Applying Extreme Value Models to Surrogate Measures for Traffic Safety Analysis

Attila Borsos

Applying Extreme Value Models to Surrogate Measures for Traffic Safety Analysis

by

Attila Borsos

to obtain the degree of Master of Science in Civil Engineering
at the Delft University of Technology,
to be defended publicly on November 13, 2019 at 11 AM.

Student number:	4746538	
Project duration:	May 1, 2018 – October 31, 2019	
Thesis committee:	Prof. Dr. M. Hagenzieker,	TU Delft, supervisor, committee chair
	Dr. H. Farah,	TU Delft, daily supervisor
	Dr. J. Juan Cai,	TU Delft, supervisor
	Dr. A. Laureshyn,	Lund University, Sweden, external supervisor

This thesis is confidential and cannot be made public until October 31, 2019.

TU Delft

LUND UNIVERSITY

To my wife Laura
and our two beautiful sons
Bendegúz and Barnabás
for all their support.

Executive summary

Improving road safety is a common global interest which can be supported in several domains, one of which is safer road infrastructure. To identify and localize the most critical elements on our road network, one has to be able to measure how (un)safe a particular location is. The most common way to evaluate safety, also widely applied in practice, is investigating the occurrence and severity of crashes using historical data. This approach however has a number of limitations, the most important of which is probably its reactive nature.

An alternative method using non-crash events has gained a lot of attention recently, especially thanks to the rapid improvement of sensing technologies. By gathering trajectory data and calculating various Surrogate Measures of Safety it has become possible to analyse safety without waiting for accidents to happen. Using these indicators combined with Extreme Value Theory (EVT) one can estimate the probability of crashes as extreme (unobserved) events. EVT offers two approaches: the Block Maxima (BM) approach divides the sample time into blocks and samples the largest value in each block, whereas the Peak-over-Threshold (POT) approach takes all peak values over a pre-defined threshold.

The primary goal of this thesis is to contribute to the research that has been done so far on the application of Extreme Value Theory to Surrogate Measures for traffic safety analysis. After a thorough literature review two main research questions with altogether four subquestions were formulated addressing issues to which less or no attention has been paid so far. These questions seek for answers to what we can learn from applying univariate EVT using indicators describing collision course and crossing course interactions, and how we can predict nearness to collision and severity using bivariate EVT models. To answer these questions data gathered at a signalized intersection in Minsk (Belarus) for left-turning and straight moving vehicle-vehicle interactions were analyzed and both univariate and bivariate EVT models were fitted.

The univariate models used and compared two temporal indicators, one of which is a collision course (TTC_{min}) and the other is a crossing course indicator (T_{2min}). Both the BM and POT approaches were applied in the univariate case and their performance were compared. TTC_{min} represents the minimum value of Time-to-Collision limited to situations when the two vehicles are on a collision course. Research has shown that when vehicles were about to miss each other by a very short time margin, they behaved as if they were on a collision course, even though technically speaking they were not. T_{2min} shows the time required for the second vehicle to arrive at the potential conflict point, thus addresses this issue by considering all crossing course interactions.

The univariate models showed that the BM approach overestimates crash probabilities and that the POT approach gives more reliable results, which could be also proved by validation using historical crash data. The validation on the other hand also showed that the crossing course indicator T_{2min} tends to overestimate crash frequencies and crash estimates based on the collision course indicator TTC_{min} are closer to reality.

Applying the Block Maxima approach in the univariate case revealed an issue related to the pre-selection of near-crash situations. This sub-sampling step is needed to select those observations that can be used as extremes for modeling. This initial threshold was set to be 3.5s for both indicators. Due to the small sample size in case of TTC_{min} the question was to which level the near-crash threshold should be increased to have enough data points and a reasonable model fit. T_{2min} on the other hand offered a more appropriate sample size and the possibility to reduce the pre-selected threshold to a more reasonable level from a traffic

a safety point of view. Similarly, the threshold selection in case of the POT approach showed that a lower value has to be used for T_{2min} compared to TTC_{min}, which practically means that one has to be "stricter" against crossing course indicators, as compared to collision course indicators when selecting thresholds.

Since POT gave more reliable results in the univariate case, the bivariate threshold excess models are used for analyzing four variable pairs, in which the two temporal indicators are accompanied with two speed related indicators, relative speed and Extended Delta-V0. The former indicates the speed of a moving vehicle relative to another moving vehicle, the latter the change of the velocity vector by a road user during a crash. Extended Delta-V0 extends the severity measure by taking into account vehicle masses as well as the angle of collision.

Taking a temporal vs. a speed related indicator a plane along with so-called severity levels can be constructed like in the Swedish Traffic Conflict Technique, where quite intuitively severity worsens as the value of the temporal indicator decreases or the value of the speed related indicator increases. Another way to interpret this plane is to investigate the joint probability of exceeding a given pair of values. Points yielding the same probability form lines, which were named probability based risk levels.

Depending on whether temporal based and/or speed related indicators exceed their thresholds, this plane can be divided into four quadrants: both indicators exceed the thresholds, both are below, or only one indicator exceeds the threshold while the other does not. In this research primarily the extreme region was analyzed but the other three regions were also briefly addressed.

Bivariate models have two components, the marginal distributions and the dependence structure. These can be modeled either simultaneously by using bivariate threshold excess models with parametric distributions describing the dependence, or separately by building a bivariate model from the univariate POT models and a copula describing the dependence structure between the two components. Both methods were investigated and it was concluded that there is weak or no dependence in between temporal and speed related indicators at extreme levels. This means that road users getting closer to each other in time do not necessarily show high relative speed or Delta-V0. In other words, judging the severity of an interaction using only a temporal based indicator can be misleading and thus it should be accompanied with another surrogate measure describing the differences in speeds/masses of vehicles.

The main result of the bivariate models is the probability based risk levels which stem from the calculated joint probabilities. It is proposed that this plane can be combined with the severity levels between temporal and speed related indicators and as a result not only the actual severity but also the probability of interactions would be known. Since these severity levels are unknown for the given variable pairs, constructing this plane could be a promising future research direction.

An important aspect related to the bivariate models is that by using a temporal indicator along with a speed related one, it is possible to estimate crash probabilities along the speed dimension on condition that the temporal indicator equals zero, that is the crash severity distribution can be constructed. An important future research direction in order to validate these models would be to investigate the speeds of vehicles at the time of collision.

List of Figures

List of Tables

List of Abbreviations

AIC	Akaike Information Criterion
BM	Block Maxima
CDF	Cumulative Distribution Function
EVT	Extreme Value Theory
GEV	Generalized Extreme Value
GPD	Generalized Pareto Distribution
PDF	Probability Density Function
PET	Post-Encroachment Time
POT	Peak-over-Threshold
SMS	Surrogate Measures of Safety
TA	Time-to-Accident
TAdv	Time Advantage
TCT	Traffic Conflict Technique
TTC	Time-to-Collision

Contents

1

Introduction

1.1. Problem statement

Road safety has a considerable impact on our society. According to the latest WHO report 1.35 million people died in 2016 in road accidents worldwide (WHO, 2018). Road crashes are the leading cause of death among young people aged between 15 and 29 years. Crashes also result in an enormous economic loss, they cost governments approximately 3% of their GDP. There is an increasing global concern for road safety - it is at the top of the agenda for both developed and developing countries.

In 2001 the European Union set an ambitious road safety objective of halving the number of road fatalities in the EU Member States in 10 years (European Commission, 2001). This objective has been renewed for another 10 years in 2011 (European Commission, 2011) with a long-term goal to reach zero fatalities by 2050. This stems from the Swedish concept of Vision Zero where "the long-term goal for Swedish road safety policy is that nobody should be killed or seriously injured in the transport system" (Johansson, 2009). Since the above indicated EU targets have not been met, recently new strategies have been adopted to reach Vision Zero in the EU (European Commission, 2019). The UN General Assembly officially proclaimed the Decade of Action for Road Safety 2011–2020 in March 2010 with an intention to save millions of lives.

In order to meet these ambitious goals, well-targeted safety measures are needed. To improve safety and to make sure that it is done in an efficient way, one must quantify safety in order to support evidence-based policy making. The most plausible way to evaluate safety is investigating the occurrence and severity of crashes using historical data. This approach however has a number of limitations (Tarko et al., 2009), which are:

- Accidents are rare events (Hauer, 1997) and are therefore associated with the random variation inherent in small numbers (Svensson and Hydén, 2006).
- In order to derive sensible implications based on historical data one needs to have a few years of observations. A typical period of time would be at least three years (Nicholson, 1985).
- The use of crash records for safety analysis is a reactive approach. There is an ethical issue related to this as well, since it requires accidents to occur, which we originally want to avoid (Songchitruksa and Tarko, 2006).
- Accident records are prone to underreporting (especially priority damage only crashes or light accidents with vulnerable road users).
- Quality of data is not always sufficient (due to missing records, false identification of location etc.).

1.2. Methodological aspects

In order to overcome the above limitations, the use of non-crash events have gained a great deal of attention especially due to the rapid improvement of sensing technologies facilitating the collection of trajectory data. Over 30 years ago Hydén (1987) pointed out that the interaction between road users can be described as a continuum of safety related events. The pyramid in Figure 1.1 shows that we usually base our safety estimates on a minute percentage of safety related events, i.e. accidents where damage, injury or death occur (Svensson and Hydén, 2006). Crashes as the rarest events are followed by conflicts of different levels of severity (serious, slight and potential). Below the conflicts, the majority of events are undisturbed passages or normal traffic processes (Laureshyn et al., 2010).

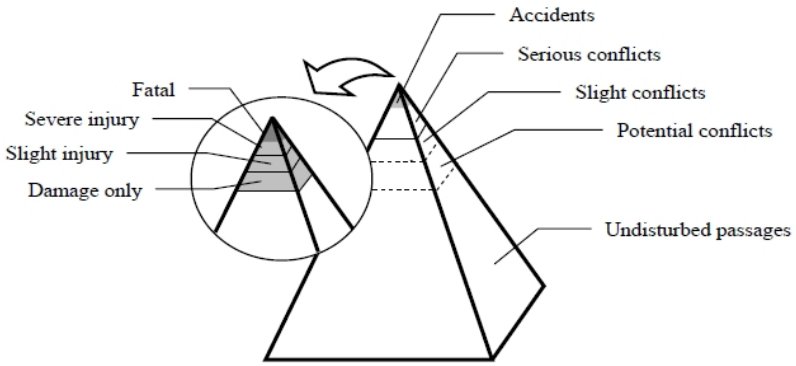


Figure 1.1: The safety pyramid (Hydén, 1987)

In the past few decades proactive methods gradually became more common. Several methodologies can be found in the literature: traffic conflict techniques such as the Swedish Traffic Conflict Technique, the Dutch Doctor method, and the use of surrogate measures of safety. Over the years, a vast number of indicators of surrogate safety have been developed to investigate traffic safety (Mahmud et al., 2017). These indicators mostly express the proximity to a crash either in time or space. As Laureshyn et al. (2017a) pointed out, very few of the existing traffic conflict indicators and techniques take into account the severity of conflicts and how to measure nearness-to-collision and severity at the same time.

Linking accident frequencies with these approaches has been also researched, however this relationship still relies on the assumption that historical accident data are accurate. An alternative approach has developed over the past few years using Extreme Value Theory to estimate the frequency of accidents using surrogate measures of safety.

Recently the Extreme Value Theory (EVT) to estimate crash probabilities using surrogate measures of safety has been applied more frequently. This theory offers two approaches to sample extreme events: the block maxima (or minima using Generalized Extreme Value distribution) and the Peak-over-Threshold (using Generalized Pareto distribution). In the former case, the method divides the sample time into blocks of a specified length and samples the largest value in each block, whereas in the latter case all peak values are sampled and the values over a certain threshold are used to model the extremes.

1.3. Thesis outline

This thesis examines Extreme Value Theory applied to Surrogate Measures of Safety by analyzing data that are gathered at a signalized intersection for vehicle-vehicle interactions, focusing on interactions between left-turning and straight moving vehicles.

The structure of the thesis is as follows. In Chapter 2 a literature review is done and the state-of-the-art is summarized. Here two aspects are dealt with, a detailed description of Surrogate Measures of Safety and an introduction of Extreme Value Theory and its application to surrogate safety indicators. This introduction places emphasis on the application of EVT to the field of traffic safety and the experiences gained up until now. Chapter 3 outlines the research framework by identifying the research gaps and formulating the research questions. Chapter 4 contains site description and an introduction to data gathered. Chapters 5 and 6 contain the univariate and bivariate models, respectively, including the presentation and discussion of results. In Chapter 7 final comments and conclusions are given indicating limitations, practical use of results, and possible future research directions. In this final chapter, a summary of publications based on the thesis is also given.

2

Literature review

This chapter focuses on providing a literature review on two clearly distinguishable aspects:

- The evolution of Surrogate Measures of Safety, indicators used, their advantages and disadvantages.
- An overview on the use of Extreme Value Theory in traffic safety, its applicability and limitations.

2.1. Surrogate Measures of Safety

In this section, a more detailed and extended interpretation of the safety pyramid presented in Chapter 1 is given followed by a detailed description of surrogate indicators used for safety evaluation.

2.1.1. The severity hierarchy

The safety pyramid (Hydén, 1987) implies that there is a relationship between the frequency and the severity of events; the tip of the pyramid shows the least frequent and most severe interactions between road users, whereas the bottom of the pyramid contains the most frequent and least severe interactions. The former are rare events that result in crashes of different severity levels (fatal, serious injury, light injury, and property damage only), the latter are a vast number of normal interactions. Critical events that do not result in a crash but are very close to that can be used as surrogate safety measures (Tarko et al., 2009).

With regard to the shape of the severity hierarchy Svensson (1998) made an important suggestion that it is not necessarily a pyramid. She proposed a diamond-shape based on the frequency of pedestrian-vehicle conflicts observed at signalized and unsignalized intersections (Figure 2.1). The idea behind the diamond shape is that at a particular site the majority of the interactions will be of moderate severity. Tarko (2012) also noted that this was the first evidence proving that there is a heterogeneity in the frequency-severity relationship due to the type of road facility influencing traffic conflicts. Other conditions, such as vehicle type, road users, collision angle and speed (Laureshyn et al., 2010) as well as weather may effect this relationship.

It has to be noted, however, that Svensson limited the events in the hierarchy only to interactions with a collision course (Svensson, 1998, Svensson and Hydén, 2006). A very important implication of this is that even low-severity interactions should be utilized because they may carry useful safety information (Tarko, 2012). This statement is highly relevant as interactions with severe conflicts usually come with low frequencies.

In an attempt to apply Svensson's reasoning, later researchers tried to adapt it by broadening the concept of traditional approaches. In Canada, St-Aubin et al. (2015) for instance

developed an approach called Probabilistic Surrogate Measures of Safety (PSMS) with a more general framework for safety analysis considering all possible paths that may lead two road users to collide. The novelty of this approach is relaxing the traffic conflict by allowing a non-zero risk of collision for road users who are not on a colliding course (Tarko, 2012). Following the same reasoning Laureshyn et al. (2010) suggested a new indicator called T_2 broadening the concept of the most common nearness-to-collision measure, time-to-collision, by including both collision and non-collision course states and allowing a smooth transition between both. (More details on this can be found in Subsection 2.1.2).

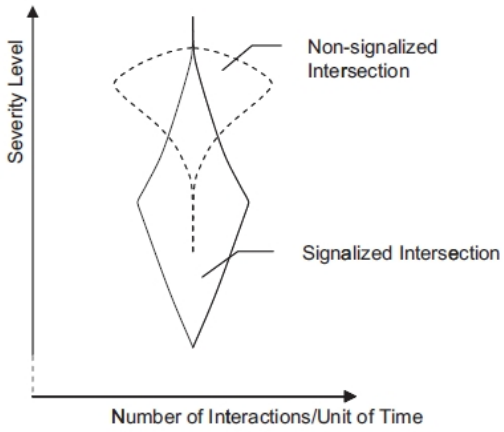


Figure 2.1: Relation between severity and frequency proposed by (Svensson, 1998) cited in (Tarko, 2012)

2.1.2. Indicators for safety evaluation

In order to overcome the shortcomings of safety analyses based on crash data (underreporting, quality issues and rare nature) a number of Surrogate Measures of Safety have been developed and proposed by various researchers. There is a consensus among researchers that observable non-crash events can be used for safety analysis as a complementary tool or may even replace analyses based on crash data (Ceunynck, 2017, Laureshyn et al., 2010).

At this point it has to be briefly mentioned that a number of traffic conflict techniques have been developed over the years. Probably the most well known is the Swedish Traffic Conflict Technique (Hydén, 1987), but other countries such as the US (Parker and Zegeer, 1989), the Netherlands (DOCTOR) (Kraay et al., 2013), the UK (Baguley, 1984), Finland (Kulmala, 1984), France (Muhlrad and Dupre, 1984), Austria (Risser and Schutzenhofer, 1984), the Czech Republic (Kočárková, 2012) also developed their own techniques. By looking at the number of studies on Traffic Conflict Techniques versus surrogate safety measures, Ceunynck (2017) pointed out that in the past few years there has been a shift towards the latter which is a clear sign that surrogate safety measures are gaining more attention. Most traffic conflict techniques are closely related to surrogate safety indicators. For instance, the Swedish traffic conflict technique uses the time-to-accident and conflicting speed, the Dutch DOCTOR technique uses TTC_{min} and Post-Encroachment Time. These techniques are well documented in the literature, however a more detailed introduction of these goes beyond the scope of this thesis.

Several papers summarized or compared a subset of indicators, for instance (Laureshyn et al., 2010) and (Laureshyn et al., 2017a) provided an overview focusing on nearness-to-collision and severity indicators. Mahmud et al. (2017) gave a more comprehensive overview on indicators by grouping them into temporal, distance based, deceleration based and other indicators, and identified 38 of them. Ceunynck (2017) in his doctoral dissertation did a literature review on the application of surrogate safety indicators and also looked into the frequency of use. He grouped indicators using the Time-to-Collision, the Post-Encroachment

Time, and the Deceleration families, plus two extra groups for other and unspecified indicators. The following paragraphs give a summary of these indicators as well as identify their advantages and disadvantages. Primarily using the above mentioned literature sources a non-exhaustive summary of indicators is given below. The interested reader is referred to the above mentioned sources for other indicators or further explanation.

One of the most widely used temporal indicator is *Time-to-Collision (TTC)* which can be calculated for any moment as long as the road users are on a collision course. TTC is defined as "the time until a collision between the vehicles would occur if they continued on their present course at the present rates" (Hayward, 1972). TTC can be easily calculated for right-angle (Equation 2.1), rear-end (Equation 2.2), and head-on collisions (Equation 2.3) (Figure 2.2). However, as two road users can meet at any angle and several collision types can occur, calculations can be more complicated (details on calculations can be found in (Laureshyn et al., 2010)). The lowest TTC value during the interaction, abbreviated as *TTCmin*, is the most commonly used indicator.

$$TTC = \frac{d_2}{v_2},\ if\ \frac{d_1}{v_1} < \frac{d_2}{v_2} < \frac{d_1 + l_1 + w_2}{v_1},\ TTC = \frac{d_1}{v_1},\ if\ \frac{d_2}{v_2} < \frac{d_1}{v_1} < \frac{d_2 + l_2 + w_1}{v_2} \tag{2.1}$$

$$TTC = \frac{X_1 - X_2 - l_1}{v_1 - v_2},\ if\ v_2 > v_1 \tag{2.2}$$

$$TTC = \frac{X_1 - X_2}{v_1 + v_2} \tag{2.3}$$

where d_1 and d_2 are distances from the front of the vehicle to the conflict area; l_1, l_2, and w_1, w_2 are lengths and widths of vehicles, respectively, v_1 and v_2 are vehicle speeds, X_1 and X_2 are vehicle positions.

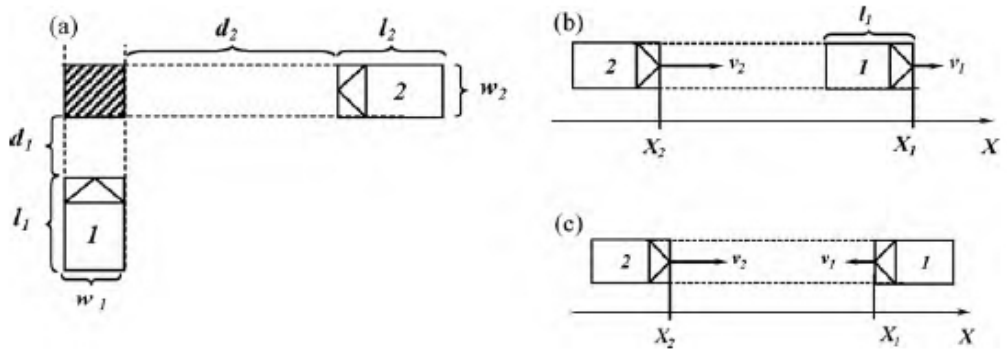


Figure 2.2: Calculation of TTC (perpendicular and parallel trajectories) (Laureshyn et al., 2010)

There is a consensus among researchers that a threshold value of TTC_{min} can be used to differentiate between severe and non-severe events. The magnitude of this value is not yet agreed upon. Generally, TTC lower than the perception and reaction time should be considered unsafe (Mahmud et al., 2017). Most common values are 1.5s, 2s and 3s (Ceunynck, 2017), probably with a more frequent use of the 1.5s value. Mahmud et al. (2017) provided a summary of threshold values known from previous researches. As for signalized intersections the threshold value ranges from 1.6s to 3s.

Notwithstanding, TTC_{min} represents only one of the two important moments in time describing the nearness-to-collision, namely it gives the lowest value of TTC during an interaction. The other important moment described by another indicator based on TTC is the Time-to-Accident (used in the Swedish Traffic Conflict Technique), which is defined as the moment when the first evasive action is taken by one of the road users (Hydén, 1987).

There are many other TTC related indicators, three further examples are given below (Time Headway, Extended Time-to-Collision, and Modified Time-to-Collision). *Time Headway* is a simple indicator taking the time that passes between two vehicles reaching the same point in space (*Time Gap* is an indicator that is highly similar to Time Headway and is used to express the distance between two consecutive vehicles in terms of time units). *Extended Time-to-Collision* comes with two possible indicators, the *Exposed Time-to-Collision (TET)* and *Time Integrated TTC (TIT)*. The former one measures the length of time a TTC-event remains below a certain threshold, whereas the latter one represents the integral of the TTC-profile during the time it is below the pre-specified threshold. To overcome the limitations of TTC (constant speeds, speed of the following vehicle is greater than that of the leading vehicle) Ozbay et al. (2008) proposed a new indicator called *Modified Time-to-Collision* taking into account all possible combinations of speed and acceleration relations of the leading and the following vehicles.

Post-Encroachment Time (PET) can be used when the two road users pass the conflict area with a time margin. PET is the time between the first road user leaving the common spatial zone and the second arriving at it (Allen and Shin, 1977). It has a single value and may be observed and measured directly. *Time Advantage (TAdv)* broadens the concept of PET as it gives for each moment the expected PET if the road users continue with the same speeds and path (Laureshyn et al., 2010). TAdv uses predicted travel lines, the lowest found TAdv has to be used, the same way as with TTC.

Laureshyn et al. (2010) proposed a supplementary indicator to Time Advantage in order to describe the nearness of the encroachment. This indicator measures the expected time that it takes for the second road user to arrive at the potential collision point, hence it is called T_2. The logic behind this indicator is that most TTC related indicators assume the two road users to be on a collision course, which however sets a limitation to the situations to be considered in safety analysis. Laureshyn et al. (2017a) argued that encounters without a collision course might have crash potential as well due to the possibility of minor changes in the spatial or temporal relationship between road users.

Looking at the possible intersections between road users (Figure 2.3) they can be classified into three groups: collision course, crossing course and non-crossing course. As it was previously mentioned there is a continuum between these, which means that for instance just a minor change in road users' trajectories or speeds can put them from crossing course into a collision course encounter. According to Svensson (1998) when road users were about to miss each other by a very short time margin, they behaved as if they were on a collision course even though strictly speaking they were not. The Dutch DOCTOR traffic conflict technique for instance (Kraay and van der Horst, 1985) includes both with and without collision course situations given that the time margin in the latter case is small enough. This difference is also present in the scope of certain indicators; PET can be measured relatively easily for all events that have a crossing course, while TTC can only be calculated for events that have a collision course (Ceunynck, 2017).

Based on the above line of reasoning, T_2 tells more about safety since the arrival at the potential collision point is the very last necessary condition for a collision to occur and it provides a smooth transfer between the collision course and crossing course situations (Laureshyn et al., 2017a). T_2 assumes unchanged speeds and planned trajectories (Figure 2.4). If the road users are on a collision course, T_2 equals TTC. In the event that the two road users pass the conflict point with a time margin, T_2 reflects the maximum time available to take evasive actions and alleviate the severity of the situation. T_2 is no longer calculated after the first road user has left the conflict zone (since the crash is no longer possible) (Laureshyn et al., 2017a). T_2 is a similar indicator to TTC in the sense that it is also continuous, therefore can be calculated for any time instance. The last possible value is when the first road user leaves the potential conflict area (the same meaning as PET). An alternative value is T_{2min} which shows the moment when the two vehicles are closest in time. These two values can be different in case of significant speed changes.

A summary of the above discussed indicators is given in Table 2.1.

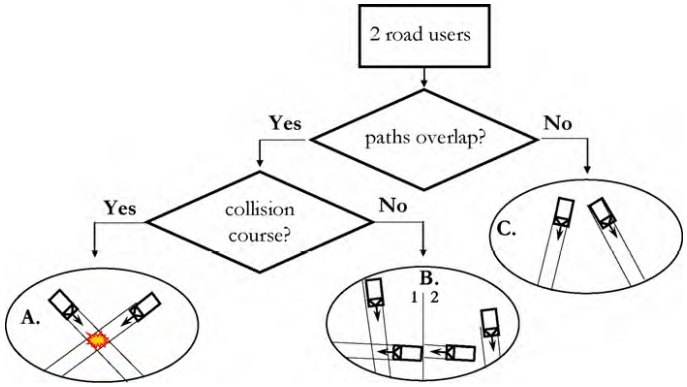


Figure 2.3: Classification of encounters (Laureshyn et al., 2010)

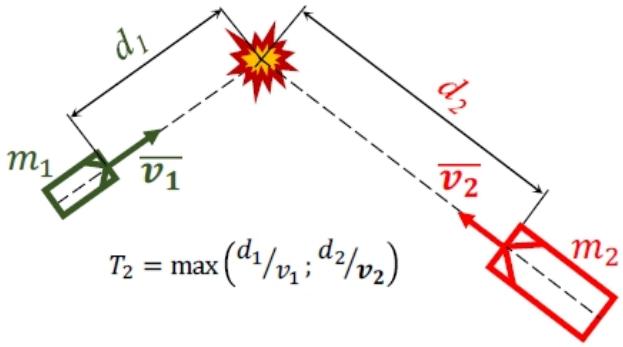


Figure 2.4: Illustration of T_2 (Laureshyn et al., 2017a)

Table 2.1: Selection of surrogate safety indicators (adapted from (Mahmud et al., 2017))

Indicator	Definition	Limitations	Advantages
Time-to-Collision (TTC)	The time until a collision between the vehicles would occur if they continued on their present course at their present speeds	Assuming same speed and direction is not practical and unlikely. Ignores many potential conflicts due to acceleration and deceleration discrepancies. Speed of following vehicle should be higher than that of the leading one. Does not account for severity. Require the road users to be on a collision course.	Frequently used, informative, also used in automobile collision avoidance systems.
Time-to-Accident (TA)	The TTC value at the moment when the first evasive action is taken.	Heavily relies on the evasive action. Other same as TTC.	Widely used, can be measured manually or by video analysis.
Time Exposed Time-to-Collision (TET)	Summation of all moments that a driver approaches a front vehicle with a TTC-value below a certain threshold.	Attainable only in a simulation environment. Does not provide information about the severity levels of different TTC values below the threshold.	Suited for application in microscopic simulation studies. Easy to include small TTC values due to time-dependent TTC values of all subjects.
Time Integrated Time-to-Collision (TIT)	Integral of the TTC-profile during the time it is below a certain threshold.	Difficult to interpret its meaning. Not preferable to use in comparative studies in which simulation tools are applied to generate trajectories. Benefits are small due to the uncertainties in driver behavior.	Level of safety of collision can be derived. Suitable for microscopic studies. Easy to include small TTC values due to time-dependent TTC values of all subjects.
Modified Time-to-Collision (MTTC)	Considers all of the potential longitudinal conflict scenarios due to acceleration or deceleration discrepancies.	Obtaining the speed of both users and the distance gap in an evolution process is difficult. Does not reflect the severity of collision.	More advanced than TTC. Considers driving discrepancies.

Continued on next page

Table 2.1 – *Continued from previous page*

Indicator	**Definition**	**Limitations**	**Advantages**
Headway	Elapsed time between the front of the lead vehicle passing a point on the roadway and the front of the following vehicle passing the same point.	Does not take into account conflicts due to lateral movement.	Easy to measure. Level of safety can be distinguished.
Post-Encroachment Time (PET)	The time between the moment a vehicle leaves the potential collision point and the other road user arrives at it.	Only useful in the case of transversal trajectories. Levels of severity as well as impact of a conflict are not taken into account.	More appropriate for intersecting conflicts, can be easily estimated.
Time advantage (TAdv)	Expected PET for each moment if the road users continue with the same speeds and paths.	Not sufficient to describe the collision risk.	Broadens the concept of PET.
T_2	The time needed for the second road user to arrive at the potential conflict point.	Also based on the prediction of the potential conflict point using planned paths and current speeds.	Provides a smooth transfer between collision course and crossing course situations.

The above discussed indicators were exclusively temporal indicators, however there are many other indicators that take into account distance (named distance based indicators looking into for instance stopping distance) or deceleration (named deceleration based indicators looking into for instance deceleration rate to avoid a crash). These will not be further discussed here, a comprehensive summary can be found in Mahmud et al. (2017).

Another aspect that of relevant interest is modeling the severity of conflicts using surrogate safety indicators. Even though the above mentioned indicators are used on their own to capture the severity of an interaction (for instance by applying a threshold value in case of TTC), these are not sufficient to describe the severity of the consequences. As Laureshyn et al. (2010) stated there is a need for the time-based indicators to be complemented with some speed-related indicator. The Swedish Traffic Conflict Technique is an exception since the conflicting speed can be taken as a measure related to the severity of crashes, where the serious conflict threshold is between severity level 25 and 26 (Figure 2.5). However, many studies on severity of crashes showed that the consequence of a crash is dependent not only on the speed but also the mass of involved road users and the angle of a collision (Zheng et al., 2014b).

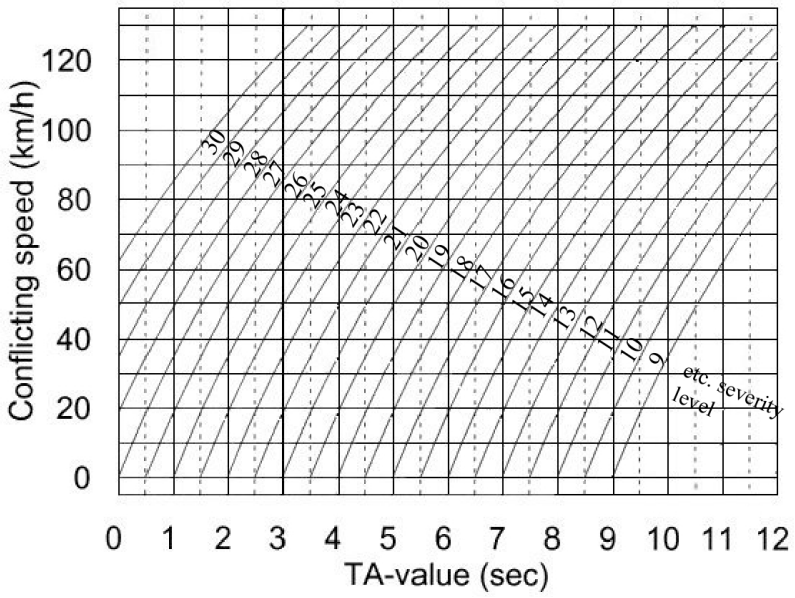


Figure 2.5: Classification of conflict severity in the Swedish TCT (Hydén, 1987)

To measure the severity outcome of a crash, the indicator Delta-V is often used. This indicator describes the change in the velocity vector experienced by a road user during a crash; the more rapid the change between pre- and post-crash is the more severe the outcome is expected to be. Delta-V values in an inelastic collision can be calculated for each of the colliding road users using Equation 2.4; the highest value of these two can be used to describe severity.

$$\Delta v_1 = \frac{m_2}{m_1 + m_2} \times \sqrt{v_1^2 + v_2^2 - 2v_1 v_2 cos\alpha}, \;\; and \;\; \Delta v_2 = \frac{m_1}{m_1 + m_2} \times \sqrt{v_1^2 + v_2^2 - 2v_1 v_2 cos\alpha} \qquad (2.4)$$

where m_1 and m_2 are vehicle masses, v_1 and v_2 are speeds and α is the approach angle.

According to Laureshyn et al. (2017a) and Ceunynck (2017) the Extended Delta-V can be calculated when assumptions are made about the road users' future movements (both vehicles will crash with the same speed as they have at a certain moment in the course of an interaction). Since Extended Delta-V becomes a continuous indicator (it can be measured in every time instant), the one associated with the lowest nearness-to-collision indicator has to be used. Extended Delta-V however assumes unchanged speeds, even though the vehicles might have time available to take an evasive action and brake before arriving at the collision point. Laureshyn et al. (2017a) therefore suggested making assumptions about the deceleration; in their paper they used 4 m/s^2 for normal and 8 m/s^2 for emergency braking.

They also proposed combining two indicators using nearness-to-collision together with the potential outcome severity. In their paper they combined T_2 with Extended Delta-V and illustrated severity lines saying the lower T_2 and the higher the Extended Delta-V are the more severe the outcome of the accident will be. One of the challenges of this method is that it is not clear how the threshold value in case of Extended Delta-V should be defined, and once defined, how it should be interpreted (Laureshyn et al., 2017a).

2.2. Extreme Value Theory and its application

This section gives a brief introduction to Extreme Value Theory as well as provides an overview of its use in traffic safety modeling.

2.2.1. Two approaches of EVT

Extreme value analysis has the objective to quantify the stochastic behavior of a process at unusually large or small levels. Basically the aim is to assess the probability of extreme events that have not been observed before. There are two distinctive methods to perform that: (1) the Generalized Extreme Value distribution used in the Block Maxima approach, and (2) the Generalized Pareto Distribution used in the Peak-over-Threshold approach. The next two subsections discuss these modeling approaches relying on Coles (2001). How to interpret model outputs and to check the goodness of the models are not discussed here, these will be explained in later chapters as models are fitted to actual data.

2.2.1.1. Block Maxima approach

EVT models focus on the behavior of

$$M_n = \max\{X_1, \dots, X_n\} \tag{2.5}$$

where $X_1,\dots,X_n$ is a sequence of independent random variables having a common distribution function F, M_n represents the maximum of the process over n time units of observation. The distribution of M_n can be derived as $Pr\{M_n \leqslant z\} = \{F(z)\}^n$. The function of F is unknown and to look for F^n a similar approach to the central limit theorem can be used, by allowing a linear renormalization of the variable M_n (Equation 2.6):

$$M_n^* = \frac{M_n - b_n}{a_n}, \tag{2.6}$$

where $\{a_n>0\}$ and $\{b_n\}$ are constants for which the appropriate values have to be found.

According to the extremal types theorem

$$Pr\left\{\frac{M_n - b_n}{a_n} \leqslant z\right\} \rightarrow G(z) \text{ as } n \rightarrow \infty, \tag{2.7}$$

where G belongs to one of the following three families (Equations 2.8, 2.9, and 2.10):

$$G(z) = \exp\left\{-\exp\left[-\left(\frac{z-b}{a}\right)\right]\right\}, \quad -\infty < z < \infty \tag{2.8}$$

$$G(z) = \begin{cases} 0, & z \leqslant b, \\ \exp\left\{-\left(\frac{z-b}{a}\right)^{-\alpha}\right\}, & z > b \end{cases} \tag{2.9}$$

$$G(z) = \begin{cases} \exp\left\{-\left[-\left(\frac{z-b}{a}\right)^{-\alpha}\right]\right\}, & z < b \\ 1, & z \geqslant b, \end{cases} \tag{2.10}$$

for parameters $a > 0$, b and, in case of Equations 2.9 and 2.10, $\alpha > 0$.

The rescaled sample maxima M_n^* converge to a variable having a distribution within one of the above three families, which are the Gumbel (Equation 2.8), Fréchet (Equation 2.9) and Weibull (Equation 2.10) families, respectively. All the three types have both a location (b) and a scale (a) parameter. The Fréchet and Weibull distributions also have a shape (α) parameter.

The above three equations can be generalized into a single distribution function (Equation 2.11).

$$G(z) = \exp\left\{-\left[1+\xi\left(\frac{z-\mu}{\sigma}\right)\right]^{\frac{-1}{\xi}}\right\}, \tag{2.11}$$

defined on the set $\{z : 1 + \xi(z\text{-}\mu)/\sigma > 0\}$, where $-\infty < \mu < \infty$, $\sigma > 0$ and $-\infty < \xi < \infty$. The three parameters that have been already mentioned before are the location parameter (μ), the scale parameter (σ), and the shape parameter (ξ). The distribution function itself determines the value of the shape parameter and vice versa. The three possible cases are:

- if $\xi > 0$, the model corresponds to a Fréchet distribution, the distribution has a heavy right tail and the right endpoint is infinite;
- if $\xi < 0$, the model corresponds to a Weibull distribution, which has a finite endpoint ($\mu\text{-}\sigma/\xi$);
- if ξ=0, the model corresponds to a Gumbel distribution, and has a light right tail in which case the model simplifies to Equation 2.12.

$$G(z) = \exp\left\{-\exp\left[-\left(\frac{z-\mu}{\sigma}\right)\right]\right\}, \quad -\infty < z < \infty. \tag{2.12}$$

The above mentioned density functions and a zoom of these functions in comparison with the standard normal density function are plotted in Figure 2.6.

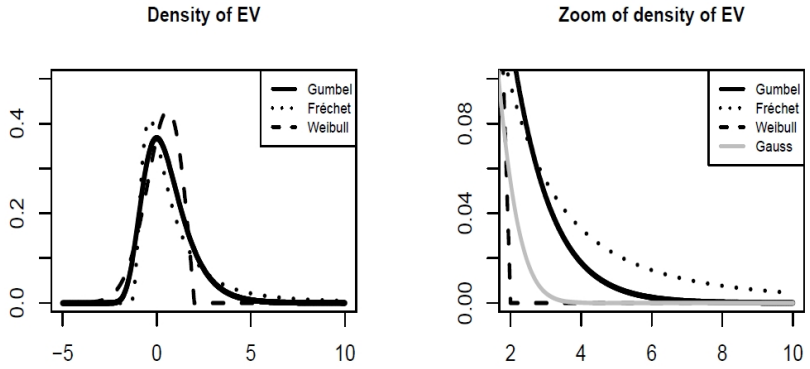


Figure 2.6: Density functions for the Gumbel, Weibull and Fréchet (left) and a zoom of these functions in comparison with the standard normal density function (right) (Penalva et al., 2013)

The GEV provides a model for the distribution of block maxima. According to the Block Maxima approach the data observed over time are split into blocks of equal intervals. The choice of the block size is not straightforward, as there is a trade-off between bias and variance. If blocks are too small, then the model approximation can be expected to be poor leading to bias in estimation and extrapolation; if blocks are too large, only a few block maxima will be generated resulting in large estimation variance. Another important aspect concerning the block size is that when time-dependent variation (e.g. within a year) is present, non-stationarity has to be accounted for.

When using likelihood methods for estimating model parameters for the GEV, certain regularity conditions are required for the maximum likelihood estimator to be valid. GEV distribution functions can have end-points $\mu\text{-}\sigma/\xi$, which is an upper end point if $\xi < 0$ (Weibull), and a lower end-point if $\xi > 0$ (Fréchet). This violation of the regularity conditions means that the applicability of results has to be checked. Smith (1985) (cited in Coles (2001) p.55) arrived at the following conclusions:

- when ξ > -0.5, maximum likelihood estimators are regular, having the usual asymptotic properties;
- when -1 < ξ < -0.5, maximum likelihood estimators are generally obtainable, but do not have the standard asymptotic properties;
- when ξ <-1, maximum likelihood estimators are unlikely to be obtainable.

GEV parameters can be estimated using maximum likelihood estimation, the log-likelihood for the parameters when $\xi \neq 0$ is

$$\ell(\mu,\sigma,\xi) = -nlog\sigma - \left(1+\frac{1}{\xi}\right)\sum_{i=1}^{n} log\left[1+\xi\left(\frac{z_i-\mu}{\sigma}\right)\right] - \sum_{i=1}^{n}\left[1+\xi\left(\frac{z_i-\mu}{\sigma}\right)\right]^{\frac{-1}{\xi}}, \quad (2.13)$$

provided that $1+\xi((z_i-\mu)/\sigma) > 0$, for i=1,..., n.

2.2.1.2. Peak-over-Threshold

An alternative - and according to many authors a better - approach to modeling the block maxima is the Peak-over-Threshold method. Block Maxima is criticized to be a wasteful approach as only the maximum value is used from each block, thus not considering other, but possibly still extreme values. Possible solutions to solve this issue is using the so-called r largest order statistic model (e.g. using the largest 5 observations) or by modeling threshold excesses. The latter one is the Peak-over-Threshold approach (POT), in which observations over a certain threshold are selected and treated as extremes.

Using the GEV distribution (Equation 2.11) for large enough threshold u, the distribution function of $(X-u)$, conditional on $X > u$, is approximately

$$H(x) = 1 - \left[1+\xi\left(\frac{x-u}{\sigma_u}\right)\right]^{-1/\xi} \quad (2.14)$$

where u is a high threshold, $x > u$, scale parameter $\sigma_u > 0$ (depending on threshold u), and shape parameter $-\infty < \xi < \infty$.

