Thesis for the Degree of Licentiate of Engineering

On Safety Validation of Automated Driving Systems using Extreme Value Theory

Daniel Åsljung

Department of Electrical Engieneering
Chalmers University of Technology
Göteborg, Sweden 2017

On Safety Validation
of Automated Driving Systems
using Extreme Value Theory
Daniel Åsljung

Technical report number: R014/2017
ISSN 1403-266X

Department of Electrical Engieneering
Mechatronics Group
Chalmers University of Technology
SE–412 96 Goteborg
Sweden
Telephone: +46 (0)31 – 772 1000

Daniel Åsljung
Telephone: +46 (0)72 – 885 2652
Email: daniel.asljung@zenuity.com
asljung@chalmers.se

Typeset by the author using LaTeX.

Chalmers Reproservice
Göteborg, Sweden 2017

to Lai Jee

Abstract

Autonomous vehicles are expected to bring safer and more convenient transports in the future. When the system in the vehicle takes care of the driving, the driver is free to spend time on other things. As the driver is no longer part of the loop and cannot be used as a fallback, the requirements that are put on safety and dependability of the system will be very high. To test the system in real traffic and measure the failure rate that leads to an accident will therefore not be feasible. However, due to the complexity of the system, it is still desirable to be able to test the safety on a complete system level.

With the emergence of automated driving systems, the vehicles will be equipped with an array of sensors that gives a representation of the environment. This opens up the possibility to use more information to estimate how safe the system behaves in real traffic. Using an area of statistics called Extreme Value Theory, the frequency of near-collision can be extrapolated into a frequency of actual collisions.

These near-collisions are measured using threat assessment methods that have been developed for active safety applications. In this thesis, two types of measures are evaluated to determine how well they can be used for extrapolation. From the results, it is clear that the measure relating to a point where a collision is unavoidable works better than the one relating to the actual collision.

Furthermore, several methods for automatically fitting the extreme value model to the data are evaluated. The result shows that all tested methods work well where some methods put emphasis on the more extreme data, which can result in a difference of the inferences drawn. This suggests that the whole process has the possibility to be automated, which is necessary when performed repeatedly on multiple large data sets.

Keywords: Automotive, Autonomous Vehicles, Verification, Performance Evaluation.

Acknowledgments

I would like to start by giving a big thank you to my colleague and industrial supervisor at Zenuity, Dr. Jonas Nilsson, for your part in my recruitment and your constant support in the research. You have continuously provided a good discussion as well as new ideas and insights. I also would like to thank my former manager at Volvo Cars, Jonas Ekmark, for hiring me and giving me the opportunity to work on this project. You have always supported me and showed much interest in my research, which has been valuable to me. A special thank you also goes to Jonathan Jonsson for supporting me in the creation of the necessary tools to process the data and for the interesting discussion we have had.

I also want to express my gratitude to my academic supervisor at Chalmers, Prof. Jonas Fredriksson, for guiding me and for always providing support when I have needed it during the course of the project. Funding for this project has been received from Volvo Cars, Zenuity and FFI under the program of Traffic Safety and Automated Vehicles, which I am grateful for.

I would like to thank all my colleagues at Zenuity and Chalmers for creating two very stimulating workplaces to come to. There have always been good discussions going on that have created an inspiring environment.

Finally, my appreciation goes to Lai Jee for always being there to support me and for inspiring me to try new things.

Daniel Åsljung
Göteborg, December 2017

List of publications

This thesis is based on the following publications:

Paper 1

D. Åsljung, J. Nilsson and J. Fredriksson, Comparing Collision Threat Measures for Verification of Autonomous Vehicles using Extreme Value Theory, in *9th IFAC Symposium on Intelligent Autonomous Vehicles*, 2016, pages 57-62, Leipzig, Germany.

Paper 2

D. Åsljung, J. Nilsson and J. Fredriksson, Validation of Collision Frequency Estimation Using Extreme Value Theory, in *Proceedings of the IEEE Intelligent Transportation Systems Conference*, 2017, pages 1857-1862†, Yokohama, Japan.

Paper 3

D. Åsljung, J. Nilsson and J. Fredriksson, Using Extreme Value Theory for Vehicle Level Safety Validation and Implications for Autonomous Vehicles, Accepted for publication in *IEEE Transactions on Intelligent Vehicles*.

Contents

II Included papers

Part I

Introductory chapters

Chapter 1

Introduction

Autonomous vehicles are expected to bring many benefits to the traffic environment. Studies show that human errors are the cause for over 90% of the traffic accidents [1]. With the human taken out of the equation, there is a possibility to significantly reduce the number of accidents. It also enables the driver to do something else with the time in the vehicle. The vehicles could also drive without any passengers to enable relocation of taxi services and delivery of goods. Currently, there is a lot of effort put into developing autonomous vehicles. Many actors promise to have vehicles on a higher level of autonomy available to be used in some way during the coming years, e.g. [2–6].

