Analysis of accident duration on urban arterials roads

Vandeyberg Nogueira de Souza[1], Francisco Moraes de Oliveira Neto[2]

[1]Universidade Federal do Ceará, Ceará, Brasil, vandeyberg@det.ufc.br
[2]Universidade Federal do Ceará, Ceará, Brasil, moraes@det.ufc.br

ABSTRACT
Traffic accidents are the major factor for the formation of non-recurrent traffic jams in urban roads, causing unexpected delays to the users. The understanding of the variability of traffic accident duration and its main explanatory factors can lead to a more efficient urban traffic management. Aiming to contribute to this problem, this paper analyzes the duration of traffic accidents occurring in the urban arterial roads of Fortaleza city, Brazil. Initially, a method to detect the duration of accidents from traffic data of the electronic surveillance system was proposed, given the unavailability of this information in the traffic accident database of the Fortaleza's traffic management agency. As results, the analysis of 324 accidents showed an average duration of 72 minutes and a standard deviation of 48 minutes. The Gamma density function was identified as the best fit for the duration of accidents, indicating that this variable is affected by events that occur over time, increasing the probability of the accidents being resolved. The analysis revealed that factors related to traffic agent services, individual behavior, number of vehicles and accident location (mid-block or intersection) may affect the accident duration. However, no evidence was found that the duration of accidents is influenced by its severity (accidents with or without victims) or by its nature.

RESUMO
Acidentes de trânsito são o principal fator para a formação de congestionamentos não recorrentes em vias urbanas e responsáveis por gerar atrasos inesperados aos usuários do sistema. Entender o comportamento da duração desses acidentes e identificar os seus fatores explicativos pode conduzir para um gerenciamento mais eficiente do tráfego urbano. Buscando contribuir com essa problemática, este trabalho teve como objetivo analisar a duração de acidentes de trânsito ocorridos em vias urbanas da cidade de Fortaleza-CE. Inicialmente, um método para detectar a duração dos acidentes a partir de dados de tráfego de equipamentos de fiscalização eletrônica foi proposto, dada a indisponibilidade dessa informação no banco de dados de acidentes de trânsito. Como resultados, a duração dos 324 acidentes detectados teve uma média de 72 minutos e uma dispersão de 48 minutos. A função densidade de probabilidade Gamma apresentou o melhor ajuste à duração dos acidentes, indicando que essa variável é influenciada por eventos que ocorrem ao longo do tempo, aumentando a probabilidade de o acidente ser resolvido. A análise mostrou que fatores relacionados ao atendimento dos acidentes, ao comportamento dos indivíduos, ao número de veículos envolvidos e ao local de ocorrência (meio de quadra ou interseção) podem estar influenciando a duração dos acidentes. Entretanto, não foram encontrados indícios de que a duração dos acidentes é influenciada pela sua gravidade (acidentes com ou sem vítimas) ou pela sua natureza.

1. INTRODUCTION

Traffic congestion is a problem faced daily by the urban population that needs to commute to perform their activities and one of the main barriers to the development of cities. In addition to the impact on urban mobility, previous work points to the negative effects of congestion on the environment (Bharadwaj *et al.*, 2017), the economy (Schrank *et al.*, 2015) and quality of life (Requia *et al.*, 2018).

Traffic congestion is caused by recurrent and non-recurrent events. Recurrent events are easy to predict since they are the result of daily commuting and restrictions caused by the urban road capacity and traffic operations. Non-recurrent events, in turn, occur at random, causing unexpected delays by temporarily reducing road capacity (Hojati *et al.*, 2014). Among the main non-recurrent events found in the literature, are traffic accidents, vehicle breakdowns, adverse weather and work zones. Out of these, traffic accidents seem to have the greatest impact on the formation of non-recurrent congestion, with up to 50% of non-recurrent congestion on urban roads attributed to this type of event (Chung, 2009; Li *et al.*, 2017).

Several studies published in the literature in the last decades highlight the importance of understanding the process behind the formation of non-recurrent congestion to effectively reduce the traffic jams on the urban roads (Li *et al.*, 2018). Acknowledging the influence of accidents in traffic congestion, Hojati *et al.* (2014) state that it is extremely important to understand the factors that influence traffic accident duration for improving the management of traffic accidents, as it would allow appropriate strategies to be implemented to mitigate the impacts of accidents through an efficient allocation of equipment and personnel. However, previous studies have mostly focused on analyzing accident duration on highways or freeways of developed countries (Li *et al.*, 2018), which have different traffic operations and policies than arterials roads of urban networks, such as in network systems of Brazilian major cities, making it impossible to use their results to understand the behavior on urban roads of developing countries.

Aiming to contribute to this issue, the main objective of this work is to analyze the duration of traffic accidents that occurred on arterials roads in the city of Fortaleza. For this purpose, due to the limited availability of data, a method was initially defined to indirectly determine the duration of accidents by relating data on accident characteristics obtained from the traffic accident database of the Fortaleza's traffic management agency and traffic data obtained from electronic surveillance equipment located on the main arterial roads of Fortaleza. The analyzes focused only on traffic accidents, disregarding other types of incidents that may impact traffic, such as vehicle breakdown and weather events.

2. LITERATURE REVIEW
2.1. Definition and detection of the duration of traffic accidents

The duration of a traffic accident is generally defined as the period between its occurrence and the moment of road clearance (Li *et al.*, 2018). This period is composed of four phases: (1) Detection time: time between the occurrence of the accident and its detection by traffic agents; (2) Preparation time: period between the end of the previous phase and the departure of agents for service; (3) Travel time: travel time of agents to the scene of the accident; (4) Clearance time: period required for agents to clear the road. However, this definition may vary according to the study performed. Some studies, for example, do not consider the detection phase (Chung, 2009; Junhua *et. al*, 2013), while other studies consider a traffic recovery phase: from the moment of road clearance to the return of traffic to its normal conditions (Hojati *et al.*, 2014; Haule *et al.*, 2018).

To determine the duration of traffic accidents, most studies use accident data provided by the traffic agency responsible for managing the road system (Chung, 2009; Li *et al.*, 2017). In general, the information provided include the duration of each accident phase and other accident characteristics, such as location, severity, nature and number of vehicles involved.