The distribution family given in Equation 2.14 is called the Generalized Pareto family, in other words, threshold excesses have a Generalized Pareto Distribution with two parameters, the shape ξ and the scale σ parameters (using the same notation as in GEV). Just like with GEV, the shape parameter ξ determines the behavior of the GPD. If $\xi < 0$ the distribution has an upper bound of u-σ/ξ; if $\xi > 0$ there is no upper limit. If $\xi = 0$, then Equation 2.14 simplifies to an exponential distribution function

$$H(x) = 1 - e^{-(x-u)/\sigma} \quad (2.15)$$

Selecting the appropriate threshold u with the POT approach is identically difficult as finding proper block sizes with the BM approach. There is an analogy between the two, namely that if the threshold value is too low, the asymptotic basis of the model is likely to be violated leading to bias, and if it is too high, just a few observations will be used leading to high variance. Two methods can be used to find an appropriate threshold level: 1) mean residual plot showing the mean of the excesses depending on the value of u, and 2) model estimation at a range of threshold values and finding stable model parameters, also called a threshold stability plot. How to use these techniques will be further elaborated when fitting models to actual data (in Chapter 5).

2.2.2. Bivariate models

A new issue in bivariate compared to univariate modeling is that as there is more than one variable one has to examine dependence. In the univariate case, one of the aims is to extrapolate outside the range of the data, however with more than one variable we have to take into account the possibility of extremes in several coordinates to occur jointly (Beirlant et al., 2004).

Bivariate (or in general multivariate) modeling has two components: the marginal distributions and the dependence structure. First, the margins are dealt with using standard univariate techniques, and second, after a transformation standardizing the margins to a common scale, the dependence (Beirlant et al., 2004).

2.2.2.1. Componentwise Maxima

Similar to the univariate case, the behaviour of multivariate extremes is based on the limiting behaviour of block maxima (Coles, 2001). Equation 2.16 is the vector of componentwise maxima.

$$M_{x,n} = \max_{i=1,\dots,n}\{X_i\} \;\; and \;\; M_{y,n} = \max_{i=1,\dots,n}\{Y_i\}, \qquad M_n = (M_{x,n}, M_{y,n}) \tag{2.16}$$

where $\{X_i\}$ and $\{Y_i\}$ are sequences of independent univariate random variables, thus standard univariate value results apply to both components. To obtain standard univariate results for each margin, the rescaled vector of M_n should be considered (that is each component in Equation 2.16 divided by n). The rescaled vector M_n^* is equal to $(M_{x,n}^*, M_{y,n}^*)$, then if

$$Pr\{M_{x,n}^* \leqslant x, M_{y,n}^* \leqslant y\} \;\xrightarrow{d}\; G(x,y), \tag{2.17}$$

where G is a non-degenerate distribution function, G has the form

$$G(x,y) = exp\{-V(x,y)\}, \;\; x > 0, \;\; y > 0 \tag{2.18}$$

where

$$V(x,y) = 2\int_0^1 max\left(\frac{w}{x}, \frac{1-w}{y}\right) dH(w), \tag{2.19}$$

and H is a distribution function on [0,1] satisfying the mean constraint

$$\int_0^1 w dH(w) = 1/2 \tag{2.20}$$

The family of distributions that arise as limits in Equation 2.18 is termed the class of bivariate extreme value distributions (Coles, 2001). When H is a measure that places mass 0.5 on w=0 and w=1, the bivariate extreme value distribution corresponds to independent variables, if H places unit mass on w=0.5, the bivariate extreme value distribution corresponds to perfectly dependent variables.

The marginal distributions can be generalized into Equations 2.21:

$$\tilde{x} = \left[1 + \xi_x\left(\frac{x-\mu_x}{\sigma_x}\right)\right]^{\frac{1}{\xi_x}}, \;\; and \;\; \tilde{y} = \left[1 + \xi_y\left(\frac{y-\mu_y}{\sigma_y}\right)\right]^{\frac{1}{\xi_y}}, \tag{2.21}$$

resulting in the distribution function

$$G(x,y) = exp\{-V(\tilde{x},\tilde{y})\}, \tag{2.22}$$

provided $[1 + \xi_x\,(x\text{-}\mu_x)/\sigma_x] > 0$ and $[1 + \xi_y\,(y\text{-}\mu_y)/\sigma_y] > 0$

Any distribution function H (also called as spectral measure) on [0,1] in Equation 2.19, satisfying the mean constraint in Equation 2.20 will result in a valid limit in Equation 2.17. It is easier to build models using parametric families for H, one standard solution is the logistic family:

$$G(x,y) = exp\left\{-\left(x^{-1/\alpha} + y^{-1/\alpha}\right)^{\alpha}\right\}, \;\; x > 0, \;\; y > 0 \tag{2.23}$$

for a parameter $\alpha \in (0,1)$. There are other parametric families available in the literature (Zheng et al., 2018), for instance asymmetric logistic, negative logistic, bilogistic, asymmetric negative bilogistic, negative bilogistic, Husler-Reiss; these will not be elaborated further. H has the density function (Equation 2.24)

$$h(w) = \frac{1}{2}(\alpha^{-1} - 1)\{w(1-w)\}^{-1-1/\alpha}\{w^{-1/\alpha} + (1-w)^{-1/\alpha}\}^{\alpha-2} \tag{2.24}$$

on $0 < w < 1$. The main reason for the logistic family lies in its flexibility. As $\alpha \to 1$ it corresponds to independent variables, as $\alpha \to 0$ it corresponds to perfectly dependent variables.

To investigate the dependence structure between two random variables copulas can be also used. With copulas one can separate the marginal distributions from the dependence structure and model them separately. They are joint cumulative distribution functions with margins uniform on the interval (0, 1). Sklar's Theorem (1959) describes the relationship between the joint distribution of a random vector, its marginal distributions, and a copula. The theorem in the bivariate case states that for any random variables X and Y with joint distribution H and marginal distributions F and G, there exists a copula, C, and for all $x, y \in R$ $H(x,y) = C\{F(x), G(y)\}$. C is unique if the marginal distributions are continuous.

2.2.2.2. Bivariate threshold excess model

The bivariate threshold excess model approximates the joint distribution $F(x,y)$ on regions of the form $x > u_x$, $y > u_y$, for large enough u_x and u_y. For suitable thresholds the marginal distributions of F each has an approximation in the form of a univariate generalized Pareto distribution (Equation 2.25), with parameter sets $(\zeta_x, \sigma_x, \xi_x)$ and $(\zeta_y, \sigma_y, \xi_y)$ where $\zeta_x = Pr(x > u_x)$ and $\zeta_y = Pr(y > u_y)$ (Coles, 2001).

$$G(x) = 1 - \zeta\left\{1 + \xi\frac{(x-u)}{\sigma}\right\}^{\frac{-1}{\xi}}, \;\; x > u, \tag{2.25}$$

The transformations

$$\tilde{X} = -\left(log\left\{1 - \zeta_x\left[1 + \frac{\zeta_x(X-u_x)}{\sigma_x}\right]^{-1/\xi_x}\right\}\right)^{-1} \tag{2.26}$$

and

$$\tilde{Y} = -\left(log\left\{1 - \zeta_y\left[1 + \frac{\zeta_y(Y-u_y)}{\sigma_y}\right]^{-1/\xi_y}\right\}\right)^{-1} \tag{2.27}$$

induce a variable $(\tilde{X}, \tilde{Y})$ whose marginal distributions are approximately standard Fréchet for $X > u_x$, and $Y > u_y$. The joint distribution $F(x,y)$ can be expressed as in Equation 2.28.

$$F(x,y) \approx G(x,y) = exp\{-V(\tilde{x},\tilde{y})\}, \;\; x > u_x, \;\; y > u_y, \tag{2.28}$$

where $V(\tilde{x},\tilde{y})$ takes the form as in Equation 2.19 and H as in Equation 2.20.

Inference of the bivariate excess models is not straightforward. The bivariate pair of indicators may both exceed their thresholds, however the specified thresholds may be exceeded by only one or neither of them. This creates four regions:

- $R_{0,0} = (-\infty, u_x) \times (-\infty, u_y)$
- $R_{0,1} = (-\infty, u_x) \times [u_y, \infty)$
- $R_{1,0} = [u_x, \infty) \times (-\infty, u_y)$
- $R_{1,1} = [u_x, \infty) \times [u_y, \infty)$

The model in Equation 2.28 applies to region $R_{1,1}$, in the other regions it is not applicable, therefore the likelihood component is censored (more details can be found in Coles (2001)).

In case of a bivariate threshold excess model a pair of optimal thresholds should be selected. The marginal distributions of excess over thresholds are modeled by a Generalized Pareto distribution and the dependence structure between the two margins by a bivariate extreme value distribution. These two can generate different thresholds, Zheng et al. (2018) proposed using both the spectral measure as well as the threshold stability plots to assist threshold selection. The practical use and application of these plots on the actual data will be illustrated in Chapters 5 and 6. In the following paragraph a brief theoretical background of the spectral measure plot is given.

After transforming the data into standard Fréchet ($x_i^* = -1/logx_i$, $y_i^* = -1/logy_i$) and using pseudo-polar coordinates ($r_i = x_i^* + y_i^*$, $w_{xi} = x_i^*/r_i$ and $w_{yi} = y_i^*/r_i$) where r_i are the radial coordinates in ascending order, the spectral measure plot (Beirlant et al., 2004) produces a graph in which the integers $k = 1, \ldots, n-1$ are plotted against $(k/n)r_{(n-k)}$. Since $\tilde{H} = [0,1]) = (k/n)r_{(n-k)}$ is an estimator of $H([0,1]) = 2$, the largest k value denoted as k_0 for which $(k/n)r_{(n-k)}$ is close to 2 determines the pair of threshold values to be used.

2.2.3. Extreme Value Theory in safety modeling

In the literature many researchers emphasized the advantages of Surrogate Measures of Safety over reactive methods and have attempted to evaluate them. As Songchitruksa and Tarko (2006) phrased it, probably the most critical issue is the validity of traffic conflicts as an alternative measure of safety. Researchers testing the validity of traffic conflicts have tried to link historical crash data with conflict frequencies. These analyses lead to inconclusive results, as some studies could confirm a relationship, some could not.

Tarek and Sany (1999) for instance arrived at the conclusion that there is statistically significant relationship between crashes and conflicts. They identified a determination coefficient (R^2) in the range of 0.70-0.77 at signalized intersections. Notwithstanding, this approach is still hampered by the fact that accident data are inaccurate, thus finding a good correlation has a limited power. Zheng et al. (2014a) also emphasized that the application of regression models is limited due to three reasons:

- the incorporation of crash counts suffers from the same quality issues as traditional road safety analysis;
- the stability of crash-to-surrogate ratio is difficult to ensure especially when mixing surrogates of varied severity levels;
- the statistical relationship between counts of crashes and surrogates hardly reflects the physical nature of crash occurrence.

An alternative approach to the traditional regression analysis without using observed crash counts was first proposed by Songchitruksa and Tarko (2006) based on the Extreme Value Theory (EVT). An important feature of the EVT is that it enables the researcher to model the stochastic behavior of unusually large or small processes. This extreme behavior is typically very rare and unobservable within a reasonable data collection time period. It often involves estimating the probability of extreme events over an extended period of time given very short and limited historical data (Songchitruksa and Tarko, 2006). The key assumption

of EVT is that the underlying stochastic behavior of the process being modeled is sufficiently smooth to enable extrapolations to unobserved levels (Coles, 2001).

EVT has been widely used in many fields, for instance in extreme weather prediction in meteorology, financial crisis prediction, wave height prediction in ocean engineering. However its use in transportation engineering is relatively limited (Songchitruksa and Tarko, 2006), (Zheng et al., 2014a); one of the earliest work is by Hyde and Wright (1986) who used it for estimating road traffic capacity. Its first application in the field of safety can be attributed to Songchitruksa and Tarko (2006); later on Tarko (2012) published a more generic application of EVT in traffic safety.

According to Tarko et al. (2009) the Extreme Value Method offers three important advantages over the traffic conflict technique:

- 1. The method abandons the assumption of a fixed coefficient converting the surrogate event frequency into the crash frequency.
- 2. The risk of crash given the surrogate event is estimated for any conditions based on the observed variability of crash proximity without using crash data.
- 3. The crash proximity measure precisely defines the surrogate event.

Songchitruksa and Tarko (2006) studied right-angle crossing events at 18 signalized intersections and measured Post-Encroachment Time. They applied the non-stationary BM method and evaluated it using historical crash data. They found a relationship between model estimates and crash data, however they drew attention to the large variance in the model estimates which they attributed to the short observation period. Farah and Azevedo (2015) used EVT for passing maneuvers and tested both the Generalized Extreme Value (Block Maxima approach - BM) and the Generalized Pareto Distribution-based (Peak-over-Threshold approach - POT) estimation (these two approaches will be further explained in Chapter 5). They concluded that the GEV approach achieved satisfactory fitting results, POT underestimated the expected number of head-on collisions. Their finding is contradicting with that of Zheng et al. (2014a), who studied the safety implication of lane change maneuvers in freeways. They demonstrated that the overall performance of the POT approach is better than that of the BM approach when data size is limited. They also added that the POT approach is more in line with the logic of traffic conflict techniques as both of them use thresholds to separate conflicts from more observable events. However, selecting the threshold for the POT approach as well as determining the block interval for the BM approach (if applicable) are of an arbitrary nature. In their paper they also drew attention to the dependency and non-stationary issues that might come with applying EVT.

Jonasson and Rootzén (2014) applied both univariate and bivariate models on data coming from naturalistic driving studies. They arrived at the conclusion that the estimate obtained from the fitted GEV distribution is much smaller than actual relative frequency of crashes (175 times difference). They also tried fitting bivariate models using TTC and nine other variables, but it was possible only for two out of nine variables (maximum speed and length overlapping glance). Cavadas et al. (2017) also tried fitting bivariate models using TTC for head-on collisions and Time headway for rear-end collisions in case of passing maneuvers. According to their results the bivariate model is suitable to estimate the joint probability of colliding with the opposite vehicle or with the passed vehicle.

Åsljung et al. (2016) applied the Peak-over-Threshold approach using TTC as a closeness to a collision indicator as well as Break Threat Number (BTN) as a closeness to an Inevitable Collision State (ICS). As previously described, TTC has the problem of being biased by speed, whereas BTN does not, which tells how much of the brake capacity that is needed in order to stop just in front of an approaching vehicle. The authors aimed for investigating the difference between the two types of measures when estimating collision frequency using EVT. Basically they compared a measure which is based on the closeness to a collision against a measure

that tells the closeness to ICS. They showed that there are significant differences between these two indicators, determining a suitable threshold was more clear for BTN than for TTC. As for TTC it was not possible to extrapolate data for estimating a collision frequency.

Wang et al. (2018) in their paper proposed a combined usage of microscopic traffic simulation and EVT to evaluate intersection safety by looking at Post-Encroachment Time for crossing, rear-end and lane change conflicts, and applied three calibration strategies (base, semi-calibration, full calibration). They applied both the BM and POT approaches, however discarded the former one due to limited data. The full-calibration strategy performed the best in terms of identifying relative safety and correlating with actual crash frequency. They also concluded that the combined usage of microscopic traffic simulation and EVT is a promising tool for safety evaluation.

Zheng et al. (2018) stressed the advantage of bivariate modeling arguing that extremal events can be characterized by several features, thus a pre-crash event can be described using different surrogate measures of safety (e.g. temporal proximity, spatial proximity, likelihood of evasive actions, and consequence of a potential collision). They applied bivariate extreme value models to estimate crashes on freeway entrance merging areas using two measures, Post-Encroachment Time and length proportion of merging. After testing several parametric distribution functions they chose models with the logistic distribution function. Through testing correlation coefficients between observed crashes and estimated crashes their results suggested that bivariate extreme value modeling increase the accuracy of safety estimates. Furthermore they also emphasized that the bivariate model can predict the number of crashes but ”an important severity dimension is missing which characterizes how serious an accident would be given that road users were close to colliding”.

A summary of the above discussed research results on the application of EVT on surrogate measures of safety is presented in chronological order in Table 2.2.

Table 2.2: Studies applying Extreme Value Theory using surrogate indicators

Paper	**Study area**	**Method(s) used**	**Surrogate indicator(s) used**	**Conclusions**
Songchitruksa and Tarko (2006)	Right-angle collisions at signalized intersections using field measurements	Non-stationary BM	Post-Encroachment Time	Method is a potential tool for safety prediction. Method would require 30-50 days of observation with good precision. Promising relationship with crashes.
Zheng et al. (2014a)	Freeway lane-changing maneuvers using field measurements.	Both BM and POT	Post-Encroachment Time	POT model performed better than BM under the condition of short time observations. Model results are comparable with crash data.
Jonasson and Rootzén (2014)	Rear-end crashes from naturalistic driving studies	Univariate BM and multivariate models using TTC and maximum speed	Time-to-Collision	Model gave underestimation (in comparison with relative frequency of crashes)
Farah and Azevedo (2015); Farah and Azevedo (2017)	Passing maneuvers using laboratory experiments for head-on collisions	Stationary BM and non-stationary BM (including several covariates), as well as POT	Time-to-Collison	BM approach yielded better results than POT. Non-stationary BM performed better than stationary BM

Continued on next page

Table 2.2 – *Continued from previous page*

Paper	**Scope of study**	**Method(s) used**	**Surrogate indicator(s) used**	**Conclusions**
Åsljung et al. (2016)	Focused on autonomous vehicles, field test of a collision avoidance system, rear-end collisions	POT	Time-to-Collision and Break Threat Number	Significant differences between the two indicators, finding a stable threshold interval for the TTC value is not clear.
Cavadas et al. (2017)	Passing maneuvers using laboratory experiments for head-on and rear-end collisions	Stationary and non-stationary BM using univariate models and joint model using bivariate distribution with copula method	Time-to-Collision (for head-on collisions) and Time-headway (for rear-end collisions)	Bivariate model can be used to link two different surrogate measures.
Wang et al. (2018)	Field measurements at urban intersections and VISSIM simulation, three conflict types: crossing, rear-end, and lane change.	Both BM and POT	Post-Encroachment Time	BM was discarded due to limited data, POT performed well.
Zheng et al. (2018)	Severity of merging events on freeway entrance merging areas using field data	Univariate and bivariate threshold excess model.	Post-Encroachment Time and length proportion merging.	Bivariate crash estimates are closer to observed crashes than univariate ones.

3

Research framework

3.1. Research gap

In the past few decades many different Traffic Conflict Techniques have been developed. However, since the validity of them has not always been the primary focus, they might give different results. The number of cross-validation studies comparing different techniques is rare. In their paper for instance Laureshyn et al. (2017b) compared three approaches, the Swedish traffic conflict technique, the Dutch conflict technique (DOCTOR) and the probabilistic surrogate measures of safety (PSMS) technique and arrived at the conclusion that the results are comparable, however ”the existing disagreements in some cases raises the issue of the validity”.

The same argument can be made for individual surrogate safety indicators (no consensus on what measures should be used), even though conceptually there might be significant differences in between them. One of these differences is whether only a collision course is considered or an extension to crossing courses is considered as we saw in Laureshyn et al. (2010). As Zheng et al. (2014b) phrased it: ”Despite decades of conceptual development and widespread application, there are still some disputes on what traffic conflict is. This finding may raise a question as to whether a traffic conflict should in fact be defined and separated from other non-conflict events.”

Previous research has proven that road users that are strictly speaking not on a collision course actually might behave and take evasive actions as if they were, thus indicating that such near-miss situations might also be relevant for safety analysis. A new indicator accounting for this abbreviated as T_2 was proposed by Laureshyn et al. (2010) allowing a smooth transfer between collision and no-collision course interactions. Thus, T_2 seems to be more suitable for detecting potentially dangerous situations, but it has not been explicitly tested and compared to other nearness-to-collision indicators so far.

Another issue with using surrogate indicators on their own is that they are limited to estimating crash risk with no further account of the possible consequences of a potential crash (Zheng et al., 2014b). There has been attempts to apply two-dimensional approaches (e.g. TTC and speed or T_2 and Extended Delta-V), however differentiating the severity levels from each other is still an issue to be solved.

Extreme Value Theory is a promising tool to evaluate safety using surrogate safety measures. Most of the research that has been done so far focused on testing the method and validating various surrogate safety indicators by comparing model estimates to actual crash frequencies. There also has been attempts to apply bivariate models using two surrogate safety indicators. However, less or no attention was paid to two issues, namely:

- comparison of various conflict indicators and their performance using EVT, especially the comparison of collision course indicators with indicators including crossing course interactions as well, and
- modeling the nearness-to-collision and severity simultaneously using EVT, especially for collision course indicators as well as indicators including crossing course interactions together with indicators describing the expected consequences of interactions.

Based on the literature survey it can also be concluded that there are contradicting results in terms of which approach (BM or POT) performs better. Some researchers claimed that the BM method comes with extensive waste of data if many of the extreme events occur in the same block and POT seems to be a better choice when having access to more continuous observations. Other researchers, however, arrived at different conclusions saying that the BM method proved to give better estimates. As there is no consensus on this question, it is also worth further investigation.

3.2. Research questions

Based on the research gaps identified in the previous section two main research questions and four subquestions are formulated as follows:

1. What can we learn from applying EVT using indicators describing collision course and crossing course interactions at signalized intersections for vehicle-vehicle interactions?
   (a) What difference is there between the two indicators TTC_{min} and T_{2min} when analysing safety using EVT and are these indicators transferable?
   (b) Which EVT approach (BM or POT) under what circumstances performs better for TTC_{min} and T_{2min} (e.g. sensitivity to sample size)?
2. How can we predict nearness to collision and severity at signalized intersections for vehicle-vehicle interactions using the Extreme Value Theory?
   (a) What can we learn from applying bivariate models using EVT for various combinations of indicators?
   (b) How can we use bivariate EVT in combination with severity levels?

Chapter 5 addresses the first main question and Chapter 6 the second main question along with their subquestions.

4

Data collection

4.1. Site description and data

The dataset was provided by Lund University and has already been used in other publications such as (Laureshyn et al., 2017a). A regular signalized intersection with two-phases in Minsk (Belarus) was analyzed (53°54'39.1"N; 27°35'44.4"E). The intersection was recorded for two days (from 6 AM till 9 PM). The video footages of two cameras set on rooftops were then analyzed in the software T-Analyst (T-Analyst, 2016) allowing the manual tracking of vehicles as well as the calculation of various surrogate measures of safety (see a screenshot in Figure 4.1).

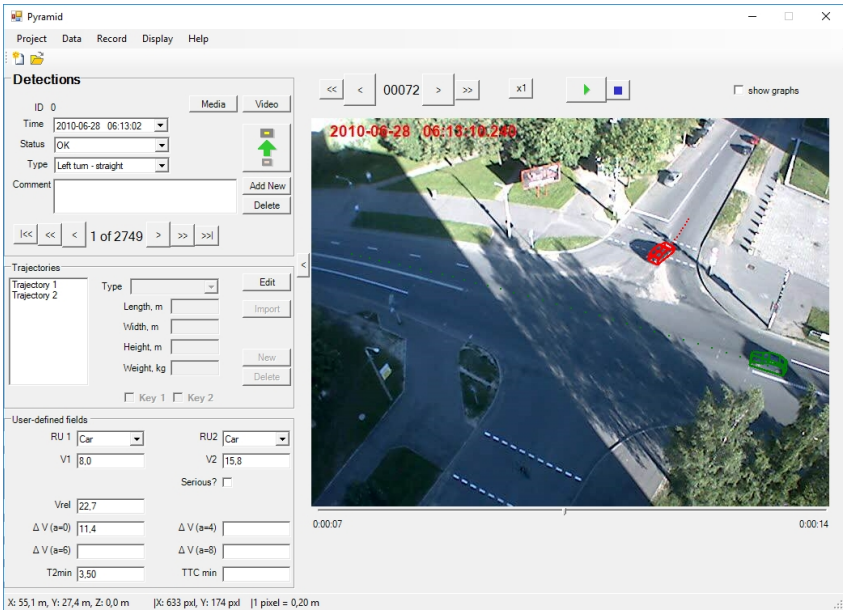


Figure 4.1: Screenshot of the T-Analyst software (T-Analyst, 2016)

Accident data were gathered for 10 years (1999-2009) before the video recordings were made. Altogether 32 accidents were recorded, out of which 5 crashes were due to the collision of left turning and straight going vehicles. The severity of all the recorded accidents were property damage only. As this type of crash severity is the most heavily prone to underreporting, unfortunately this historical accident dataset cannot be used for validation. In the course of video recordings no accidents were observed.

4.2. Descriptive statistics

Altogether 2749 interactions were detected. A subset of interactions between straight going and left turning vehicles was created (n=792), when the left turn was done in front of the straight moving vehicle. Whenever an indicator cannot be calculated the software indicates -1 as entry. These entries were not considered when compiling the descriptive statistics. All the statistical analyses including the content of later chapters are done in R (R Core Team, 2013).

4.2.1. Temporal indicators

This subsection gives an overview of the two temporal indicators TTC_{min} and T_{2min} used in the analysis. Table 5.1 shows the descriptive statistics for the two indicators.

Table 4.1: Descriptive statistics

Indicator	TTC_{min}	T_{2min}
Sample size	194	792
Min	0.79	0.06
Max	182.50	35.12
Mean	6.45	3.61
Stdev	14.25	2.22
Skewness	10.83	5.49
Kurtosis	123.76	61.25

Cumulative distribution functions of unfiltered T_{2min} and TTC_{min} are shown in Figure 4.2. Figure 4.3 shows a zoomed in version of Figure 4.2 for values smaller than 20s. This figure reveals that the cumulative distribution function for TTC_{min} is less steep than that of T_{2min} showing that the observed TTC_{min} values are more spread out and that the share of observations in the lower range (between 0 and 5s) is smaller than for T_{2min}.

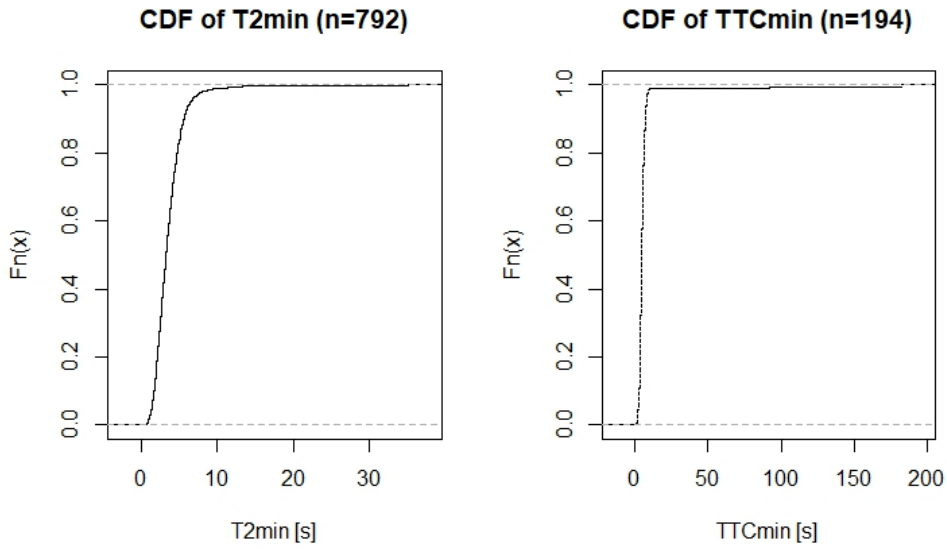


Figure 4.2: Cumulative distribution functions of T_{2min} and TTC_{min} (unfiltered data)

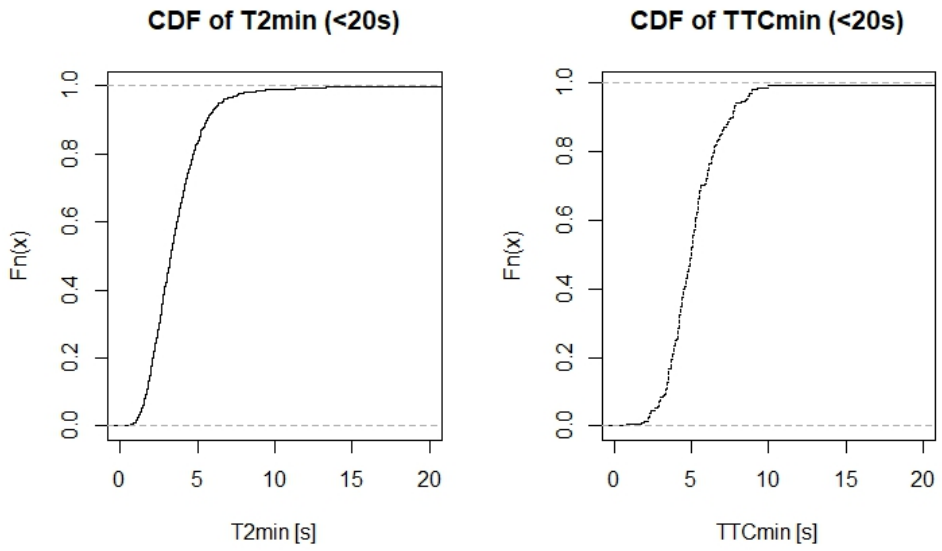


Figure 4.3: Cumulative distribution functions of T_{2min} and TTC_{min} (<20s)

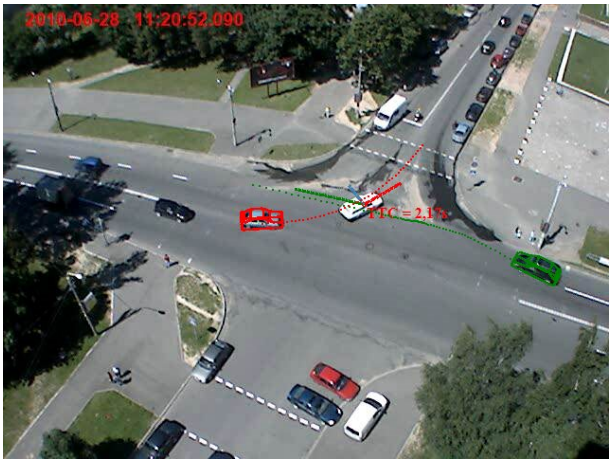
The underrepresentation of TTC_{min} compared to T_{2min} is due to the nature of these indicators, as TTC_{min} can be measured only for collision course interactions, whereas T_{2min} can be measured for both collision as well as crossing course interactions. Figure 4.4 illustrates this difference for a given interaction showing its distinctive moments. In the course of the interaction the straight going (green) and the left turning (red) vehicles are on a collision course, so both T_{2min} and TTC_{min} can be measured. The straight going vehicle then changes its path as well as speed due to which TTC_{min} cannot be measured anymore as the two vehicles are no longer on a collision course. However, the two vehicles are still on a crossing course and the time required for the second vehicle to arrive at the potential conflict point can be expressed and T_{2min} can be measured. T_2 can be measured until the first vehicle leaves the conflict point.

4.2.2. Speed related indicators

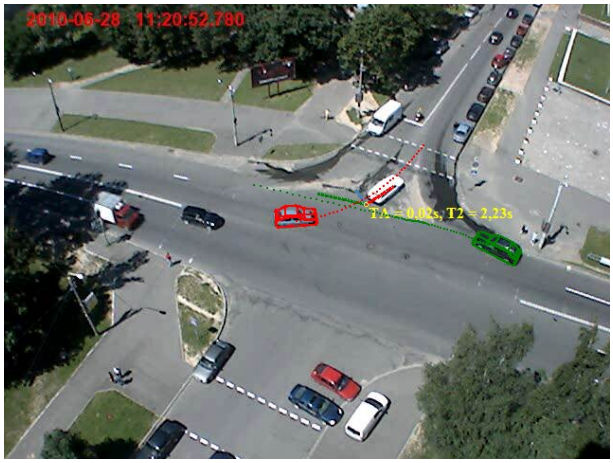
This subsection gives an overview of the indicators to be used in the bivariate models to account for crash severity. Both of them are speed related indicators:

- relative speed using the speeds of interacting vehicles, and
- Extended Delta-V0.

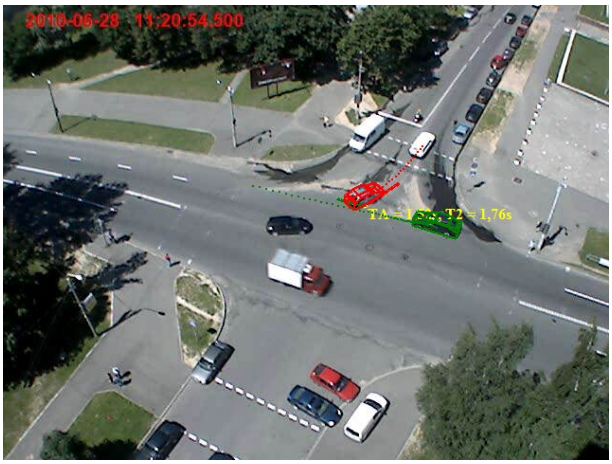
These indicators are associated with the minimum values of temporal indicators TTC_{min} and T_{2min}. Table 4.2 shows the descriptive statistics and Figure 4.5 the cumulative distribution functions of left turning and straight moving vehicles' speed as well as their relative speed. As TTC_{min} comes with a smaller sample size its CDFs are naturally less smooth. Mean speeds of interacting vehicles at TTC_{min} are slightly smaller than those of T_{2min} which is due to the collision course nature of TTC. Left-turning vehicles have much smaller speeds in general, as they have to give priority for straight moving vehicles.



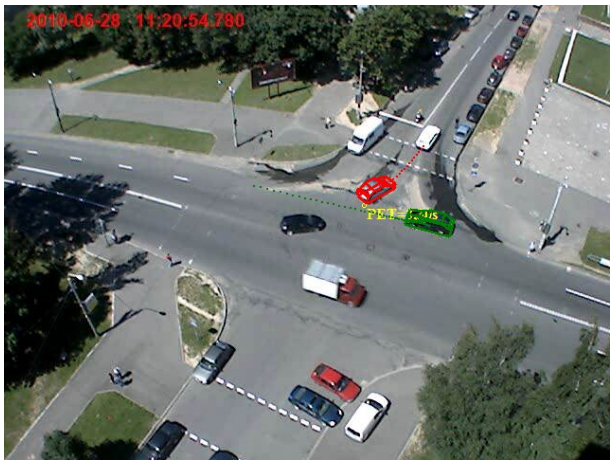
(a) Collision course: TTC is measured (and equal to T_2), TTC_{min} shows its lowest value



(b) Vehicles are no longer on collision course: TTC ceases to exist, T_2 is measured



(c) Crossing course: T_{2min} as the lowest value of T_2 is shown



(d) First vehicle leaves the conflict point, PET is measured

Figure 4.4: TTC_{min} and T_{2min} in an interaction

Table 4.2: Descriptive statistics of speeds associated with TTC_{min} and T_{2min}

Indicator	Statistics	Speed of left turning vehicle (m/s)	Speed of straight moving vehicle (m/s)	Relative speed (m/s)
TTC_{min} (n=193)	min	0.01	0.11	0.17
	max	14.47	19.47	22.37
	mean	2.93	10.46	12.41
	st. dev.	2.82	4.66	3.61
T_{2min} (n=789)	min	0.48	0.81	3.31
	max	16.30	22.08	31.41
	mean	5.72	11.91	16.14
	st. dev.	2.35	4.08	4.16

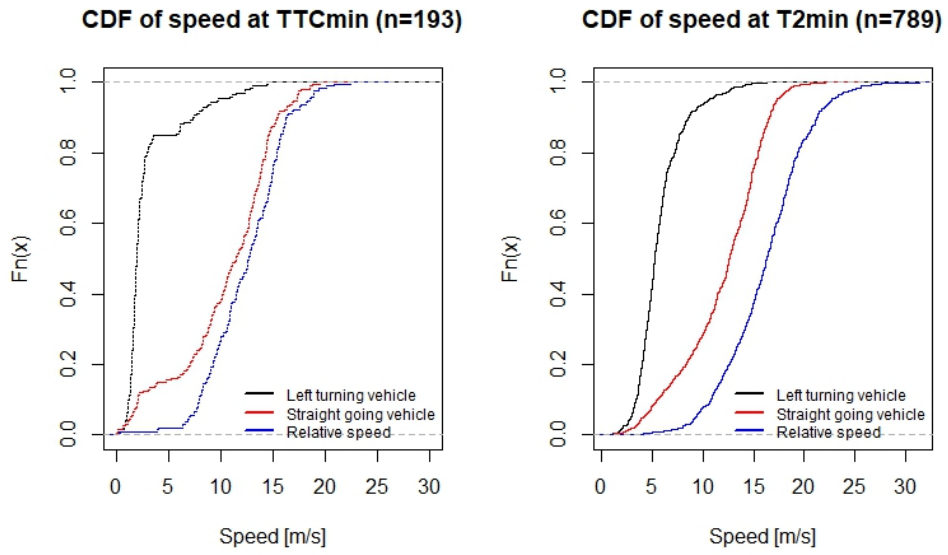


Figure 4.5: Cumulative distribution functions of speeds at TTC_{min} and T_{2min}

As for Delta-V the descriptive statistics of the base value (no braking) indicated as Delta-V0 are given in Table 4.3 and the frequency of vehicle types per traffic stream and per temporal indicator is given in Table 4.4. In the calculation of Delta-V values four different vehicle masses (car, minivan, bus, truck) were considered. Each road user has its own Extended Delta-V value, to describe severity the highest has to be considered (Laureshyn et al., 2017a); the CDFs of that are shown for both temporal indicators in Figure 4.6. Due to lower speeds at TTC_{min} compared to T_{2min}, Delta-V0 values are also smaller for the former temporal indicator.

Sample sizes have marginal differences compared to the dataset introduced in the previous section and used for the univariate case in Chapter 5. Through analyzing the relative speeds three observations were found with unrealistically small values. These interactions were between left-turning and right turning vehicles and coded incorrectly as a left-turning vs. straight moving interaction, thus were removed from the sample (one observation in case of TTC_{min} and three observations in case of T_{2min}). These were not extreme observations as far as the temporal indicators are concerned, so the univarite analysis were not redone as no significant changes were expected to the results.

Table 4.3: Descriptive statistics of Delta-V0 associated with TTC_{min} and T_{2min}

Indicator	Statistics	Left turning vehicle (m/s)	Straight moving vehicle (m/s)	Highest Delta-V0 (m/s)
TTC_{min} (n=193)	min	0.13	0.03	0.13
	max	12.21	13.96	13.96
	mean	6.56	5.84	7.23
	st. dev.	2.36	2.60	2.22
T_{2min} (n=789)	min	0.55	0.54	1.65
	max	19.62	18.38	19.62
	mean	8.43	7.70	9.27
	st. dev.	2.95	3.17	2.73

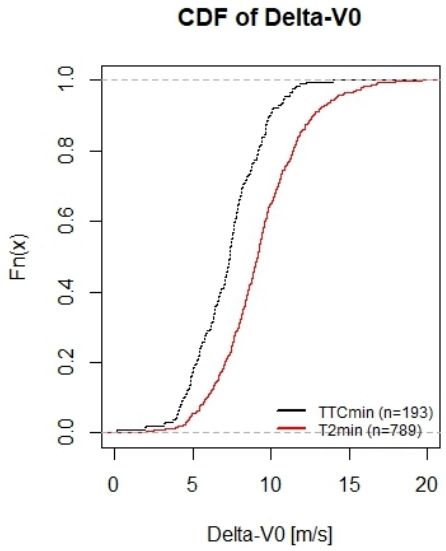


Figure 4.6: Cumulative distribution functions of Delta-V0 at TTC_{min} and T_{2min}

Table 4.4: Vehicle types associated with the TTC_{min} and T_{2min} samples

Indicator	Vehicle type	Left turning vehicle	Straight moving vehicle
TTC_{min} (n=193)	bus	0	3
	car	160	136
	van	18	21
	truck	15	33
T_{2min} (n=789)	bus	4	32
	car	665	608
	van	56	52
	truck	64	97

5

Univariate models

In this chapter univariate models are used with an intention to investigate the differences in using T_2 versus TTC. For this analysis the initial sample of 1616 interactions was narrowed down into two subsets where T_2 and TTC values were available. For both indicators the minimum values are used (T_{2min} and TTC_{min}) representing the moment in time when the two vehicles are closest to each other.

Both the Block Maxima as well as the Peak-over-Threshold approaches are used as modeling techniques. In this analysis observations with low values close to zero can be considered extremes, thus in both cases the negated values of observations were used. This is the standard technique used in all the traffic safety literature mentioned in the review section (Chapter 2.2.3). In the current research an alternative approach was tested using $1/TTC_{min}$ and $1/T_{2min}$ instead of the negated values. Modeling results can be found in Appendix A along with a brief analysis proving that this approach did not yield sensible results and was not further elaborated.

5.1. Block Maxima approach

Each interaction can be considered as a block in which the minimum value of T_2 and TTC are used. Since the minimum values are determined per interaction for both indicators, they can also be high and therefore irrelevant occurrences (e.g. a TTC_{min} value of 10 seconds can not be considered as a near crash, hence an extreme value). Therefore a selection of near-crash events is needed, which can be considered as ”sub sampling of maxima” (Jonasson and Rootzén, 2014).

Mahmud et al. (2017) gave an overview of minimum and desirable TTC threshold values from a selection of studies for different conditions. As far as signalized intersections are concerned Mahmud et al. (2017) did not indicate any minimum values, however he cited two references (Huang et al. (2013) and Sayed et al. (2013)) where desired values of 1.6s and 3s were given. Figure 5.1 depicts the cumulative distribution functions for values of TTC_{min} and T_{2min} smaller than 5s showing that for the former there are very few observations in the range of 0~2s.

Taking 3s as a threshold value for near-crashes would only result in 15 observations. Based on what the literature suggests and considering the range of observed values near-crashes were selected using a threshold value of 3.5 seconds for the first run. The above problem does not hold for T_{2min} thanks to its bigger sample size, but for the sake of comparability the same threshold value was applied for the first run. The results of these two model runs were evaluated in detail and followed by several other runs using different threshold values for near-crash situations.

Descriptive statistics for the filtered data are given in Table 5.1. As for TTC_{min} a positive Kurtosis indicates possibly a heavy tail and in the case of T_{2min} a negative Kurtosis slightly below zero suggests a light tail. Scatter plots are given in Figure 5.2, where values are plotted against their corresponding observation number (index).