The driver of an autonomous vehicle is effectively put out of the loop and cannot be used as a fallback plan when things go wrong. As a consequence, there will be very high dependability requirements in relation to safety. To know what these requirements are in practice, it has to be understood what safe behavior actually means. The vehicle needs to be able to handle traffic laws, but also rare road hazards that are hard to predict. Then there must be a strategy how to validate that the vehicle actually has reached the required level of safety. It is argued that to solve this problem, a large effort across many different domains has to be made [7].

1.1 Driver Assistance and Automated Driving

Advanced Driver Assistance Systems (ADAS) supports the driver and automates some type of control. The driver is still responsible for the vehicle and often have the possibility to override the function. There is also a limit of what the automated task can perform in order to ensure safe control in cooperation with the driver. The driver must monitor the system and also acts as a fallback in case there is a failure to the system. The most simple type of assistance systems has to the role of relieving the driver of one specific driving task. These are referred to as Level 1 automation according to the SAE J3016 standard [8]. An overview of the different levels of automation can be seen in Figure 1.1.

SAE Level	Name	Control of steering and acceleration	Monitors driving and environment	Fallback responsible	Capability of system
1	Driver Assistance	Human driver and system	Human driver	Human driver	n/a
2	Partial Automation	System	Human driver	Human driver	Limited scope
3	Conditional Automation	System	System	Human driver	Limited scope
4	High Automation	System	System	System	Limited scope
5	Full Automation	System	System	System	Full scope

Figure 1.1: A table illustrating the five levels of automation from the SAE J3016 standard. The different columns highlight where the responsibility lies within different areas for the respective level.

An example of a Level 1 system is Adaptive Cruise Control (ACC), which job is to control the acceleration and braking to maintain a certain gap to the vehicle in front. This relieves the driver of a substantial part of the driving. Lane Keeping Assistance (LKA) is another Level 1 function, which instead focuses on the steering and makes sure that the vehicle remains in the lane. If the vehicle detects that it is about to leave the lane, it can automatically steer the vehicle back into the lane. A detailed description of ACC and LKA, as well as other ADAS systems, can be found in [9].

ACC and LKA can be combined to one function controlling acceleration, deceleration, and steering. There are systems of this type that are in production in e.g. Mercedes' Drive Pilot, Tesla's Autopilot and Volvo's Pilot Assist. These systems are referred to as Level 2 automation or partial automation because the driver still needs to monitor the system and the environment.

1.1.1 Unsupervised automated driving

By moving to Level 3 and higher, you remove the driver's responsibility to monitor, which opens up the possibility to do other things while the car is driving. This is referred to as an unsupervised automated driving system. The function could be limited to special conditions such as weather and traffic. An example of this is a system which handles the driving in traffic jam scenarios during certain conditions.

When the vehicle is about to exit the scope of the function it hands back the control to the driver. If the driver does not take over, the system needs to have a backup plan that it can execute to put the vehicle in a safe state. The operational design domain (ODD) can be expanded to increase the capability of the system and include more driving scenarios. Ultimately, the vehicle can be driven autonomously without a driver present in all situations and conditions. This opens up for new models on how transportation can be carried out in the future.

1.1.2 Implication for system design

In the case of a simple ADAS function such as ACC, the scope is limited to keeping a certain distance to a vehicle in front. If there is no vehicle in front, the system should act as a regular cruise control, keeping a set speed. This function can be realized with a single radar sensor in the front, measuring position and speed of a possible vehicle. Out of the possible objects that are detected, it has to be selected which of them that is the target vehicle. Based on that, an action is taken to keep the set distance to that vehicle.

Suppose that the same function with the same ODD is to be developed, but now as an unsupervised function. The driver is no longer responsible for monitoring and not available as a fallback option. This would result in much higher requirements on perception to detect all possible objects that could be in front of the vehicle. That is because there is no longer a driver that monitors the road that can intervene if an object is missed. This might result in added sensors for redundancy that also has to be handled in the perception. Decision-making will also have higher requirements on interpreting the situation correctly, choosing the right target to follow. There will also be a requirement on vehicle control that guarantees the execution of a braking maneuver. To fulfill this it might be necessary to add a redundant braking system.

When the function's scope expands towards unsupervised automated driving and a complete ODD, the function needs to handle many more types of situations compared to the ACC case. This means that the environment, that the system should be designed to act in, will be much more complex. The implications of this on perception is that there will be high requirements to detect objects all around the vehicle and at long distances. To fulfill these requirements many more sensors need to be added that give a surround view of the environment around the vehicle. There will also be a need for redundant sensors at many places to reach the high level of robustness needed. For decision-making, there will be a lot more scenarios that should