However, this type of source does not provide the traffic detection and recovery phases (Li *et al*., 2018). For such cases, or in studies where the database does not provide information on the duration of accidents, other sources are used to obtain this information indirectly, such as traffic data recorded by loop detectors.

Previous studies have therefore suggested to determine the duration of accidents by detecting anomalies in traffic conditions near the reported accident site and time (Hojati *et al*. 2014; Haule *et al*., 2018). The main traffic variables adopted to detect the duration of accidents on highway systems are speed, flow and density (Li *et al*., 2018). Since these variables were defined as part of methods developed and applied to determine the duration of accidents on uninterrupted roadways, they may not be suitable for detecting accidents on urban arterial roads. The development of methods for determining the duration of accidents on arterial roads is still a challenge, since these roads are expected to have a greater variability of traffic conditions.

2.2. Analysis of the duration of traffic accidents

Prior studies have showed that duration of traffic accidents has a high variance (standard deviation of approximately 70% of the mean) and an asymmetrical distribution with a long right tail (Junhua *et al*., 2013). Considering this behavior, some probability distributions have been suggested to represent the variability of the accident duration variable. Some authors have suggested the Log-normal distribution as the best fit for the duration of traffic accidents (Chung and Yoon, 2012; Xie *et al*., 2015), while other researchers have indicated that the duration of accidents is better represented by the Log-logistics (Nam and Mannering, 2000; Junhua, *et al*., 2013) or Weibull distribution (Kaabi *et al*., 2012; Hojati *et al*., 2013). Based in these results, Li *et al*. (2018) pointed out that the appropriate definition of a probability distribution that fits the characteristics of the duration of traffic accidents is one of the key tasks in the analysis. This definition allows the adjustment of parametric models to analyze the influence of explanatory factors on the duration of accidents.

As highlighted by Haule *et al*. (2018), the duration of traffic accidents is influenced by several factors, among which are: accident characteristics (severity, nature, location, number of vehicles involved), weather conditions, road geometry (road layout, number and width of lanes, roadway shoulder availability), traffic control signals (vertical and horizontal signals), traffic conditions (flow, speed, occupancy), temporal factors (day, time, month), among others. Hojati *et al*. (2014), for example, analyzed factors that could influence the duration of road traffic accidents in Queensland, Australia, and suggested that the variability of accidents duration is affected by the severity and type of accident, the need for medical service, roadway shoulder availability, time of day and traffic conditions. In another study, on expressways of Shanghai, China, Shi (2014) identified that the number of vehicles involved, type of accident, number of lanes, location of the accident, and time period significantly influence the duration of traffic accidents.

Despite these advances by previous studies in the understanding of the factors that determine the duration of traffic accidents, Li *et al*. (2018) indicate that most studies focused on accidents occurring on highways or expressways (uninterrupted flow roads). Thus, there is a lack of studies that analyze the variability of the duration of traffic accidents in interrupted flow roads. Hojati *et al*. (2014) also point out that many of the research studies cannot be generalized to other localities because the characteristics of the regions may be different, as well as the data available for the studies.

3. DATA COLLECTION AND PRE-PROCESSING

The accident data were obtained from the Traffic Accident Information System of Fortaleza (SIATFOR - Sistema de Informações de Acidentes de Trânsito de Fortaleza), which is an accident database kept by the Traffic Management Agency of Fortaleza (AMC - Autarquia Municipal de Trânsito e Cidadania). The traffic data used in this research were also provided by the AMC through the Surveillance System of Fortaleza, which detects vehicle passage on the main arterial roads by a set of inductive loops. Both accident and traffic datasets were for the period between January 2015 and June 2018. The SIATFOR dataset has 52,503 records containing the time, location, type and severity of each accident, as well as the number and type of vehicles involved. While the traffic dataset has the date and time information of each vehicle's passage, lane used, vehicle type, and detected speed. The records for each vehicle passage were obtained from 251 loop detectors spread across the main arterial roads of the city.

Accidents that occurred in periods of low traffic, before 06:00 am and after 07:00 pm, were removed from the dataset given that in these periods is not possible to detect any anomaly in traffic conditions due to traffic accidents. In addition, only accidents that occurred no farther than 50 meters from a loop detector were considered, assuming that a longer distance could make it impossible to estimate the duration of the accident. Furthermore, accidents with incomplete or inconsistent data were removed from the sample. After this pre-processing, an integrated dataset with 1,785 traffic accidents was obtained.

It is noteworthy that motorcycle records were removed from the traffic dataset. In the urban streets of Brazil, Motorcyclists usually move between other vehicles and can reach higher speeds even in congested conditions, thus they are not affected by the occurrence of accidents. Because of this behavior, to include motorcycle records on the dataset may compromise traffic accident detection from traffic data records.

4. METHODOLOGY

The proposed methodology to achieve the objectives of this study consists of two main steps: i) determination of the duration of traffic accidents based on traffic conditions, and ii) analysis of the variability of the duration of traffic accidents, based on an exploratory analysis of the duration of accidents and the inference of the possible explanatory factors of this duration. Figure 1 outlines the proposed method.