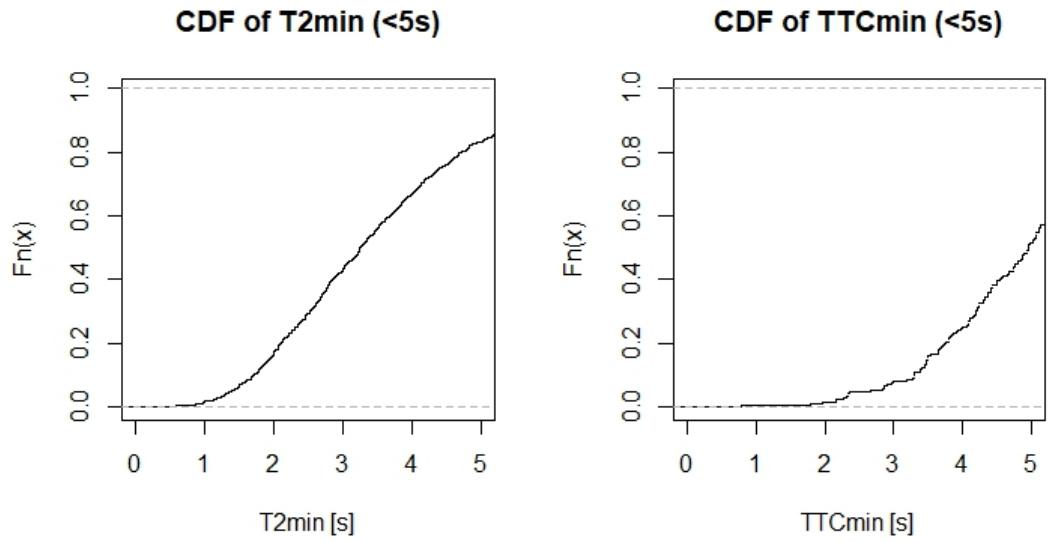


Figure 5.1: Cumulative distribution functions of T_{2min} and TTC_{min} (<5s)

Table 5.1: Descriptive statistics of near-crash values <3.5s

Indicator	TTC_{min}	T_{2min}
Sample size	31	443
Min	0.79	0.06
Max	3.49	3.50
Mean	2.88	2.38
Stdev	0.65	0.69
Skewness	-1.19	-0.38
Kurtosis	1.12	-0.54

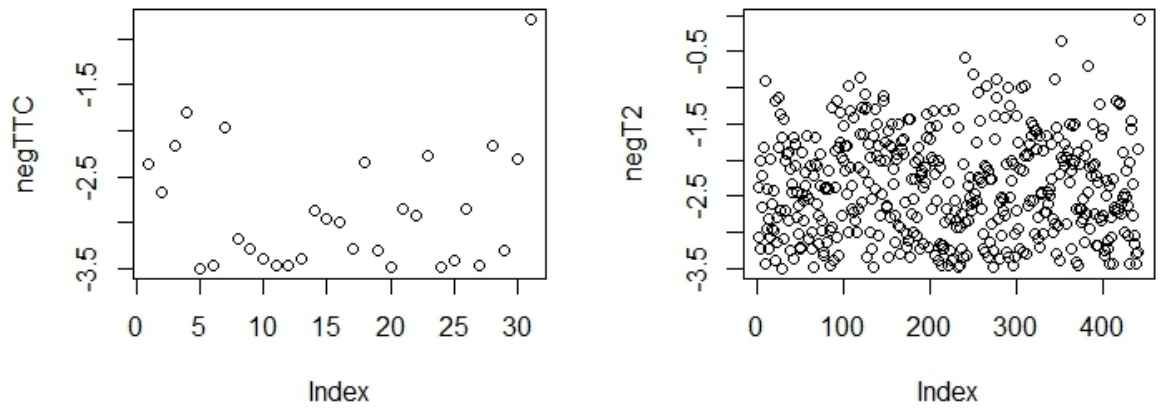


Figure 5.2: Scatter plots of negated T_{2min} and TTC_{min} smaller than 3.5s

5.1.1. Analysis of model results for TTC

Table 5.2 gives a summary of the model results of a fitted GEV using a near-crash threshold of 3.5s.

Table 5.2: Model results of GEV for TTC_{min} (near-crash threshold 3.5s)

Indicator	Location	Scale	Shape
Estimated parameter	-3.336	0.230	1.099
Standard error	0.064	0.078	0.462
Lower bound of confidence interval	-3.461	0.078	0.193
Upper bound of confidence interval	-3.210	0.382	2.004
AIC	47.292		
BIC	51.594		
Deviance	41.292		
Log-likelihood	20.646		

The confidence intervals (95% as well as 99%) of the shape parameter does not include zero, thus we can accept the Frechet distribution as the shape parameter is greater than zero. (In case of a confidence interval including zero the Gumbel distribution can be fitted and the two model fits compared using the log-likelihood ratio test.) Notwithstanding, a greater accuracy for the confidence intervals is usually attained by the profile likelihood (Figure 5.3), which yields similar results.

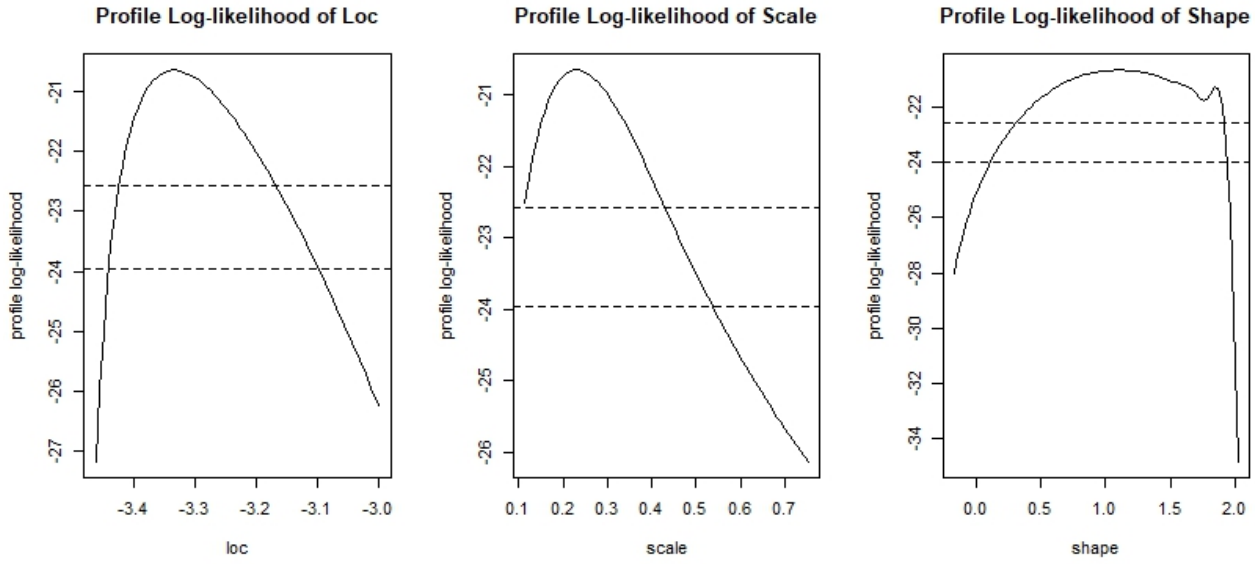


Figure 5.3: Profile log likelihood plots of model parameters for TTC_{min}

Substituting the model estimates into the GEV function (Equation 5.1) using a given value for z_i one can calculate its probability. We are interested in the probability of crash occurrence, that is, when $TTC_{min} < 0$ (z_i=0 in Equation 5.1). This calculation yields a probability of 0.0733 (1-G(z_i)).

Using a given return level z_i one can also obtain the return period, which is 1/p. Equation 5.2 shows how the return level and return period are associated with each other. This means that the level z_i is expected to be exceeded on average once every 1/p. If each block corresponds to one year, then the return period can be interpreted in years. In this particular case each block is an individual near-crash interaction.

$$G(z_i) = \exp\left\{-\left[1+\xi\left(\frac{z_i-\mu}{\sigma}\right)\right]^{\frac{-1}{\xi}}\right\}, \tag{5.1}$$

$$z_i = \begin{cases} \mu - \frac{\sigma}{\xi}\left[1-\left\{-\log(1-p)\right\}^{-\xi}\right], & for\ \xi \neq 0 \\ \mu - \sigma\log\left\{-\log\left(1-p\right)\right\}, & for\ \xi = 0 \end{cases} \tag{5.2}$$

Using Equation 5.2 and the previously calculated probability of crash occurrence (0.0733) one can calculate the return period, which is 1/0.0733=13.65. In other words it means that one out of every 14 near crash interactions (with a TTC_{min} smaller than 3.5s) will result in a crash. A profile log-likelihood plot for the return level of 0s is given in Figure 5.4; with this plot one can also gain confidence intervals. From Figure 5.4 it is clear that there is a wide confidence interval associated with the given return level of 0s.

To further check the goodness of fit of the model probability, quantile, return level and density plots can be used (Figure 5.5). The probability plot is a comparison of the empirical and fitted distribution functions, in the quantile plot their quantiles against each other are plotted. Both can can be visually checked, if in both cases the points are sufficiently close to linearity, the model can be accepted. In order to further investigate the probability plot and to compare the fitted and the empirical distributions, a Kolmogorov-Smirnov test was used, of which the null hypothesis is that the sample is drawn from the fitted distribution. As the p-value is 0.8235 we cannot reject the null hypothesis that our sample deviates from the GEV distribution.

Nevertheless, as Coles (2001) points out the weakness of the probability plot is that it provides the least information in the region of large values (in this case small values). Thus it is important to look at the quantile plot. In this case the probability plot does not reveal any model deficiencies, however the quantile plot indicates model failure. The return level plot shows the return period against return level together with confidence bands (in blue colour). As ξ is greater than zero, the return level plot is a concave curve and has no finite bound. This plot also reveals that the model does not provide a good extrapolation in the region of extreme values (below the lowest observed TTC_{min} value of 0.79). The density plot shows the probability density function of the fitted model together with the histogram of observed data; this plot is less informative for checking the goodness of fit of the model.

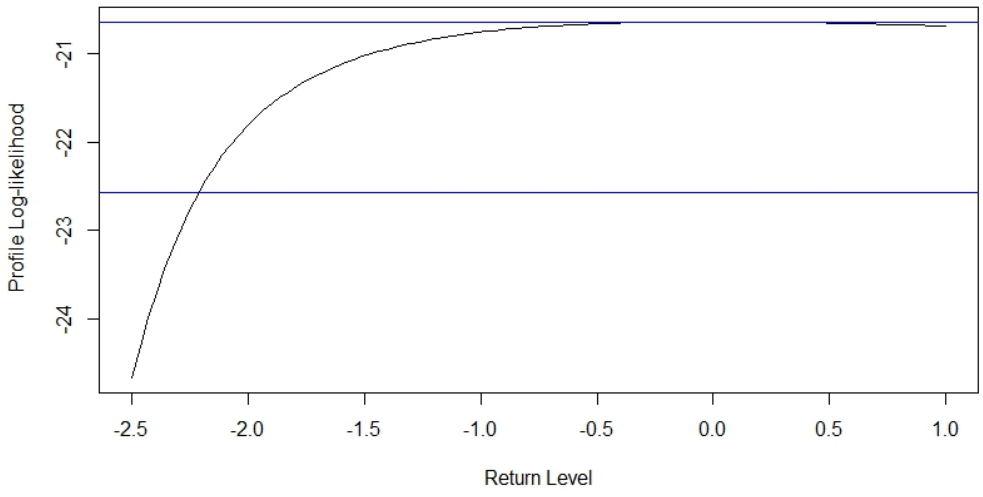


Figure 5.4: Profile log-likelihood plot for the return period of 13.65 associated with the return level of TTC_{min}=0s

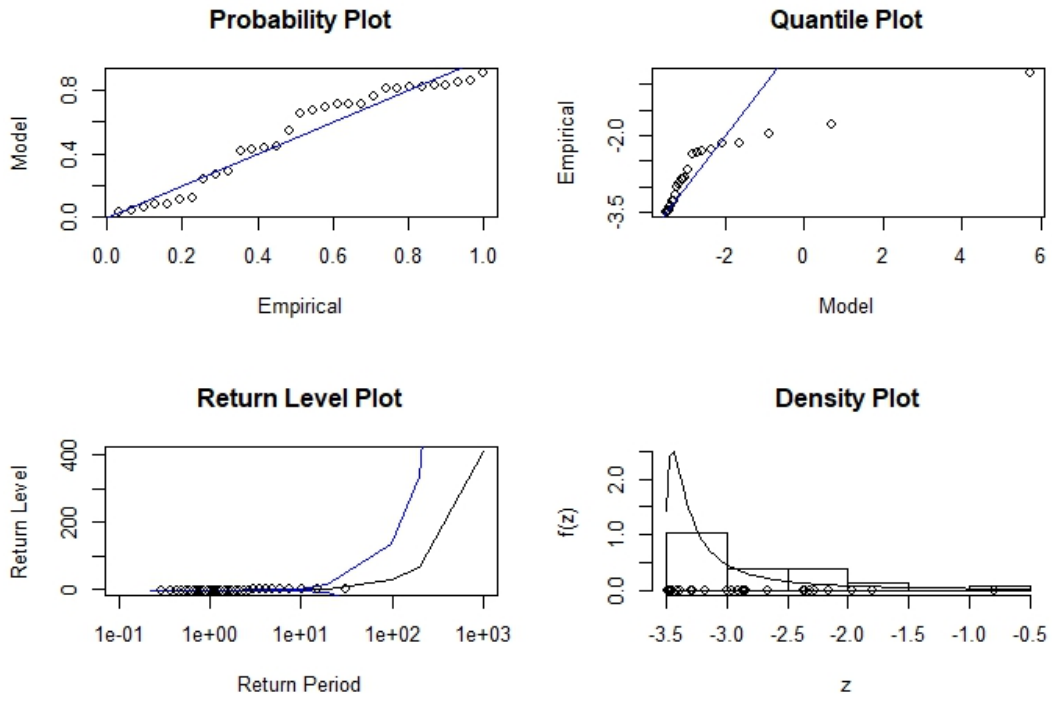


Figure 5.5: Diagnostic plots for GEV fit to TTC_{min} (near crash threshold < 3.5s)

The above analysis reveals that 3.5s as a threshold value for near-crash situations lead to unsatisfactory model results and irrationally high crash probability. This is due to the combined effect of practical as well as statistical reasons. The initially small sample size of TTC_{min} is due to the fact that we are looking at left turning and straight moving vehicle interactions, where in many cases TTC cannot be interpreted due to stopped left-turning vehicles waiting for straight moving ones to pass. From a statistical point of view the small sample size results in unreliable extrapolation and large variance.

For the above reasons several models were tested using different threshold values for the pre-selection of near-crash situations. The threshold was gradually increased with a 0.5s increment. Results are presented in Table 5.3.

Diagnostic plots reveal that by increasing the near-crash threshold the model fit improves gradually, Figure 5.6 shows the results using 5s as a threshold for near crash situations.

Table 5.3: Results of GEV for TTC_{min} with different thresholds for near-crash situations

Indicator	TTC_{min}**<3.5s**	TTC_{min}**<4s**	TTC_{min}**<4.5s**	TTC_{min}**<5s**
Sample size	31	48	76	100
Location parameter	-3.336	-3.552	-3.977	-4.277
Scale parameter	0.230	0.378	0.466	0.605
Shape parameter	1.099	0.327	0.244	0.087
Shape p. lower bound of conf. int.	0.193	0.010	-0.024	-0.109
Shape p. upper bound of conf. int.	2.004	0.644	0.513	0.284
Probability of crash TTC_{min}<0	0.073	0.014	0.010	0.004
Return period for TTC_{min}<0	13.65	73.92	101.04	246.57
Kolmogorov-Smirnov test p-value	0.824	1	0.974	0.994

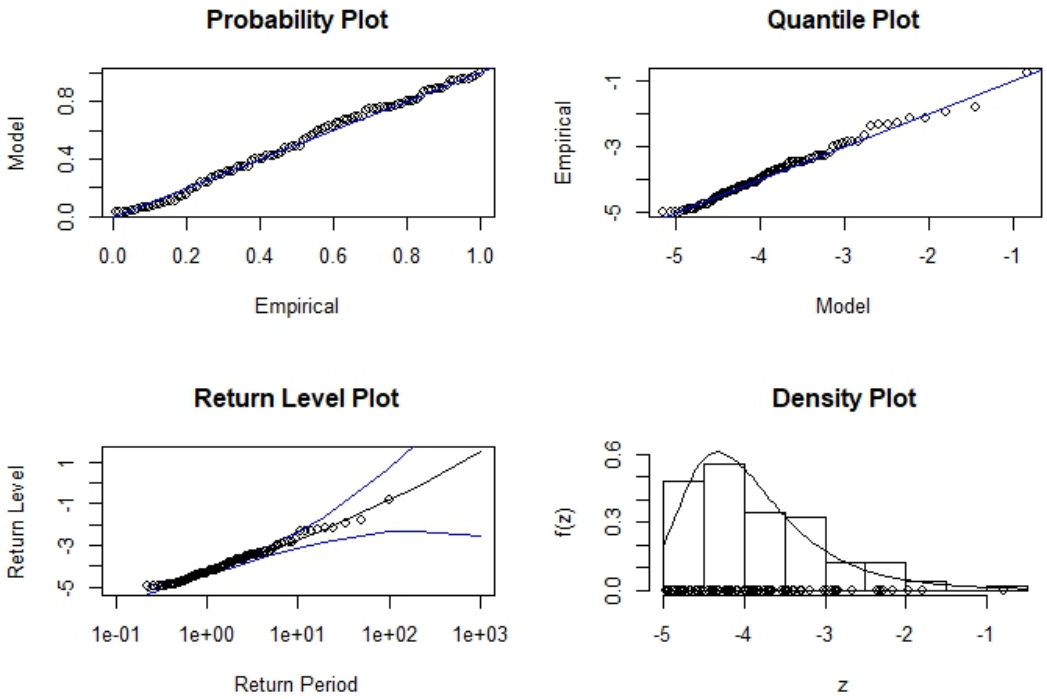


Figure 5.6: Diagnostic plots for GEV fit to TTC_{min} (near crash threshold < 5s)

From Table 5.3 it can be seen that as the near-crash threshold increases (resulting in bigger sample size) the shape parameter converges to zero. With 4.5s threshold the 95% confidence intervals include zero. Setting the shape parameter to zero the Gumbel distribution can be fitted and an analysis of deviance between the two models can reveal whether it is more appropriate (Penalva et al., 2013). The results of these analyses are given in Table 5.4. The results obtained for 3.5, 4, and 4.5 seconds show significant differences between the two models, however, for 5s there is no significant difference so the Gumbel model with two parameters is a good choice for modeling these data.

Table 5.4: Deviance analysis

Model	M.Df	Deviance	df	Chisq	Pr(>Chisq)
M1 <3.5s	3	41.292			
M2 <3.5s (ξ=0)	2	50.147	1	8.855	0.003**
M1 <4s	3	75.840			
M2 <4s (ξ=0)	2	81.178	1	5.338	0.021*
M1 <4.5s	3	144.99			
M2 <4.5s (ξ=0)	2	149.20	1	4.208	0.040*
M1 <5s	3	225.77			
M2 <5s (ξ=0)	2	226.63	1	0.864	0.353

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

There is a specialty in fitting the GEV distribution to the observed negated values, since the GEV distribution is conditional on the values being smaller than zero ($G_0(z)$=G(z)/G(0)) (already noted by Jonasson and Rootzén (2014)). Therefore parameter estimation given in Equation 2.13 has to be slightly modified as an additional term is added (Equation 5.3). Jonasson and Rootzén (2014) also noted that omitting this term leads to only slight changes in the estimated parameters of the GEV distribution.

$$\ell(\mu,\sigma,\xi) = -nlog\sigma - \left(1+\frac{1}{\xi}\right)\sum_{i=1}^{n} log\left[1+\xi\left(\frac{z_i-\mu}{\sigma}\right)\right] - \sum_{i=1}^{n}\left[1+\xi\left(\frac{z_i-\mu}{\sigma}\right)\right]^{\frac{-1}{\xi}} + \sum_{i=1}^{n}\left(1-\frac{\xi-\sigma}{\mu}\right)^{\frac{-1}{\xi}} \quad (5.3)$$

In order to see whether adding this extra term would modify the model parameters, the in-built $gev.fit$ function in R was modified and the above models in Table 5.3 were rerun. Model parameters and crash probabilities are compared in Table 5.5. In general the location as well as the scale parameters do not change, slight changes in the shape parameter can be seen, which tend to disappear by increasing the near-crash threshold and thus the sample size. Also the calculated crash probabilities and return periods are comparable.

Table 5.5: Comparison of GEV model results for TTC_{min} with the condition G(z)/G(0)

Indicator	TTC_{min}**<3.5s**	TTC_{min}**<4s**	TTC_{min}**<4.5s**	TTC_{min}**<5s**
Location parameter	-3.336	-3.552	-3.977	-4.277
Scale parameter	0.230	0.378	0.466	0.605
Shape parameter	1.099	0.327	0.244	0.087
Probability of crash TTC_{min}<0	0.073	0.014	0.010	0.004
Return period for TTC_{min}<0	13.65	73.92	101.04	246.57
Modified location parameter	-3.307	3.547	-3.969	-4.271
Modified scale parameter	0.241	0.374	0.467	0.606
Modified shape parameter	0.681	0.260	0.181	0.055
Modified prob. of crash TTC_{min}<0	0.032	0.008	0.006	0.003
Modified return period for TTC_{min}<0	31.50	119.73	172.56	386.76

5.1.2. Analysis of model results for T2

Table 5.6 gives a summary of the model results of a fitted GEV using the initial near-crash threshold of 3.5s. Neither the 95% nor the 99% confidence intervals include zero, thus we can accept the Weibull distribution as the shape parameter ξ is below zero. Figure 5.7 also confirms this.

Table 5.6: Model results of GEV for T_{2min} (near-crash threshold 3.5s)

Indicator	**Location**	**Scale**	**Shape**
Estimated parameter	-2.674	0.615	-0.129
Standard error	0.034	0.025	0.042
Lower bound of confidence interval	-2.740	0.566	-0.211
Upper bound of confidence interval	-2.608	0.664	-0.047
AIC	914.401		
BIC	926.682		
Deviance	908.401		
Log-likelihood	454.200		

The probability of a crash occurrence (T_2=0) is 0.0016, and the return period (1/p) is therefore 596.03 meaning that one out of 596 near-crash interactions (with a T_2 smaller than 3.5s) will result in a crash. The profile log-likelihood plot for this particular return level (Figure 5.8) shows a much narrower confidence interval than for TTC_{min} (Figure 5.4).

The probability and quantile plots (Figure 5.9) both show a linear pattern so the model can be considered as a good fit. The return level plot shows a convex curve, which is due to ξ being smaller than zero. In this case the return level plot has a finite bound (i.e. due to its convex nature it has a plateau), which can be calculated using the equation μ-σ/ξ yielding 2.079. This value however cannot be interpreted from a traffic safety point of view.

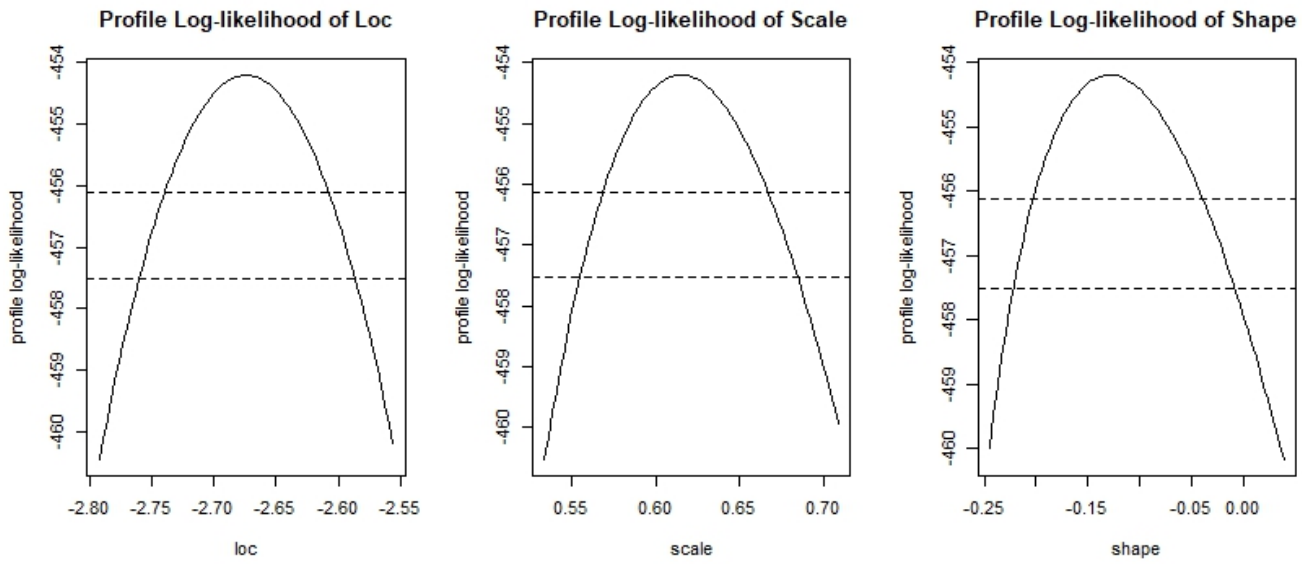


Figure 5.7: Profile log likelihood plots of model parameters for T_2

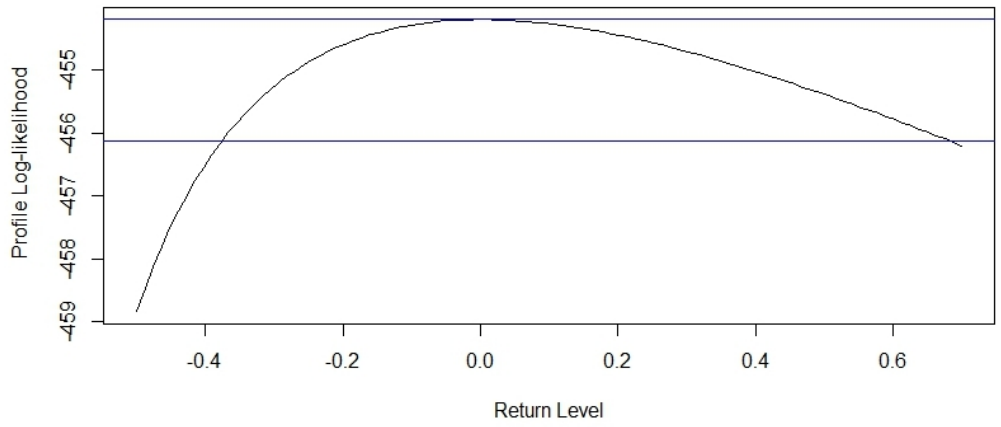


Figure 5.8: Profile log-likelihood plot for the return period of 596.03 associated with the return level of T_2=0s

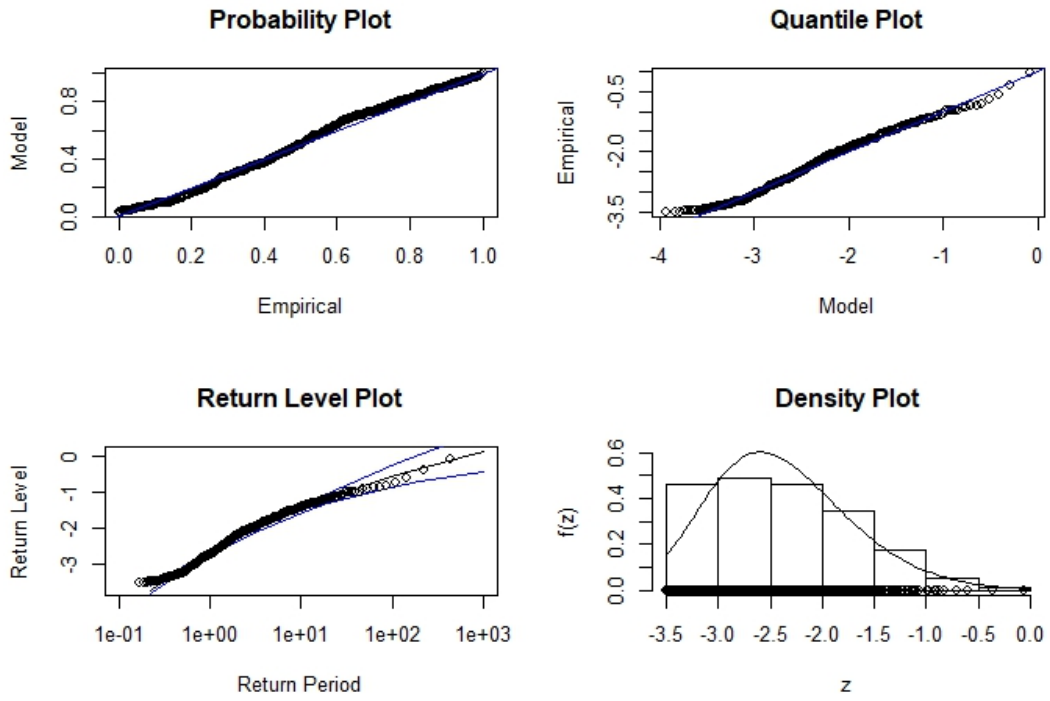


Figure 5.9: Diagnostic plots for GEV fit to T_2 (near-crash threshold < 3.5s)

As for T_2 further steps in model checking are just the opposite as compared to those of TTC_{min} in terms of changing the near-crash threshold. As previously noted, for a critical value of near-crash situations the literature actually suggests a lower threshold than 3.5s, as low as 1.5s. Thus, it is interesting to check how the model fit and output values change as we gradually decrease the near-crash threshold level. In Table 5.7 the most important results are summarized for four different near-crash thresholds. The shape of the distribution changes from a Weibull type (ξ<0) to a Fréchet (ξ>0) as the near-crash threshold levels as well as the sample sizes decrease. Crash probability is gradually increasing by decreasing near-crash thresholds. At a near-crash threshold of 2s a crash probability of 0.0098 is calculated associated with a return level of 101.96 meaning that one crash would happen out of 102 near-crash interactions. Almost the same result is given in case of TTC_{min} using a near-crash threshold of 4.5s (Table 5.4). The model fit associated with the 2s near-crash threshold still gives acceptable results (Figure 5.10) as both the probability as well as quantile plots show a linear pattern, however the return level plot gives wider confidence bands compared to Figure 5.9. The Kolmogorov-Smirnov test shows in all four cases that we cannot reject the null hypothesis that our sample deviates from the GEV distribution.

Table 5.7: Results of GEV for T_{2min} with different thresholds for near-crash situations

Indicator	T_{2min}**<3.5s**	T_{2min}**<3s**	T_{2min}**<2.5s**	T_{2min}**<2s**
Sample size	443	341	232	130
Location parameter	-2.674	-2.382	-2.050	-1.712
Scale parameter	0.615	0.473	0.352	0.246
Shape parameter	-0.130	-0.045	0.026	0.166
Shape p. lower bound of conf. int.	-0.211	-0.150	-0.098	-0.031
Shape p. upper bound of conf. int.	-0.047	0.059	0.150	0.364
Probability of crash T_{2min}<0	0.002	0.003	0.004	0.010
Return period T_{2min}<0	596.02	302.40	226.10	101.96
Kolmogorov-Smirnov test p-value	0.589	0.537	0.982	0.992

Again it is reasonable to check whether the Gumbel distribution is a good choice for models where ξ falls within a confidence interval including zero. The output results for the deviance analysis can be found in Table 5.8 showing that with 3 and 2.5s as near-crash threshold levels the Gumbel model is actually not significantly different and thus a linear approximation of the return levels is appropriate.

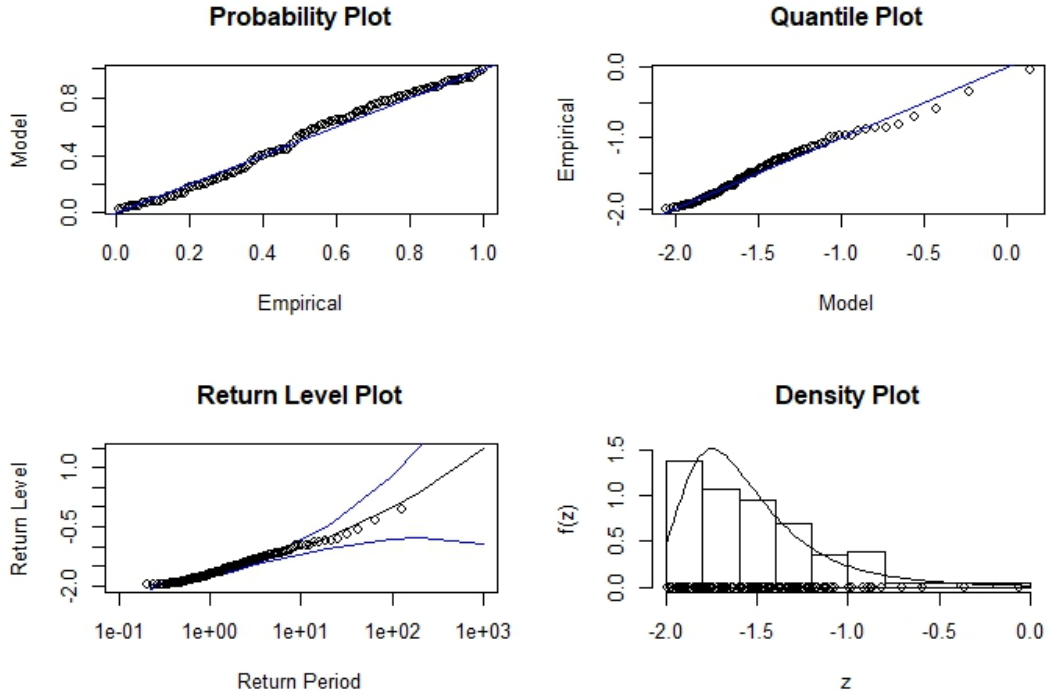


Figure 5.10: Diagnostic plots for GEV fit to T_2 (near-crash threshold < 2s)

Table 5.8: Deviance analysis

Model	M.Df	Deviance	df	Chisq	Pr(>Chisq)
M1 <3.5s	3	908.40			
M2 <3.5s (ξ=0)	2	915.96	1	7.556	0.006**
M1 <3s	3	552.97			
M2 <3s (ξ=0)	2	553.64	1	0.672	0.413
M1 <2.5s	3	257.01			
M2 <2.5s (ξ=0)	2	257.19	1	0.178	0.6733
M1 <2s	3	71.232			
M2 <2s (ξ=0)	2	74.625	1	3.393	0.066

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

In order to check again whether conditioning on G(z)/G(0) modifies the model parameters, the above models in Table 5.7 were rerun. Model parameters and crash probabilities are compared in Table 5.9. Similar results as for TTC_{min} are received, showing that the location and scale parameters do not change, there are only minor differences in the shape parameters. Also the calculated crash probabilities and return periods are comparable.

Table 5.9: Comparison of GEV model results for T_{2min} with the condition G(z)/G(0)

Indicator	T_{2min}**<3.5s**	T_{2min}**<3s**	T_{2min}**<2.5s**	T_{2min}**<2s**
Location parameter	-2.674	-2.382	-2.050	-1.712
Scale parameter	0.615	0.473	0.352	0.246
Shape parameter	-0.130	-0.045	0.026	0.166
Probability of crash T_{2min}<0	0.002	0.003	0.004	0.010
Return period T_{2min}<0	596.02	302.40	226.10	101.96
Modified location parameter	-2.669	-2.376	-2.046	-1.708
Modified scale parameter	0.618	0.476	0.353	0.247
Modified shape parameter	-0.148	-0.075	-0.007	0.105
Modified prob. of crash T_{2min}<0	0.001	0.002	0.003	0.006
Modified return period for T_{2min}<0	991.63	525.86	370.29	179.93

5.2. Peak-over-Threshold Approach

The POT approach offers a different solution to modeling extreme events. It is necessary to choose a threshold over which extreme events are considered. "It is important to choose a sufficiently high threshold in order that the theoretical justification applies thereby reducing bias. However, the higher the threshold, the fewer available data remain. Thus, it is important to choose the threshold low enough in order to reduce the variance of the estimates." (Gilleland and Katz, 2016).

Selecting the appropriate threshold using POT is as important as selecting the block size using BM. There are basically two methods available for that:

- Mean residual plot: this plot shows the mean of the excesses depending on the value of the chosen threshold level u. Above a certain value the GPD provides a valid approximation to the excess distribution (Coles, 2001). Here a threshold has to be selected where the graph is linear within uncertainty bounds. This is, however, not always straightforward, and based on a subjective choice.
- Model estimation: The model is estimated at a range of threshold values with the intention to find stable model parameters. Again, above a certain level of u the GPD is valid, if estimates of the shape parameter ξ are constant, while estimates of the scale parameter σ is linear in u. This point can be read from the plot by checking linearity, in other words estimates will not change much within uncertainty bounds, as the threshold increases.

5.2.1. Analysis of model results for TTC

The lowest threshold where the mean residual plot (Figure 5.11) becomes linear within uncertainty bounds is a value around -4s. There is a lot of fluctuation at the right hand side of the plot due to fewer observations. The parameter estimates against thresholds (Figure 5.12) also show relatively stable results for the selected value. The GPD model results are given in Table 5.10. The shape parameter is below zero resulting in a convex return level plot with a finite upper bound.

The probability of crash occurrence, namely when TTC_{min} equals zero, can be calculated using Equation 5.4 substituting the model parameters ξ and σ, the threshold u=-4 and x=0. This calculation yields a crash probability of 0.00017 (see Equation 5.4). If we account for the ratio of near-crashes by multiplying this value with ζ_u (48/194) (see Equation 5.5), the estimated probability is 0.000042.

$$Pr\{X > x | X > u\} = \left[1 + \xi\left(\frac{x-u}{\sigma}\right)\right]^{\frac{-1}{\xi}}, \tag{5.4}$$

From Equation 5.4 it follows that

$$Pr\{X > x\} = \zeta_u\left[1 + \xi\left(\frac{x_m-u}{\sigma}\right)\right]^{\frac{-1}{\xi}}, \tag{5.5}$$

where ζ_u=k/n $(\zeta_u$=$Pr\{X > u\})$, with k being the number of excesses over u, n being the number of observations, in other words, the sample proportion of points exceeding u, and x_m being the return level (that is exceeded on average once every m observations). Using Equation 5.5, which is equal to $1/m$, the m-observation return period can also be calculated.

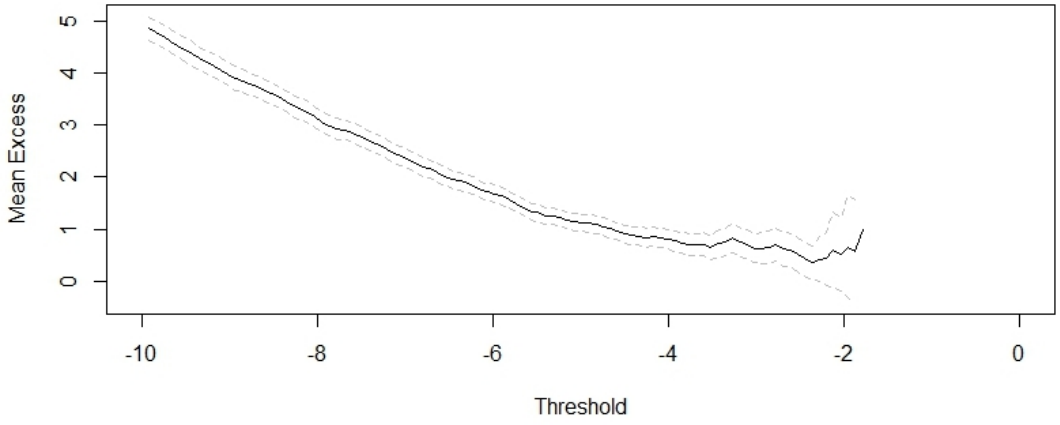


Figure 5.11: Mean residual plot for TTC_{min}

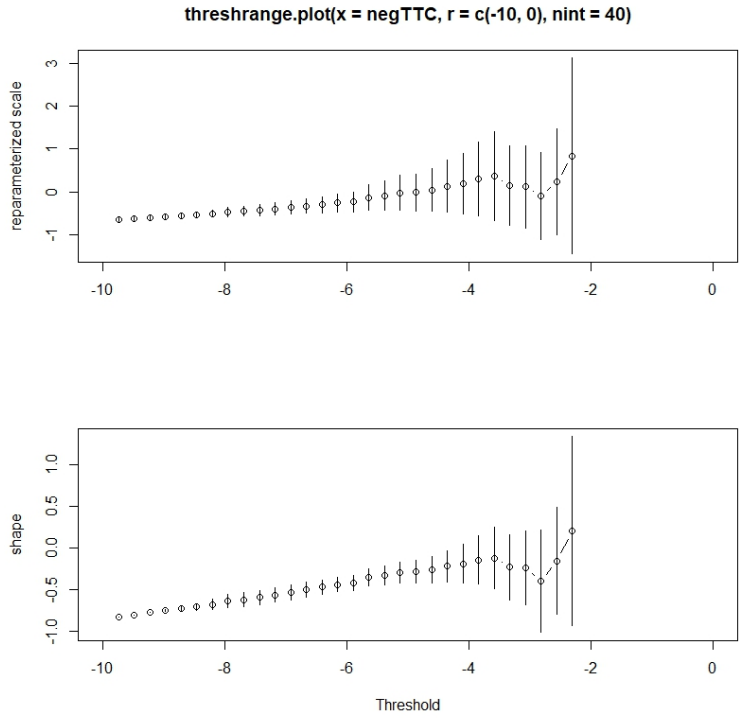


Figure 5.12: Parameter estimates against threshold for TTC_{min}

Table 5.10: Model results of GPD for TTC_{min} (threshold 4s)

Indicator	**Scale**	**Shape**
Sample size	48	
Estimated parameter	0.970	-0.200
Standard error	0.180	0.120
Lower bound of confidence interval	0.617	-0.436
Upper bound of confidence interval	1.323	0.036
AIC	77.906	
BIC	81.649	
Log-likelihood	36.953	
Kolmogorov-Smirnov test p-value	0.997	

The GPD diagnostic plots suggest a reasonable model fit, also the Kolmogorov-Smirnov test is not significant and thus we cannot reject that the sample deviates from the GPD distribution. However the return level plot shows that as the return period increases the return level confidence bends tend to be wider meaning that the prediction of unobserved extreme values comes with uncertainty. The return period associated with TTC_{min} < 0 is 5,884.8 (1/0.00017) (using Equation 5.4), meaning that one out of 5885 near-crash interactions results in a crash. Accounting for the ratio of near-crashes in the full dataset the return period is 23,784.4 (1/0.000042) (using Equation 5.5), which means that one out of every 24 thousand interactions results in a crash. (The return level plot in Figure 5.13 is set to years assuming that one observation corresponds to one day. Therefore in the return level plot zero return level is associated with a value of 65.16 (23,784.4/365). Confidence interval bounds of the parameters as well as the previously calculated return level can be seen in Figure 5.14.