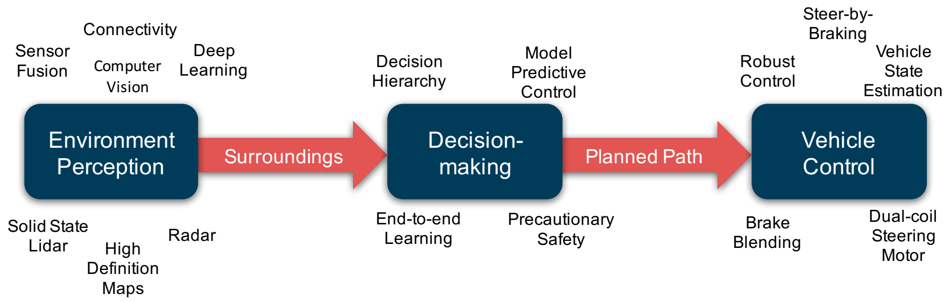

Figure 1.2: An illustration of the main building blocks for an automated driving system. Some parts that belongs to the blocks are mentioned to illustrate the complexity of safely bringing everything together.

be correctly interpreted and complex traffic scenarios with many different participants, which behavior needs to be predicted. There also needs to be decisions on multiple levels taking care of strategic as well as operational planning with logic determining what is currently the most important to safely reach the target. For vehicle control, the scope now also include steering, which probably needs to be redundant to guarantee high enough availability. The scope of actions that should be possible to actuate has also now increased to include a large variety of highly dynamical maneuvers. The end result is a highly complex system with very tough safety requirements that need to be handled by every part of the whole system, as illustrated in Figure 1.2.

1.2 Safe system design

In order to develop a complex system such as an automated driving system, one needs to define what needs to be developed, how it going to be implemented and to make sure that the system is doing what it is supposed to. This process falls under an area called systems engineering, which deals with how to design and manage this type of complex systems.

The process usually contains the steps of refining requirements, functional allocation, and physical implementation. Each of these steps then has to be verified at each level and validated against the top level requirements. In the automotive industry, this is done according to a framework called the V-model, shown in Figure 1.3, which is also a part of the ISO 26262 standard for functional safety [10].

To make the system behave in a safe way, possible failures have to be detected and mitigated. These failures include both hardware and software related faults and it needs to be shown that these are sufficiently rare

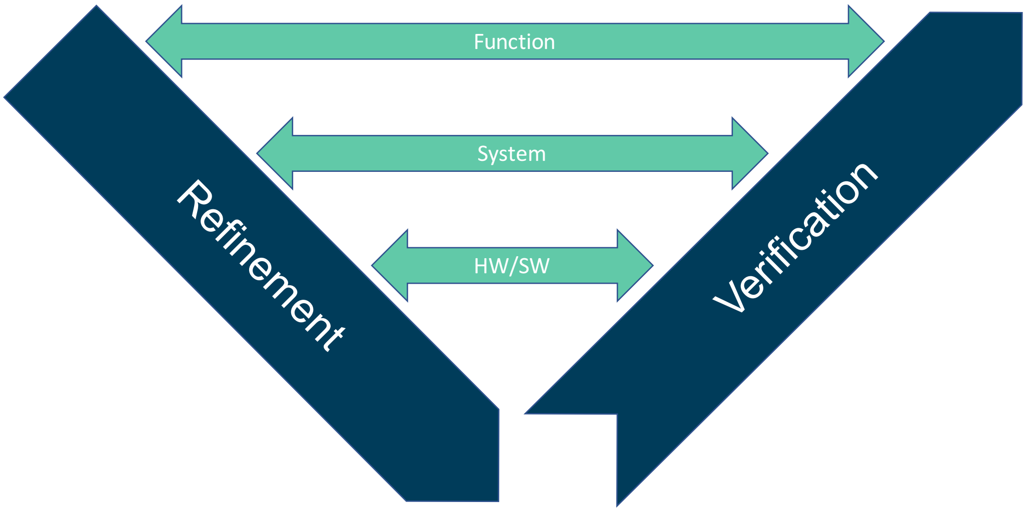

Figure 1.3: Figure of the V-model for software development. The left leg represent the refinement of requirements into implementation in hardware and software. The right leg consist of the verification of each step of refinement on the different levels of abstraction.

events. For an automated driving system, it is also important to ensure that the nominal performance of the system is good enough to ensure a safe operation. It must be designed to be safe when everything is working as intended.

The function could, for example, be designed so that the host should always keep a minimum distance to the vehicle in front. This distance could be insufficient in some situations in order to drive in a safe manner. Another critical area is the sensor performance, which includes, for example, technological limitations. An example of this is when a vision sensor has been trained on a data set that does not contain a certain type of object and therefore fails to classify it. The ISO 26262 standard does not explicitly describe how to extract and verify this type of requirements.

The development process of ISO 26262 starts with defining the item in question, which could represent a function. From the basis of the function, everything that can go wrong is investigated. These hazards are considered without the possible causes for these events and are classified a certain criticality level, called Automotive Safety Integrity Level (ASIL). The level which the hazard is classified as depends on the severity, exposure, and controllability of the situation. From the hazard analysis, safety goals are derived and they also inherit the respective ASIL classifications. All the safety goals need to completely cover all hazardous events for the respective item. The safety goals forms the vehicle level safety requirements that should be met in order to ensure a safe function. For an automated driving function, the safety goals might be more general to more broadly cover all situations, but that leads to more abstract formulations that are more difficult to verify [11]. This might require adding more abstraction layers in order to be able to show completeness between each layer.