4.1. Detection of the duration of traffic accidents

The basic idea for detecting the duration of traffic accidents is to compare the daily profile of each traffic variable (aggregated at time intervals of size t along the day) on the day of the accident with its historical profile and to identify, for a certain level of confidence, significant differences between the two profiles. For this purpose, daily traffic profiles are defined based on traffic data on the day of the accident and an expected or historical traffic profile is obtained using data from the 20 days before and after the accident day (aiming to smooth the effect of seasonality). The analysis period was set to be two hours before and four hours after the reported accident time.
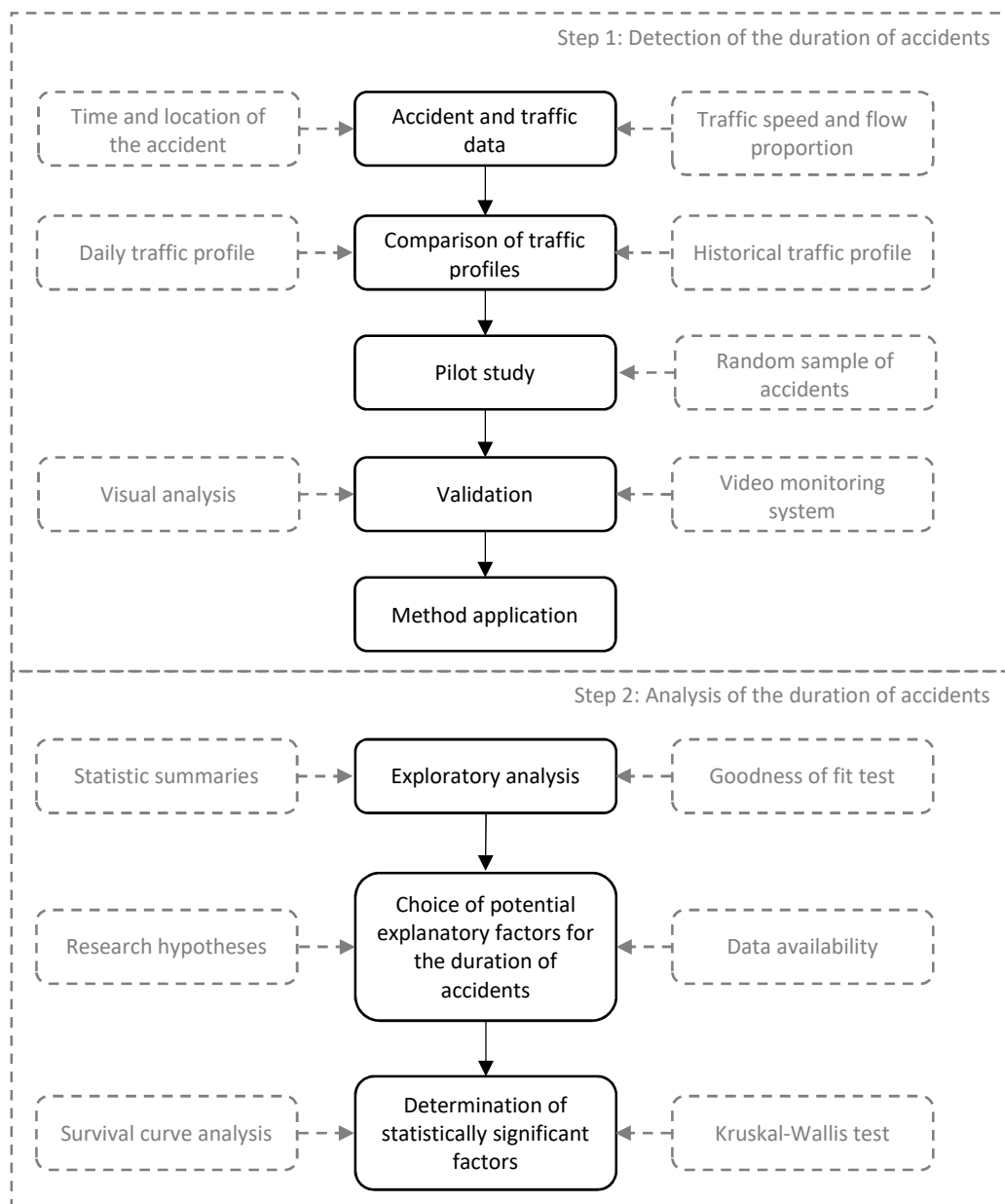
Figure 1. Framework for the proposed methodology

The traffic profiles were constructed for two variables: traffic speed (s) on an arterial road and traffic flow proportion (p) by lane of a multiple lane street. In the case of a single lane street, only the traffic speed profile was determined. The traffic flow proportion (p_i) of a given lane i is defined by Equation 1, where f is the total traffic flow on a given direction of a arterial road; and f_i is the traffic flow on the lane i. The main reason of using p_i to detect the duration of traffic accidents was that, under typical traffic conditions, p_i has less variability than the traffic flow f_i and is more sensitive to any traffic interruption, enhancing the power of detection method proposed.

$$p_i = 100 * \frac{f_i}{f} \tag{1}$$

4.1.1. Development of traffic profiles and detection criteria

To build each historical profile, the observations of the variables traffic speed and traffic flow proportion were aggregated at time intervals of t minutes. Thus, each interval t consisted of 40

observations (one for each day of historical data). For each time interval t, a truncated mean of the 40 observations was calculated, disregarding 10% of the sample to minimize outlier effects, and a Prediction Interval (PI_t), for a given confidence level of 100 (1-α)%, was determined. The prediction interval seeks to predict a single future value of a variable and is calculated by Equation 2. Where, PI is the prediction interval, n is the sample size, s is the sample standard deviation, α the significance level and $t_{a/2;n-1}$ is the critical value of the t-student distribution for ($n-1$) degrees of freedom.

$$PI = t_{\alpha/2; n-1} * s \sqrt{1 + \frac{1}{n}} \tag{2}$$

The profiles of traffic speed and traffic flow proportion by lane on the day of the accident were generated with the same aggregation interval as the historical profile. Due to the high variability of (s) and (p) profiles, even under conditions without the occurrence of a traffic accident, the moving average technique was applied to smooth the generated profiles and minimize the variability of these variables.

The daily and historical traffic profiles were then compared to indentify the period when the road was obstructed by the vehicles involved in the accident. To determine this interruption period or the accident duration, the first step was to identify a minimum sequence of (m) aggregation intervals (t) in which the values of both variables (s) and (p_i) on the day of the accident were outside the limits of the prediction interval defined by the historical profile. After identifying this period in which traffic profiles are outside the expected traffic conditions, the moment when the traffic accident occurred was determined by analyzing the 15 minutes prior to the first aggregation interval (t_1) of the sequence of m intervals in which traffic conditions were outside the prediction interval. It was assumed that the accident occurred at the interval before t_1 where it was observed an abrupt change (d) in the distance between the daily and historical profiles. Finally, the moment of road clearance was identified by analyzing the aggregation intervals between t_1 and t_m to search for the moment when the traffic variables on the day of the accident seemed to return to the expected conditions for that period of day.