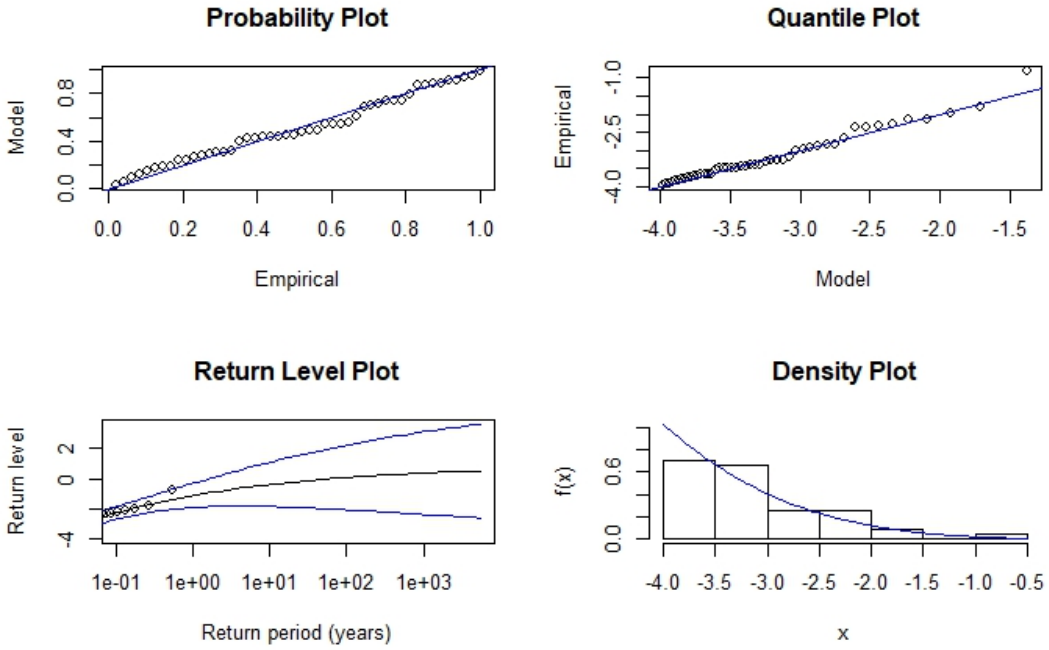


Figure 5.13: Diagnostic plots for GPD fit to TTC_{min} (threshold 4s)

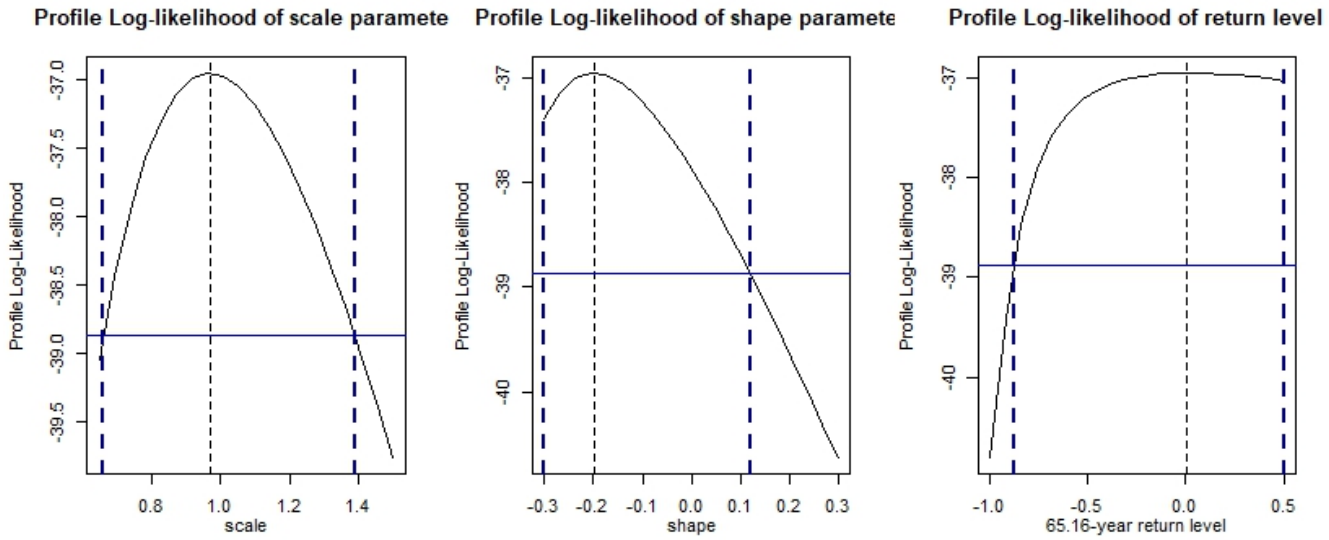


Figure 5.14: Profile Log-likelihood plots for GPD fit to TTC_{min} (threshold 4s)

Having tested one threshold value it is worth investigating what results other threshold values would yield. To that end a zoomed-in version of the threshold stability plot without confidence limits is shown in Figure 5.15. This plot actually proves that the previously chosen threshold of 4s is justifiable. We can also try testing further threshold values thus we need to search for other stable regions in the plot. The region between 4.0 and 4.2s seems to be stable, thus further models are also fitted using threshold values of 4.1, and 4.2s. Results are presented in Table 5.11. Based on the diagnostic plots, and especially the density plot (these are not presented here) it can be concluded that the original model with the 4s threshold gives the best result, therefore this model was kept.

Table 5.11: Results of GPD for TTC_{min} with different thresholds

Indicator	**u=4s**	**u=4.1s**	**u=4.2s**
Sample size	48	54	59
Scale parameter	0.970	0.960	0.997
Shape parameter	-0.200	-0.183	-0.192
AIC	77.906	87.826	98.930
BIC	81.648	91.804	103.086
Log-likelihood	36.953	41.913	47.465
Probability of crash TTC_{min}<0	0.00017	0.00024	0.00018
Return period for TTC_{min}<0	5884.8	4160.2	5618.1
Kolmogorov-Smirnov test p-value	0.997	0.999	0.999

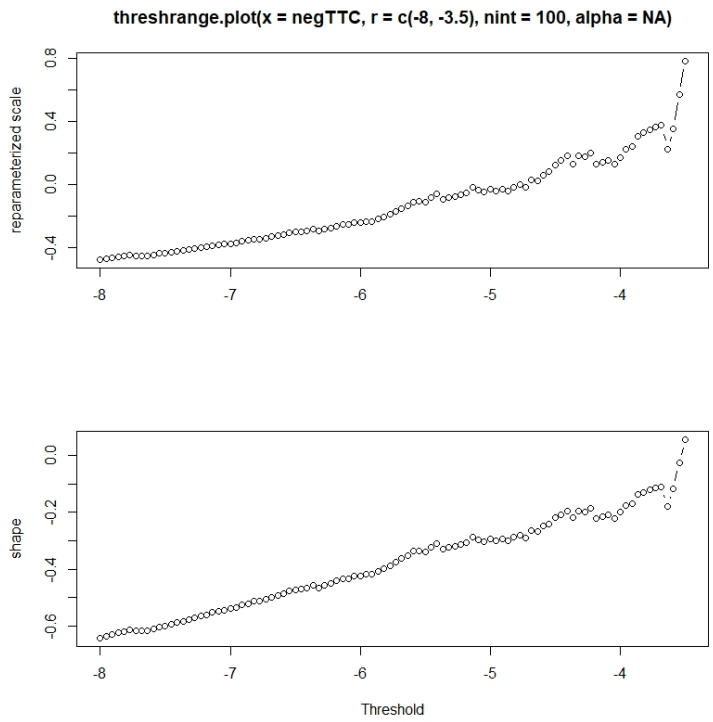


Figure 5.15: Parameter estimates against threshold for TTC_{min}

5.2.2. Analysis of model results for T2

As for T_{2min} a different threshold was chosen, both the mean residual plot (Figure 5.16) and the plots of parameter estimates against thresholds (Figure 5.17) suggest a threshold of -2s to be used. Model results are presented in Table 5.12. The crash probability associated with this model is 0.00055. If we account for the ratio of near-crashes by multiplying this value with ζ_u (130/792) (see Equation 5.5), the estimated probability is 0.000091.

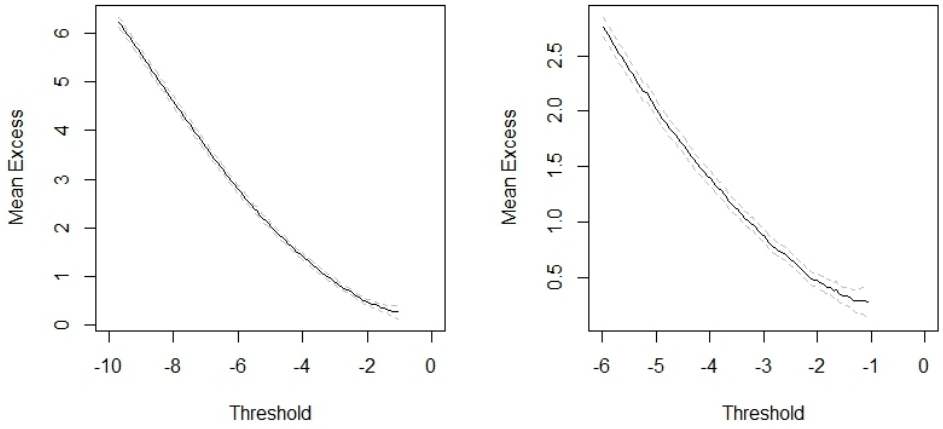


Figure 5.16: Mean residual plots for T_{2min} (left: T_{2min} > -10s, right: T_{2min} > -6s)

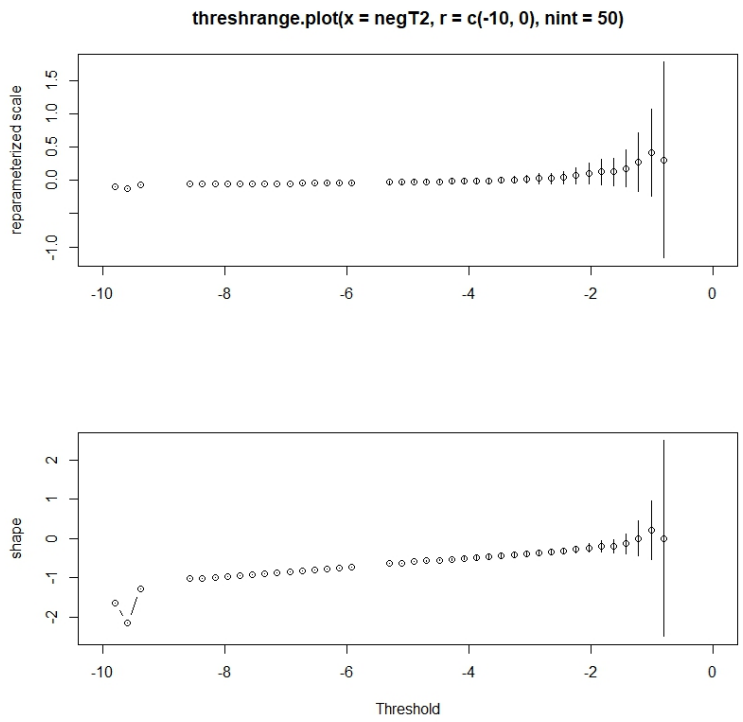


Figure 5.17: Parameter estimates against threshold for T_{2min}

The GPD diagnostic plots (Figure 5.18) suggest a good model fit with favorable confidence bends in the return level plot. The return period associated with T_{2min} < 0 is 1,807.3 (1/0.00055) (using Equation 5.4), meaning that one out of 1800 near-crash interactions results in a crash. Accounting for the ratio of near-crashes in the full dataset the return period is 11,010.4 (1/0.000091) (using Equation 5.5), which means that one out of every 11 thousand interactions results in a crash (this value is associated with 30.16 in the return level plot). Confidence interval bounds of the parameters as well as the previously calculated return level can be seen in Figure 5.19.

Table 5.12: Model results of GPD for T_{2min} (threshold 2s)

Indicator	**Scale**	**Shape**
Sample size	130	
Estimated parameter	0.585	-0.246
Standard error	0.060	0.058
Lower bound of confidence interval	0.467	-0.360
Upper bound of confidence interval	0.703	-0.132
AIC	60.489	
BIC	66.224	
Log-likelihood	28.24	
Kolmogorov-Smirnov test p-value	0.992	

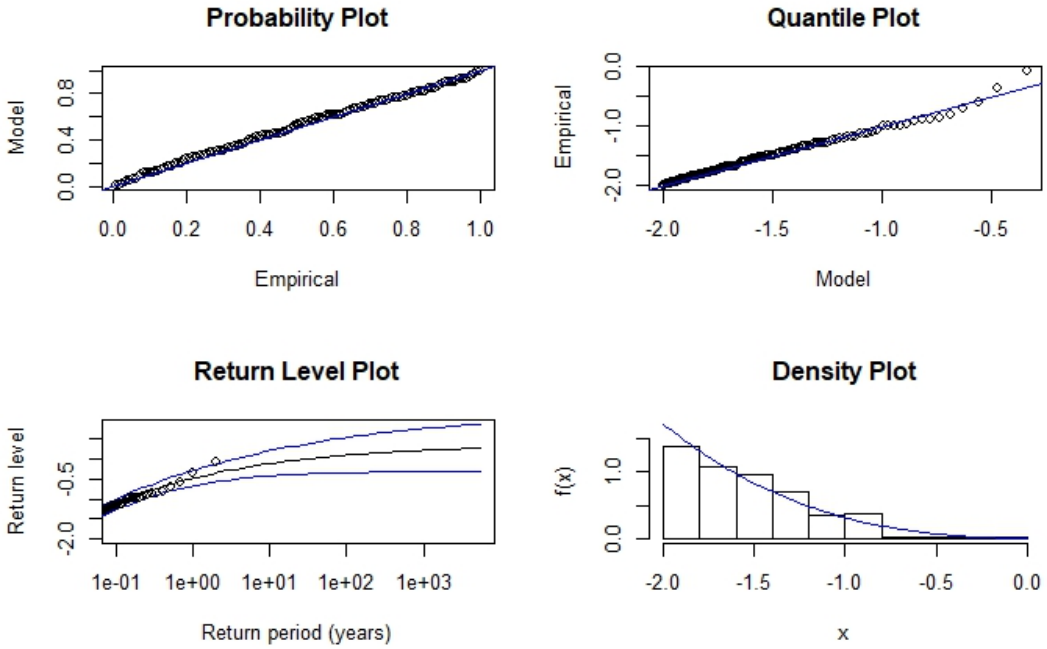


Figure 5.18: Diagnostic plots for GPD fit to T_{2min} (threshold 2s)

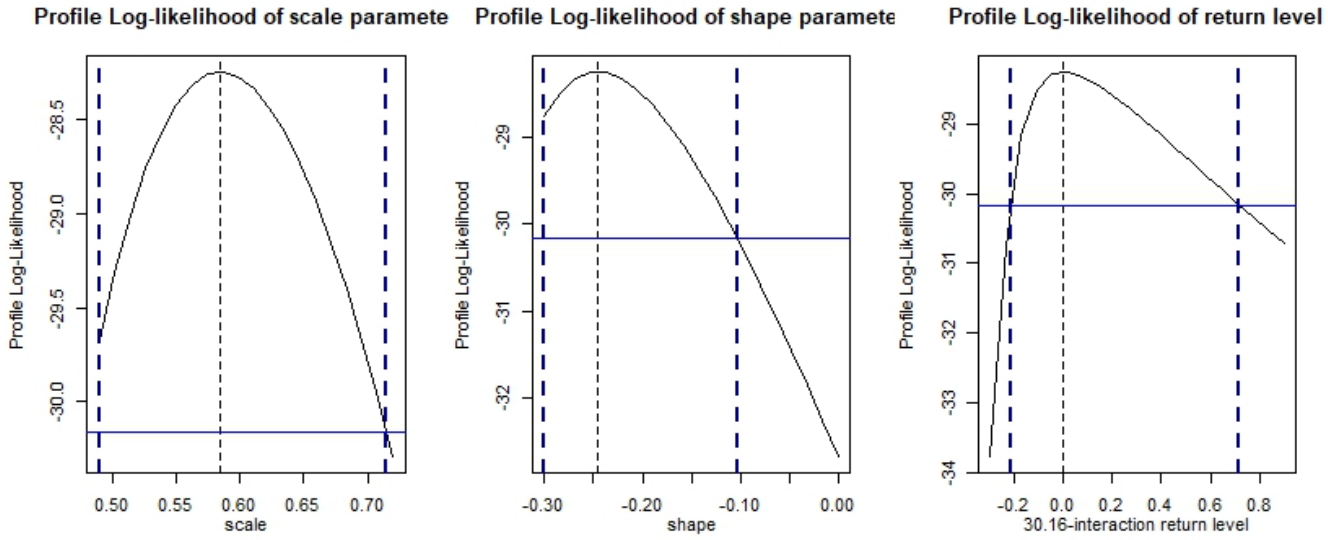


Figure 5.19: Profile Log-likelihood plots for GPD fit to T_{2min} (threshold 2s)

To further investigate whether other thresholds would yield similar results a zoomed-in version of the threshold stability plot without confidence limits is shown in Figure 5.20. This plot indicates that the range of values between 2.2 and 1.9s seems to be stable. Three further models using threshold values of 2.2, 2.1 and 1.9s are fitted; results are given in Table 5.13. Diagnostic plots for the extra models showed slight differences in the density plots. Models with threshold values of 2 and 2.2s proved to provide a good fit, thus the original models with a 2s threshold was kept.

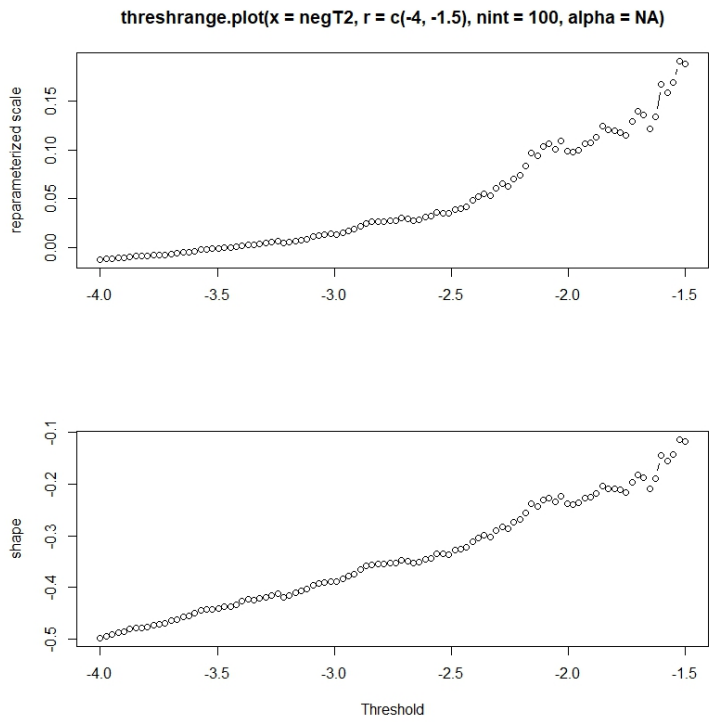


Figure 5.20: Parameter estimates against threshold for T_{2min}

Table 5.13: Results of GPD for T_{2min} with different thresholds

Indicator	**u=2.2s**	**u=2.1s**	**u=2s**	**u=1.9s**
Sample size	176	159	130	113
Scale parameter	0.662	0.586	0.585	0.530
Shape parameter	-0.266	-0.229	-0.246	-0.221
AIC	117.044	79.177	60.489	36.794
BIC	123.385	85.315	66.224	42.248
Log-likelihood	56.522	37.589	28.244	16.397
Probability of crash T_{2min}<0	0.0003	0.0005	0.0005	0.0008
Return period for T_{2min}<0	3384.6	1808.6	1807.2	1193.2
Kolmogorov-Smirnov test p-value	0.939	0.999	0.914	0.997

5.3. Discussion

When applying the Block Maxima approach the selection of near-crash situations as a sub-sampling step proved to be a critical issue. As for TTC_{min} the question was to which level the near-crash threshold should be increased to have a reasonable model fit, whereas for T_{2min} to which - from a traffic a safety point of view - more reasonable level can we decrease the near-crash threshold in such a way that we still have a good model fit.

In the former case with TTC_{min} increasing the near-crash threshold resulted in better model fits, however from a traffic safety point of view these high thresholds can not actually be considered as near-crash events. In the latter case with T_{2min} the threshold value could be further decreased with the disadvantage of slightly less well performing models. Overall, obviously there is a trade-off between a good model fit and reasonable threshold values.

Applying the POT approach seems to give more reasonable results in terms of crash probabilities and return periods (Table 5.14), which were in the hundreds with BM but in the

thousands with POT. If we accept a few assumptions we can attempt to validate these probabilities. These assumptions are as follows:

- the number of interactions used in the analysis (194 for TTC_{min} and 792 for T_{2min}) was all the interactions observed in the 2-day period between 6AM and 9PM, and no interactions were left out;
- the observation period (6AM-9PM) is a good representation of the entire day and accidents did not happen outside this time period;
- accident data provided are accurate, namely 5 crashes due to the collision of left turning and straight going vehicles in a 11-year period, which is approximately 800 days/accident occurrence (one accident happened in 800 days on average).

Accepting the above assumptions and comparing the model results we can actually state that indeed the POT results are much closer to the actual crash frequency. The POT model for TTC_{min} (245.20 days/accident) gives the best prediction, especially if we accept the assumption that property damage accidents are in general underreported.

It can also be concluded that models for T_{2min} provide more reliable results as the return level plots come with narrower confidence bands, thus the prediction of probabilities of extreme events is more accurate. On the other hand, validation results show that T_{2min} tends to overestimate crash frequencies, e.g. for POT T_{2min} estimated one crash in every 28 days, which value was more realistic (245) for TTC_{min}.

Table 5.14: Summary of results for both BM and POT with different thresholds

Method	Indicator (threshold)	Probability	Probability of crash	Return period	Sample (2 days)	Days / accident
BM	TTC (<3.5s)	Pr{TTC<0|TTC<3.5}	0.0733	13.65	31	0.88
	TTC (<4s)	Pr{TTC<0|TTC<4}	0.0135	73.92	48	3.08
	TTC (<4.5s)	Pr{TTC<0|TTC<4.5}	0.0098	101.04	76	2.66
	TTC (<5s)	Pr{TTC<0|TTC<5}	0.0040	246.57	100	4.93
	T2 (<2s)	Pr{T2<0|T2<2}	0.0098	101.96	130	1.57
	T2 (<2.5s)	Pr{T2<0|T2<2.5}	0.0044	226.10	232	1.95
	T2 (<3s)	Pr{T2<0|T2<3}	0.0033	302.40	341	1.77
	T2 (<3.5s)	Pr{T2<0|T2<3.5}	0.0016	596.02	443	2.69
POT	TTC (<4s)	Pr{TTC<0|TTC<4}	0.00017	5,884.80	48	245.20
	T2 (<2s)	Pr{T2<0|T2<2}	0.00055	1,807.26	130	27.80

The above results illustrate that the near-crash threshold value affecting sample size is a critical issue, especially with the BM approach. Figure 5.21 illustrates this by further refining near-crash threshold values. Here the block-maxima approach was used and sub-samples were created using near-crash threshold values by using a 0.05 increment. As for T_{2min} 75 models were fitted for near-crash values ranging from 1.3s to 5s and for TTC_{min} 31 models were fitted for near-crash values ranging from 3.5s to 5s. Crash probabilities were calculated for all the models. As the sub-sampling near-crash threshold is increased, sample sizes become bigger resulting in better model fits, even though with less pragmatic near-crash thresholds; so the previously mentioned trade-off is clearly illustrated.

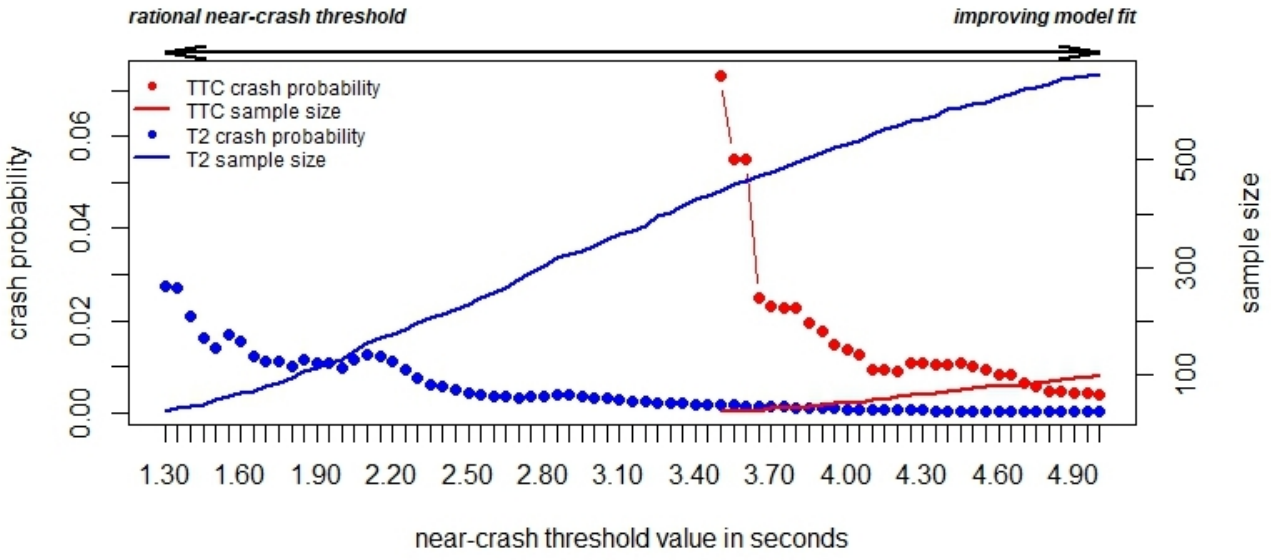


Figure 5.21: Results of Block Maxima approach for a sequence of threshold values

Figure 5.21 also illustrates that for different near-crash thresholds different crash probabilities are predicted for the two indicators. As we are analysing two indicators of different nature (a collision course (TTC_{min}) and a crossing course (T_{2min}) indicator), it is worth investigating for what near-crash threshold values would they predict similar crash probabilities, in other words whether there is transferability in between them. This would provide further insight into the applicability of collision and crossing course indicators. Based on the results shown in Figure 5.21 threshold values were selected for which both indicators yielded the most similar crash probability (i.e. the difference between the predicted crash probabilities was marginal). This plot (Figure 5.22) actually describes the relationship between these two indicators, saying under what near-crash thresholds we receive almost the same crash probability. There is some fluctuation in the graph, but the pattern clearly shows that for T_{2min} lower near-crash thresholds would yield the same crash probability (e.g. 3.5s for TTC_{min} and 1.25s for T_{2min}). A Pearson correlation test was used to determine how strong the relationship is. This test was highly significant (p-value = 8.96e-14) and indicated a strong correlation (0.93).

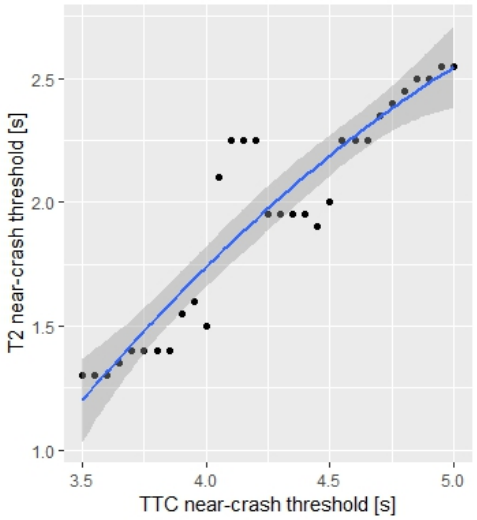


Figure 5.22: Temporal indicators with similar crash probabilities

Judging which indicator is better could be done by validation using a proper accident dataset which is unfortunately not available for this study. Notwithstanding, it has to be noted that this is not an exceptional case, and even with available historical data its applicability can be questioned in general (for reasons associated with the drawbacks of accident records already outlined previously). A judgment can only be made by considering the rationalism of crash probabilities and return periods as well as the goodness of the models.

Having compared the estimated crash probabilities for different near-crash thresholds and by checking the correlation between them we can conclude that collision and crossing course indicators are transferable. Crash probabilities calculated using EVT showed that one has to be ”stricter” against crossing course interactions, as compared to collision course interactions, lower near-crash values would yield similar probabilities. It also has to be noted, that in this study exclusively straight moving and left-turning vehicle-vehicle interactions were analyzed and thus results cannot be generalized to other interaction types. However, this analysis revealed that for this type of interaction, in comparison with a crossing course indicator T_{2min}, a collision course indicator TTC_{min} is of limited use because of its smaller sample size that can be gathered in a given time period.

A possible step to refine the models is using motion prediction. As the above investigated indicators both assume constant speeds and unchanged paths, which is not realistic, it is worthwhile considering a probabilistic approach to predict trajectories and speeds of interacting vehicles. This approach would result in different values with different probabilities for a single interaction, thus providing an increased sample size for both indicators. St-Aubin et al. (2015) developed an approach called Probabilistic Surrogate Measures of Safety (PSMS) considering all possible paths that may lead to two road users to collide. At the time of writing there are also initiatives at Lund University to apply a probabilistic framework.

Using the model results one can also calculate the probability of non-crash events, in other words when a temporal indicator is higher than zero. The closeness of a temporal indicator to zero can also explain the severity of the interaction. For this purpose the POT model parameters were used as they yielded more rational results. Figure 5.23 and Figure 5.24 show the predicted probabilities associated with the range of values between zero and u, which was 4s and 2s for TTC_{min} and T_{2min}, respectively. The predicted probabilities plotted are conditional on the temporal indicator equal to the threshold u used in the model. As it was mentioned in the literature review section one of the approaches to determine the severity of conflicts is simply setting thresholds separating serious and light conflicts. This approach certainly has limitations, as the choice of this threshold value is arbitrary and using a temporal indicator on its own does not account for the possible consequences. The bivariate approach in Chapter 6 addresses this issue.

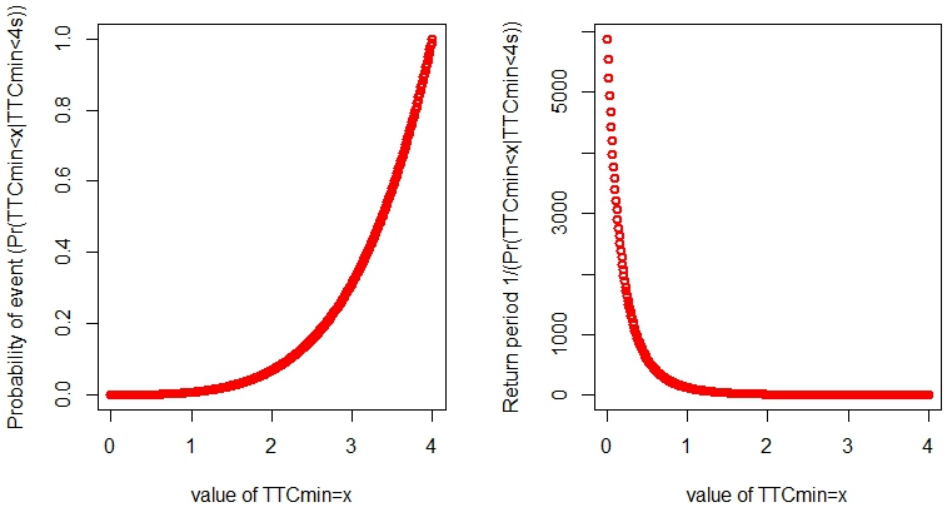


Figure 5.23: Probability of events and return periods for TTC_{min} values (POT approach)

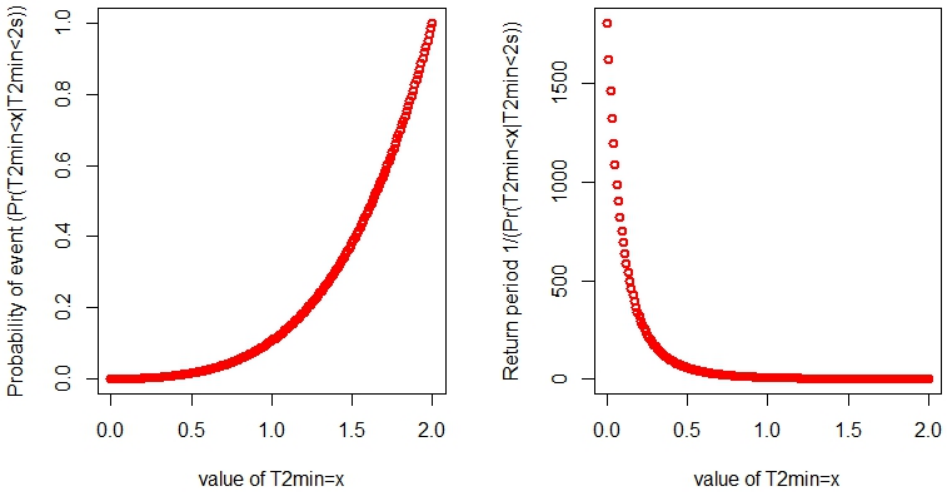


Figure 5.24: Probability of events and return periods for T_{2min} values (POT approach)

6

Bivariate models

In the previous chapter univariate models were used with the objective to estimate crash probabilities. As it was outlined in the literature review a few researchers (for instance Zheng et al. (2014b), Laureshyn et al. (2017a), Zheng et al. (2018)) already stressed the relevance of the joint modeling of nearness-to-collision and a severity indicator of some sort. The primary goal of this chapter is to provide a solution for that using bivariate models. As for the temporal indicators TTC_{min} and T_{2min} will be used with the secondary intention of analyzing and comparing a collision and a crossing course indicator. As for severity two speed related indicators are used:

- Relative speed, which is the speed of a moving vehicle relative to another moving vehicle. If for instance the vehicles are moving in the opposite direction, their relative speed is the sum of their individual speeds.
- Delta-V0, which shows the change of the velocity vector by a road user during a crash. It extends the severity measured by relative speed by taking into account the masses of vehicles colliding and the angle of collision (see Equation 2.4). Of the individual Delta-V0 values of the two vehicles the highest can be used to describe severity. Further assumptions can be made about the deceleration of the vehicles, however this leads to a gradual decrease in the sample of interactions. For this reason in this analysis the base value will be used assuming no braking.

As it was already mentioned at the end of Chapter 4 the dataset used in this chapter is slightly different from the one used for the univariate case. Three observations due to their different type of interaction were removed.

In this chapter altogether four cases are investigated using bivariate models:

- negated TTC_{min} vs. relative speed;
- negated T_{2min} vs. relative speed;
- negated TTC_{min} vs. DeltaV0;
- negated T_{2min} vs. DeltaV0.

As POT gave more realistic and reliable results for univariate models, it is preferred over the Block Maxima approach and will be used for the bivariate case. Depending on whether temporal based and/or speed related indicators exceed their thresholds (u_1 and u_2, respectively), the plane can be divided into four quadrants. These quadrants are indicated in Figure 6.1 (note that in the course of bivariate modeling for TTC_{min} the negated values are used, thus the relation signs are actually the opposite):

- $R_{0,0} = (TTC > u_1, RS < u_2)$: both indicators are below their thresholds;
- $R_{0,1} = (TTC > u_1, RS > u_2)$: speed related indicator exceeds the threshold, temporal based stays below;
- $R_{1,0} = (TTC < u_1, RS < u_2)$: temporal based indicator exceeds the threshold, speed related indicator stays below;
- $R_{1,1} = (TTC < u_1, RS > u_2)$: both indicators exceed their thresholds.

If temporal based indicators are equal to 0, a traffic conflict results in a crash. Speed related indicators explain the consequence (severity) of a crash. A speed related indicator being high does not necessarily mean a crash, however it increases the probability of a more severe outcome.

As it was earlier suggested by Laureshyn et al. (2017a) the combination of temporal and speed related indicators into one plane creates so-called severity levels. These lines show events of equal severity. Quite intuitively severity increases as the value of the temporal indicator goes down and the value of the speed related indicator goes up. Even though these lines represent encounters of equal severity the probability of points along these lines can be different. Another way to interpret this plane is that we look for points (events) located along such lines that have the same or very similar probability. The shape of these probability based risk levels has to be investigated as it is not necessarily the same as that of the severity levels. By means of bivariate threshold excess models this chapter also aims for investigating these probability based risk levels.

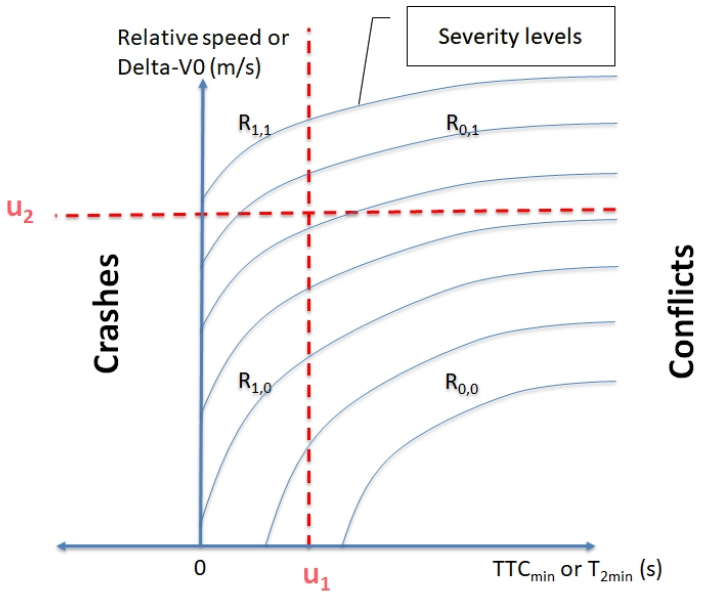


Figure 6.1: Severity levels and the four regions of the bivariate extreme value model

6.1. Modeling procedure

The steps of the modeling procedure are outlined and explained as follows:

1. Correlation analysis (Pearson, Spearman, Kendall)

   Traditional tests are run to see the correlation between the above mentioned temporal and speed related variables. For all the three correlation tests their coefficients and p-values are given. In cases when the p-value is smaller than the significance level α, we can reject the null hypothesis and conclude that the correlation between the two indicators is significant.

2. Independence Test for Bivariate Copula Data

   This test returns the p-value of a bivariate asymptotic independence test based on Kendall's τ. The test is performed using pseudo observations [0,1], which are normalized ranked data. A test yielding a significant result shows dependence between the two variables.

3. Coefficient of extremal dependence

   Bivariate extreme value distributions assume that the margins are either asymptotically dependent or perfectly independent. They cannot account for situations where the dependence between variables disappears at extreme levels (Stephenson, 2018).

   One measure to investigate tail dependence is the coefficient of extremal dependence; $\chi \in [0;1]$ is the tendency for one variable to be large given that the other is large. If χ=0 the two variables are asymptotically independent. Suppose that F_X and F_Y are the marginal distribution functions of X and Y respectively (Coles, 2001), and define

$$\chi = \lim_{u \to 1} Pr\{F_Y(Y) > u | F_X(X) > u\}, \tag{6.1}$$

   For 0 <u<1,

$$\begin{aligned} \chi(u) &= 2 - \frac{\log Pr\{F_X(X) < u, F_Y(Y) < u\}}{\log Pr\{F_X(X) < u\}} \\ &= 2 - \frac{\log Pr\{F_X(X) < u, F_Y(Y) < u\}}{\log u} \end{aligned} \tag{6.2}$$

   and

$$\chi = \lim_{u \to 1} \chi(u) \tag{6.3}$$

   χ provides a measure of extremal dependence within the class of asymptotically dependent distributions, where χ increases with strength of dependence at extreme levels (Coles, 2001). For asymptotically independent variables, $\chi(u) = 0$ for all u in (0,1). For perfectly dependent variables, $\chi(u) = 1$ for all u in (0,1).

   χ fails to provide a measure for asymptotically independent distributions (Coles, 2001) and thus a second measure $\bar{\chi}$ is needed. For 0 <u<1,

$$\begin{aligned} \bar{\chi}(u) &= \frac{2 \log Pr\{F_X(X) > u\}}{\log Pr\{F_X(X) > u, F_Y(Y) > u\}} - 1 \\ &= \frac{2 \log(1-u)}{\log Pr\{F_X(X) > u, F_Y(Y) > u\}} - 1 \end{aligned} \tag{6.4}$$

   and

$$\bar{\chi} = \lim_{u \to 1} \bar{\chi}(u) \tag{6.5}$$

   $\bar{\chi}$ provides a measure of extremal dependence within the class of asymptotically independent distributions, where $\bar{\chi}$ increases with strength of dependence at extreme levels (Coles, 2001). For asymptotically dependent variables $\bar{\chi}$ = 1, and for independent variables $\bar{\chi}$ = 0.

   These two measures together provide a summary of extremal dependence as summerized in Table 6.1.

Table 6.1: Characteristics of extremal dependence

Measure	Class	=0	(0,1)	=1
χ	asymptotically dependent variables	variables are asymptotically independent	behaviour of χ and $\bar{\chi}$ to be investigated as $u \to 1$	variables are perfectly dependent
$\bar{\chi}$	asymptotically independent variables	variables are independent		variables are asymptotically dependent

Using observed proportions the empirical estimates of χ and $\bar{\chi}$ can be plotted and used for several purposes. As for the bivariate threshold model they can be used for distinguishing between asymptotic dependence and asymptotic independence, for determining suitable thresholds and for validating the choice of a particular model for V (see Equation 2.19) (Coles, 2001).