These safety goals are then refined in multiple steps until there are requirements on specific hardware and software components. In each of these steps, it has to be verified that the requirements on the lower level fulfill the scope of the higher level, showing the completeness of the requirements. Each abstraction level also needs a verification strategy. This includes how to prove that the relationship between input and output of the model is implemented correctly in the product. For lower levels of the implementation, it is possible to show completeness and check all relationships. However, at the higher abstraction levels and for more complex systems, it becomes less practically possible to do so [9].

1.3 Problem formulation

The challenge of assuring safety for an automated driving function has given rise to the following questions: How to make sure that all safety goals are fulfilled? Are the safety goals correct and complete? The first question addresses the verification of the safety goals and also makes sure that the refinement of requirements is done correctly. By answering the second question, the safety goals are also validated that they cover all hazards which are connected to the item definition.

1.4 Delimitation

In this thesis, only the validation of vehicle level requirements called safety goals is considered. It is assumed that the refinement and verification of the lower level requirements are already performed. The validation of these safety goals is in this thesis delimited to only consider the situation of rear-end collisions. In order to validate the method, data based on human drivers have been used in order to be able to compare with a reference.

1.5 Contributions

This thesis presents a method to estimate the collision frequency of a vehicle using Extreme Value Theory (EVT). To enable this, a measure of the closeness to a collision is needed and in this thesis, two types of measures are evaluated, see Paper 1 and 3. The method also generates confidence intervals that take into account the uncertainty of the extrapolation, which can be used for safety validation purpose. Using data gathered from human drivers, the method is validated by comparing the results with data from crash statistics, see Paper 2 and 3. Several methods for automatically applying the EVT model on the data have been evaluated in Paper 3.

1.6 Outline

This thesis is made up out of two parts where Part I acts as an introduction to what is presented in Part II. In Part II there are three scientific papers, which are the base of the thesis. Part I provides background information and puts the appended papers into context with the following structure. In Chapter 1, the setting of the thesis is introduced by first describing an unsupervised automated driving function. It is then described what it takes to design this type of system in a safe way. This background is followed by a formulation of the problem that this thesis addresses and what delimitations have been made. In Chapter 2, different types of verification and validation methods are described. Chapter 3 provides an introduction to EVT and describes how it can be applied to traffic safety. In Chapter 4, the papers included in Part II are briefly summarized and in Chapter 5 the thesis is concluded with suggestions for further research.

Chapter 2

Verification and validation methods

In Chapter 1 it is described how the refined requirements of an automated driving function need to be verified on different levels. This is in order to ensure that the requirements have been implemented correctly and thereby create a safe function. There are several different approaches to verifying requirements that could be used to verify and validate the vehicle level safety goals. In this chapter, some methods are presented together with a description of their respective strengths and weaknesses.

2.1 Formal methods

Formal methods use mathematical models to verify that the system fulfills the requirements. They can be used in the whole development process from requirements engineering to implementation [12]. At the implementation level, the software is connected to mathematical contracts between input and program variables. With these mathematical models present, the code can also be automatically generated. In [13] reachability analysis and viability theory are used to formally verify a collision avoidance system. Unsafe and safe sets are computed to determine if an ideal system should intervene or not. Similarly, in [14] the safety of an autonomous vehicle has been verified using reachability analysis. The set all possible occupancies of the ego and surrounding vehicles are predicted. Mathematical models are used to consider all possible behaviors and uncertainties of sensors and actuators.

+ Powerful to mathematically prove that requirements are always fulfilled

− Need validated mathematical models of every part of the system

2.2 Statistical methods

To capture the stochastic behavior of the system due to the uncertainty of the sensor information, one can use stochastic verification methods. For estimating the frequency of failures, the system is often modeled as a Poisson process for the number of failures during a certain time. A confidence interval can be created to verify with a certain confidence that the failure rate is lower than the requirement. This is the basis for the proven in use argument in ISO 26262 [10]. An autonomous driving function has very tough requirements on failure rates, which leads to that a large amount of driving data is needed in order to verify them [15]. In order to get a representative sample of the driving, a real-world user profile is used as in [16, 17]. In these examples, statistical methods are used to verify that the false positive rate is sufficiently low. This can be done in a similar way to verify false negatives for sensor detection in the case of missed objects.

In the included papers, a statistical method using the theory presented in Chapter 3, is presented. This method utilizes more of the available data compared to Poisson statistics to verify similar requirements and therefore needs less amount of driving data.

+ Possibility of having a high content validity

− Requires a large amount of data for each new version of the system

2.3 Directed testing

For testing the performance of collision avoidance systems, directed testing on test tracks have been used in [16,17]. There, a number of scenarios based on real-world driving situations are tested. This is also done in several different weather and light conditions together with variations of similar situations. A benefit of using this method is that the whole system from sensors to actuators are used as it is implemented. It is also possible to test rare difficult scenarios repeatedly, which is not possible in real traffic.