4.1.2. Pilot study and validation

To define the values of t, m, and d a pilot study was conducted with 80 randomly chosen accidents. Eight scenarios were tested based on two values of aggregation intervals, 3 or 5 minutes, two values of minimum number of subsequent aggregation intervals, 3 or 5 intervals, and two values for the abrupt change in distance between daily and historical profiles for each variable, 10% or 15% for profiles based on lane flow proportions and 15% or 25% for speed profiles. To determine the best scenario, the results provided by the proposed method were compared with a visual analysis of graphs of traffic speed and traffic flow propotion. The best scenario was chosen based on two criteria: sensitivity (probability of the method identifying an accident that was verified by the visual analysis) and specificity (probability of the method not identifying an accident that was not verified by the visual analysis).

Finally, the method was validated by comparing, for a validation sample of accidents, the phases of the duration of accidents estimated by the proposed method with those observed by the video monitoring system of the Traffic Control Center of Fortaleza (CTAFOR - Controle de Tráfego em Área de Fortaleza). Using the video monitoring system is possible to record the moments in which the traffic agents detect the accident, arrive and leave the scene of the accident. Thus, the start time of a given accident detected by the method was considered

satisfactory if it was earlier than that recorded by CTAFOR. While the estimated moment of road clearance was considered correct if it was between the arrival and departure times of agents from the accident site.

4.2. Analysis of the duration of traffic accidents

Innitially, an exploratory analysis of the duration of accidents was carried out in which measures of central tendency, dispersion and form of distribution of the variable were calculated. Goodness of fit tests were performed to test whether the variable follows any of the following probability distribution proposed in the literature: Weibull, Gamma, Lognormal, or Loglogistic. Possible explanatory factors for the duration of accidents were defined based on the literature review and data availability. Based on this analysis, hypotheses of the effect of these factors on accident duration were defined, which were initially analyzed by the development of survival curves of accident duration for each factor. The survival curve is defined as a probability function of a given individual surviving for a further period of time given that it has survived up to a certain point. Events that occur over time (e.g., arrival of agents) may or may not extend the survival, or in the case of accidents, their duration. In addition to the visual analysis of the survival curves, the Kruskal-Wallis nonparametric test was performed, since the analyzed variable does not fit the assumptions of homoscedasticity and normality requested by the classic analysis of variance (ANOVA) test.

5. RESULTS AND ANALYSIS

This section presents the application of the method to detect and analyze the duration of accidents in the city of Fortaleza.

5.1. Detection of the duration of traffic accidents

The detection method proposed in Section 4.1 was applied to the dataset of 1,785 traffic accidents. It is worth noting that for a portion of these accidents the method was not applicable, due to some limitations, such as: i) the accident occurred on the opposite direction of the traffic flow detected by the surveillance equipment; ii) some accidents were probably resolved quickly and had no significant impacts on traffic, making it impossible to detect their duration from traffic data. These accidents could not be removed from the dataset during the pre-processing due to limitations on the information obtained.

5.1.1. Pilot study

The pilot study to define the values of *t*, *m*, and *d* was performed with 80 traffic accidents randomly drawn from the dataset. Through a visual analysis of the curves of traffic speed and traffic flow proportion for different aggregation intervals, it was possible to manually detect the duration of 22 accidents (approximately 30% of the calibration sample). After this visual analysis, the method was applied for the 80 accidents and the results obtained for each scenario (presented in Table 1) were compared with those of the visual analysis. Scenario 4 presented the best result by detecting the duration of 18 of the 80 accidents analyzed. Out of these 18 accidents, 15 were among the 22 detected by visual analysis, resulting in a sensitivity of 68%, while 3 were erroneously detected, resulting in a specificity of 95%.

Table 1 – Results obtained in the pilot study for the eight scenarios analyzed

| Scenario	t (min)	m	d (%)		Number of detected accidents		Sensitivity (%)	Specificity (%)
			s	p	Total	Correct		
1	3	3	10	15	28	16	73	79
2	3	3	15	25	25	15	68	83
3	3	5	10	15	23	13	59	83
4	3	5	15	25	18	15	68	95
5	5	3	10	15	27	16	73	81
6	5	3	15	25	24	15	68	84
7	5	5	10	15	23	15	68	86
8	5	5	15	25	17	14	64	95

5.1.2. Validation

To validate the method, a sample of 10 traffic accidents detected by the proposed method and observed by the video monitoring system of the CTAFOR was analyzed. This small number of accidents used for validation was due to the limited number of locations equipped with both loop detectors and video monitoring system. Analyzing this sample, it was observed that the time detected by the method as the beginning of the accident was prior to the detection time recorded by CTAFOR for all 10 accidents. As for the road clearance phase, in 6 out of 10 accidents the time estimated by the method was between the arrival and departure time of traffic agents at the accident site, as expected. While for the other four accidents, the estimated road clearance happened either right before the agent arrival, suggesting that the drivers reached an agreement, or right after the agent departure, suggesting that the drivers left the scene after the agents. The differences between estimated road clearances and the recorded arrivals or departures of traffic agents were not more than 3 minutes. In spite of the small sample used for validation, the events estimated by the method are reasonable compared to real events, so the validation results obtained suggest that the method is adequate to detect the duration of traffic accidents that occur on interrupted flow roads.

5.1.3. Method application

The proposed method was applied to all 1,785 accidents in the dataset, with the parameters defined in the pilot study. The method was able to estimate the duration of 324 accidents, approximately 18% of the total. This reduction in the sample was because the method is not applicable for all accidents as explained at the beginning of Subsection 5.1. Figures 2 and 3 illustrate how the method detects an accident based on the variation of both profiles of traffic speed by lane and traffic flow proportion by lane, respectively. The time of detection reported by the traffic officers for the observed accident was 04:30 pm. However, from the analyzed profiles it can be observed that at 04:15 pm the speed at both lanes reduced abruptly, while the traffic flow proportion on the lane 1 increased from approximately 50% to almost 100%. These facts indicate that the accident occurred at this time, blocking lane 2 of the road. At 04:30 pm, the speed of vehicles and the traffic flow proportion begun to return to normal conditions, indicating that the vehicles involved in the accident have been removed from the road. At 04:55 pm, traffic conditions (traffic speed and traffic flow proportion) returned to their normal conditions. The fact that the time of detection reported by the traffic officers coincides with the