4. Threshold selection

   Threshold selection is done by using the threshold stability plot and mean excess plot along with the spectral measure plot (or bivariate threshold choice plot). These plots were already explained previously, in Chapter 5.2 and 2.2.2, respectively. These diagnostic plots used for threshold selection might point to a different range of thresholds. The principle (Zheng et al., 2018) is to select a pair of thresholds whose order is not greater than k_0, the largest k value where $H([0,1]) = 2$ (in other words the largest n-k_0 observations are taken as exceedances) and meanwhile they should be within the stable range of each marginal distribution.

5. Fitting POT models

   After determining appropriate thresholds for both variables two univarite models are fitted, followed by bivariate threshold excess modeling. As for the bivariate models several parametric families are available to model the dependence structure in between variables. A widely used parametric model is the logistic (see Equation 2.23), there are many other parametric families available in the literature such as asymmetric logistic, negative logistic, bilogistic, asymmetric negative bilogistic, negative bilogistic, Husler-Reiss. In a recent study Zheng et al. (2018) tested several bivariate extreme value distributions and concluded that the logistic distribution function performed the best since in most of the cases it yielded the lowest AIC values. Therefore a bivariate POT model using logistic distribution is fitted, but other distribution functions are also tested. Another advantage of the logistic distribution is its compact form, it has one (Equation 2.24) parameter α, if $\alpha \to 1$ it corresponds to independent variables, if $\alpha \to 0$ it corresponds to perfectly dependent variables.

6. Calculating probabilities

   The joint probability of a certain range of temporal and speed related indicators is calculated. We are primarily interested in the region $R_{1,1}$ where both indicators are above their pre-selected thresholds. A matrix of combinations for the two indicators is constructed and the joint probabilities of these events are calculated and probability based risk levels yielding the same probability are constructed.

7. Testing Copula

   An alternative to bivariate threshold excess modelling is using a copula capturing the joint behaviour, for which the appropriate family has to be selected. This can be done using the in-built $BiCopSelect$ function in R, for which pseudo observationts are used.

The suggested Copula family with its known parameters then can be fitted to actual data. The estimated and fitted parameters of the copula as well as Kendall's τ can be compared. A goodness of fit test for fitted copula can be run using the $gofCopula$ function in R. Sample data from the fitted joint distribution and actual data can be plotted together and compared.

6.2. Results

Results are presented separately for all the four cases (variable pairs) outlined at the beginning of this chapter.

6.2.1. TTC vs. relative speed

The scatter plot in Figure 6.2 shows the relationship between negated TTC_{min} and relative speed. After removing the two extreme outliers (two observation points on the left in Figure 6.2), correlation tests were run to determine how strong the relationship is. These tests were all significant and indicated a moderate correlation (Table 6.2).

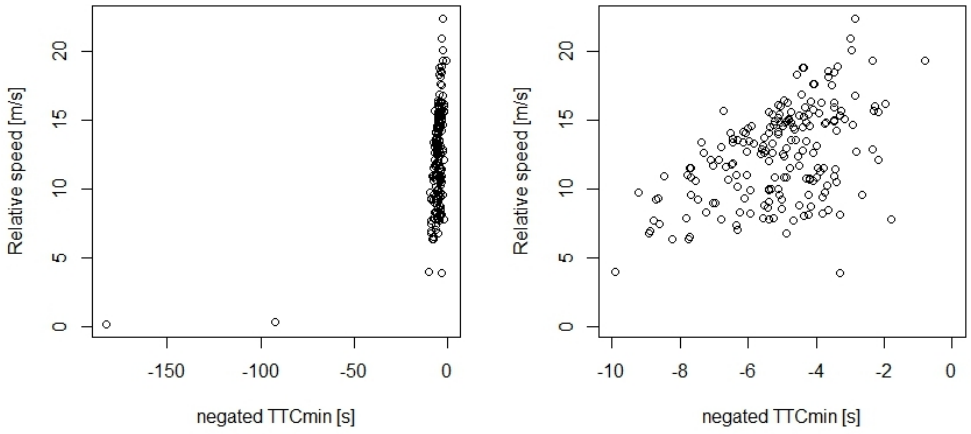


Figure 6.2: Negated TTC_{min} vs. relative speed: a) all observations; b) filtered data TTC_{min}<15s

Table 6.2: Correlation analysis between TTC_{min} and relative speed

Test	Statistic	p-value
Pearson	-0.490	*5.7e-13*
Spearman's rho	-0.453	*6.5e-11*
Kendall's tau	-0.328	*1.5e-11*

The asymptotic independence test for bivariate copula using pseudo observations (Figure 6.3) gives a significant result showing dependence (statistic 6.74, p-value 1.517431e-11).

Threshold stability plot for TTC_{min} can be found in Chapter 5 (Figures 5.12 and 5.15), the one for relative speed at TTC_{min} is in Figure 6.4. Mean excess plot for TTC_{min} can be found in Chapter 5 (Figure 5.11), the one for relative speed at TTC_{min} is Figure 6.5. Spectral measure plot is given in Figure 6.6 also indicating the value of k_0.

Threshold selection for TTC_{min} and relative speed:

- TTC_{min}: a range of -3.5s to -4s seems to be rational based on the mean excess plot and the threshold stability plot (in Chapter 5 a value of -4 s was used as a threshold value).
- Relative speed at TTC_{min}: both the mean excess plot and the threshold stability plot suggest a range of 15 m/s and 15.5 m/s.

- The threshold values associated with k_0=47 are -3.97 s for TTC_{min} and 15.01 m/s for its relative speed, which are in line with the above indicated values.

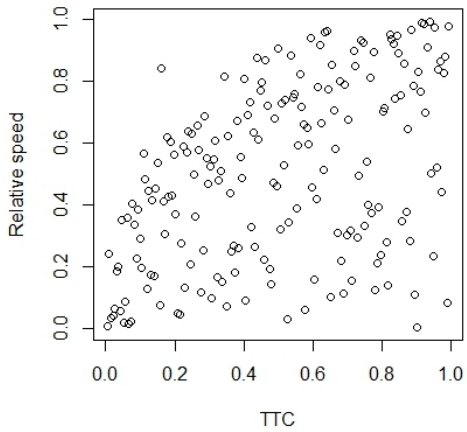


Figure 6.3: TTC_{min} vs. relative speed data transformed to uniform distribution

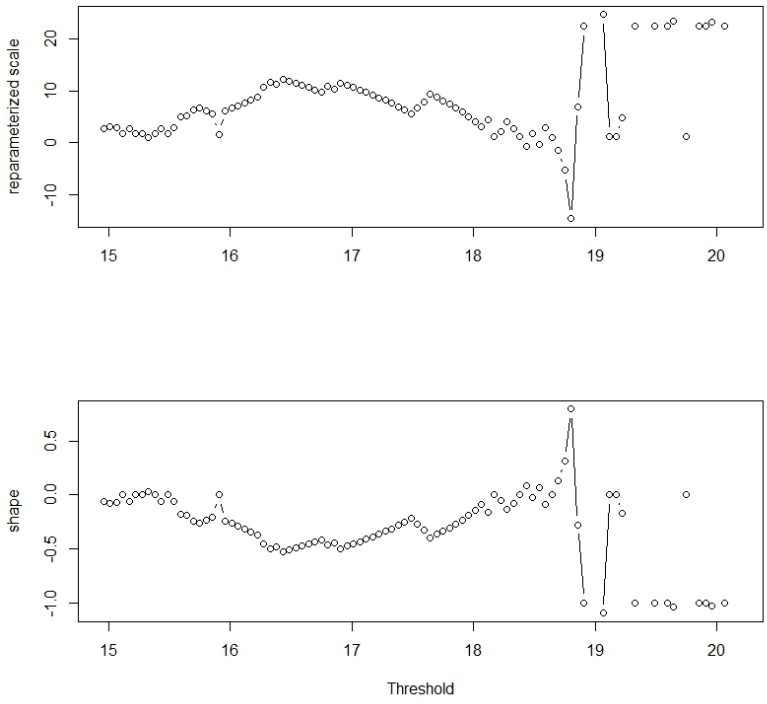


Figure 6.4: Threshold stability plot for relative speed at TTC_{min}

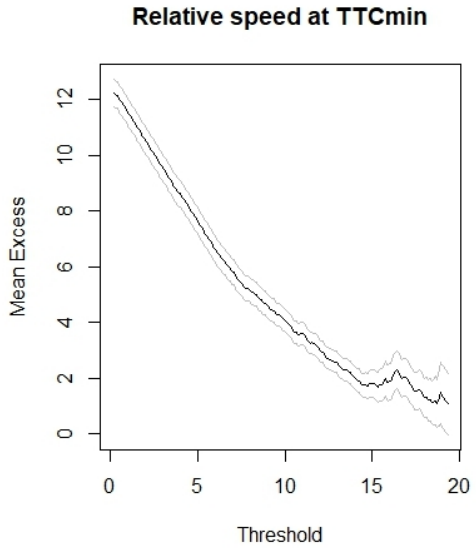


Figure 6.5: Mean excess plots for relative speed at TTC_{min}

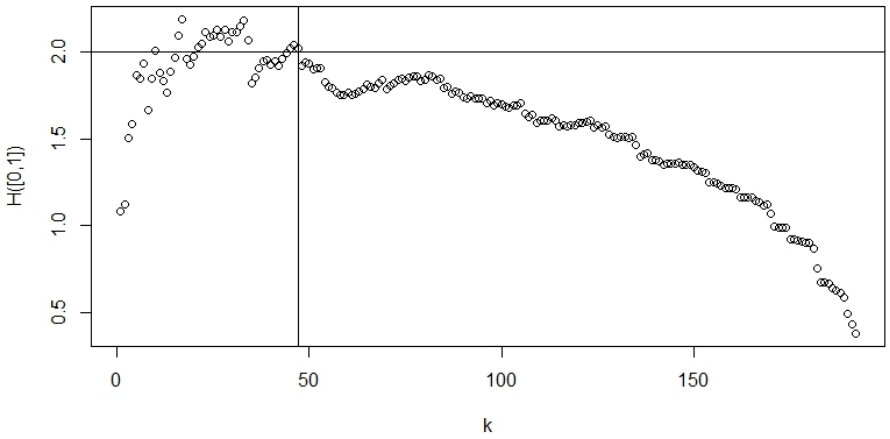


Figure 6.6: Spectral measure plot for negated TTC_{min} vs. relative speed (k_0=47)

Based on the above threshold values -4s for negated TTC_{min} and 15m/s for relative speed are used to fit the univariate POT models first (marginal number above the threshold is 47 for both indicators). These models yielded $\sigma = 0.989$ and $\xi = -0.208$ for negated TTC_{min} and $\sigma = 1.879$ and $\xi = -0.058$ for relative speed.

As for the bivariate models margins are modelled using a GPD for points above the threshold and an empirical model for those below. In the censored likelihood method the number of points lying below both thresholds are used, but the locations of those points are not (with the $fbvpot$ command from the evd package (R Core Team, 2013)).

The fitted bivariate threshold excess model (joint number above the threshold is 23) with logistic parametric distribution function using the censored maximum likelihood estimation method yielded $\sigma = 0.941$ and $\xi = -0.128$ for negated TTC_{min} and $\sigma = 1.866$ and $\xi = 0.008$ for relative speed, with an α value of 0.746 indicating a weak dependence between the two variables. Figure 6.7 shows the resulting diagnostic plots. The spectral density appears to be flat with peaks at zero and one. These two sides of the plot correspond to cases in which only one component is extreme, whereas the middle of the plot at 0.5 corresponds

to data for which both indicators are extreme. This weak dependence is also suggested by the dependence function, which takes the value of 1 at complete independence and 0.5 at complete dependence. In this plot both the parametric and non-parametric functions are plotted, which are actually quite close proving a good fit. Suppose there was a strong dependence in between the two variables, then the spectral density would show a peaked curve and the convex dependence function would get closer to the bottom of the triangular region, at complete dependence $A(1/2) = 0.5$.

Lines in the density and quantile curves plots at extreme levels group the combinations of the two indicators having the same risk level. The density plot is associated with the probability density function, whereas the quantile curves plot is associated with the cumulative distribution function. Both plots show the contour lines of either the bivariate PDF or CDF, respectively.

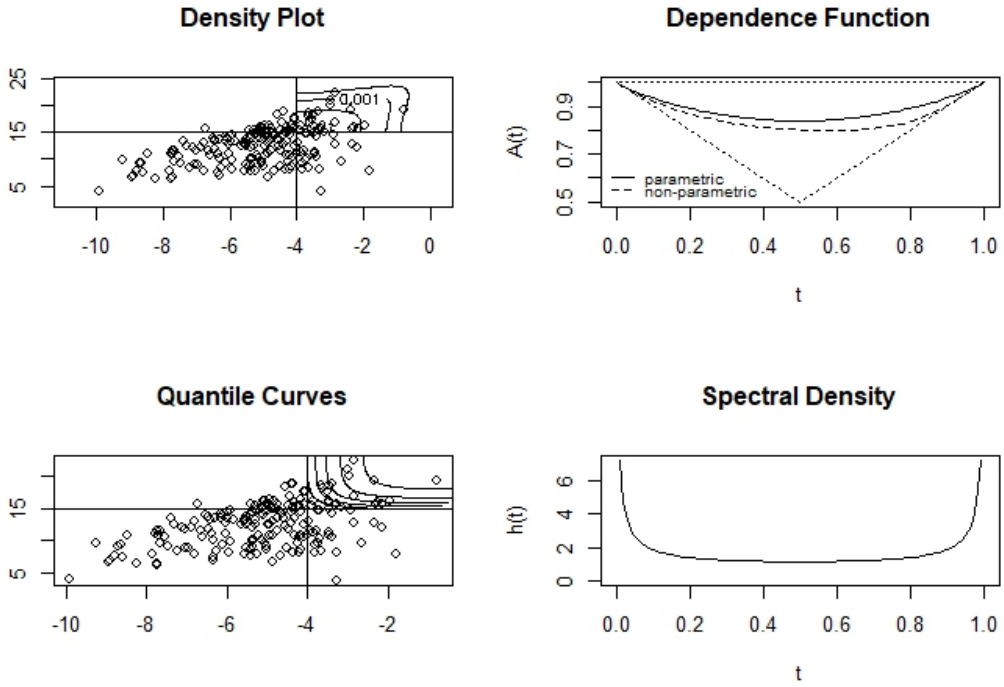


Figure 6.7: Diagnostic plots for bivariate GPD for negated TTC_{min} vs. relative speed

Bivariate threshold excess models using other parametric distributions were fitted and compared by their Akaike Information Criterion (AIC), where the lower this value the better the model is. Results in Table 6.3 in an increasing order show that there are marginal differences in between the various parametric distribution functions.

Table 6.3: Comparison of parametric bivariate extreme value distributions for TTC_{min} vs. relative speed

Name	**AIC**
Husler-Reiss	632.0
Negative logistic	632.9
Negative bilogistic	633.2
Bilogistic	633.5
Logistic	634.6
Asymmetric negative logistic	638.2
Asymmetric logistic	639.1

Using the results of the bivariate threshold excess model with logistic distribution the joint probabilities of events for a range of values (for TTC_{min} between 0 and 4s with an increment of 0.1s, for relative speed between 15 and 25 m/s with an increment of 0.5 m/s) are calculated using Equation 2.23. The resulting matrix is plotted in Figure 6.8. These joint probabilities actually show the probability of exceeding a given pair of values conditional on exceeding the threshold values 4s and 15m/s (TTC<x|TTC<4s and RS>y|RS>15m/s). As we take higher values closer to 4s for TTC_{min} and smaller values closer to 15m/s for relative speed, the joint probability approaches to 1. Taking the same risk levels the probability based risk levels can be plotted in a two-dimensional plane (Figure 6.9).

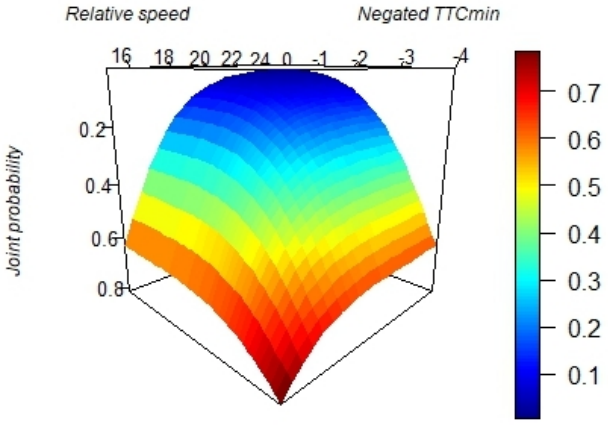


Figure 6.8: Cumulative distribution function of the fitted bivariate GPD for negated TTC_{min} vs. relative speed

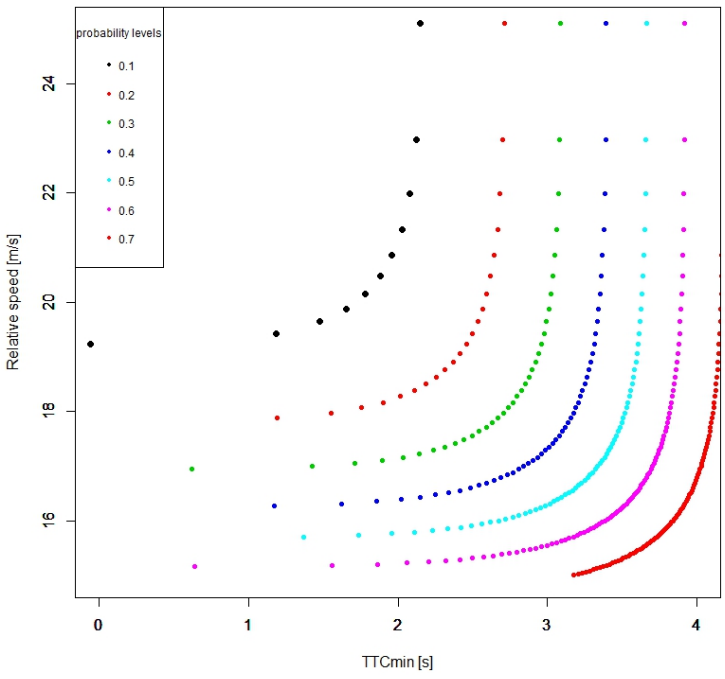


Figure 6.9: Probability based risk levels of the fitted bivariate GPD for negated TTC_{min} vs. relative speed

To further investigate the dependence between the two variables χ and $\bar{\chi}$ plots are used (Figure 6.10). In the figure dotted lines correspond to 95% confidence intervals. Neither of

the plots reveals strong dependence and based on these plots it is hard to explicitly state that the two variables show asymptotic dependence at extreme levels as u→1. A closer look at the region $u > 0.8$ (Figure 6.11) leads to the same conclusion.

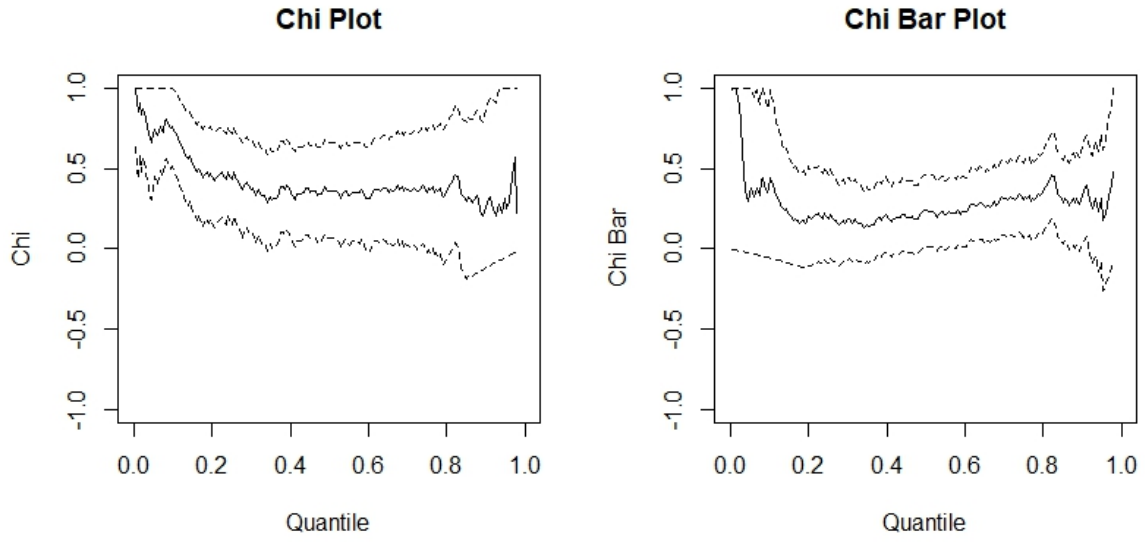


Figure 6.10: Empirical estimates of $\chi(u)$ and $\bar{\chi}(u)$ for TTC_{min} - relative speed data

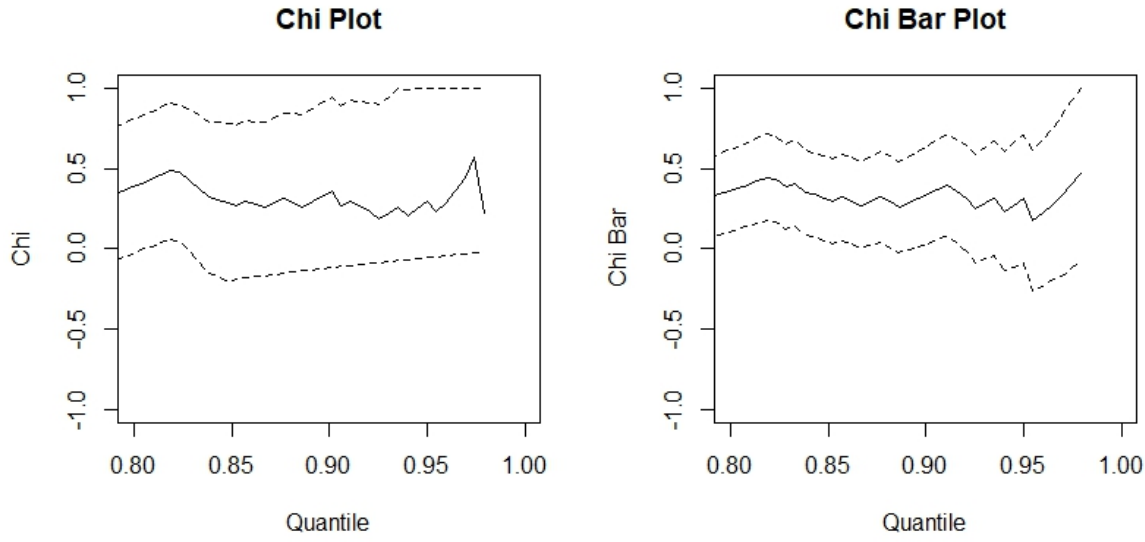


Figure 6.11: Empirical estimates of $\chi(u)$ and $\bar{\chi}(u)$ $u > 0.8$ for TTC_{min} - relative speed data

Previously dependence and marginal distributions were modeled simultaneously with bivariate threshold excess models using parametric distribution functions. Another way is to model the dependence and marginal distributions separately, where a fitted parametric copula describes the dependence and univariate POT models the marginal distributions. These two then are incorporated into one bivariate model.Thus the next step is to build the bivariate distribution using a parametric copula as well as the previously fitted univariate POT models as marginal distributions.

The appropriate family of copula was chosen using the in-built $BiCopSelect$ function in R. For given bivariate copula data this command selects an appropriate bivariate copula family from 40 copulas. In this particular case the selected copula was the Rotated Tawn type 2 180 degrees (Family: 214) with parameters 2.65 and 0.39. This copula was then fitted to actual data resulting in parameters 2.648 and 0.386, which are exactly the same as the ones estimated before. The Kendall's tau for this particular copula using the above parameters is 0.30, which is in line with the value indicated in Table 6.2. As a goodness of fit the empirical copula with a parametric estimate of the copula can be compared. Approximate p-values are obtained using parametric bootstrap. The test resulted in p-values higher than 0.05, thus we cannot reject the null hypothesis (the resulting sample is from the copula), that is, the fitting process was successful.

Random sample observations can be generated from this bivariate distribution which then can be compared to observed data (Figure 6.12). This plot offers a possibility to visually check whether simulated data are in line with observed data. This is not quite straightforward here as we analyze the extreme region and have only a few observations. In this particular case, however, simulated data seem to describe observed values well.

Furthermore, probability density (Figure 6.13) as well as cumulative distribution functions (Figure 6.14) along with their contour lines can be plotted. The PDF flattens rapidly as both indicators approach their extreme values (i.e. TTC_{min} converges to zero and relative speed to 20 m/s). The CDF also proves that the dependence in between the two variables is weak; keeping one indicator constant (e.g. relative speed at 15.5 m/s) and by changing the value of TTC_{min} does not really affect the probability (contour line at approximately 0.1). The contour lines in Figure 6.14 can actually be interpreted the same way as in Figure 6.9, where the probability based risk levels were constructed using the bivariate threshold excess model results. In Figure 6.14 however the negated TTC values are shown and the probabilities refer to the area below the contour lines (not the area above as in Figure 6.9).

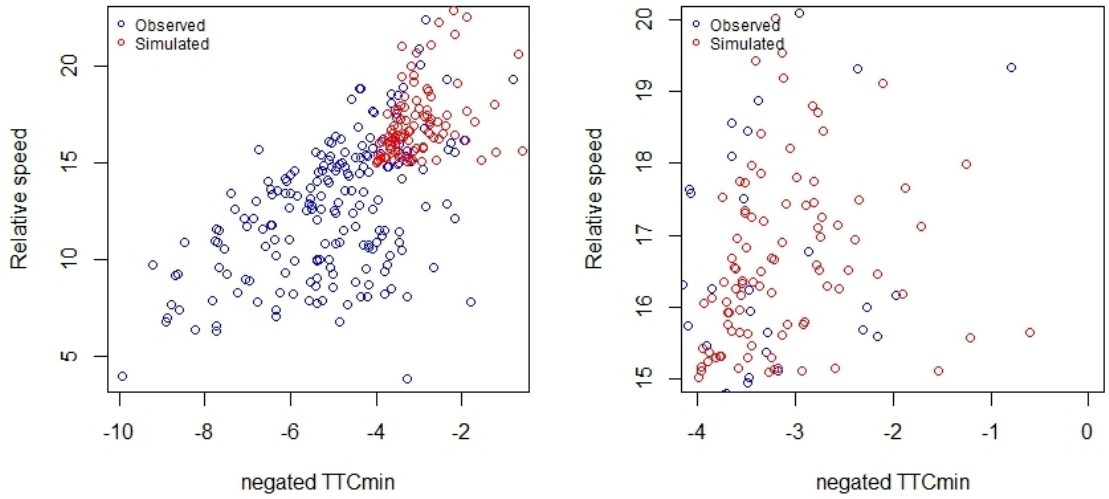


Figure 6.12: Testing the copula with a sample of 100 against actual data for TTC_{min} vs. relative speed

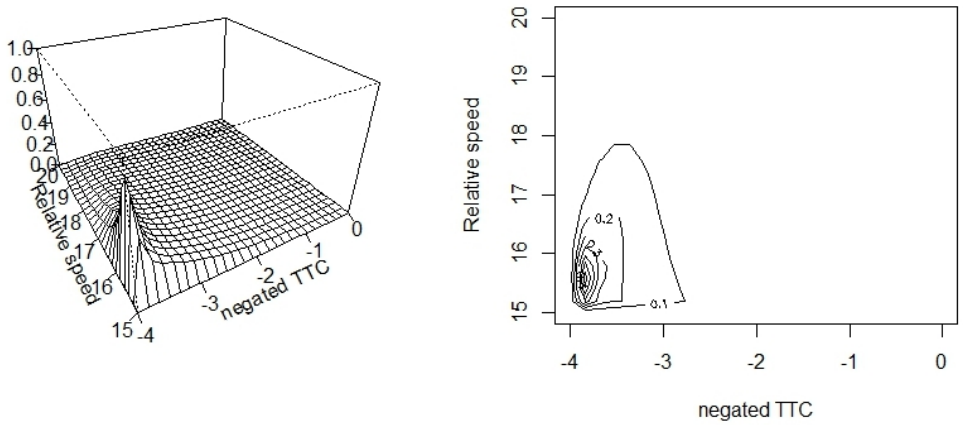


Figure 6.13: Probability density function (left) and its contour plot (right) of the fitted copula for TTC_{min} vs. relative speed

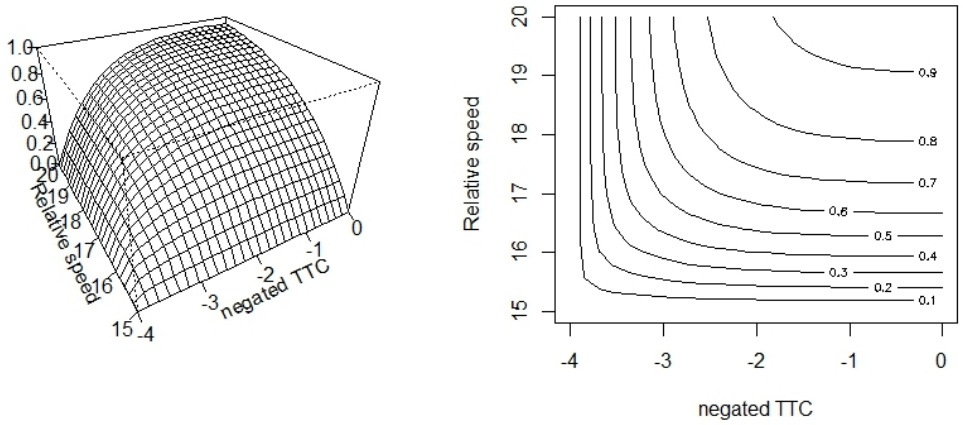


Figure 6.14: Cumulative distribution function (left) and its contour plot (right) of the fitted copula for TTC_{min} vs. relative speed

6.2.2. T2 vs. relative speed

The scatter plot in Figure 6.15 shows the relationship between negated T_{2min} and relative speed. After removing the three extreme outliers (observation points on the left in Figure 6.15), correlation tests (Table 6.4) were all significant and indicated a very weak correlation.

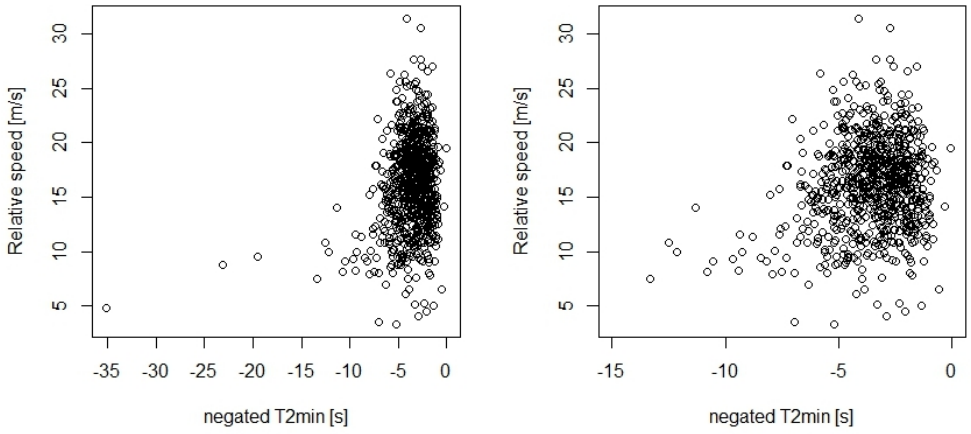


Figure 6.15: Negated T_{2min} vs. relative speed: a) all observations; b) filtered data T_{2min}<15s

Table 6.4: Correlation analysis between T_{2min} and relative speed

Test	Statistic	p-value
Pearson	-0.183	*2.0e-07*
Spearman's rho	-0.099	*0.005*
Kendall's tau	-0.065	*0.006*

The asymptotic independence test for bivariate copula using pseudo observations (Figure 6.16) gives a significant result showing dependence (statistic 2.73, p-value 0.00616).

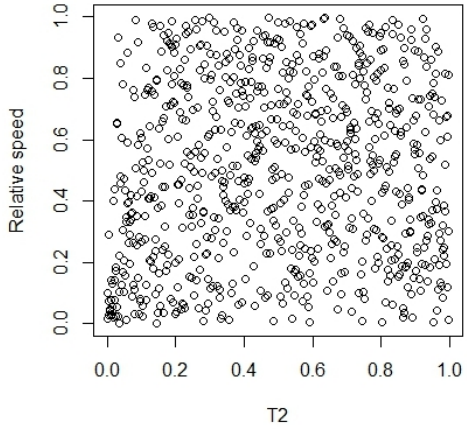


Figure 6.16: $T2_{min}$ vs. relative speed data transformed to uniform distribution

Threshold stability plot for T_{2min} can be found in Chapter 5 (Figures 5.17 and 5.20), the one for relative speed at T_{2min} is in Figure 6.17. Mean excess plot for T_{2min} can be found in Chapter 5 (Figure 5.16), the one for relative speed at T_{2min} is in Figure 6.18. Spectral measure plot is given in Figure 6.19 indicating the value of k_0.

Threshold selection for T_{2min} and relative speed:

- T_{2min}: a range of -1.9s to -2.2s seems to be rational based on the mean excess plot and the threshold stability plot (in Chapter 5 a value of -2 s was used as a threshold value).
- Relative speed at T_{2min}: the threshold stability plot suggests a stable range in between 18 and 20 m/s, which is line with the mean excess plot.
- The threshold values associated with k_0=441 are -3.5 s for T_{2min} and 15.8 m/s for its relative speed, both values are lower than what the previous plots suggest (resulting in more excesses).

This is one of the cases when marginal and dependence considerations suggest different thresholds. As it was suggested by Zheng et al. (2018) the intersection of these can be chosen taking threshold values smaller than k_0 in the appropriate range suggested by the mean excess plot and threshold stability plot. Considering these limitations a threshold value of 2s for T_{2min} and 18 m/s for relative speed are used (marginal number above threshold is 129 and 259, respectively, joint number above is 38).

The simple univariate models yielded $\sigma = 0.580$ and $\xi = -0.242$ for negated T_{2min} and $\sigma = 2.916$ and $\xi = -0.122$ for relative speed. The fitted bivariate threshold excess model with logistic parametric distribution function yielded $\sigma = 0.572$ and $\xi = -0.231$ for negated T_{2min} and $\sigma = 2.788$ and $\xi = -0.143$ for relative speed, with an α value of 0.999 indicating perfect independence between the two variables. Diagnostic plots are presented in Figure 6.20, the

above mentioned independence can also be captured in the spectral density and dependence functions. Bivariate threshold excess models using other parametric distributions were also fitted, results in Table 6.5 show that again there are marginal differences in between them and the logistic model performs rather well.

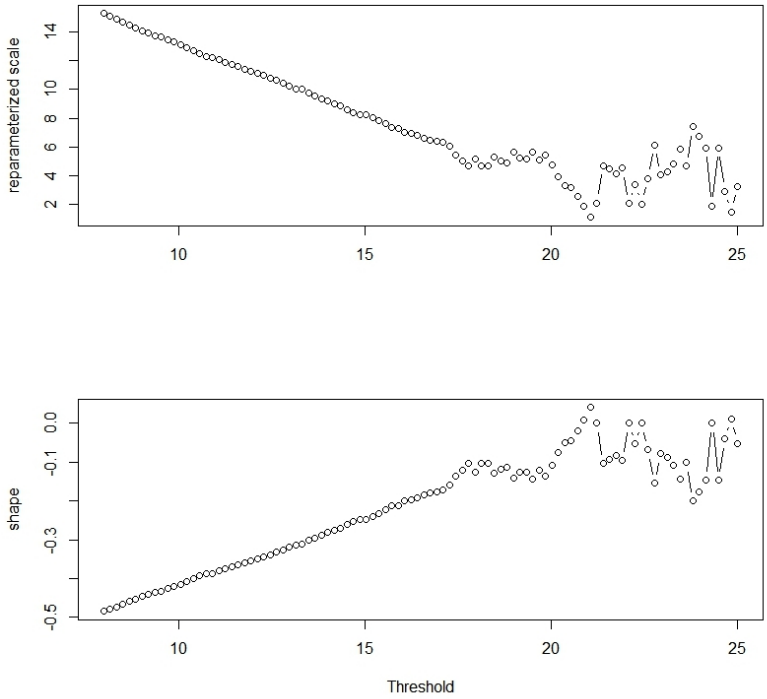


Figure 6.17: Threshold stability plot for relative speed at T_{2min}

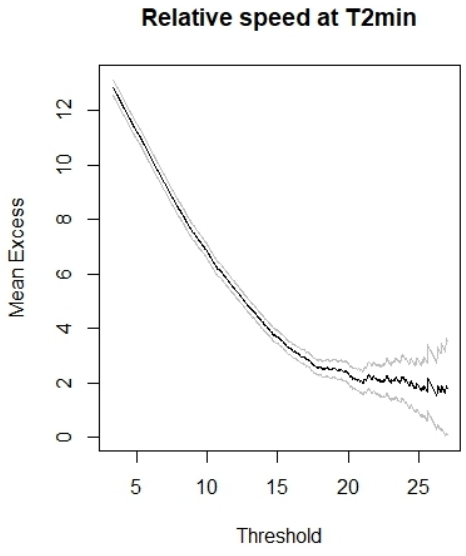


Figure 6.18: Mean excess plots for relative speed at T_{2min}

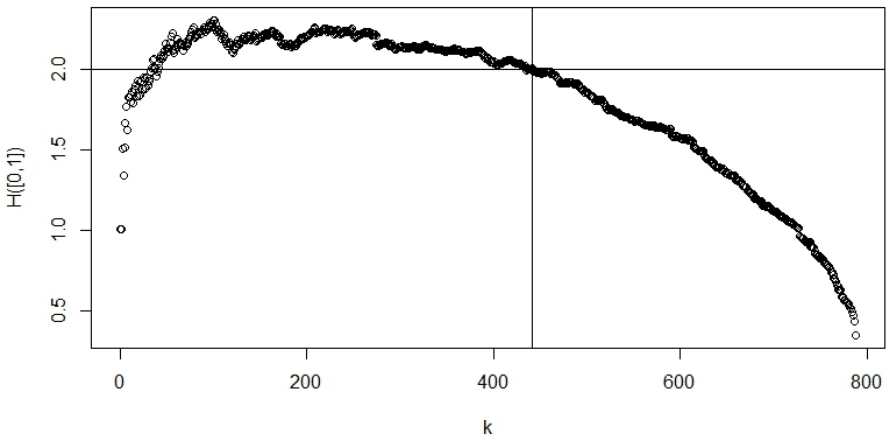


Figure 6.19: Spectral measure plot for negated T_{2min} vs. relative speed (k_0=441)

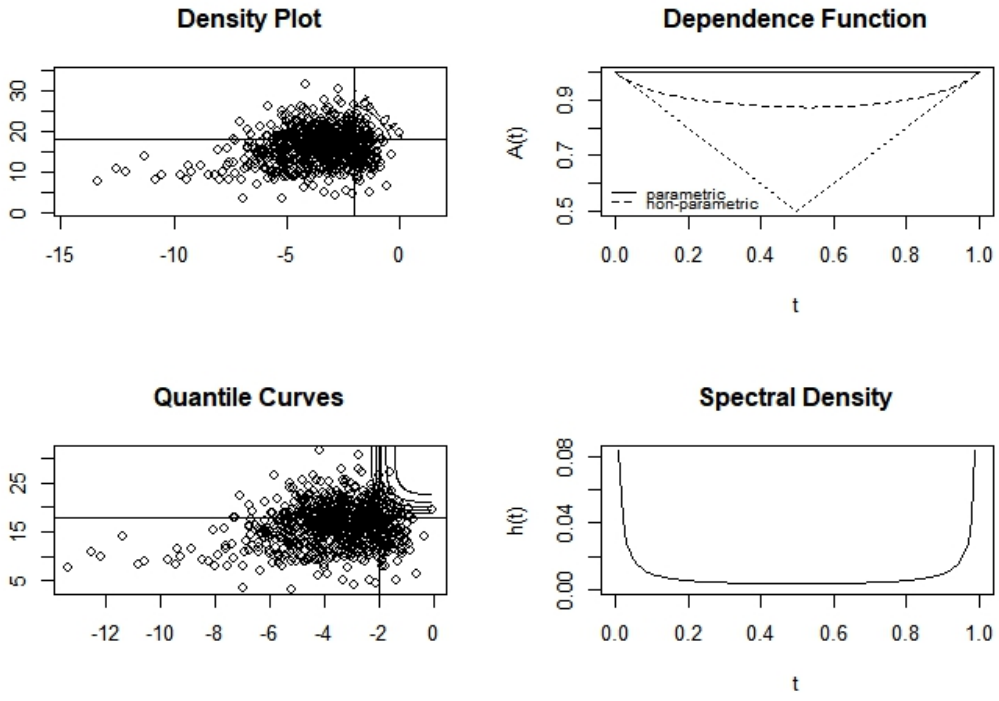


Figure 6.20: Diagnostic plots for bivariate GPD for negated TTC_{min} vs. relative speed

Table 6.5: Comparison of parametric bivariate extreme value distributions for T_{2min} vs. relative speed

Name	**AIC**
Husler-Reiss	2772.2
Negative bilogistic	2774.1
Logistic	2774.3
Bilogistic	2775.8
Negative logistic	2777.2
Asymmetric negative logistic	2779.1
Asymmetric logistic	2783.7

Using the results of the bivariate threshold excess model with logistic distribution the joint probabilities of events for a range of values (for T_{2min} between 0 and 2s with an increment of 0.1s, for relative speed between 18 and 35 m/s with an increment of 0.5 m/s) are calculated. The resulting matrix is plotted in Figure 6.21. This plot shows the probability of exceeding a given pair of values conditional on exceeding the threshold values 2s and 18m/s (T2<x|T2<2s and RS>y|RS>18m/s). Probability based risk levels are plotted in a two-dimensional plane in Figure 6.22.