With directed testing at a test track, it is hard to recreate variations of situations realistically. When using directed testing on a test track for verification, the worst-case scenarios are often tested. An example of how worst-case scenarios can be defined for a collisions avoidance system is found in [18]. It is in those situations where a system error is most likely and from there it can be argued that less challenging scenarios are also handled. However, for an autonomous vehicle, it is not obvious in many situations what is the worst-case situation and how to argue that all other situations are handled.

+ Effective when testing the system in extreme scenarios

− Difficult to define a complete set of test cases

2.4 Simulation

The aim of using simulation for verification is to test the system in closed-loop based on computer-generated inputs. Some parts of the system and the environment are then modeled to create as close to the real experience as possible. One type of simulation is Model-In-the-Loop (MIL), where the whole system is a model of what is implemented. Another type of simulation is called Software-In-the-Loop (SIL), which uses the actual implementation of the system in the simulation. Examples of implementations of MIL and SIL can be found in [17, 19, 20]. In both these types of simulation virtually generated scenarios are sent as input to the system. The scenarios can be generated from the specifications, but also based on what has been experienced in real traffic, as seen in [21]. One benefit of using this type of simulation instead of in real traffic is that it can be performed offline and done multiple times faster than real-time. It is also possible to control the process and test multiple variations of the same situation in a simple way.

Another type of simulation is called Hardware-In-the-Loop (HIL), which is the case when the software is run on the actual hardware in the vehicle, as seen in [22]. Thereby it is possible to test the system performance with both the software and actuators working together. However, the sensors still need to be modeled, which is a difficult task.

+ Can perform tests of scenarios much faster than in real-world and also test variations that have not been seen

− Needs to have validated models for the system and the environment

Chapter 3

Extreme Value Theory

Extreme Value Theory (EVT) is an area of statistics which focuses on the rare instead of the common events. It was first applied in the area of civil engineering to better understand the requirements for what structures need to be able to handle over a long period of time [23]. It provided a framework to describe the magnitude of forces that could be expected based on historical data. The framework of EVT contains a set of models that enable the usage of observed levels of data and extrapolate that into estimates of unobserved levels.

An example of how EVT is being used today is in the design of coastal defense barriers. You may have data on the sea level at the specific location for the last 10 years, but the barriers should be able to protect against high sea levels for maybe the next 100 years. EVT can then be used to model the observed sea levels from the last 10 years in order to estimate the highest expected sea level during the expected lifetime of the barrier.

3.1 Block Maxima

The statistical behavior that is modeled in the classical extreme value theory is the maximum, M_n, of a sequence of independent random variables.

$$M_n = \max\{X_1, ..., X_n\} \tag{3.1}$$

These measurements, $X_1, .., X_n$, could, for example, be daily measurements of sea-level, as visualized in Figure 3.1. The value M_n is then the maximum of these measurements during a certain time, for example, one year.

If the cumulative distribution F of the max value is known, this could be used to estimate the frequency of more rare events. In practice, the distribution F is unknown but can be approximated to a set of models based only on the extreme data [23]. This is similar to the normal approximation

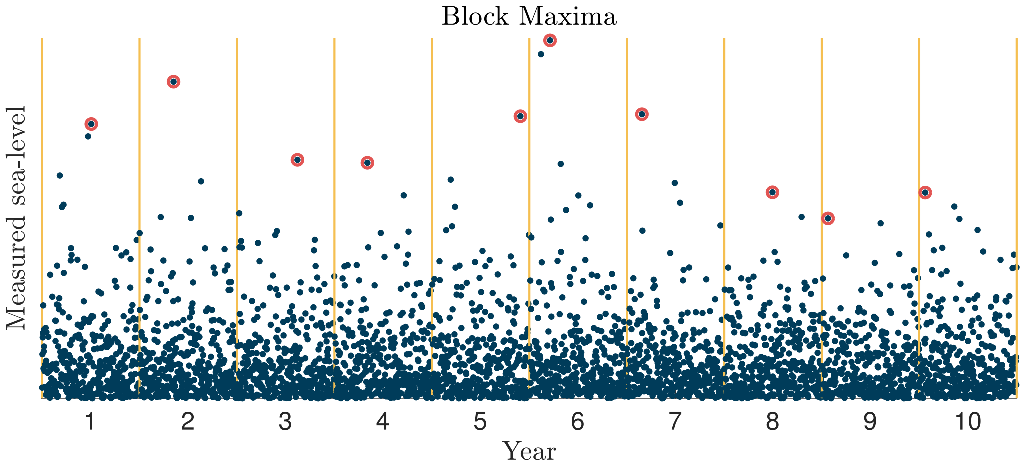

Figure 3.1: This figure illustrates how the block maxima values are selected in the example of daily sea-level measurements. The selected maximum values of each block are highlighted with a red ring.

of sample means, using the central limit theorem. The set of models can be represented by the Generalized Extreme Value (GEV) distribution, as seen in Figure 3.2.