time when traffic conditions begun to return to normal conditions indicates that the traffic officers reported the accident right after they finished the occurrence report and requested the drivers to release the road.
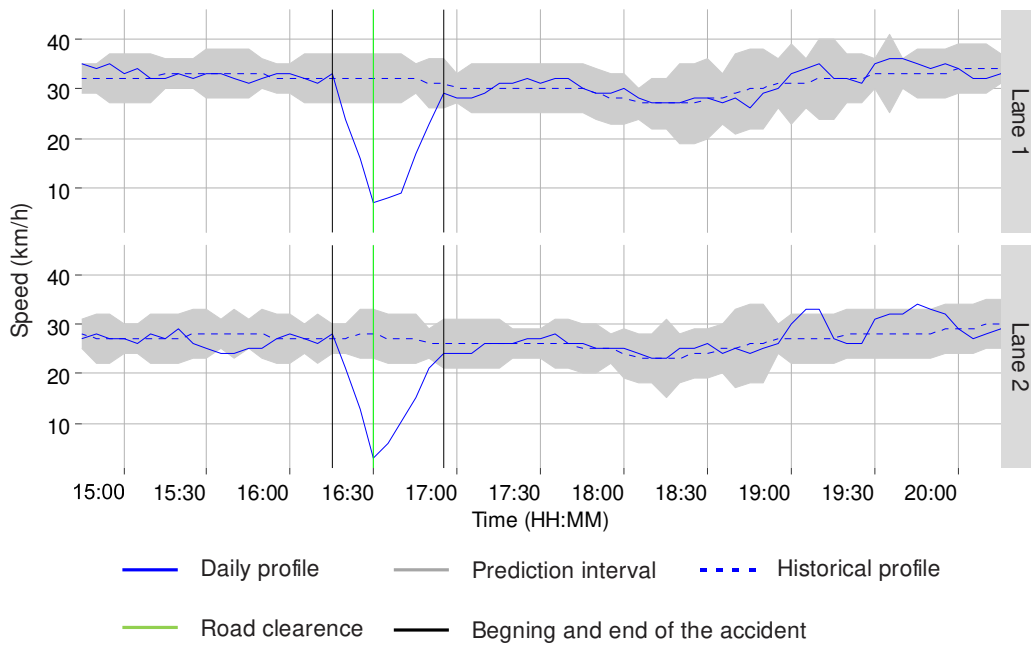
Figure 2. Estimated accident duration from the traffic speed profile by lane
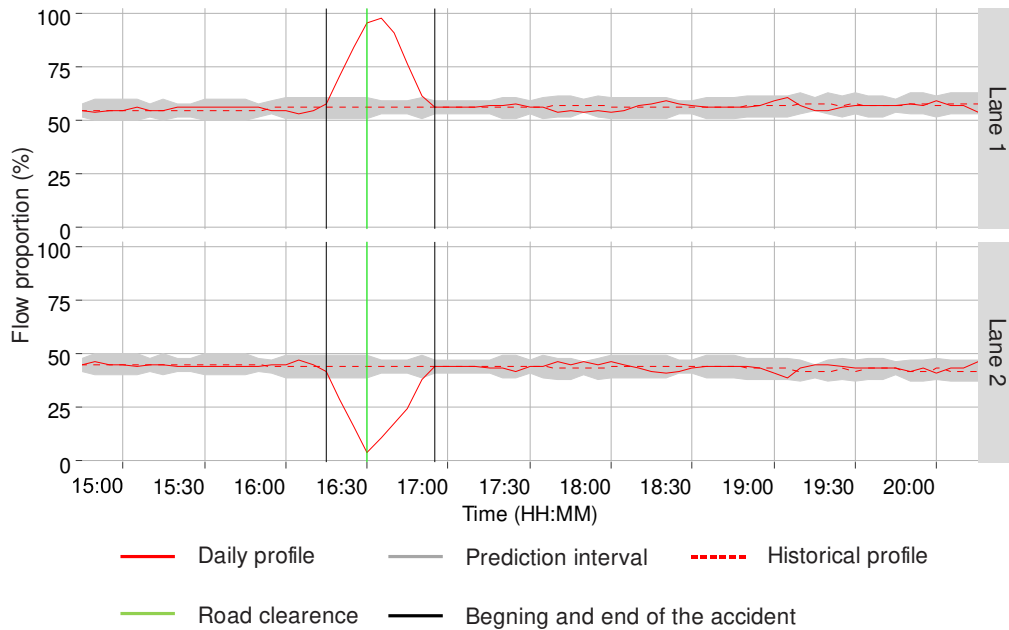
Figure 3. Estimated accident duration from the profile of traffic flow proportion by lane

5.2. Analysis of the duration of traffic accidents

Figure 4 shows the spatial distribution of the 324 accidents detected in the previous section, which shows that the accidents are spread out over the main arterial roads of Fortaleza city. As will be seen, it is possible to analyze the difference in accident duration for different regions of the network system.
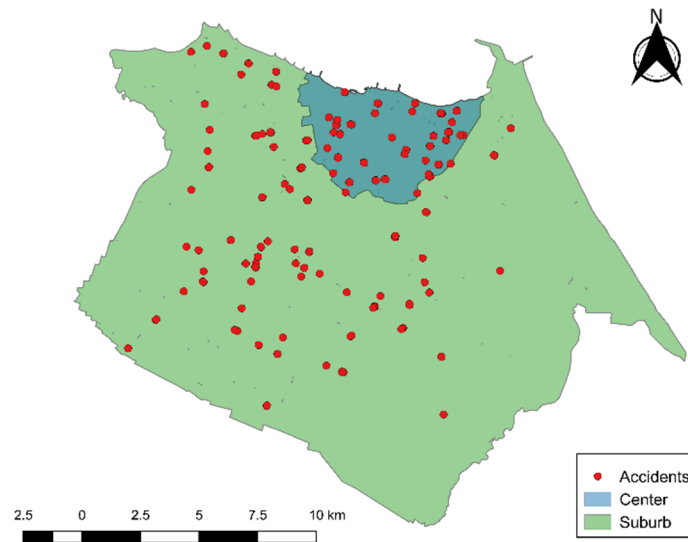
**Figure 4.** Spatial distribution of detected accidents

5.2.1. Exploratory analysis

Table 2 presents the statistics for the duration of the 324 accidents detected. The results showed that accidents had an average duration of 72 minutes and a high variability (standard deviation of 48 minutes), ranging from about 10 minutes to more than 3.5 hours, which suggests that the duration of accidents may be impacted by some factors.