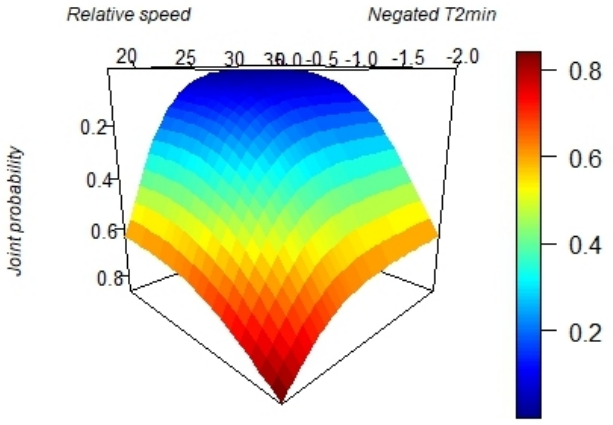


Figure 6.21: Cumulative distribution function of the fitted bivariate GPD for negated T_{2min} vs. relative speed

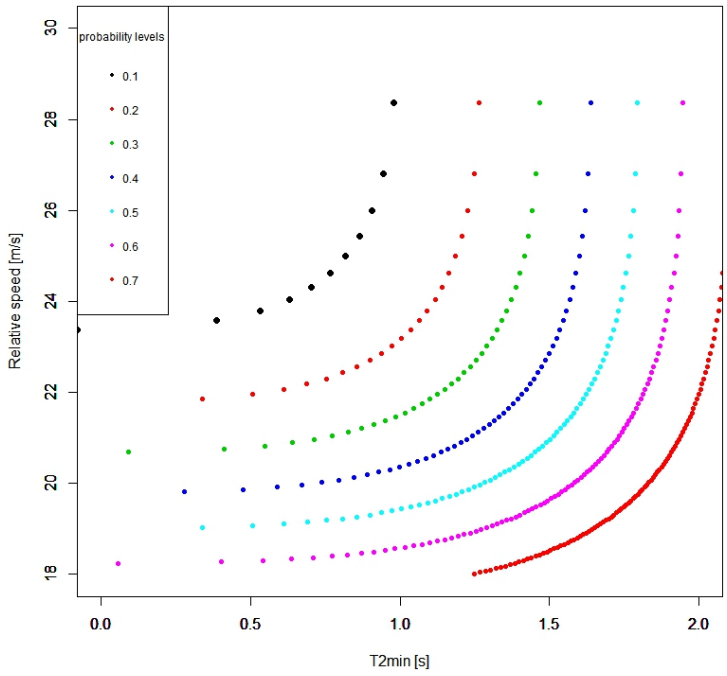


Figure 6.22: Probability based risk levels of the fitted bivariate GPD for negated T_{2min} vs. relative speed

Neither of the χ and $\bar{\chi}$ plots suggests asymptotic dependence at extreme levels as u→1 and it is more reasonable to say that they are perfectly independent (Figure 6.23 and Figure 6.24).

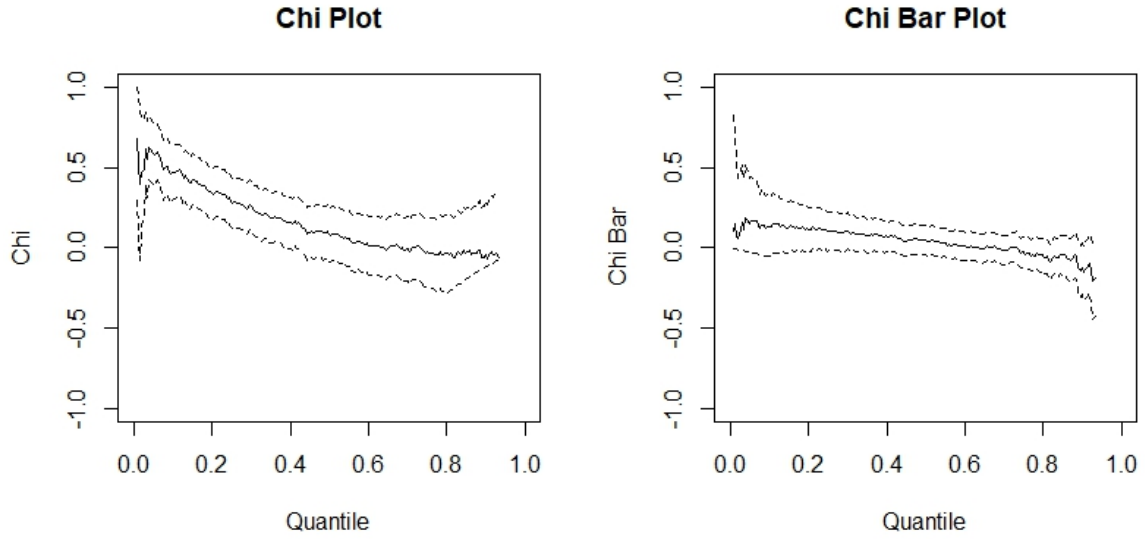


Figure 6.23: Empirical estimates of $\chi(u)$ and $\bar{\chi}(u)$ for T_{2min} - relative speed data

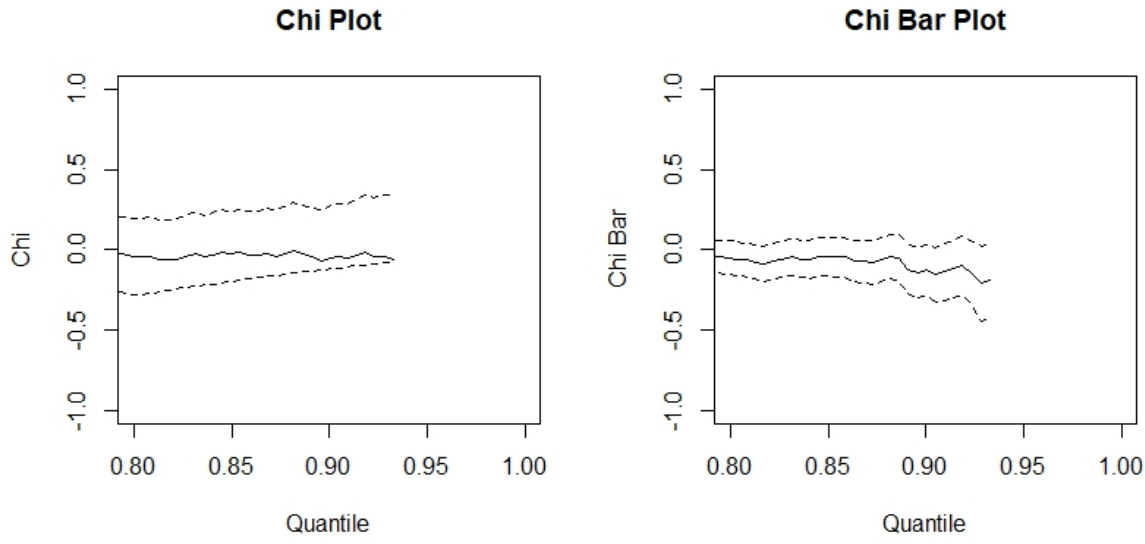


Figure 6.24: Empirical estimates of $\chi(u)$ and $\bar{\chi}(u)$ $u > 0.8$ for T_{2min} - relative speed data

The appropriate family of copula selected was the Rotated Tawn type 2 180 degrees (Family: 214) with parameters 1.65 and 0.15. This copula was then fitted to actual data resulting in parameters 1.650 and 0.145, which are exactly the same as the ones estimated before. The Kendall's tau for this particular copula using the above parameters is 0.10, which is just slightly different from the value of 0.07 indicated in Table 6.4. The goodness of fit test (comparison of the empirical copula with a parametric estimate of the copula) yielded p-values smaller than 0.05, thus we reject the null hypothesis (the resulting sample is from the copula), that is, the fitting process was not successful. Random sample observations are generated from this bivariate distribution and compared to observed data (Figure 6.25). Since the two indicators seem to be perfectly independent, there are no observations in the extreme value region, however simulated data show it otherwise.

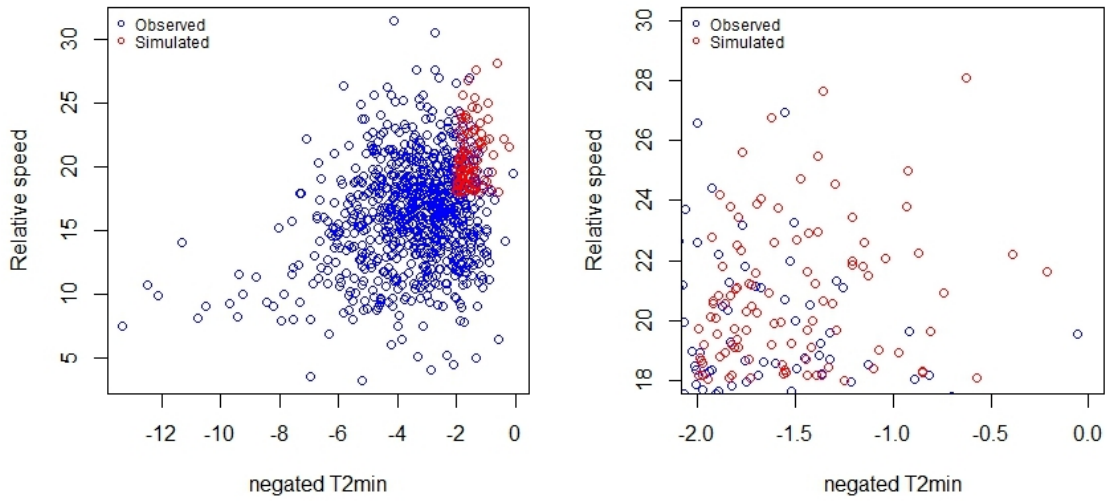


Figure 6.25: Testing the copula with a sample of 100 against actual data for T_{2min} vs. relative speed

This copula fitting actually suggests that the two variables are indeed perfectly independent at extreme levels.

6.2.3. TTC vs. Delta-V0

The scatter plot in Figure 6.26 shows the relationship between negated TTC_{min} and Delta-V0. The two extreme outliers (two observation points on the left in Figure 6.26) were again removed and correlation tests were run which were all significant and indicated a weak correlation (Table 6.6).

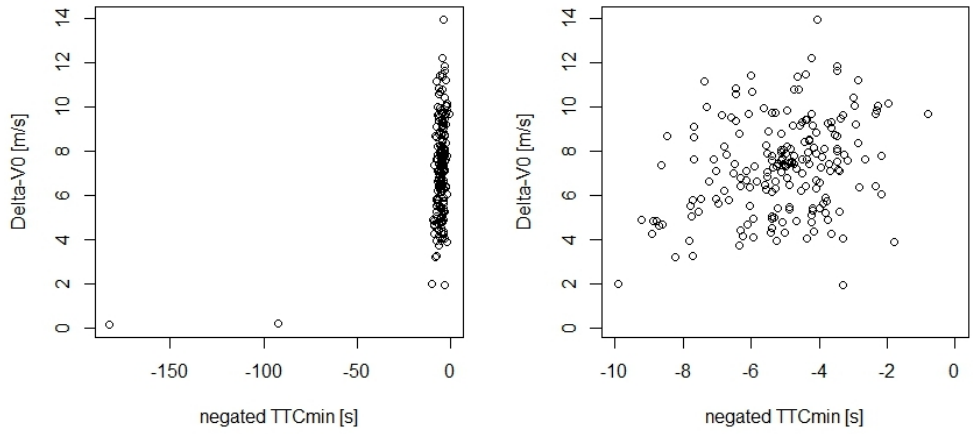


Figure 6.26: Negated TTC_{min} vs. Delta-V0: a) all observations; b) filtered data TTC_{min}<15s

Table 6.6: Correlation analysis between TTC_{min} and Delta-V0

Test	Statistic	p-value
Pearson	-0.286	*6.0e-05*
Spearman's rho	-0.255	*0.0003*
Kendall's tau	-0.183	*0.0001*

The independence test for bivariate copula using pseudo observations (Figure 6.27) gives a significant result showing dependence (statistic 3.77, p-value 0.0001).

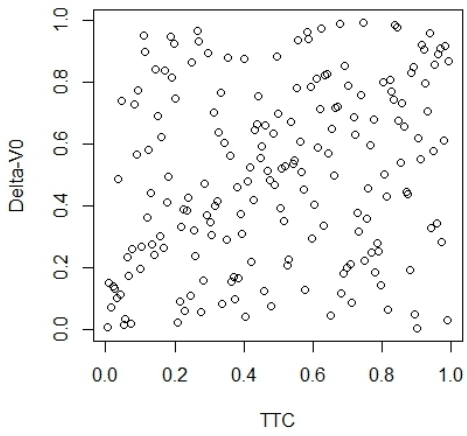


Figure 6.27: TTC_{min} vs. Delta-V0 data transformed to uniform distribution

Threshold stability plot for TTC_{min} can be found in Chapter 5 (Figures 5.12 and 5.15), the one for Delta-V0 at TTC_{min} is Figure 6.28. Mean excess plot for TTC_{min} can be found in Chapter 5 (Figure 5.11), the one for Delta-V0 at TTC_{min} is Figure 6.29. Spectral measure plot is given in Figure 6.30 also indicating the value of k_0.

Threshold selection for TTC_{min} and Delta-V0:

- TTC_{min}: as previously already noted a range of -3.5s to -4s seems to be rational, earlier a value of -4 s was used as a threshold.
- Delta-V0 at TTC_{min}: both the mean excess plot and the threshold stability plot suggest a range of 8 m/s and 10 m/s.
- The threshold values associated with k_0=57 are -4.18 s for TTC_{min} and 8.26 m/s for its Delta-V0, which are in line with the above indicated values.

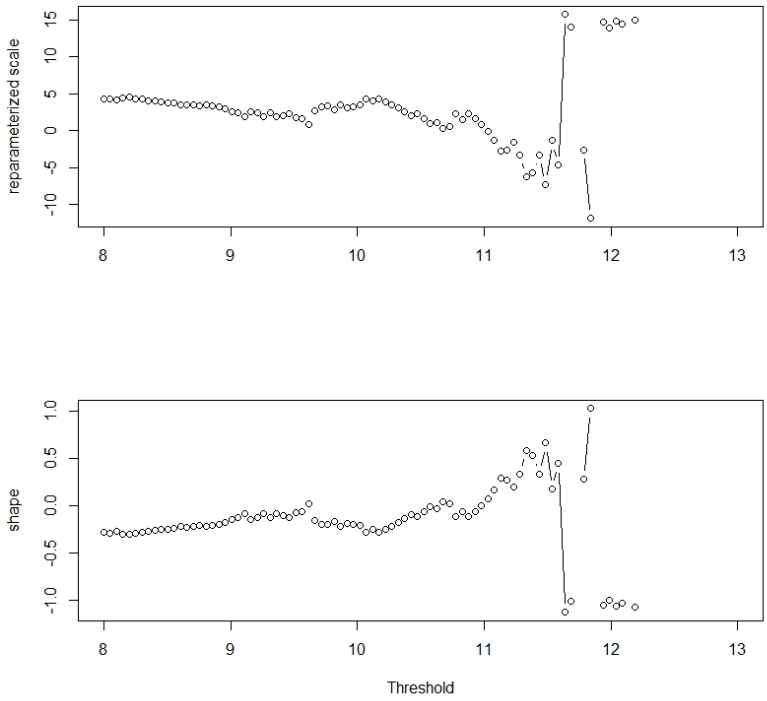


Figure 6.28: Threshold stability plot for Delta-V0 at TTC_{min}

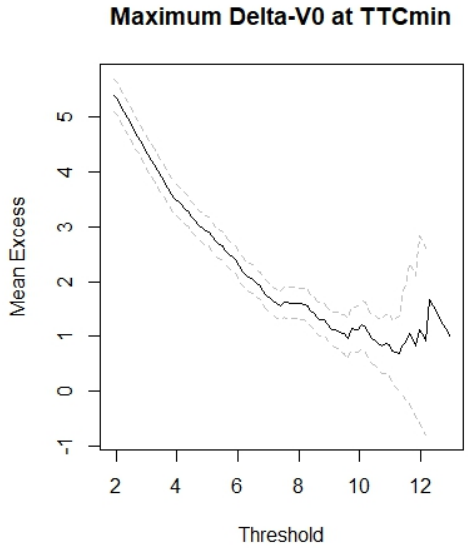


Figure 6.29: Mean excess plots for Delta-V0 at TTC_{min}

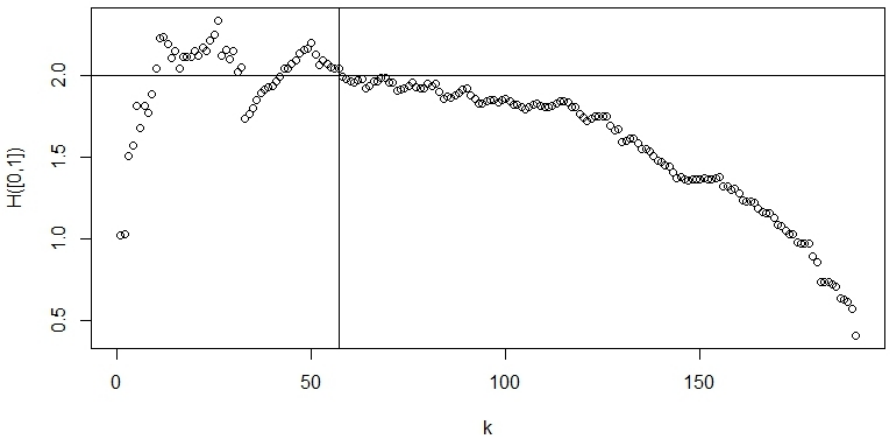


Figure 6.30: Spectral measure plot for negated TTC_{min} vs. relative speed (k_0=57)

Based on the above threshold values -4s for negated TTC_{min} and 9m/s for Delta-V0 are used to fit the univariate POT models first (marginal number above threshold is 47 and 44, respectively, joint number above is 16). These models yielded $\sigma = 0.989$ and $\xi = -0.208$ for negated TTC_{min} and $\sigma = 1.339$ and $\xi = -0.146$ for Delta-V0.

The fitted bivariate threshold excess model with logistic parametric distribution function using the censored maximum likelihood estimation method yielded $\sigma = 0.996$ and $\xi = -0.191$ for negated TTC_{min} and $\sigma = 1.313$ and $\xi = -0.108$ for Delta-V0, with an α value of 0.897 indicating a weak dependence between the two variables. Figure 6.31 shows the resulting diagnostic plots. Both the spectral density as well as the dependence functions suggest weak dependence.

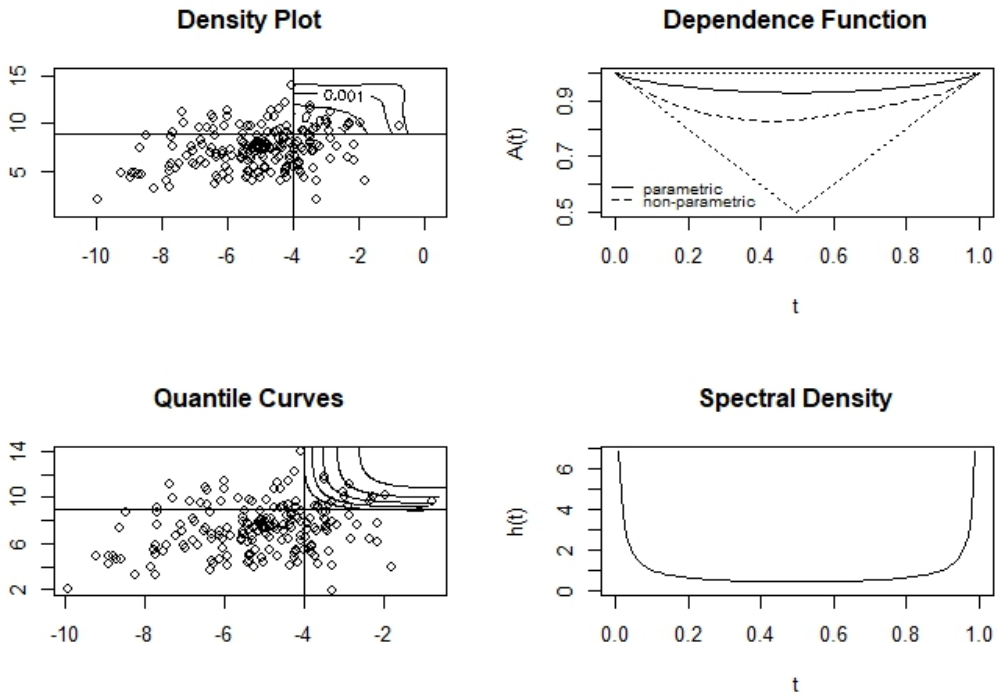


Figure 6.31: Diagnostic plots for bivariate GPD for negated TTC_{min} vs. Delta-V0

Results of bivariate threshold excess models using other parametric distributions are given in Table 6.7 show that the logistic distribution resulted in a good model and again there are only marginal differences in between the various parametric distribution functions.

Table 6.7: Comparison of parametric bivariate extreme value distributions for TTC_{min} vs. Delta-V0

Name	AIC
Negative bilogistic	599.8
Husler-Reiss	600.2
Bilogistic	600.2
Logistic	600.3
Negative logistic	600.3
Asymmetric negative logistic	602.0
Asymmetric logistic	602.7

Using the results of the bivariate threshold excess model with logistic distribution the joint probabilities of events for a range of values (for TTC_{min} between 0 and 4s with an increment of 0.1s, for Delta-V0 between 9 and 15 m/s with an increment of 0.5 m/s) are calculated using Equation 2.23. The resulting matrix is plotted in Figure 6.32. These joint probabilities show the probability of exceeding a given pair of values conditional on exceeding the threshold values 4s and 9m/s (TTC<x|TTC<4s and Delta-V0>y|Delta-V0>9m/s). As we take higher values closer to 4s for TTC_{min} and smaller values closer to 9m/s for Delta-V0, the joint probability approaches to 1. Taking the same risk levels the probability based risk levels can be plotted in a two-dimensional plane (Figure 6.33).

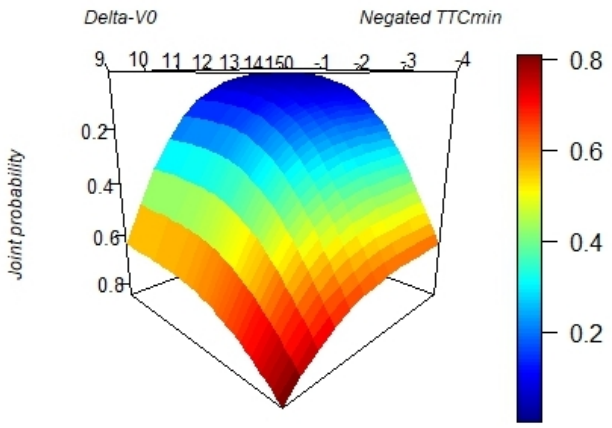


Figure 6.32: Cumulative distribution function of the fitted bivariate GPD for negated TTC_{min} vs. Delta-V0

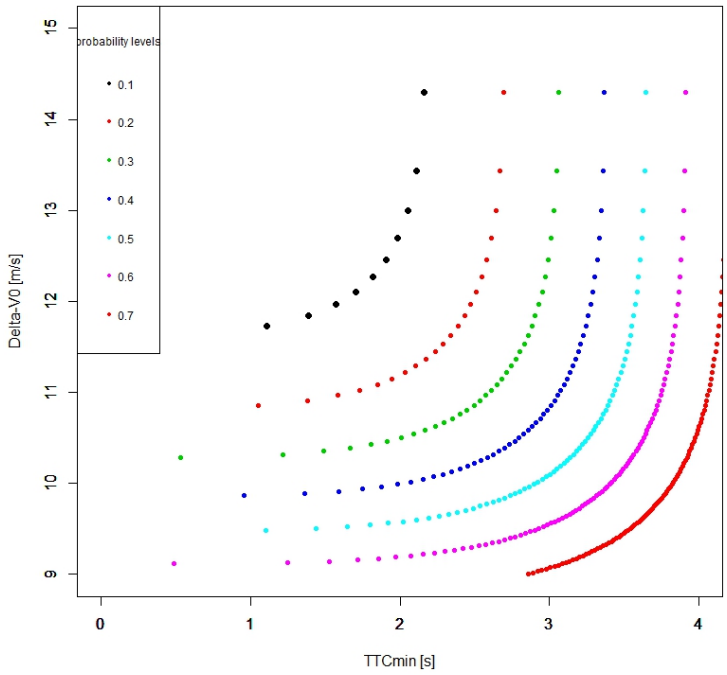


Figure 6.33: Probability based risk levels of the fitted bivariate GPD for negated TTC_{min} vs. Delta-V0

As far as the dependence between the two variables is concerned the χ and $\bar{\chi}$ plots (Figure 6.34 and Figure 6.35) show independence at extreme levels as u→1.

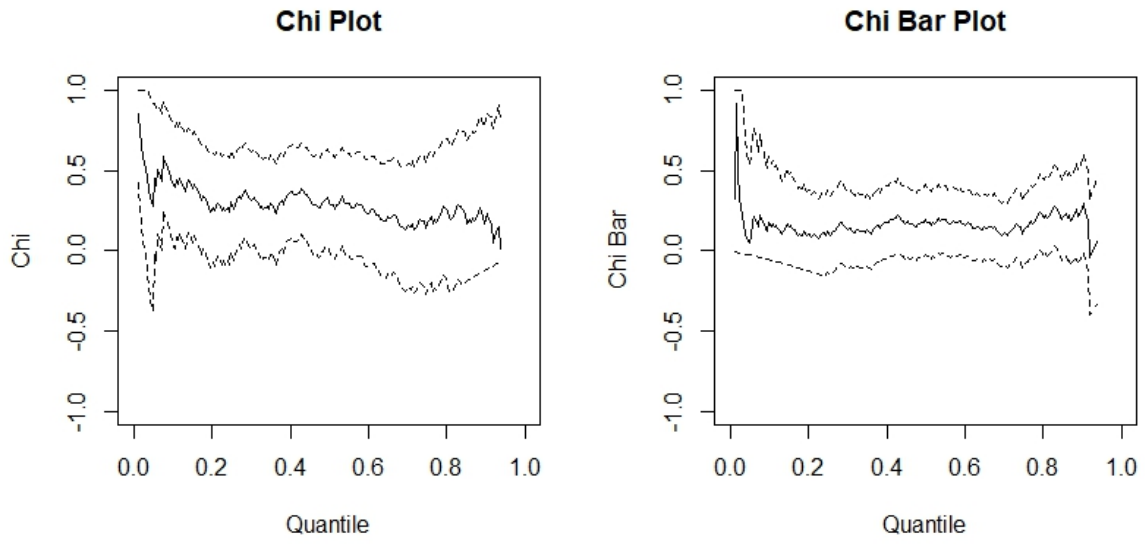


Figure 6.34: Empirical estimates of $\chi(u)$ and $\bar{\chi}(u)$ for TTC_{min} - Delta-V0 data

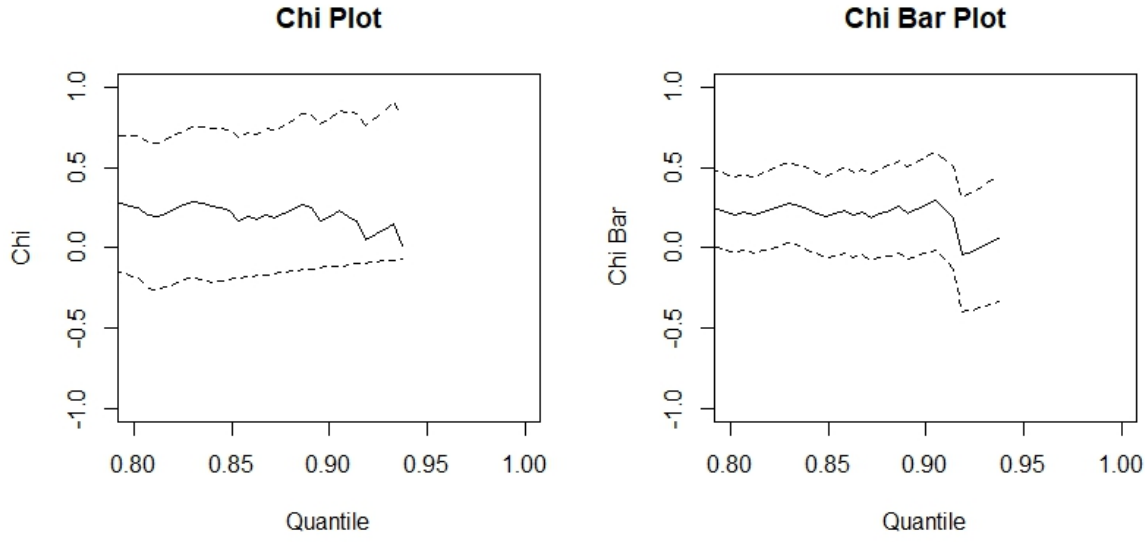


Figure 6.35: Empirical estimates of $\chi(u)$ and $\bar{\chi}(u)$ $u > 0.8$ for TTC_{min} - Delta-V0 data

The copula selected was the Survival Gumbel (Family: 14) with one parameter 1.24. After fitting this copula to actual data the resulting parameter was 1.236, exactly the same as the estimated value. The Kendall's tau (0.193) is also in line with the value indicated in Table 6.6 (0.183). Approximate p-values were higher than 0.05, thus we can conclude that the resulting sample is from the copula and the fitting process was successful.

Following the same procedure as before the bivariate distribution using the copula was built and a random sample was generated and compared to observed data (Figure 6.36). Probability density (Figure 6.37) as well as cumulative distribution functions (Figure 6.38) along with their contour lines are plotted.

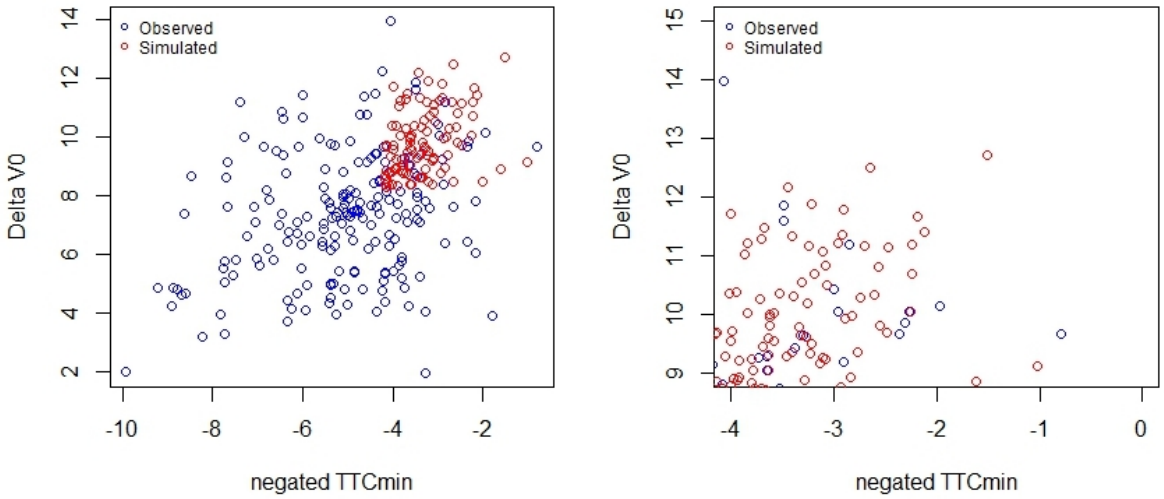


Figure 6.36: Testing the copula with a sample of 100 against actual data for TTC_{min} vs. Delta-V0

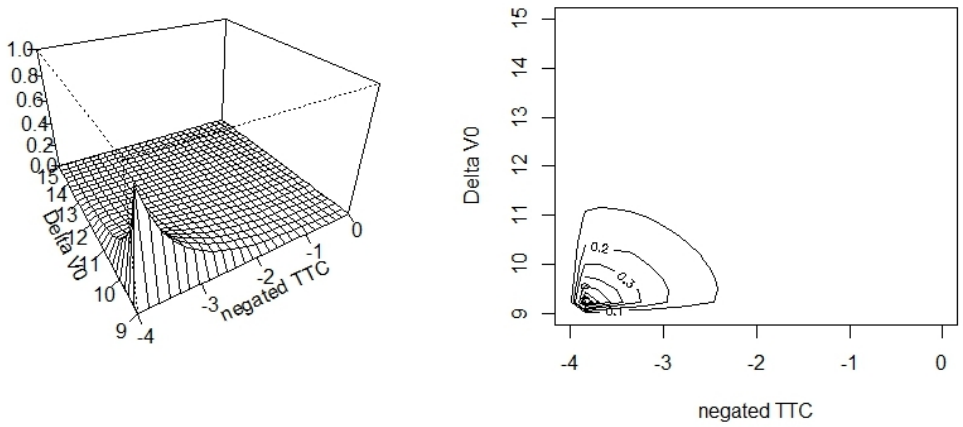


Figure 6.37: Probability density function (left) and its contour plot (right) of the fitted copula for TTC_{min} vs. Delta-V0

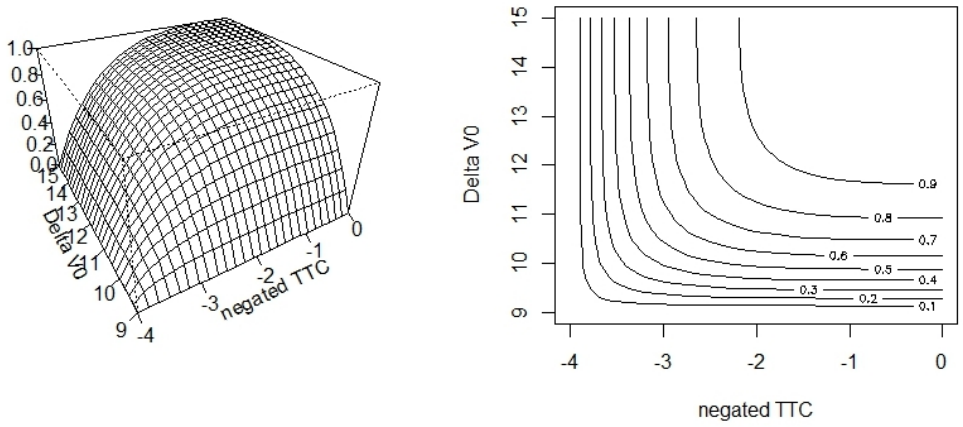


Figure 6.38: Cumulative distribution function (left) and its contour plot (right) of the fitted copula for TTC_{min} vs. Delta-V0

6.2.4. T2 vs. Delta-V0

The scatter plot in Figure 6.39 shows the relationship between negated T_{2min} and Delta-V0. Three extreme outliers were again removed and three correlation tests were run, two of which were not significant and all indicated a weak correlation (Table 6.8).

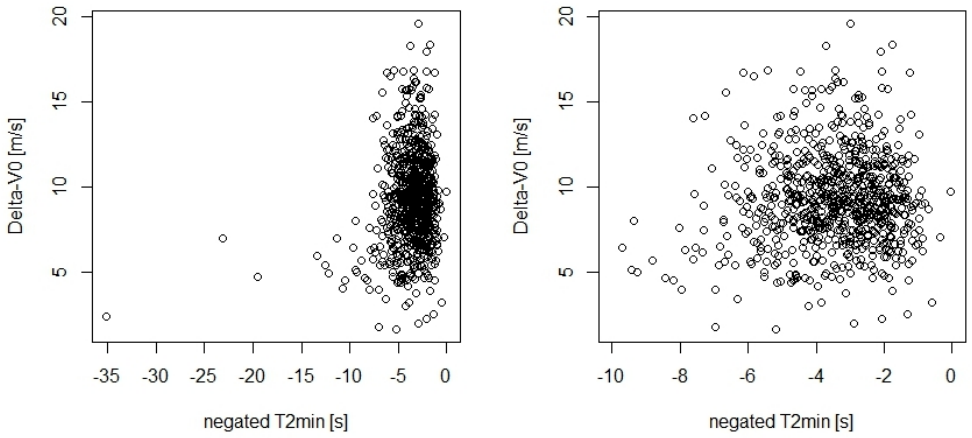


Figure 6.39: Negated T_{2min} vs. Delta-V0: a) all observations; b) filtered data T_{2min}<15s

Table 6.8: Correlation analysis between T_{2min} and Delta-V0

Test	Statistic	p-value
Pearson	-0.117	*0.0009*
Spearman's rho	-0.051	0.145
Kendall's tau	-0.033	0.165

The independence test for bivariate copula using pseudo observations (Figure 6.40) did not give a significant result (statistic 1.39, p-value 0.165).

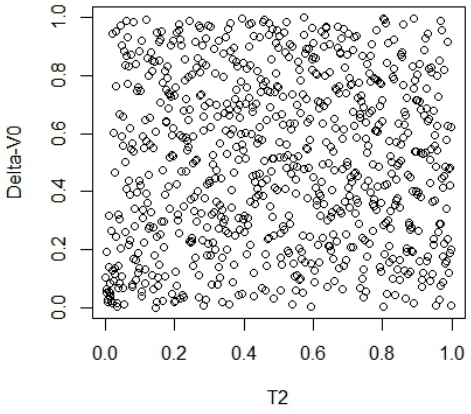


Figure 6.40: T_{2min} vs. Delta-V0 data transformed to uniform distribution

Threshold stability plot for T_{2min} can be found in Chapter 5 (Figures 5.17 and 5.20), the one for Delta-V0 at T_{2min} is in Figure 6.41. Mean excess plot for T_{2min} can be found in Chapter 5 (Figure 5.16), the one for Delta-V0 at T_{2min} is in Figure 6.42. Spectral measure plot is given in Figure 6.43 indicating the value of k_0.

Threshold selection for T_{2min} and Delta-V0:

- T_{2min}: a range of -1.9s to -2.2s seems to be rational based on the mean excess plot and the threshold stability plot (previously a value of -2 s was used).
- Delta-V0 at T_{2min}: the threshold stability and mean excess plots suggest a stable range in between 10 and 11 m/s.
- The threshold values associated with k_0=439 are -3.50 s for T_{2min} and 8.85 m/s for its Delta-V0, both values lower than the ones mentioned above.

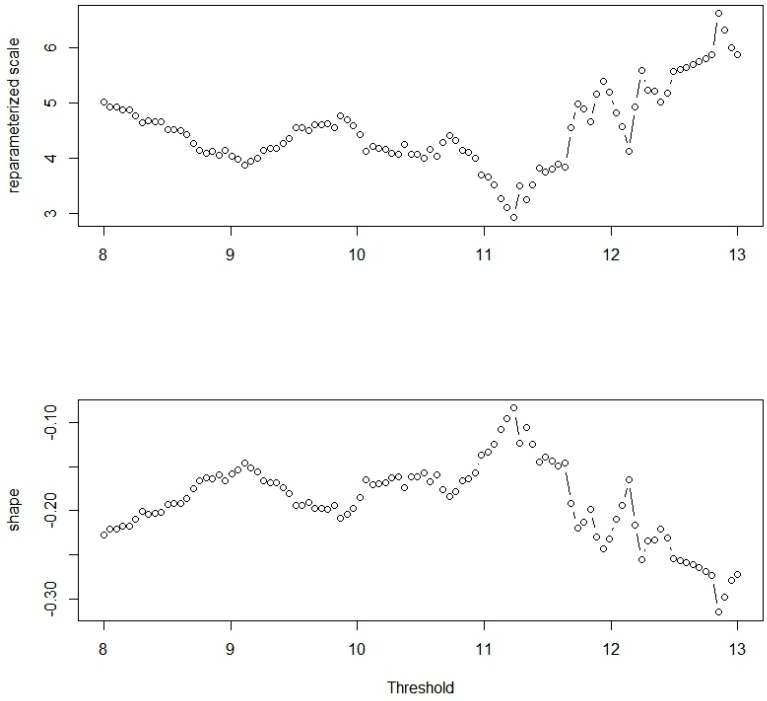


Figure 6.41: Threshold stability plot for Delta-V0 at T_{2min}

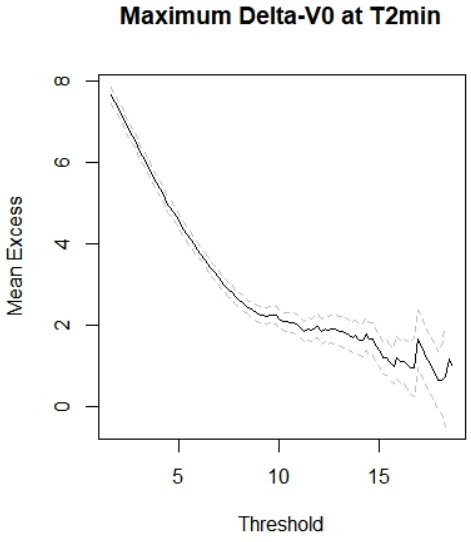


Figure 6.42: Mean excess plots for Delta-V0 at T_{2min}

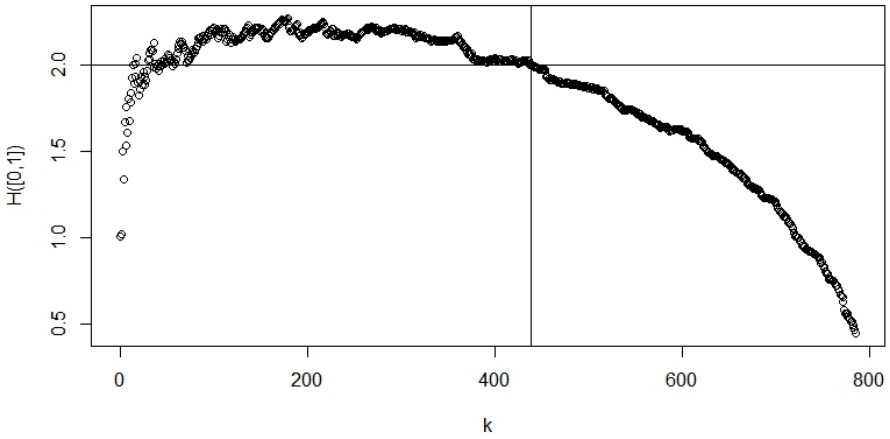


Figure 6.43: Spectral measure plot for negated T_{2min} vs. relative speed (k_0=439)

Based on the above threshold values -2 s for negated T_{2min} and 10.5 m/s for Delta-V0 are used to fit the univariate POT models first (marginal number above threshold is 129 and 232, respectively, joint number above is 30). These models yielded $\sigma = 0.580$ and $\xi = -0.242$ for negated T_{2min} and $\sigma = 2.345$ and $\xi = -0.154$ for Delta-V0.