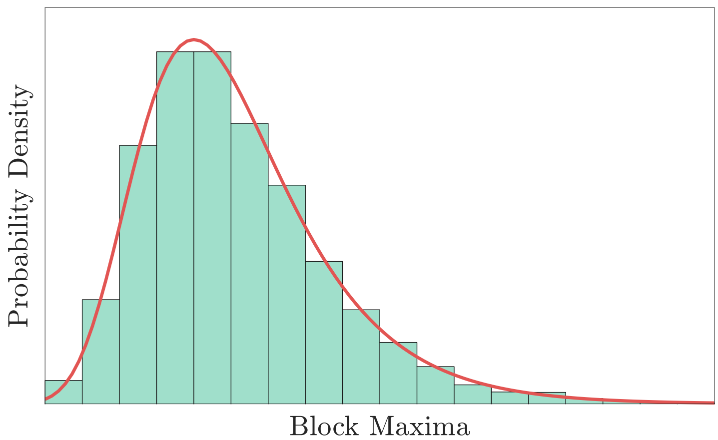

Figure 3.2: This figure illustrates how the GEV distribution is fitted to data. The probability density function for the distribution is shown as the red solid line. The values on the x-axis represent the maximum measurement from each block.

The distribution consists of the three parameters location (μ), shape (ξ) and scale (σ) with the following probability density function:

$$f(x|\xi,\sigma,\mu) = \frac{1}{\sigma}\exp\Big(-(1+\xi\frac{(x-\mu)}{\sigma})^{-\frac{1}{\xi}}\Big)(1+\xi\frac{(x-\mu)}{\sigma})^{-1-\frac{1}{\xi}}. \tag{3.2}$$

If data is collected for multiple years, a series of block maxima, $M_{n,1}, ..., M_{n,m}$, can be used to fit a GEV distribution. Then the probability that a yearly maximum is exceeding the value x_p can be found using the inverse cumulative distribution function:

$$p = 1 - F(x_p). \tag{3.3}$$

When implementing this model on a data set, the choice of block size can have a significant impact on the result. Choosing a too small block size leads to bias in the estimation due to the poor approximation of the limit theorem. A large block size will instead lead to few maxima and thereby large variance of the estimation. Another important aspect in choosing the block size is that maxima need to be equally distributed. Therefore, if there are seasonal differences in the measured variable, these need to have the same conditions in each block. Using block maxima could mean that a large part of the available data is wasted. This is especially true if many of the extreme events occur in the same block.

3.2 Peak Over Threshold

Another method is to avoid the blocking and instead only model the most extreme events that exceed some threshold, u, which is visualized in Figure 3.3. The k values that are exceeding the threshold, $x_i : x_i > u$, are called exceedances and are labeled $x_{(1)}, ..., x_{(k)}$.

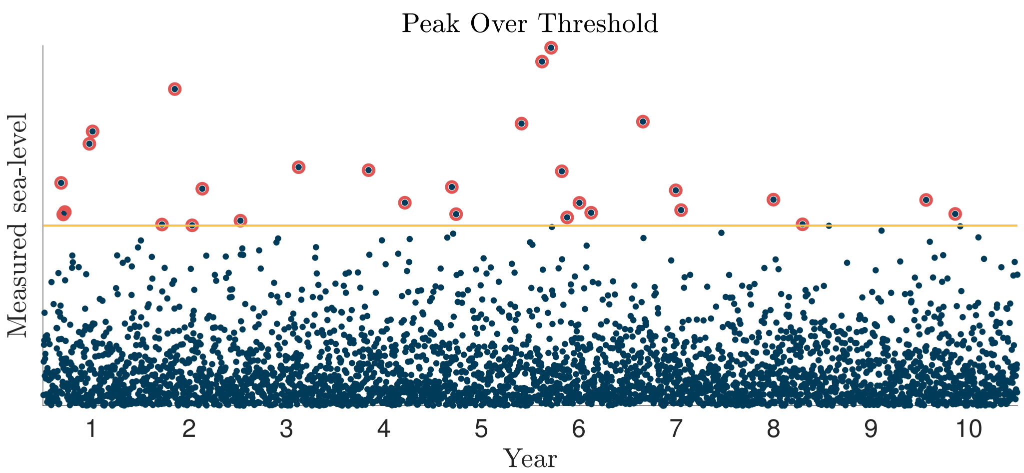

Figure 3.3: This figure illustrates how the exceedances are selected in the example of daily sea-level measurements. The selected peak values that exceed the threshold are highlighted with a red circle. The threshold is represented with a horizontal yellow line.