Table 2 – Statistic summaries for the sample of accident duration

Statistic	Value	Statistic	Value
Sample size	324	1º quartile	39 (min)
Mean	72 (min)	Median	60 (min)
Standard deviation	48 (min)	3º quartile	93 (min)
Minimum	9 (min)	Skewness	1.27
Maximum	210 (min)	Kurtosis	2.02

Figure 5 shows the accident duration histogram, which has an asymmetric form to the right, with a considerable proportion of accidents below the mean. The Anderson-Darling (A-D) statistical test was applied to verify whether the sample data was drawn from any of the main probability distributions suggested in previous studies: Weibull, Lognormal, LogLogistic and Gamma (as shown in Figure 5). The null hypothesis of the A-D test is that a given variable follows a known probability distribution. According to the test results, at a significance level of 5%, the null hypothesis was rejected for three of the four distributions (Lognormal, Loglogistic and Weibull), and could not be rejected for the Gamma distribution (p-value > 0.25) with shape parameter of 2.36 and scale of 30.74. The best adjustment of the Gamma function indicates that the likelihood that the accident will be finished and the road will be cleared soon increases over time. This increased behavior on the probability of accident ending is probably related to some events (e.g., traffic congestion, arrival of agents, weather conditions, and so on) that can occur over the duration of the accident and change drivers' attitude.
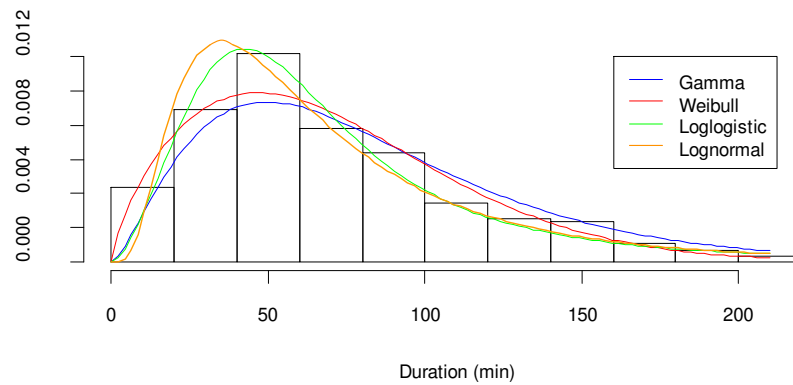
Figure 5. Histogram and fitted probability distributions for accident duration

5.2.2. Analysis of potential explanatory factors for the duration of traffic accidents

The possible explanatory factors for the duration of traffic accidents in Fortaleza were proposed based on the literature and hypotheses about the phenomenon for the study region, which are presented below:

I. Accidents with non-injuries last less than accidents with injuried victims, since the Brazilian Traffic Code (CTB), in its Article 178, requires that drivers involved in an accident with minor damages must clear the road as quickly as possible, being subject to fines for violation of this traffic law;

II. Accidents involving many vehicles last longer than the other types of accidents since they usually involve more drivers and may lead to more damages or injuried people;

III. Accidents that occur in the central region of the city have a shorter duration compared to accidents in more peripheral regions, since the traffic agent crew is more concentrated on the central region due to the higher traffic flows in this region, what makes the agent attendance to the traffic occorrences faster in the central areas of the city;

IV. Accidents occurring during peak hours last more than accidents during off-peak periods, since during peak hours the planned number of traffic officers may not be sufficient to handle all occurrences and congestion may also delay the arrival of these officers to the accidente site;

V. Accidents occurring at intersections may be shorter compared to mid-block accidents because they generally have a greater impact on traffic, thus having a higher priority by the traffic officers;

VI. Rear-end collision accidents are shorter in duration than side-impact collision accidents because the traffic laws state that a driver of the car that hit a leading car is usually considered the faulty driver, what makes rear-end accidents easier to solve, most of the time, with no need for help of traffic agents in the scene;

VII. Accidents in congested roads tend to be shorter because they have a higher priority by the traffic officers due to their greater impact on traffic.

Based on the hypotheses listed above and data availability nine factors were considered, namely: Day of the week, Period of the day, Zone of the city, Location of the accident, Volume/ capacity ratio (V/C), Severity, Type of collision, Number of vehicles and Heavy vehicle involvement. Table 3 presents the levels of the analyzed factors and the p-values for the

Kruskal-Wallis tests. According to the test results, two factors, Period of the day and Zone of the city, had a significant influence on the traffic accident duration, at a significance level of 5%, while two other factors, Location and Number of vehicles involved, were significant at a significance level of 10%. By analyzing the survival curve for each factor (Figure 6), it can be seen that accidents occurring either during off-peak hours or in the city center tend to last less than accidents occurring during peak hours or in the suburbs, respectively, supporting the suggested hypotheses III and IV. The survival curves also show that accidents involving up to two vehicles or occurring at intersections tend to end earlier, when compared to accidents with more than two vehicles involved or that occurred in the mid-block, suggesting that the hypotheses II and V might be right.

However, some hypotheses were not supported by the analysis. The main unexpected result was that hypothesis I, that accidents with non-injuries usually have a shorter duration, was not confirmed by the analysis. Three reasons may explain this result: 1) the classification of accidents into only two levels of severity (i.e., accidents with or without injuries) in the dataset is not enough to represent all possible types of injuries in traffic accidents, as for example, minor injuries may not affect accident duration, since the victims usually do not need any medical attention; 2) drivers usually do not follow the guidelines of the CTB and do not clear the road after an accident with non-injuries, what makes this type of accidents last longer than expected; 3) other factors may has a greater influence than severity on accident duration (for example, pile-up collisions are expected to have long durations even without injuries).