The fitted bivariate threshold excess model with logistic parametric distribution function resulted in parameters $\sigma = 0.580$ and $\xi = -0.242$ for negated TTC_{min} and $\sigma = 2.345$ and $\xi = -0.154$ for Delta-V0, with an α value of 0.999 indicating perfect independence between the two variables, this is also supported by the diagnostic plots in Figure 6.44.

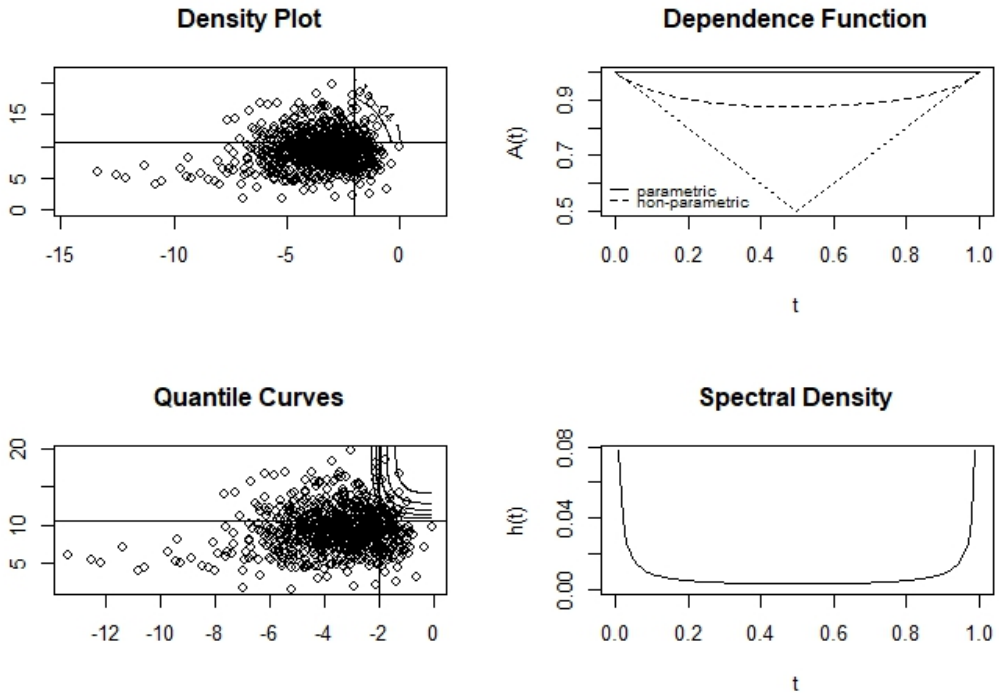


Figure 6.44: Diagnostic plots for bivariate GPD for negated T_{2min} vs. Delta-V0

Results of bivariate threshold excess models using other parametric distributions given in Table 6.9 show that the logistic distribution resulted in almost the best model.

Table 6.9: Comparison of parametric bivariate extreme value distributions for negated T_{2min} vs. Delta-V0

Name	AIC
Negative logistic	2508.5
Logistic	2508.6
Husler-Reiss	2509.4
Negative bilogistic	2510.5
Asymmetric negative logistic	2512.5
Bilogistic	2521.9
Asymmetric logistic	2525.5

Using the results of the bivariate threshold excess model with logistic distribution the joint probabilities of events for a range of values (for T_{2min} between 0 and 2s with an increment of 0.1s, for Delta-V0 between 10.5 and 20 m/s with an increment of 0.5 m/s) are calculated. The resulting matrix is plotted in Figure 6.45. These joint probabilities show the probability of exceeding a given pair of values conditional on exceeding the threshold values 4s and 9m/s (T2<x|T2<2s and Delta-V0>y|Delta-V0>10.5m/s). As we take higher values closer to 2s for T_{2min} and smaller values closer to 10.5m/s for Delta-V0, the joint probability approaches to 1. Taking the same risk levels the probability based risk levels are plotted in a two-dimensional plane (Figure 6.46).

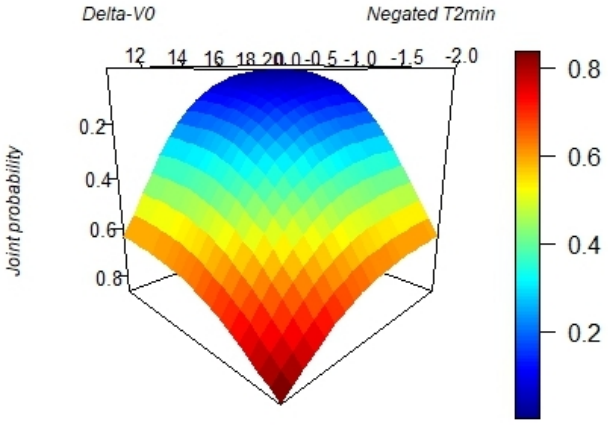


Figure 6.45: Cumulative distribution function of the fitted bivariate GPD for negated T_{2min} vs. Delta-V0

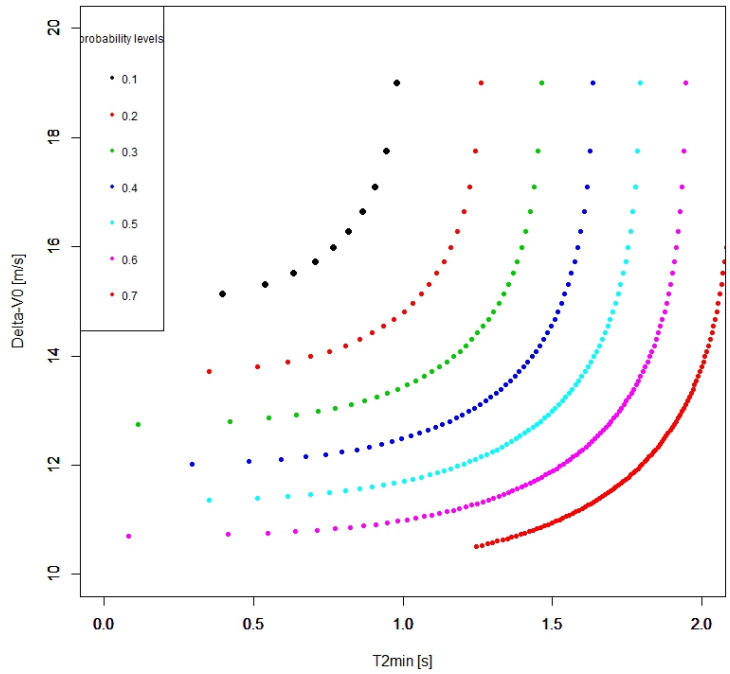


Figure 6.46: Probability based risk levels of the fitted bivariate GPD for negated T_{2min} vs. Delta-V0

As far as the dependence between the two variables is concerned the χ and $\bar{\chi}$ plots (Figure 6.47 and Figure 6.48) show independence at extreme levels as u→1.

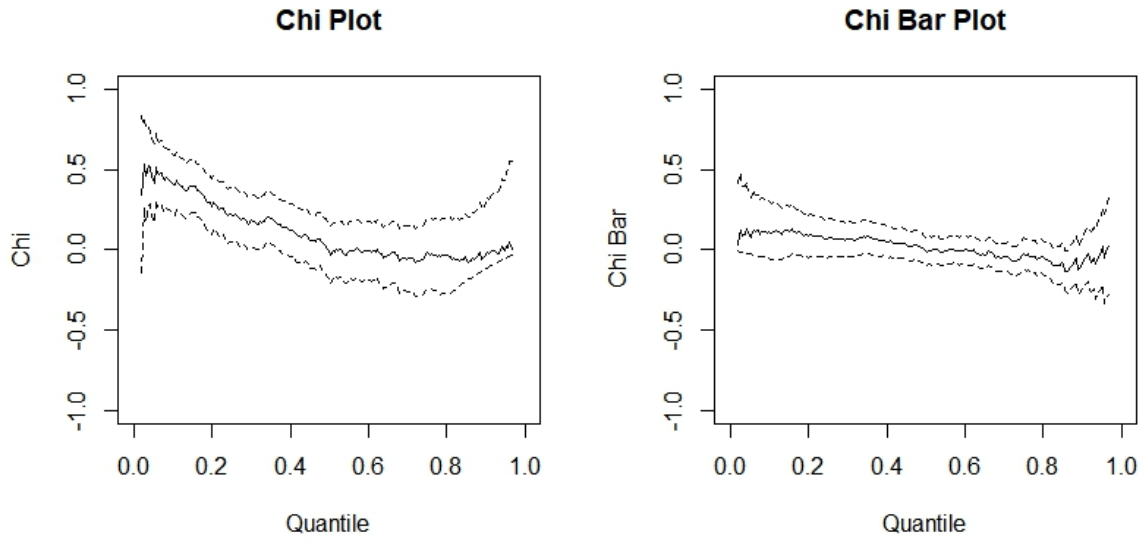


Figure 6.47: Empirical estimates of $\chi(u)$ and $\bar{\chi}(u)$ for T_{2min} - Delta-V0 data

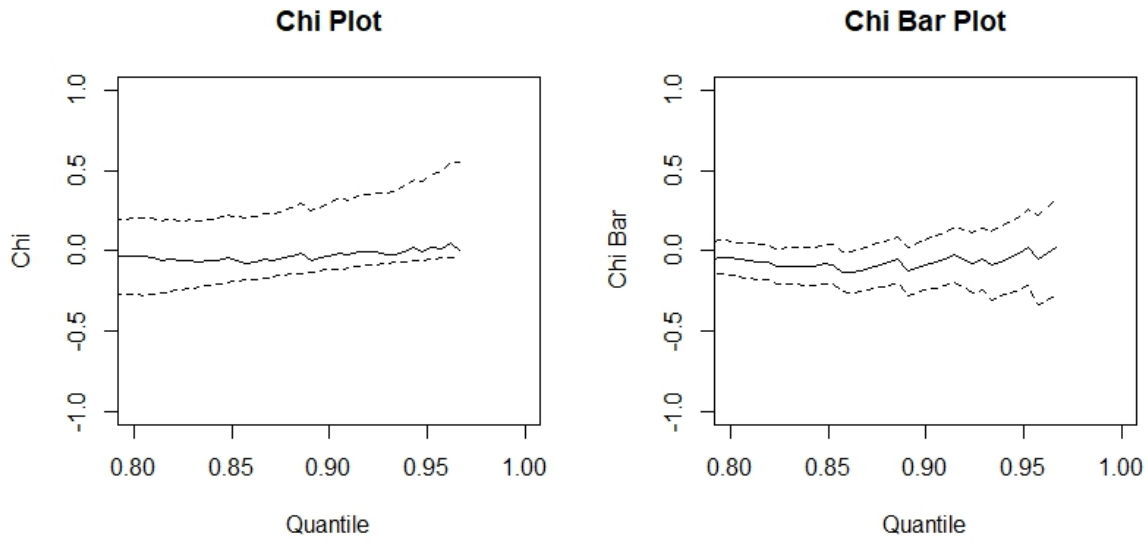


Figure 6.48: Empirical estimates of $\chi(u)$ and $\bar{\chi}(u)$ $u > 0.8$ for T_{2min} - Delta-V0 data

The appropriate family of copula selected was the Survival Joe (Family: 16) with a parameter 1.15. This copula was then fitted to actual data resulting in the parameter value exactly the same as the estimated one. The Kendall's tau for this particular copula using the above parameters is 0.08, which is just slightly different from the value of 0.03 indicated in Table 6.8. The goodness of fit test (comparison of the empirical copula with a parametric estimate of the copula) yielded p-values smaller than 0.05, thus we reject the null hypothesis (the resulting sample is from the copula), that is, the fitting process was not successful. Random sample observations generated are compared to observed data (Figure 6.49), simulated data do not seem to give a good estimation.

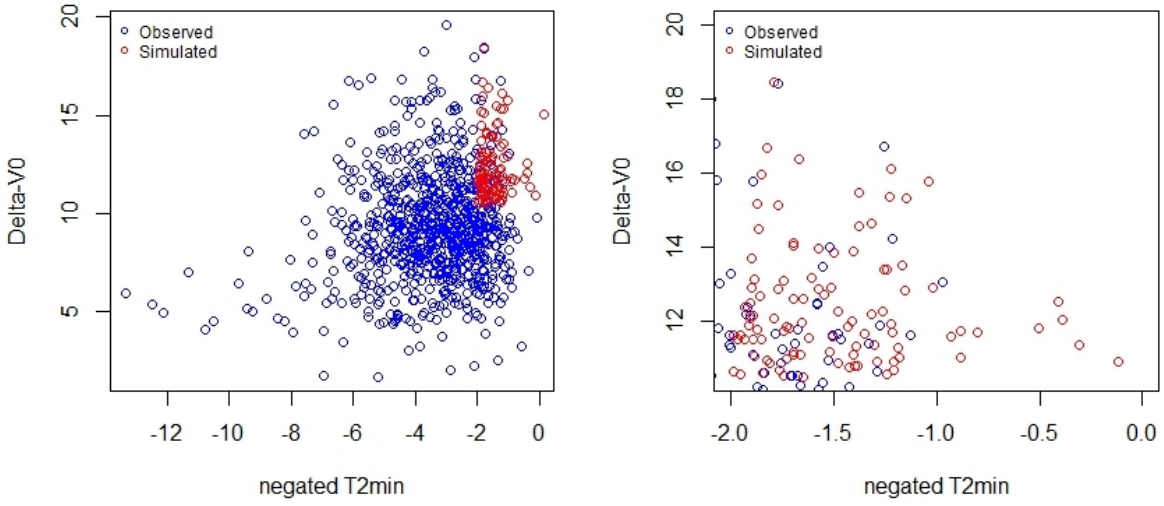


Figure 6.49: Testing the copula with a sample of 100 against actual data for T_{2min} vs. Delta-V0

6.3. Discussion

Bivariate extreme value models consist of two marginal distributions and a dependence function. The bivariate threshold excess models using a logistic distribution estimates the marginal distributions and the dependence structure in between the two variables simultaneously. The dependence parameters are between 0.746 and 0.999 showing weak dependence or perfect independence (Table 6.10). TTC_{min} and relative speed was the only pair of variables that showed the signs of asymptotic dependence. In the other three cases temporal and speed related indicators showed perfect independence at extreme levels, which means that road users getting closer to each other in time do not necessarily show high relative speed or Delta-V0. If the dependence parameter was close to 0 showing strong dependence, it would mean that the two variables are actually dependent at extreme levels. In other words if the temporal indicator in between two vehicles is very small (close to a collision) it would likely be accompanied with high relative speed or Delta-V0.

Table 6.10: Comparison of bivariate extreme value modeling

Aspect	TTC_{min} **vs. Relative speed**	$T2_{min}$ **vs. Relative speed**	TTC_{min} **vs. Delta-V0**	T_{2min} **vs. Delta-V0**
Threshold selection (based on spectral measure plot and univariate POT)	Plots indicate the same threshold values	Plots suggest different threshold values	Plots indicate the same threshold values	Plots suggest different threshold values
Selected thresholds	-4s and 15m/s	-2s and 18m/s	-4s and 9m/s	-2s and 10.5m/s
Excess values selected	Equal number of observations (47)	Unequal number of observations (129, 259)	Almost equal number of observations (47, 44)	Unequal number of observations (129, 232)
Dependence parameter of the logistic bivariate threshold excess model	Weak dependence ($\alpha = 0.746$)	Perfect independence ($\alpha = 0.999$)	Weak dependence ($\alpha = 0.897$)	Perfect independence ($\alpha = 0.999$)
Extremal dependence based on χ and $\bar{\chi}$ plots	Asymptotic dependence cannot be unequivocally proven	Perfect independence	Perfect independence	Perfect independence
Copula fitting	Appropriate fit using the Rotated Tawn type 2 180 degrees copula	Fitted copula does not describe the sample well	Appropriate fit using the Survival Gumbel copula	Fitted copula does not describe the sample well

The above reasoning has implications on how we look at the relationship between speed and safety. Studies in the past few decades have all shown a close correlation between speed, crash frequency and severity (Vadeby and Forsman, 2018). Higher speeds lead to a higher risk of a crash and a more severe outcome (Nilsson, 2004). The relationship between speed and severity of conflicts is more complex. Many studies used one single surrogate indicator to express the severity of an encounter by measuring the proximity to a crash in space or time (Zheng et al., 2014b). Setting a threshold value for TTC_{min} to differentiate between serious and light conflicts is a perfect example of that. This approach has a limitation, namely that even if the two vehicles are very close to each other in time, the severity of a possible collision

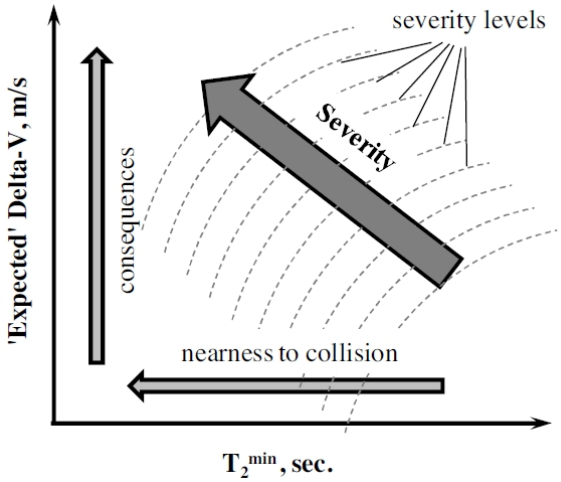
can be still mild if for instance their speeds are very low. Laureshyn et al. (2017a) pointed out that the severity of a near-miss between two vehicles also depends on the differences in their masses. Overall, the severity of a conflict is higher if either the proximity to a crash gets lower, or vehicle speeds or the difference in vehicle masses get higher. Research results of the bivariate models presented in this thesis actually proved that temporal and speed related indicators are independent (one being small does not lead to the other being high), thus they should be combined in order to properly predict severity; a temporal indicator on its own is indeed not enough to make inferences about the severity of encounters.

The probability based risk levels (i.e. the contours of the fitted bivariate CDFs) show similar patterns for all the four variable pairs. As for the two temporal indicators TTC_{min} and T_{2min} the difference lies in the selected thresholds. In case of T_{2min} the calculated conditional probabilities are based on smaller thresholds for the temporal and higher thresholds for the speed related indicator. This is due to the nature of T_{2min} being a crossing course indicator, namely that compared to TTC_{min} low temporal values are more frequent and higher relative speeds are observed (as already mentioned in Chapter 4). Extended Delta-V0 and relative speed do not seem to affect the shape of the probability based risk levels differently (i.e. for instance if we compare TTC_{min} vs. relative speed or Extended Delta-V0).

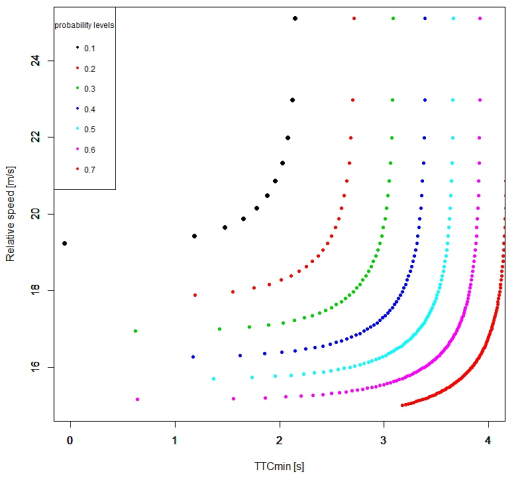
In the course of threshold selection the spectral measure plot and the plots used for univariate models (threshold stability and mean excess plots) led to the same threshold values for TTC_{min}, however suggested different thresholds in case of T_{2min} making the selection slightly arbitrary. Modeling results also showed that the various parametric distributions yielded quite similar AIC values and the logistic distribution was a well-performing one.

Jonasson and Rootzén (2014) (one of the first authors to investigate bivariate models) highlighted in their paper that a future possibility in bivariate models is crash severity estimation. They actually suggested using the minimum of a crash proximity measure and speed at the time when this minimum is attained, however they did not really clarify what they meant by speed (e.g. the maximum of the two speeds or relative speed). In the current chapter this research direction was well investigated and by doing so, such bivariate models actually indicated an additional problem in terms of validation. In the bivariate setting using a temporal indicator along with a speed related one it is possible to estimate crash probabilities along the speed dimension on condition that the temporal indicator equals zero, in other words crash severity distribution can be constructed. In order to validate these models using historical crash data one would also need to know the speeds of vehicles at the time of collision, which type of information is usually not registered in accident records and was unfortunately not available for the current study, either.

It is known from the literature (Hydén, 1987, Laureshyn et al., 2017a) that by combining temporal and speed related indicators one can construct severity levels. The Swedish TCT provides the exact location of these severity levels in the Time-to-Accident and conflicting speed plane (as shown in Figure 2.5). For other possible variable pairs the actual position of these severity levels is not known, however the conceptual illustration in Figure 6.50(a) follows the same logic (severity increases with lower values for temporal indicators and higher values for speed related indicators). Furthermore, we can assume that this severity level plane will be different for relative speed and Extended Delta-V. Once these severity levels are known they can be combined with the probability based risk levels (Figure 6.50(b)). By overlaying these two planes not only the severity of a given combination of surrogate indicators but also its probability could be known (to be more exact the probability of exceeding both values).



(a) Conceptual illustration of severity levels (Laureshyn et al., 2017a)



(b) Probability based risk levels

Figure 6.50: Severity levels and probability based risk levels

To investigate the dependence structure between variable pairs separately from their marginal distributions copulas were fitted where possible. The built-in function in R selects an appropriate bivariate copula family which can then be tested against empirical data. The parametric copula fitting procedure was only possible in two cases, both for TTC_{min}, where the two components showed at least a weak dependence. As for T_{2min}, where the bivariate extreme value models indicated perfect independence, the fitted copula was not able to describe the sample well.

A methodological aspect is to what range of data the copula is fitted. Previously Cavadas et al. (2017) used copula for their componentwise maxima bivariate model, in which a sample with both indicators above given thresholds was used, that is the length of the two variables was the same. Copulas require the two components to have an equal length, that is the marginal number of observations above thresholds should be identical. As for the bivariate GPD models this can be satisfied when thresholds are set based on the bivariate threshold choice plot resulting in an equal number of observations above thresholds for both components. However, if the threshold selection plots used for univariate POT (mean excess and threshold stability plots) indicate other thresholds to be used, the equal length of components can not be guaranteed. This is relevant, because bivariate models using copulas incorporate marginal distributions, which in this case are the two univariate POT models using all the observations above individual thresholds (marginal number above). In the current research all the observations were used in the course of parametric copula fitting and therefore the resulting copula describes the joint distribution in the entire range of values, whereas the marginal distributions (GPD) using the parameters of the univariate models were limited to the values above thresholds. Another possibility to further investigate the role of copulas could be to limit the analysis only to those observations which are extreme for both components (joint number above). However, a drawback of this approach is that we lose data points, which is actually the main criticism against the componentwise maxima approach (Rakonczai and Tajvidi, 2010).

All the analyses so far focused on the extreme region ($R_{1,1}$). In this quarter of the plane either the bivariate threshold excess models or - if there is some sort of dependence at extreme levels - parametric copula can be tested. As for the other three regions different methods can be used.

For the two regions $R_{1,0}$ and $R_{0,1}$ where one indicator exceeds the threshold and the other stays below, a modified version of the bivariate GPD is proposed by Rakonczai and Tajvidi

(2010), Rakonczai and Zempléni (2011). The usual way to fit a bivariate GPD is to use only those observations which exceed the threshold in all components, this method is called as Type I by Rakonczai and Tajvidi (2010). They proposed an alternative (called as bivariate GPD Type II), the main advantage of which is that it includes all observations that are extreme in at least one component. Practically speaking, in their approach the probability measure is positive in the upper three quarters and zero in the bottom left one. This is illustrated in Figure 6.51.

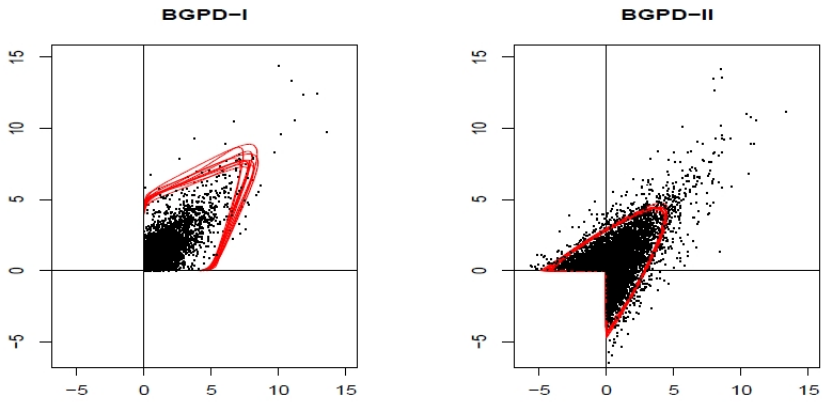


Figure 6.51: Prediction regions for simulated data from a logistic model. Bivariate GPD Type I (left), Type II (right) (Rakonczai and Zempléni, 2011)

In case of region $R_{0,0}$, where both indicators are below their thresholds, the most plausible way is to use the empirical bivariate copula. Since the empirical copula is not an appropriate tool for extrapolation in the extreme region, it is not possible to use it when either of the indicators exceeds its threshold. Below the empirical bivariate cumulative distributions are plotted (Figures 6.52 - 6.59). A 3D surface over these data points can be smoothed through bivariate interpolation using the *akima* package in R. The contour lines of this smoothed surface show the risk levels (joint probability) for the entire range of observations. These plots can actually draw attention to further characteristics of the relationship between the indicators. Two very distinctive differences can be found in between TTC_{min} and T_{2min}:

1. The cumulative distribution functions of T_{2min} are steeper than the ones of TTC_{min} resulting in a higher density of risk levels closer to the extreme region. A slight change in the crossing course indicator T_{2min} closer to the extreme region results in a swift change in the joint probability.

2. On the other hand joint probabilities in the plane between T_{2min} and speed related indicators are less sensitive about changes in speeds in the extreme region. The range of values in relative speeds for T_{2min} is much wider, the maximum value is slightly over 30 m/s, whereas for TTC_{min} it is approximately 22 m/s. As for T_{2min} there is a clear pattern that the probability lines are more or less vertical above the relative speed of 20m/s meaning that in the extreme region of speeds a change in the speed would not really affect the joint probability. A similar pattern can be seen in case of Delta-V0, where above 14 m/s all the probability lines are vertical.

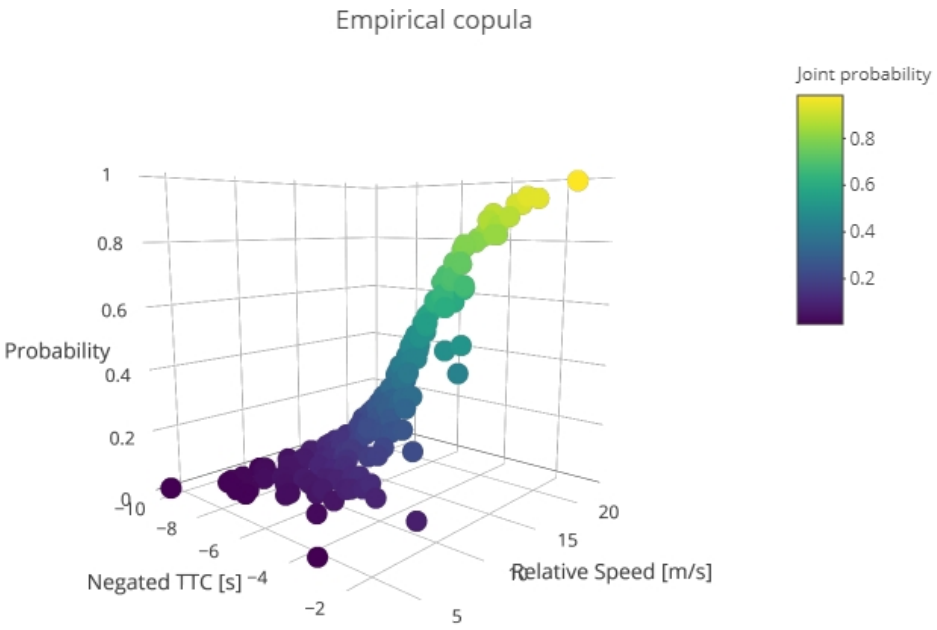


Figure 6.52: Empirical copula for TTC_{min} and relative speed

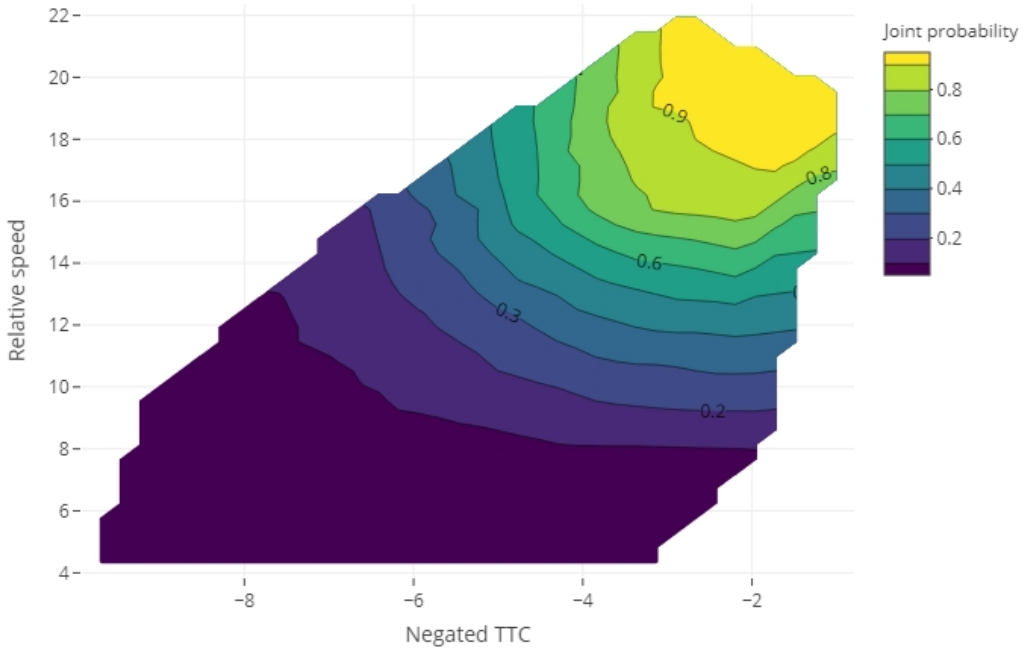


Figure 6.53: Contour plot of empirical copula for TTC_{min} and relative speed

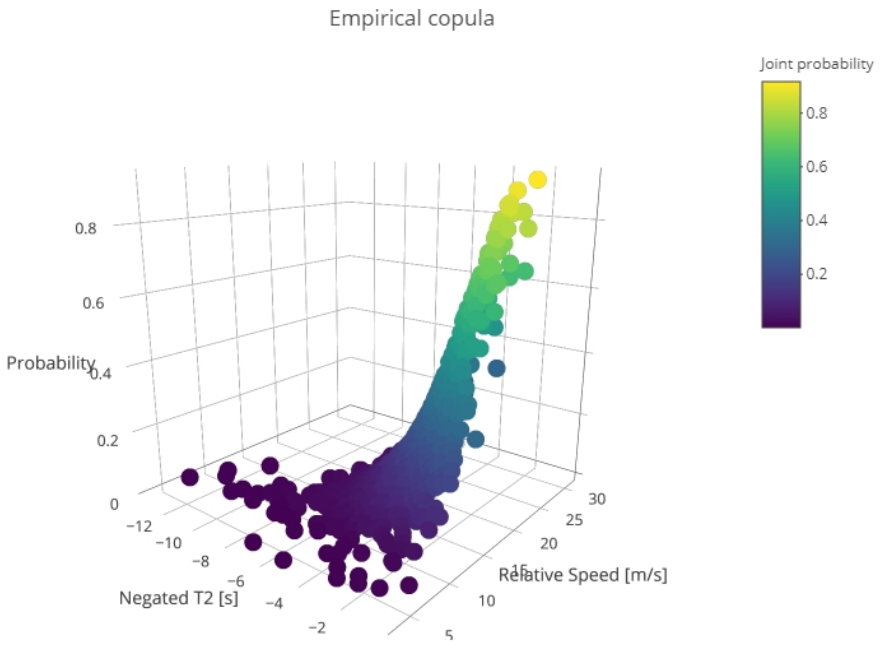


Figure 6.54: Empirical copula for T_{2min} and relative speed

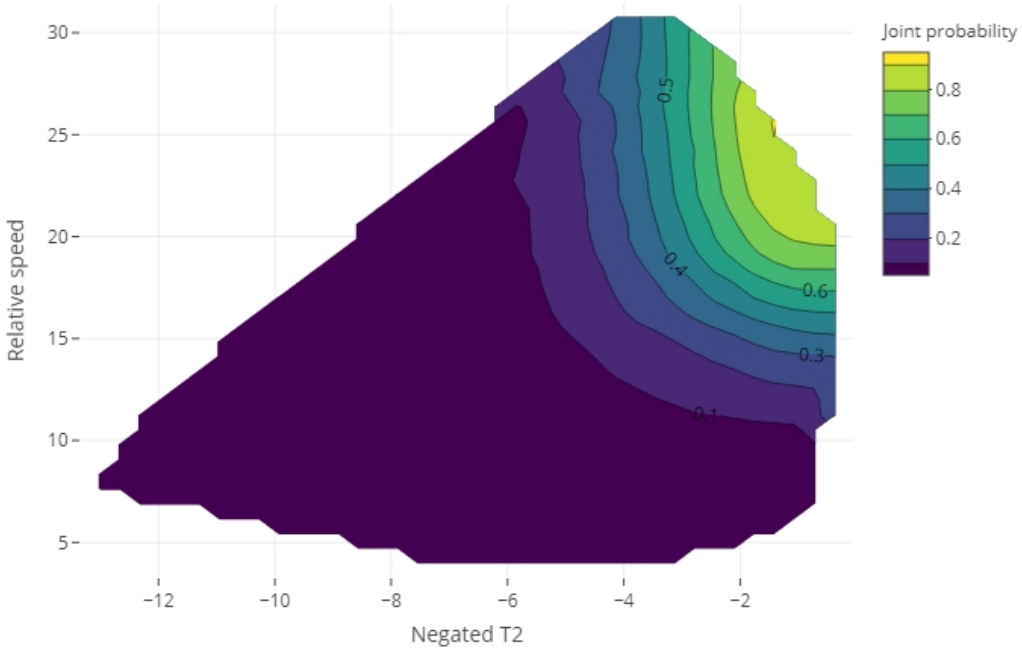


Figure 6.55: Contour plot of empirical copula for T_{2min} and relative speed

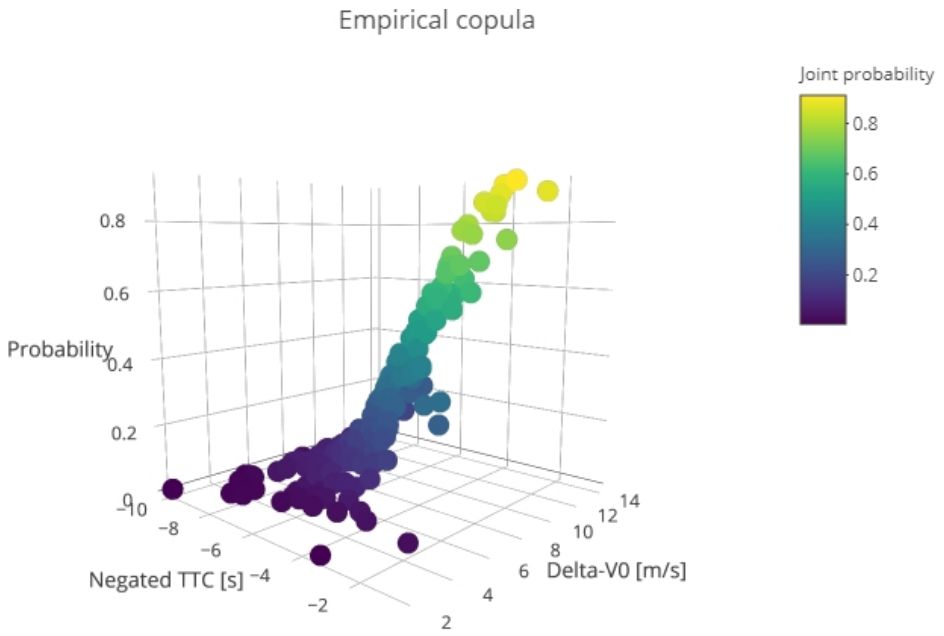


Figure 6.56: Empirical copula for TTC_{min} and Delta-V0

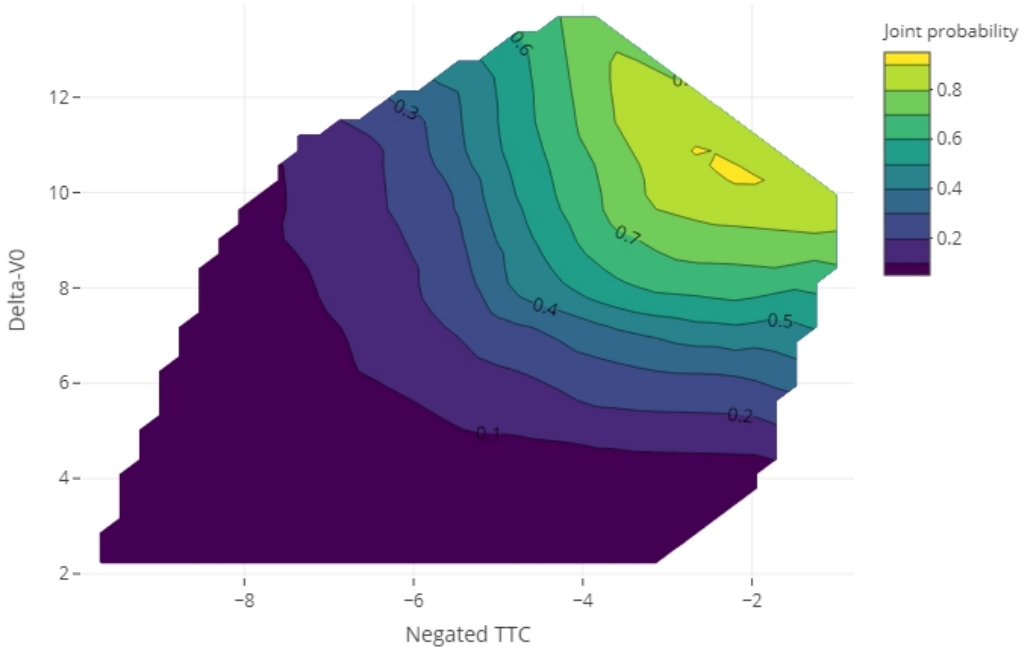


Figure 6.57: Contour plot of empirical copula for TTC_{min} and Delta-V0

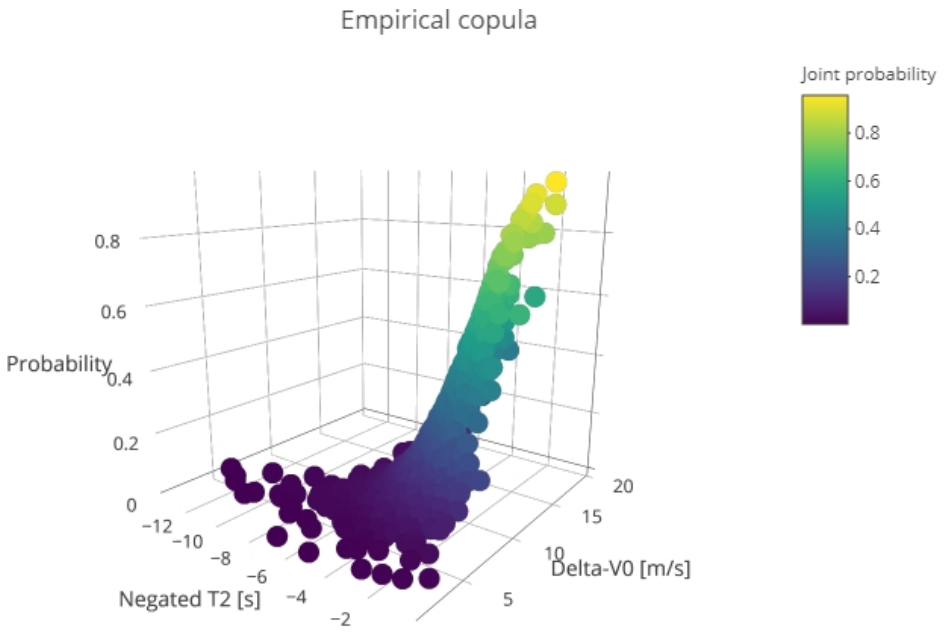


Figure 6.58: Empirical copula for T_{2min} and Delta-V0

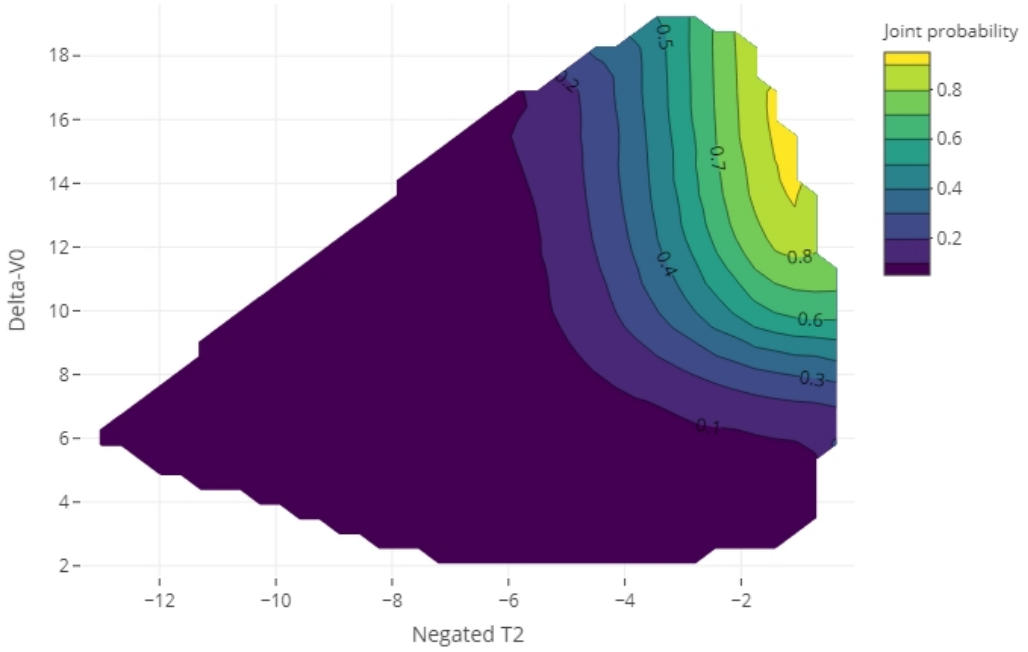


Figure 6.59: Contour plot of empirical copula for T_{2min} and Delta-V0

7

Final discussion and conclusions

7.1. Conclusions

The most important conclusions of the thesis research are summarized and structured according to the research questions below.