These values then belong to a distribution family called the Generalized Pareto (GP) Distribution as shown in Figure 3.4. The GP distribution consists of similar parameters as the GEV distribution, with shape (ξ), scale (σ) and threshold (μ). It has the following probability density function:

$$f(x|\xi,\sigma,\mu) = \frac{1}{\sigma}\big(1+\xi\frac{x-\mu}{\sigma}\big)^{-(1/\xi+1)}. \tag{3.4}$$

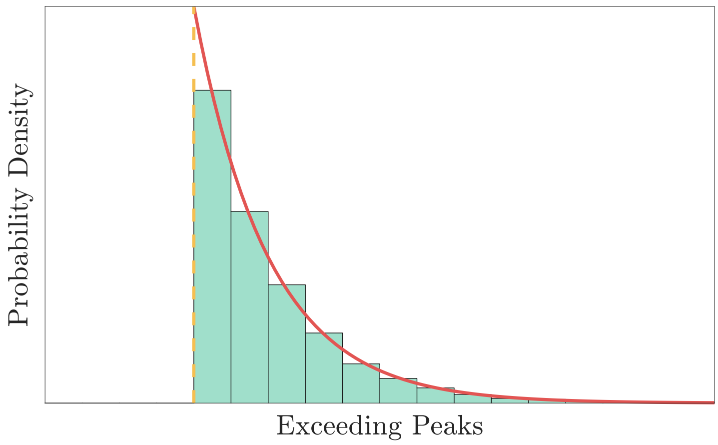

Figure 3.4: This figure illustrates how the GP distribution is fitted to all values exceeding a certain threshold. The threshold is represented by the dashed yellow line and the probability density function by the red solid line.

To avoid bias or high variance of the estimation, the threshold, u, is chosen as low as possible while still having a good fit to the model [23]. This is often done by manually inspecting the shape parameter for different choices of thresholds. When the shape parameter is constant, the estimation is stable, which indicates a good fit to the model. Finding a good threshold in practice can be difficult and often relies on experience.

The probability that a specific value is exceeded can be calculated similarly to the block maxima method. Suppose that $\zeta_u = \Pr\{X > u\}$, then the probability, p, that the value x_p is exceeded is:

$$p = \zeta_u\left(1 - F(x_p)\right). \tag{3.5}$$

3.3 Return Level

The probability, p, that is received for a certain value, x_p, can be used to find the average time between measurements that exceed this value. In EVT, this time is referred to as the return period and the corresponding sea-level value is called return level. Given a probability, the return period, t_p, can be found using the following formula:

$$t_p = \frac{t_{tot}}{np}, \tag{3.6}$$

where t_{tot} is the total time of data gathering and n is the number of blocks for the BM method or the total number of measurements for the POT method.

When the return level is plotted against different return periods, you get something similar to what is seen in Figure 3.5. You can also create confidence intervals of these estimates which takes into account the uncertainty of more extreme return levels that have not yet occurred.

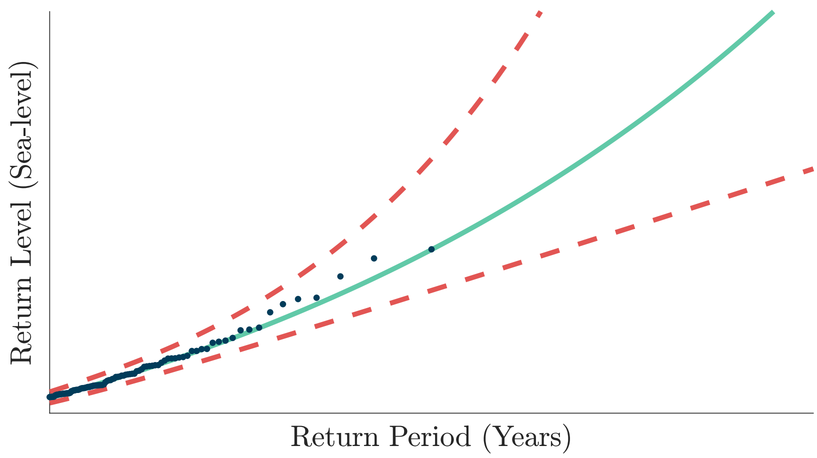

Figure 3.5: The figure illustrates how EVT can be used to estimate the sea-level that is expected to be exceeded once in a certain time interval (return period). The green solid line represents the most likely estimate, while the red dashed lines corresponds to a confidence interval of this estimate. The blue dots correspond to the measurements used to fit the EVT model, which are plotted along the estimate to show how well the model fits the data.

If one is interested in how often a certain value is exceeded, the answer would be the corresponding return period. This could be of interest to evaluate the effectiveness of a certain height for a seawall. The return period would then correspond to an estimate of how often the barrier would be flooded.

3.4 Application to vehicle safety

Extreme value methods have the possibility to estimate the frequency of events that have not yet occurred. This is done by extrapolating from the models fitted to the extreme data that has been recorded. For this to be possible for vehicle safety, there is a need for a measure that reflects the closeness to an accident for each given time instance. The measure also needs a definite value where a collision happens or is unavoidable.

Such measures have been developed in the active safety area for avoiding, for example, rear-end collisions with an auto-braking system. These measures are called threat assessment since they are used to decide if the situation is threatening enough for the collision avoidance system to activate. The main differences between these threat assessment methods are what dynamic model that is used for the host vehicles and the objects around it, and how their respective future actions and motions are predicted [24].

3.4.1 Deterministic threat assessment

Generally, a collision can be avoided in many different ways. The vehicle has the possibility to steer, brake and accelerate and there is a lot of combinations of these inputs. Therefore, threat assessment is often simplified for computational reasons. Deterministic threat assessment assumes a given model which gives one prediction that result in one specific value of the threat for a given moment. This often done for one of the vehicle's possible actions at a time. Below follows a description of some common deterministic threat assessment methods.