Table 3 – Description of possible explanatory factors for the duration of accidents

Factor	Levels	Value	Kruskal-Wallis test (p-value)
Day of the week	5	1 – Monday 2 – Tuesday 3 – Wednesday 4 – Thursday 5 – Friday	0.779
Period	2	1 – Peak period 2 – Off-peak period	0.026
Zone of the city	2	1 – Center 2 – Suburb	0.005
Location	2	1 – Intersection 2 – Mid-block	0.083
V/C	3	1 – V/C < 0,50 2 – 0,50 ≥ V/C < 0,75 3 – V/C ≥ 0,75	0.200
Severity	2	0 – with non-injuries 1 – with injuries	0.332
Type of collision	2	1 – Rear-end collision 2 – Side-impact collision	0.769
Number of vehicles	2	1 – Up to two vehicles 2 – Three or more vehicles	0.100
Heavy vehicle involvement	2	1 – No; 2 - Yes	0.354

The factors Day of the week, Type of collision, Heavy vehicle involvement and Congestion levels (V/C) were also not supported, according to the Kruskal-Wallis test. However, when analyzing the survival curve for different V/C levels it is clear that accidents occurring in more

congested places tend to end earlier, with a higher probability of ending earlier after the first 50 minutes. One hypothesis for this divergence in results is that other factors may confound the congestion effect on accident duration, which violates the sample independence assumption of the Kruskal-Wallis test.
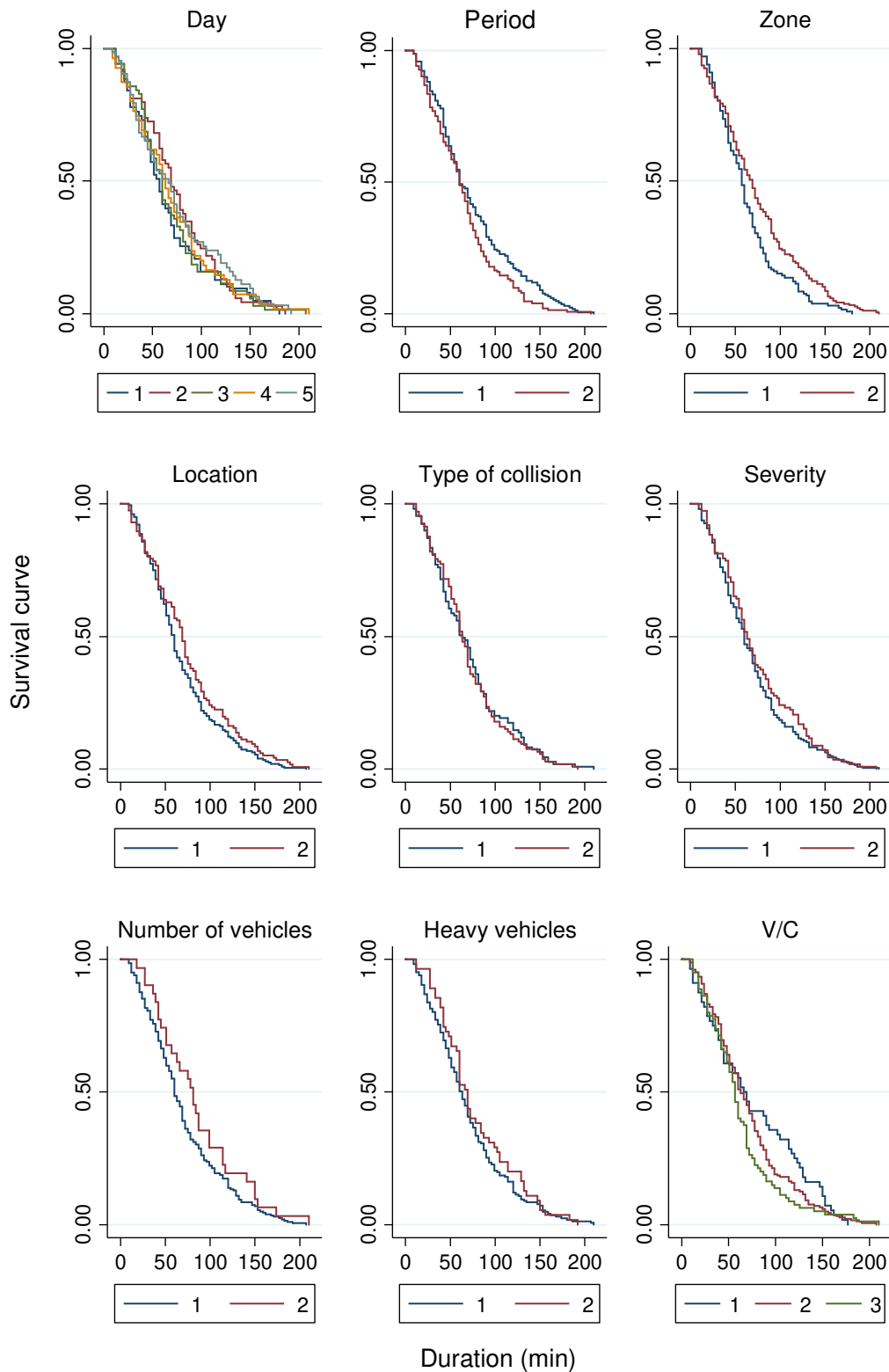
Figure 6. Survival functions of accident duration for different factors

6. CONCLUSIONS

This paper presents an analysis of the duration of traffic accidents that occurred on arterial roads in the city of Fortaleza, Brazil. A method for detecting the duration of traffic accidents on urban roads was developed using traffic data from loop detectors and accident information dataset. A new traffic variable (traffic flow proportion) has been proposed to detect the duration of accidents, since it presents less variability than the variables used in previous works and a greater sensitivity to the occurrence of traffic accidents.

The exploratory analysis of the duration of the 324 accidents detected by the method showed that the average duration of an accident in Fortaleza's arterial roads was 72 minutes, with a standard deviation of 48 minutes. The Gamma probability distribution with estimated shape parameter of 2.36 and scale of 30.74 presented the best fit to the analyzed data, indicating that the probability of the accident being finished soon increases over time, probably due to events that may occur over the accident duration (e.g., arrival of agents, traffic congestion, weather conditions, and son) that may change drivers' attitude.