1. What can we learn from applying EVT using indicators describing collision course and crossing course interactions at signalized intersections for vehicle-vehicle interactions?

   (a) What difference is there between the two indicators TTC_{min} and T_{2min} when analysing safety using EVT and are these indicators transferable?

   The analysis showed that for straight moving vs. left turning vehicle interactions collision course indicator TTC_{min} results in a smaller sample size gathered in a given time period in comparison to the crossing course indicator T_{2min}. Thus, for both EVT approaches higher near-crash values (BM approach) and thresholds (POT approach) had to be used for TTC_{min} to have a sufficient sample, on the other hand for T_{2min} even lower limiting values resulted in appropriate models. Models for T_{2min} provided more reliable results since the prediction of probabilities of extreme events came with narrower confidence intervals. The validation on the other hand showed that T_{2min} tends to overestimate crash frequencies and crash estimates based on TTC_{min} were closer to reality. The transferability of these indicators was also investigated (for the BM approach) and it was concluded that in order to have similar crash probabilities (conditional on the selected thresholds) for collision and crossing course indicators one has to set lower near-crash thresholds for the latter (i.e. T_{2min}).

   (b) Which EVT approach under what circumstances performs better for TTC_{min} and T_{2min} (e.g. sensitivity to sample size)?

   As for the two EVT approaches it can be concluded that overall applying the POT approach gave more reliable and pragmatic results in terms of both crash probabilities and return periods. By accepting a few assumptions regarding historical accident data the POT results were much closer to the actual crash frequency. The application of the BM approach came with a critical issue, namely the pre-selection of near-crash situations as a sub-sampling step showed a classical trade-off between variance and bias.

2. How can we predict nearness to collision and severity at signalized intersections for vehicle-vehicle interactions using the Extreme Value Theory?

   (a) What can we learn from applying bivariate models using EVT for various combinations of indicators?

   The bivariate threshold excess models using logistic distribution yielded dependence parameters between 0.746 and 0.999 showing weak dependence or perfect independence in between the two temporal and speed related indicators. TTC_{min} and relative speed was the only variable pair that showed the signs of asymptotic dependence. Overall, however, this weak or no dependence at extreme levels means that road users getting closer to each other in time do not necessarily show high relative speed or Delta-V0. The bivariate models showed an additional problem in terms of validation. Since with these models one can estimate the crash severity distribution, for validation it would be necessary to know the speeds of vehicles at the time of collision. Without this information these types of bivariate models combining temporal and speed indicators cannot be validated. The independence structure between the variable pairs was also investigated by means of copula. Fitting parametric copula was only successful for TTC_{min} which indicator showed at least some sort of - even though weak - dependence on speed indicators.

   (b) How can we use bivariate EVT in combination with severity levels?

   The resulting risk levels (based on calculated joint probabilities) of the EVT models can be combined with the severity levels between temporal and speed related indicators. The latter plane shows the actual severity of the interactions, whereas the former plane gives further information on the probability of each combination of indicators. In order to overlay these two planes (severity levels with probability based risk levels) the actual severity levels would need to be known first. This information is only available in the conflict diagram of the Swedish Traffic Conflict Technique so far for surrogate indicators (Time-to-Accident and conflicting speed) other than what were used in this research. An important future research direction is thus to either do bivariate modeling for the Swedisch TCT indicators or to construct the severity levels for the indicators used in the current research.

7.2. Limitations

As every research, this one also comes with certain limitations, the most important of those are explained in this section.

The basis of the analysis was a dataset collected in Minsk, Belarus. All the data came from one location, thus cross comparison in between different locations was not possible. Using several locations can actually give an added value, Zheng et al. (2018) for instance used 16 merging areas and could also provide a comparison of models across these locations. Nevertheless, the observation periods at these locations were quite short (56-88 minutes) and the authors also admitted that this short time frame can hardly claimed to be representative for a five-year period for which accident data was gathered and compared with the estimated number of crashes. In the current research one location with data for two days was used (6AM - 9PM), which in that sense is more representative, even though estimated results were compared with accident data gathered for a ten-year period.

Besides being representative in terms of the length of time period analysed it can also be questioned whether the results based on observations in Minsk can be generalized or they are location specific. In order to answer this question one would need to know certain local characteristics such as road user behaviour (e.g. priority giving or surrendering attitude), gap acceptance etc. Unfortunately this type of information was not at hand.

The type of interaction analysed was exclusively left-turning vs. straight moving vehicle to vehicle interactions. The results gained in this research are therefore restricted and applicable to this interaction type only. Nevertheless, even this type of interaction could be further examined by looking into other possible conflicts such as the possibility of rear-end crashes in the straight going stream.

Probably the most important limitation was the uncertainty or lack of validation. As for the univarite models validation was possible along with certain assumptions. Even if accident data are available one has to be cautious about its accuracy, especially if mostly low severity or property damage accidents are present (underreporting). As for the bivariate models the second dimension describing the potential consequences in case of a crash has to be some sort of speed related indicator. This causes two further issues: 1) speeds of interacting vehicles in accidents are usually not known thus breaking down accidents according to speed for validation is not possible; and 2) the different severity thresholds for various speed-related indicators should be pre-defined.

7.3. Further research directions

This section gives a response to the limitations detailed previously as well as defines other research directions. One possible way to extend the use of EVT models is to provide a comprehensive analysis to assess the safety level of a given intersection. This would include all possible interactions in between vehicles as well as other road users (e.g. pedestrians, bicycles) and as a result all (or most) intersection accidents could be used for validation. This is however very labour intensive since all the vehicle movements (trajectories) should be constructed and all possible conflicts be analyzed. Since the relation between surrogate measures and crashes is different for different types of maneuvers, they should be analyzed separately (Laureshyn et al., 2017c). A fully automated software could greatly help to reduce the time and effort needed for that.

In order to improve validation using proper accident data two sets of video footage could be used. One set using a much longer observation time period would merely focus on providing an accurate number of accidents. Another set, similar to what was used for this study, could then be used for calculating surrogate measures and for model estimation.

To combine severity lines with probability based risk levels one would need to construct actual severity lines first. So far these severity lines are only available in the conflict diagram of the Swedish Traffic Conflict Technique using time-to-accident and conflicting speed (Laureshyn and Varhelyi, 2018). A plausible way to combine these two planes would have been to start with the indicators used in the Swedish TCT and combine it with probability based risk levels, as also noted by Jonasson and Rootzén (2014). These indicators were however not provided by the T-Analyst software. After constructing a plane similar to that of the Swedish TCT using other pairs of indicators such as TTC vs. relative speed or Delta-V0 the probabilities along these severity lines could be calculated using the modeling results. For the bivariate case it would be also very important to first set the threshold values in speed related indicators for the various layers of severity. By knowing the estimates of these severity outcomes one could compare them with the severity of observed accidents. Besides comparing the estimated number of accidents with the observed counts this could be a further step in the validation process.

7.4. Application in practice

Even though this thesis research is highly theoretical it offers a few possibilities for application in practice. EVT in general could be used at locations which are newly built or reconstructed facilities recently given over to traffic. Since these locations at the stage of early operation usually do not have accidents, EVT together with surrogate measures could be a complimentary tool to on-site road safety audits (or inspections).

For road authorities with limited funds it is a common task to allocate money over the road network for road safety purposes. To that end it is usual to evaluate the safety level of sites (or sites with promise) and rank them. This ranking is heavily based on accident data and could be supplemented with crash estimates coming from EVT models. This would be more relevant in an urban setting for two reasons: 1) accidents usually tend to be less severe and thus suffer more from underreporting; and 2) accidents usually happen due to the interaction between various road users (for which surrogate measures can be calculated), single vehicle accidents are less likely.

Another possibility for road authorities would be the use of EVT models for before-after studies. The safety effects of road safety measures could then be analyzed through video footages recorded before and after the intervention is put in place. A significant improvement in time-to-collision values can for instance indicate that accidents are less likely to occur, and following the same line of reasoning a decrease in speed can indicate that even if accidents happen they will tend to be less severe.

7.5. Dissemination

In the table below (Table 7.1) all the past and future publications based on the thesis and their current status are summarized.

Table 7.1: List of publications

Title	Type	Venue / Journal	Status
Comparison of two nearness-to-collision surrogate indicators at a signalized intersection in Minsk using Extreme Value Theory	Oral presentation	31st ICTCT Conference, Porto, Portugal, October 25–26 2018	Presented
Surrogate safety indicators combined with extreme value theory for road safety analysis	Oral presentation / Full paper	Conference on Transport Sciences, Győr, Hungary, March 21-22, 2019	Presented / published
Probability based severity of conflicts using bivariate Extreme Value models	Oral presentation	32nd ICTCT Conference, Warsaw, Poland, October 24–25 2019	Presented
Are collision and crossing course surrogate safety indicators transferable? A probability based approach using Extreme Value Theory	Poster presentation / Full paper	Transportation Research Board Annual Meeting, January 12-16, 2020, in Washington, D.C.	Accepted
Are collision and crossing course surrogate safety indicators transferable? A probability based approach using Extreme Value Theory	Full paper	Accident Analysis and Prevention	Submitted, under review
Application of Bivariate Extreme Value models to describe the joint behaviour of temporal and speed related surrogate measures of safety	Full paper	Accident Analysis and Prevention	In writing

Acknowledgment

I would like to thank all the committee members for supporting this research project in different ways. Most importantly I owe special thanks to my daily supervisor, Dr. Haneen Farah, who was always very enthusiastic to discuss research results and kept providing fresh ideas for further improvement. Professor Marjan Hagenzieker as the committee chair assisted me in defining a good structure for the thesis and in keeping a healthy balance between theory and practical implications. Dr. Aliaksei Laureshyn provided all the data that were needed to embark on this research and also gave invaluable technical knowledge on Surrogate Measures of Safety. Dr. Juan Juan-Cai as an expert in Extreme Value Theory made sure that all the statistical aspects are thoroughly dealt with and by that she greatly contributed to the quality of the thesis.

I would not have been able to pursue the Transport and Planning programme at TU Delft without my family. My wife Laura and our two sons Bendegúz and Barnabás always showed patience even in the hardest times. Last but not least I would like to express my profound gratitude to my good Dutch friend, Dr. Frank van Vliet, for all his invaluable support which made my participation in the Transport and Planning programme possible.

Bibliography

Allen, B. L. and Shin, B. T. (1977). Analysis of Traffic Conflicts and Collisions. *Transportation Research Record*, 667:67–74.

Åsljung, D., Nilsson, J., and Fredriksson, J. (2016). Comparing Collision Threat Measures for Verification of Autonomous Vehicles using Extreme Value Theory. *IFAC-PapersOnLine*, 49-15:57–62.

Baguley, C. J. (1984). The British traffic conflicts technique. In *NATO advanced research workshop on international calibration study of traffic conflict techniques*, Copenhagen, Denmark.

Beirlant, J., Goegebeur, Y., Segers, J., and Teugels, J. (2004). *Statistics of Extremes: Theory and Applications*. John Wiley and Sons, San Francisco, USA.

Cavadas, J., Azevedo, C. L., and Farah, H. (2017). Extreme value theory approach to analyze safety of passing maneuvers considering drivers' characteristics. *IATSS research*, 41(1):12–21.

Ceunynck, d. T. (2017). *Defining and applying surrogate safety measures and behavioural indicators through site-based observations (Doctoral dissertation)*. PhD thesis, Lund University, Sweden.

Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*.

European Commission (2001). White Paper, European transport policy for 2010: Time to decide.

European Commission (2011). White Paper, Roadmap to a Single European Transport Area – Towards a competitive and resource efficient transport system.

European Commission (2019). EU Road Safety Policy Framework 2021-2030 – Next step towards "Vision Zero". Technical report, Brussels.

Farah, H. and Azevedo, C. L. (2015). Using Extreme Value Theory for the Prediction of Head-On Collisions during Passing Maneuvres. In *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*.

Farah, H. and Azevedo, C. L. (2017). Safety analysis of passing maneuvers using extreme value theory. *IATSS Research*, 41(1):12–21.

Gilleland, E. and Katz, R. W. (2016). extRemes 2.0: An Extreme Value Analysis Package in R. *Journal of Statistical Software*, 72(8).

Hauer, E. (1997). *Observational before-after studies in road safety*. Emerald Group Publishing Limited, Oxford, UK.

Hayward, J. C. (1972). Near-miss determination through use of a scale of danger. *Highway Res. Rec.*, 384:22–34.

Huang, F., Liu, P., Yu, H., and Wang, W. (2013). Identifying if VISSIM simulation model and SSAM provide reasonable estimates for field measured traffic conflicts at signalized intersections. *Accident Analysis and Prevention*, 50:1014–1024.

Hyde, T. and Wright, C. C. (1986). Extreme value methods for estimating road traffic capacity. *Transportation Research Part B*, 20B(2):125–138.

Hydén, C. (1987). *The development of a method for traffic safety evaluation: The Swedish Traffic Conflicts Technique*. PhD thesis, Lund University, Sweden.

Johansson, R. (2009). Vision Zero - Implementing a policy for traffic safety. *Safety Science*, 47(6):826–831.

Jonasson, J. K. and Rootzén, H. (2014). Internal validation of near-crashes in naturalistic driving studies: A continuous and multivariate approach. *Accident Analysis and Prevention*, 62:102–109.

Kočárková, D. (2012). Traffic Conflict Techniques in Czech Republic. *Procedia - Social and Behavioral Sciences*, 53:1029–1034.

Kraay, J. and van der Horst, A. (1985). The Trautenfels Study: A diagnosis of roaad safety using Dutch Conflict observation Technique DOCTOR. Technical report, Institute for Road Safety Research SWOV, The Netherlands.

Kraay, J. H., van der Horst, R., and Oppe, S. (2013). Manual conflict observation technique DOCTOR - Dutch Objective Conflict Technique for Operation and Research (No. 2013–1). Technical report, The Netherlands: Foundation Road safety for all, Voorburg.

Kulmala, R. (1984). The Finnish traffic conflict technique. In *NATO advanced research workshop on international calibration study of traffic conflict techniques*, Copenhagen, Denmark.

Laureshyn, A., De Ceunynck, T., Karlsson, C., Svensson, A., and Daniels, S. (2017a). In search of the severity dimension of traffic events: Extended Delta-V as a traffic conflict indicator. *Accident Analysis and Prevention*, 98:46–56.

Laureshyn, A., Goede, M. d., Saunier, N., and Fyhri, A. (2017b). Cross-comparison of three surrogate safety methods to diagnose cyclist safety problems at intersections in Norway. *Accident Analysis and Prevention*, 105:11–20.

Laureshyn, A., Johnsson, C., Kidholm, T., Madsen, O., Várhelyi, A., De Goede, M., Svensson, A., Saunier, N., and Van Haperen, W. (2017c). Exploration of a method to validate surrogate safety measures with a focus on vulnerable road users. Technical report.

Laureshyn, A., Svensson, A., and Hydén, C. (2010). Evaluation of traffic safety, based on micro-level behavioural data: Theoretical framework and first implementation. *Accident Analysis and Prevention*, 42:1637–1646.

Laureshyn, A. and Varhelyi, A. (2018). The Swedish Traffic Conflict Technique - Observer's manual. Technical report, Lund University, Lund.

Mahmud, S. S., Ferreira, L., Hoque, M. S., and Tavassoli, A. (2017). Application of proximal surrogate indicators for safety evaluation: A review of recent developments and research needs. *IATSS Research*, 41:153–163.

Muhlrad, N. and Dupre, G. (1984). The French conflict technique. In *NATO advanced research workshop on international calibration study of traffic conflict techniques*, Copenhagen, Denmark.

Nicholson, A. J. (1985). The variability of accident counts. *Accident Analysis and Prevention*, 17(1):47–56.

Nilsson, G. (2004). *Traffic Safety Dimensions and the Power Model to Describe the Effect of Speed on Safety*. PhD thesis, Lund Institute of Technology.

Ozbay, K., Yang, H., Bartin, B., and Mudigonda, S. (2008). Derivation and Validation of New Simulation-Based Surrogate Safety Measure. *Transportation Research Record: Journal of the Transportation Research Board*, 2083:105–113.

Parker, M, R. and Zegeer, C, V. (1989). Traffic conflict techniques for safety and operation - observers manual (No. FHWA-IP-88-027). Technical report, Federal Highway Administration, U.S. Department of Transport.

Penalva, H., Neves, M., and Nunes, S. (2013). Topics in data analysis using R in extreme value theory. *Metodoloski Zvezki*.

R Core Team (2013). R: A language and environment for statistical computing.

Rakonczai, P. and Tajvidi, N. (2010). On Prediction of Bivariate Extremes. *The Thailand Econometrics Society*, 2(2):174–192.

Rakonczai, P. and Zempléni, A. (2011). Bivariate generalized Pareto distribution in practice: models and estimation. *Environmetrics*, 23(3):219–227.

Risser, R. and Schutzenhofer, A. (1984). Application of traffic conflict technique in Austria. In *NATO advanced research workshop on international calibration study of traffic conflict techniques*, Copenhagen, Denmark.

Sayed, T., Zaki, M. H., and Autey, J. (2013). Automated safety diagnosis of vehicle-bicycle interactions using computer vision analysis. *Safety Science*, 59:163–172.

Smith, R. L. (1985). Maximum likelihood estimation in a class of nonregular cases. *Biometrika*, 72(1):67–90.

Songchitruksa, P. and Tarko, A. P. (2006). The extreme value theory approach to safety estimation. *Accident Analysis and Prevention*, 38:811–822.

St-Aubin, P., Saunier, N., and Miranda-Moreno, L. (2015). Large-scale automated proactive road safety analysis using video data. *Transportation Research Part C: Emerging Technologies*, 58:363–379.

Stephenson, A. (2018). Statistics of Multivariate Extremes.

Svensson, A. (1998). *A method for analysing the traffic process in a safety perspective (Doctoral dissertation)*. PhD thesis, Lund University, Sweden.

Svensson, A. and Hydén, C. (2006). Estimating the severity of safety related behaviour. *Accident Analysis and Prevention*, 38:379–385.

T-Analyst (2016). Software for semi-automated video processing.

Tarek, S. and Sany, Z. (1999). Traffic conflict standards for intersections. *Transportation Planning and Technology*, 22(4):309–323.

Tarko, A., Davis, G., Saunier, N., Sayed, T., and Washington, S. (2009). White Paper SURROGATE MEASURES OF SAFETY ANB20(3) Subcommittee on Surrogate Measures of Safety ANB20 Committee on Safety Data Evaluation and Analysis.

Tarko, A. P. (2012). Use of crash surrogates and exceedance statistics to estimate road safety. *Accident Analysis and Prevention*, 45:230–240.

Vadeby, A. and Forsman, □. (2018). Traffic safety effects of new speed limits in Sweden. *Accident Analysis and Prevention*.

Wang, C., Xu, C., Xia, J., Qian, Z., and Lu, L. (2018). A combined use of microscopic traffic simulation and extreme value methods for traffic safety evaluation. *Transportation Research Part C: Emerging Technologies*, 90:281–291.

WHO (2018). *Global status report on road safety 2018*. Geneva.

Zheng, L., Ismail, K., and Meng, X. (2014a). Freeway safety estimation using extreme value theory approaches: A comparative study. *Accident Analysis and Prevention*, 62:32–41.

Zheng, L., Ismail, K., and Meng, X. (2014b). Traffic conflict techniques for road safety analysis: open questions and some insights. *Canadian Journal of Civil Engineering*, 41:633–641.

Zheng, L., Ismail, K., Sayed, T., and Fatema, T. (2018). Bivariate extreme value modeling for road safety estimation. *Accident Analysis and Prevention*, 120:83–91.

A

Appendix - Alternative approach to univariate models

In this subsection an alternative method to using negated values is tested. The motivation for this analysis was to better zoom in the region of low temporal indicator values. To that end instead of negated values $1/TTC_{min}$ and $1/T_{2min}$ are used (Figure A.1). As model results did not provide convincing results, they are very briefly summarized.

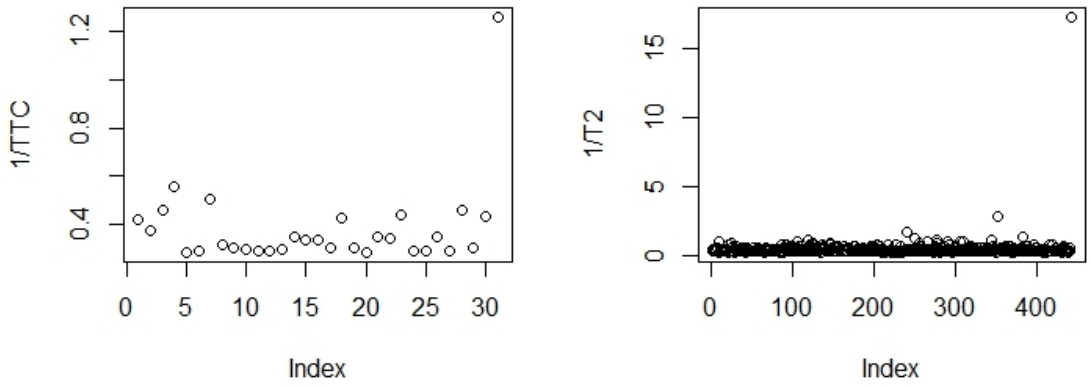


Figure A.1: Scatter plots of $1/T_{2min}$ and $1/TTC_{min}$ smaller than 3.5s

A.1. Block Maxima

As for $1/TTC_{min}$ with a threshold value of 3.5s convergence is not reached, thus 3.6s is used (Table A.1).

Table A.1: Model results of GEV for $1/TTC_{min}$ (near-crash threshold 3.6s)

Indicator	**Location**	**Scale**	**Shape**
Estimated parameter	0.300	0.024	1.232
Standard error	0.007	0.009	0.376
Lower bound of confidence interval	0.287	0.007	0.479
Upper bound of confidence interval	0.311	0.039	1.991
AIC	-88.958		
BIC	-84.561		
Deviance	-94.919		
Log-likelihood	-47.479		

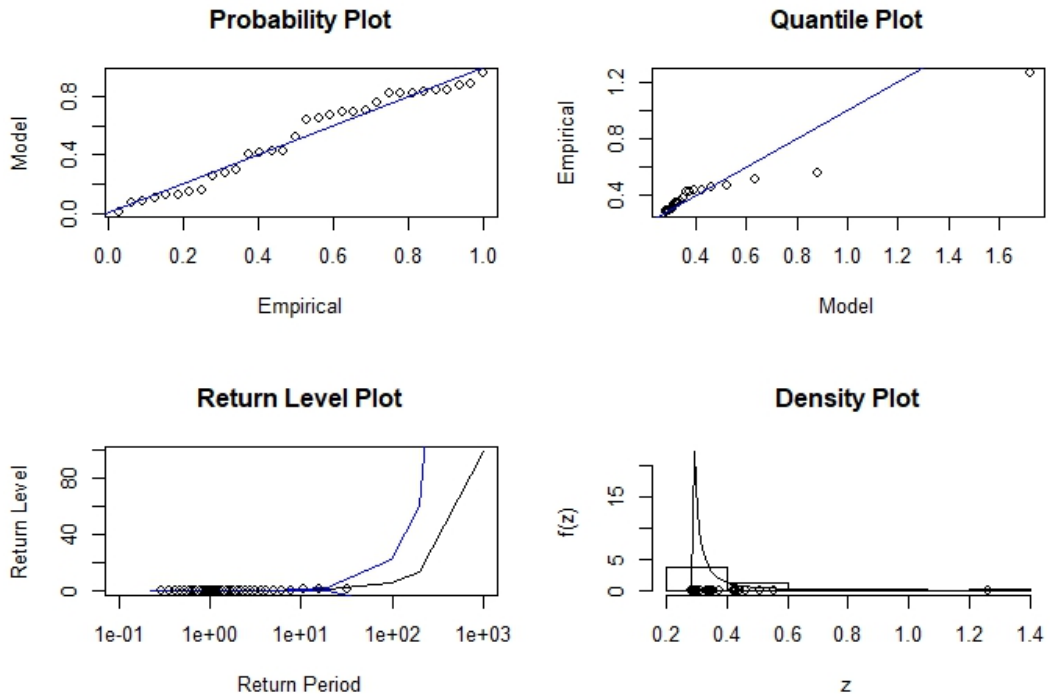


Figure A.2: Diagnostic plots for GEV fit to $1/TTC_{min}$ (near crash threshold < 3.6s)

Crash probability is calculated using $1/100$ (0.01 second) for the indicators. Models give more reasonable prediction in terms of crash probabilities as compared to using negated values, however, the return level plots indicate that for return levels 100 (i.e. when TTC equals to $1/100$s) the model is not reliable due to wide confidence intervals (Figure A.2). Model results with other near-crash threshold values can be found in Table A.2, diagnostic plots for a near-crash threshold of 5s is shown in Figure A.3.

Table A.2: Results of GEV for $1/TTC_{min}$ with different thresholds for near-crash situations

Indicator	TTC_{min}**<3.6s**	TTC_{min}**<4s**	TTC_{min}**<4.5s**	TTC_{min}**<5s**
Sample size	32	48	76	100
Location parameter	0.300	0.280	0.250	0.232
Scale parameter	0.0244	0.031	0.030	0.033
Shape parameter	1.232	0.702	0.672	0.536
Shape p. lower bound of conf. int.	0.479	0.355	0.372	0.290
Shape p. upper bound of conf. int.	1.991	1.053	0.974	0.781
Probability of crash TTC_{min} <1/100	0.0009	1.6728e-05	1.0213e-05	1.0308e-06
Return period for TTC_{min} <1/100	1011.86	59778.86	97910.06	970041.50

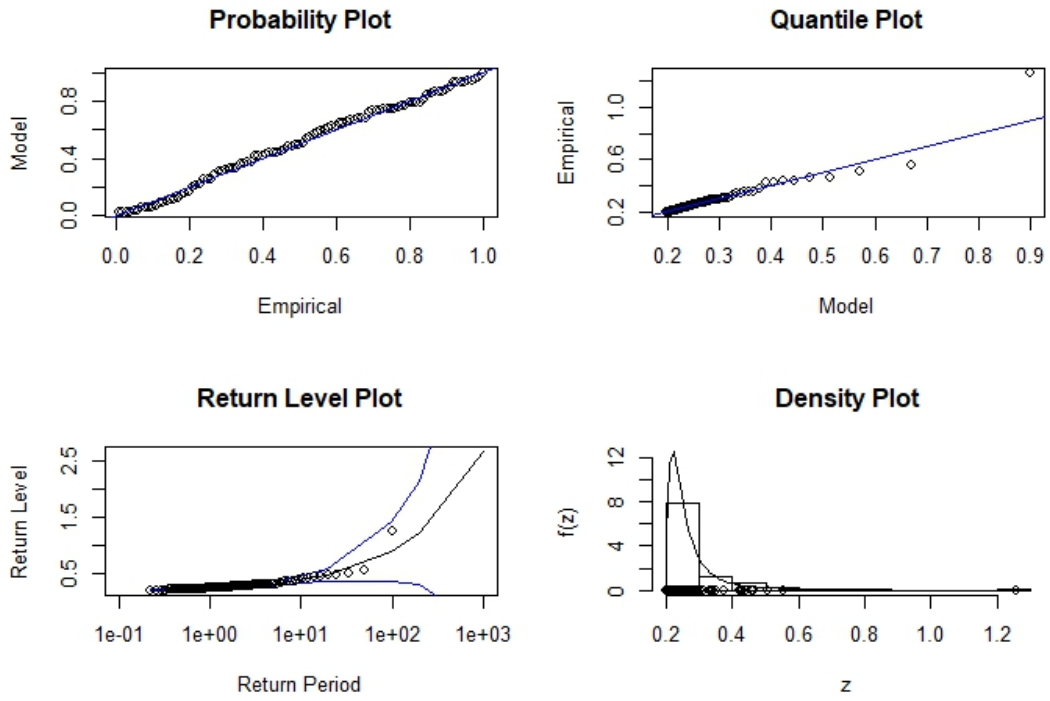


Figure A.3: Diagnostic plots for GEV fit to $1/TTC_{min}$ (near crash threshold < 5s)

As for $1/T_{2min}$ the same near-crash threshold values are used and model results are presented in Table A.3 and A.4, diagnostic plots are given in Figures A.4 and A.5. Here similar conclusions can be drawn as previously for TTC_{min}.

Table A.3: Model results of GEV for $1/T_{2min}$ (near-crash threshold 3.5s)

Indicator	**Location**	**Scale**	**Shape**
Estimated parameter	0.367	0.083	0.535
Standard error	0.005	0.005	0.061
Lower bound of confidence interval	0.357	0.074	0.416
Upper bound of confidence interval	0.376	0.092	0.654
AIC	-539.543		
BIC	-527.262		
Deviance	-545.543		
Log-likelihood	-272.772		

Table A.4: Results of GEV for $1/T_{2min}$ with different thresholds for near-crash situations

Indicator	T_{2min}**<3.5s**	T_{2min}**<3s**	T_{2min}**<2.5s**	T_{2min}**<2s**
Sample size	443	341	232	130
Location parameter	0.367	0.413	0.482	0.577
Scale parameter	0.083	0.081	0.084	0.084
Shape parameter	0.535	0.574	0.570	0.711
Shape p. lower bound of conf. int.	0.416	0.439	0.415	0.483
Shape p. upper bound of conf. int.	0.654	0.709	0.725	0.939
Probability of crash T_{2min} <1/100	5.5784e-06	1.0808e-05	1.0716e-05	7.6419e-05
Return period for T_{2min} <1/100	179262.20	92523.52	93311.37	13085.60

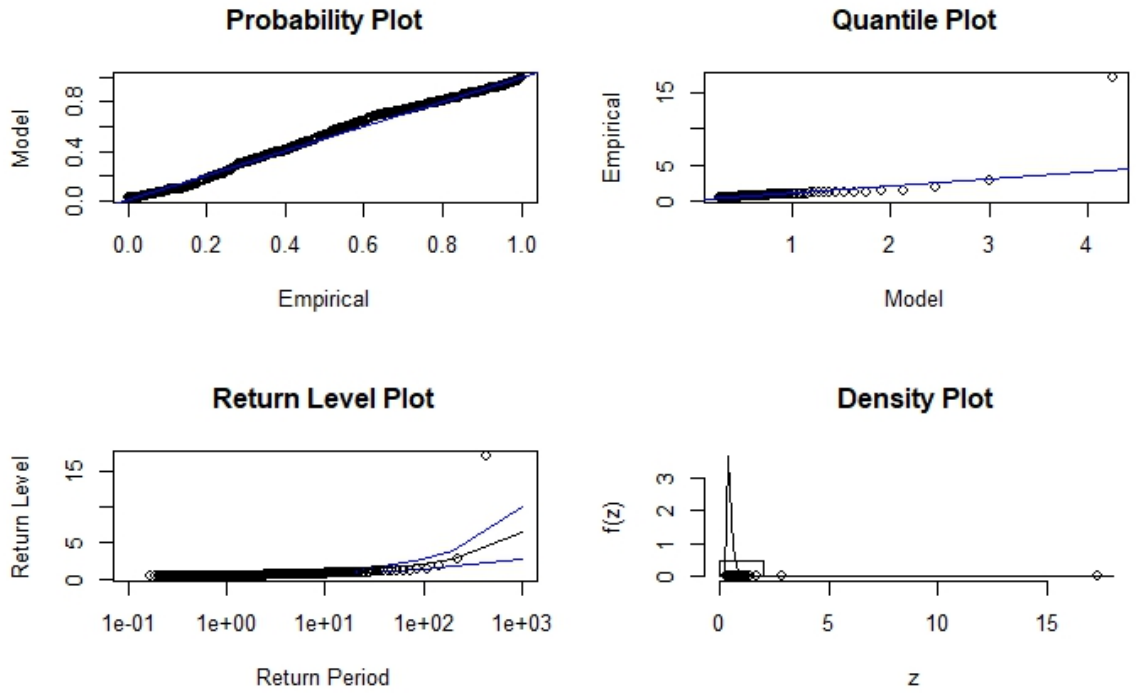


Figure A.4: Diagnostic plots for GEV fit to $1/T_{2min}$ (near crash threshold < 3.5s)

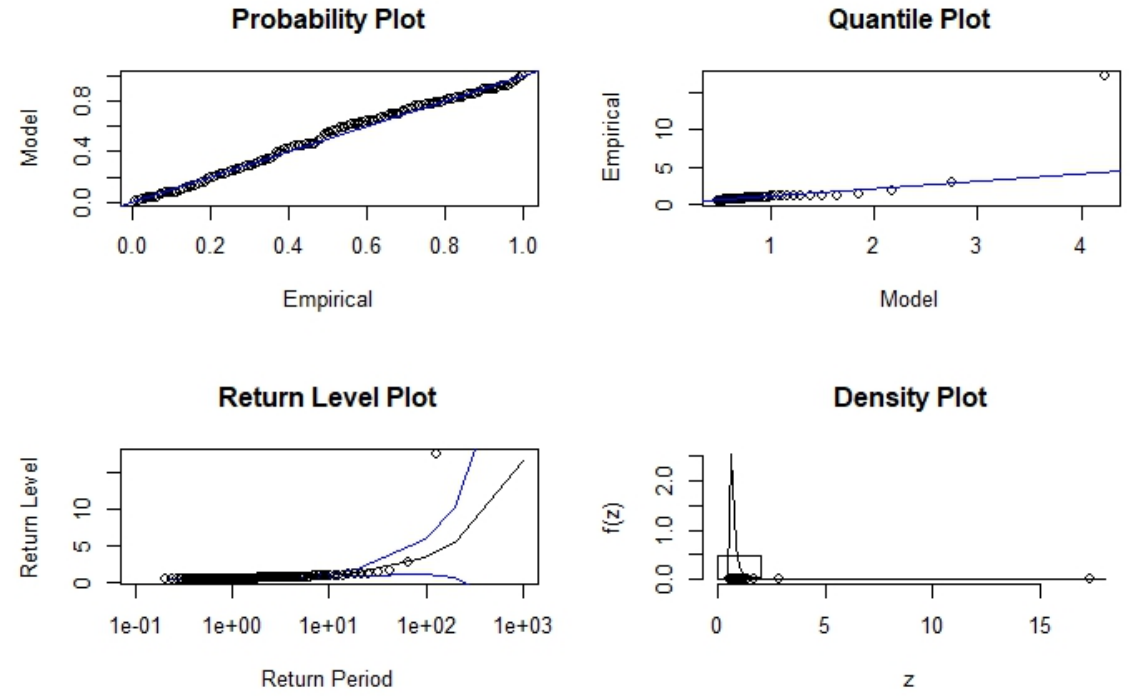


Figure A.5: Diagnostic plots for GEV fit to $1/T_{2min}$ (near crash threshold < 2s)

A.2. Peak over Threshold

Threshold selection based on both the threshold stability plots (Figure A.6 for TTC_{min} and Figure A.7 for T_{2min}) as well as the mean excess plot (Figure A.8) seems to be a difficult task as no stable region can be identified for either indicators, thus here no models are fitted.

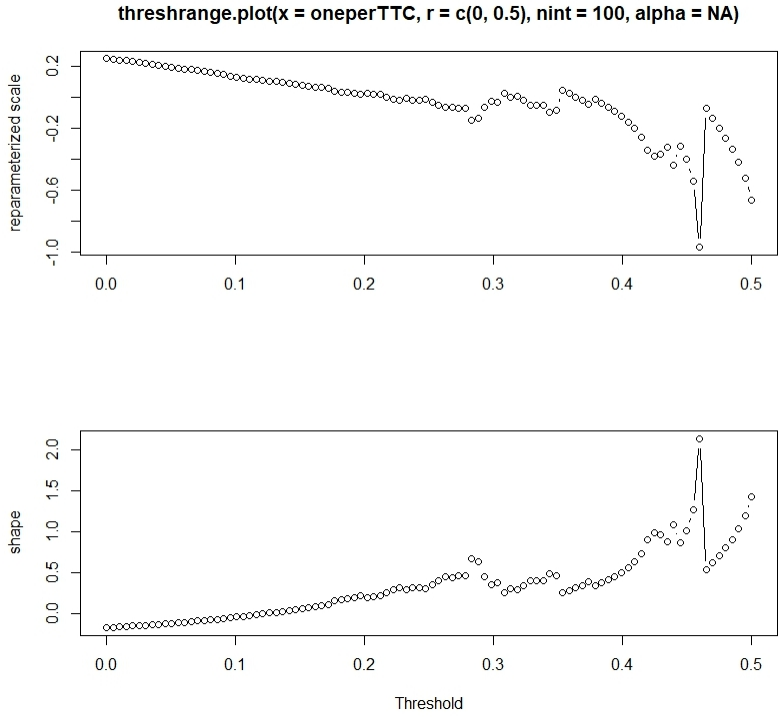


Figure A.6: Parameter estimates against threshold for $1/TTC_{min}$

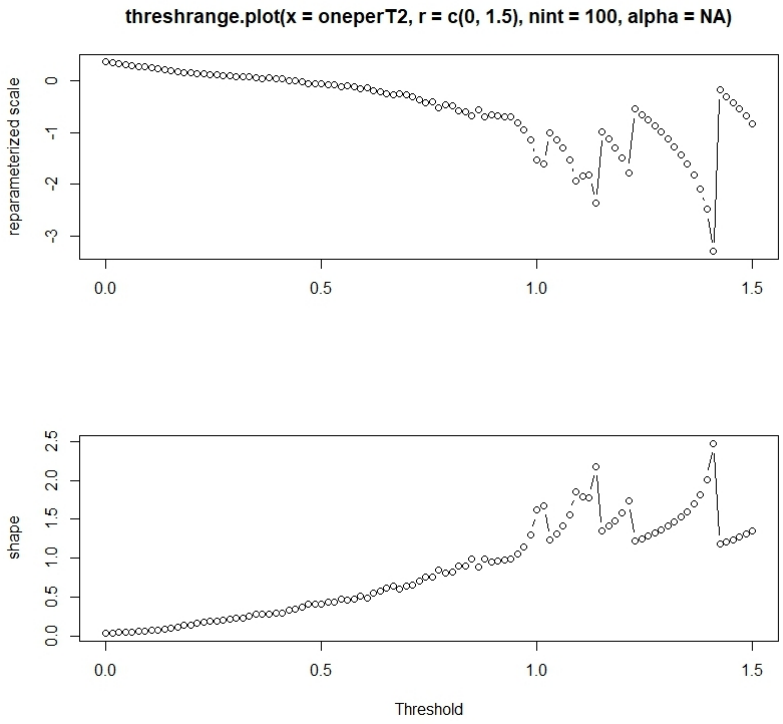


Figure A.7: Parameter estimates against threshold for $1/T_{2min}$

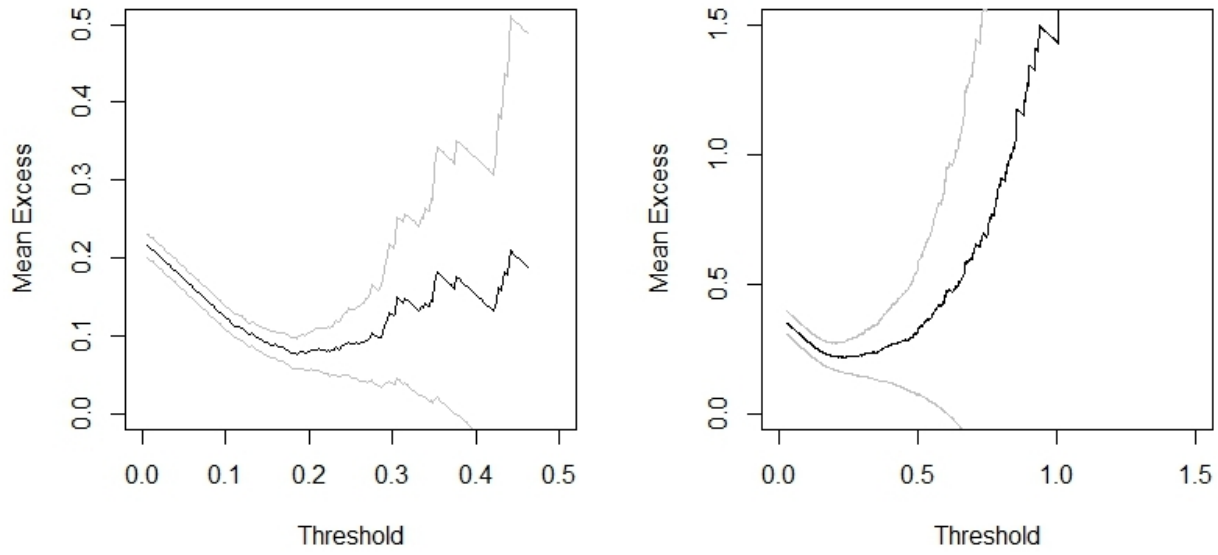


Figure A.8: Mean residual plots for $1/TTC_{min}$ (left) and $1/T_{2min}$ (right)