ 1.46 [sec] to classify conflicts in high accident risk.

The equivalence among the cut off points obtained for TE and TC, can be confirmed by observing the dispersion graphs in Figure 5, which clearly distinguishes the similarity of groups generated with a critic value TC = 1.46 [sec] and a TE = 2.45 [sec], either HR or LR, which discontinuity matches with a data structural breakdown. Finally, validating TE as a surrogate safety measurement, allows the simplification of risk evaluation for the studied intersections, since TE calculation only requires the SA and BF measurement from field sites, without needing to conduct an incident time line analysis for each traffic conflict.

7. Conclussion

Quantitative and Qualitative classification models provide criteria that allow making risk decisions on a given intersection, based on adequate cut off points and, applying the same model to information from other intersection at low error rates. Therefore, it is confirmed that Model of Intersection I2 matches other studied intersections, so this model and critical values are representative for those intersection in T shape with similar features to the ones employed in this research. In the same way, the employed methodology may be applied for the creation of risk classification models at any intersection enabling the assessment of BD, AS, TE, TC and PET variables in traffic conflicts. For later studies and based on new measurement from other "T" intersections, the adjustment of the proposed index in this exploratory stage will be adjusted.

An identified advantage for these models is the evaluation of an intersection risk after introducing infrastructure improvements and/or transit control devices, as well as a Before-After analysis, but not comparing from accidents occurrences, however only considering traffic conflicts. This is done by comparing the average variation of conflicts classified as "High Accident Risk", and after the intersection is already improved.

8. Referencias / References