The non-parametric Kruskal-Wallis tests and the survival curve analyzes of the sample of accident duration pointed to the significant influence of four factors (Period of the day, Zone of the city, Location of the accident and Number of vehicles involved) on the duration of traffic accidents, showing mainly that accidents occurring during peak hours and in congested locations tend to be resolved faster. The analysis of the influence of factor V/C on accident duration also corroborates this result. Although this factor did not have a significant influence on the duration of accidents by the Kruskal-Wallis test, the survival curves for different levels of congestion showed that accidents on roads with higher traffic flow tend to end faster after a certain length of time. The main hypotheses suggested for these results was firstly that locations with higher congestion levels usually are better served by traffic agents, and secondly that drivers may be more concern to clear the road when the impact on traffic is greater.

The results obtained did not confirm some of the proposed hypotheses, such as that the existence of injuries, as well as that the involvement of heavy vehicles, would result in longer accidents. This may suggest that the existence of other factors (e.g., data limitations or other variables not considered) that may have confounded the effects of the variables considered in the analysis, pointing to the need to analyzing the effects of several factors altogether.

It is noteworthy that other factors were not analyzed in this work, such as weather conditions and characteristics of the drivers involved. In addition, the accident data presented some limitations regarding to the identification of severity, the type of collision, and the identification of the exact location of the accident and the direction of traffic that was affected by the accident. Furthermore, only a few accidents observed by CTAFOR's video monitoring system could be used to validate the detection method.

Finally, the next steps for the development of this work are to estimate and analyze the joint effect of several explanatory factors on the duration of traffic accidents, using classical statistical models proposed in the literature, and to investigate the impact of unobserved effects or heterogeneity on parameter estimates.

REFERENCES

Bharadwaj, S.; S. Ballare; Rohit e M. K. Chandel (2017) Impact of congestion on greenhouse gas emissions for road transport in Mumbai metropolitan region. *Transportation Research Procedia*, v. 25, p. 3538-3551. Doi:10.1016/j.trpro.2017.05.282

Chung, Y. (2009) Development of an accident duration prediction model on the Korean Freeway Systems. *Accident Analysis and Prevention*, v. 42, n. 1, p. 282–289. Doi:10.1016/j.aap.2009.08.005

Chung, Y. e B. J. Yoon (2012) Analytical method to estimate accident duration using archived speed profile and its statistical analysis. *KSCE Journal of Civil Engineering*, v. 16, n. 6, p. 1064–1070. Doi:10.1007/s12205-012-1632-3

Haule, H. J.; T. Sando; R. Lentz; C. H. Chuan e P. Alluri (2018) Evaluating the impact and clearance duration of freeway incidents. *International Journal of Transportation Science and Technology*, v. 8, n. 1, p. 13–24. Doi:10.1016/j.ijtst.2018.06.005

Hojati, A. T.; L. Ferreira; S. Washington e P. Charles (2013) Hazard based models for freeway traffic incident duration. *Accident Analysis and Prevention*, v. 52, p. 171–181. Doi:10.1016/j.aap.2012.12.037

Hojati, A. T.; L. Ferreira; S. Washington; P. Charles e A. Shobeirinejad (2014) Modelling total duration of traffic incidents including incident detection and recovery time. *Accident Analysis and Prevention*, v. 71, p. 296–305. Doi:10.1016/j.aap.2014.06.006

Junhua, W.; C. Haozhe; e Q. Shi (2013) Estimating freeway incident duration using accelerated failure time modeling. *Safety Science*, v. 54, p. 43–50. Doi:10.1016/j.ssci.2012.11.009

Kaabi, A. Al; D. Dissanayake e R. Bird (2012) Response Time of Highway Traffic Accidents in Abu Dhabi: Investigation with Hazard-based Duration Models. *Transportation Research Record: Journal of the Transportation Research Board*. v. 2278, n. 1, p. 95-103. Doi:10.3141/2278-11

Li, R.; M. Guo; H. Lu e M. Guo (2017) Analysis of the Different Duration Stages of Accidents with Hazard-Based Model. *International Journal of Intelligent Transportation Systems Research*, v. 15, n. 1, p. 7–16. Doi:10.1007/s13177-015-0115-6

Li, R.; F. C. Pereira e M. E. Ben-Akiva (2018) Overview of traffic incident duration analysis and prediction. *European Transport Research Review.* v. 10, n. 22. Available in:< https://etrr.springeropen.com/track/pdf/10.1186/s12544-018-0300-1> (Consulted on 08/02/2020).

Nam, D. e F. Mannering (2000) An exploratory hazard-based analysis of highway incident duration. *Transportation Research Part A: Policy and Practice*, v. 34, n. 2, p. 85–102. Doi:10.1016/S0965-8564(98)00065-2

Requia, W. J.; C. D. Higgins; M. D. Adams; M. Mohamed e P. Koutrakis (2018) The health impacts of weekday traffic: A health risk assessment of PM2.5 emissions during congested periods. *Environment International*, v. 111, p. 164-176. Doi:10.1016/j.envint.2017.11.025

Schrank, D.; T. Lomax; R. Fellow e J. Bak (2015) 2015 URBAN MOBILITY SCORECARD. Texas, *Texas A&M Transportation Institute e INRIX*, Available in: <https://pdfs.semanticscholar.org/2efb/66f800cd8f3218c348fd83ce8443f8138c0d.pdf?_ga=2.116744134.1536304838.1580230472-1854347183.1580230472> (Consulted on 08/02/2020)

Shi, Y. (2014) Survival Analysis of Urban Traffic Incident Duration : a Case Study at Shanghai Expressways. *Journal of Computers,* v. 26, n. 1, p. 29-39. Available in: < http://www.csroc.org.tw/journal/JOC26-1/JOC26-1-3.pdf > (Consulted on 08/02/2020)

Xie, K.; K. Ozbay e H. Yang (2015) Spatial analysis of highway incident durations in the context of Hurricane Sandy. *Accident Analysis and Prevention*, v. 74, p. 77–86. Doi:10.1016/j.aap.2014.10.015