One of the most simple measures is the distance to an obstacle in the host's path ahead. This measure is called headway, p_{HW}, and for a straight road, it is equal to the radial distance. For a curved road, it is the distance that has to be traveled along the middle of the road to reach the object. This measure can also be expressed in time headway, t_{HW}, which is the time it takes for the host to reach the object's position. If the host's acceleration is zero, then:

$$t_{HW} = \frac{p_{HW}}{v_{0,host}}, \tag{3.7}$$

where $v_{0,host}$ is the initial speed of the host vehicle.

The headway measure relates to the exposure of a hazardous situation, i.e. how sensitive the host vehicle is to sudden events. However, the measure does not predict the future motions of the object, which becomes a problem if there is a high relative speed. A measure that handles this is the time to collision, t_{TTC}. It is often assumed that the acceleration of the host and the object is constant. This means that the t_{TTC} is found by solving:

$$0 = p_{x,0} + v_{x,0}t_{TTC} + \frac{a_{x,0}t_{TTC}^2}{2}, \tag{3.8}$$

where t_{TTC} is the lowest positive solution. This measure is directly related to the point of a collision. There are also measures such as required longitudinal acceleration, a_x, that reflects how much effort is needed to avoid a collision. This type of measure can also be related to the capacity of braking or acceleration. Assuming constant acceleration for both the host vehicle and the object, the required acceleration can be found by solving the following system of equations:

$$\begin{cases} 0 = v_{x,0} + a_x t, \\ 0 = p_{x,0} + v_{x,0}t + \frac{a_x t^2}{2}. \end{cases} \tag{3.9}$$

There is a difference between the measures presented here in how they characterize a threatening situation. The measure of TTC reflects the closeness in time of a predicted collision. Time headway does not predict a collision but instead relates to an obstacle-free distance, which is a conservative measure of the closeness to a collision. In the case of a standstill object or an object that stops instantly, these measures are very similar. The measure of required acceleration is different to the other two measures since it does not relate to a possible collision. Instead, it measures the action needed to avoid a collision and hence when a collision is practically unavoidable. Required acceleration, therefore, gives an earlier indication when a collision is happening compared to the other two measures.

3.4.2 Advanced threat assessment

Threat assessment methods such as these can be extended to include more detailed models for actuation of actions such as braking to make them more realistic. The simple models presented here only takes into account one target at a time, which sometimes underestimates the threat since some paths might be blocked by other objects. By including multiple objects in the threat assessment this can be mitigated but at the cost of increased complexity. There are also a lot of uncertainties in state measurement and prediction. This can be countered by introducing safety margins in the deterministic models or by using stochastic models instead.

Stochastic models of the uncertainties can give a more realistic measurement of the current risk. This can include both measurement uncertainties as well as to consider multiple future trajectories. Stochastic models can be applied to the measures presented in section 3.4.1. For TTC that would mean that the result will be a distribution of values instead of a single one, as seen in [24]. The result of using stochastic models can also be a probability of collision for each given instance, as shown in [25, 26]. This can be done by assuming stochastic models of the future paths and calculating the risk that an object will occupy the same place as the ego vehicle at the same time in the future. Another approach is to model the uncertainties of the measurements together with a model of the other traffic participants as in e.g. [27]. Then stochastic reachable sets can be used to predict the probability of collision for a certain path of the ego vehicle.

Chapter 4

Summary of Included Papers

This chapter provides a brief summary of the papers included in the thesis and also describes the contributions to each paper by the author of this thesis. Full versions of the papers are included in Part II.

Paper 1

D. Åsljung, J. Nilsson and J. Fredriksson, Comparing Collision Threat Measures for Verification of Autonomous Vehicles using Extreme Value Theory, in *9th IFAC Symposium on Intelligent Autonomous Vehicles*, 2016, pages 57-62, Leipzig, Germany.

As described in Chapter 3, there is a need for a measure that reflects the closeness to a collision in order to use EVT to estimate the collision frequency. The measure needs to be able to continuously show the closeness to a collision and comparable between different situations.

This paper investigates how different threat measures affect the inferences drawn from EVT. Two different types of threat measures are compared and a subset of a larger field test is used as input data, where the vehicles are driven by humans. The results show that there is a clear difference between the two types, especially when looking at the estimated collision frequency. The measure which shows the closeness to the point where a collision is unavoidable looks much more promising in that regard.

The thesis author was responsible for the problem formulation, implementation, analysis and writing the paper.

Paper 2

D. Åsljung, J. Nilsson and J. Fredriksson, Validation of Collision Frequency Estimation Using Extreme Value Theory, in *Proceed-*

ings of the IEEE Intelligent Transportation Systems Conference, 2017, pages 1857-1862†, Yokohama, Japan.

In Paper 1 it was shown that one type of measure showed greater promise of being able to estimate the collision frequency using EVT. In order to be used as a validation method for safety requirements, as described in Chapter 2, the method needs to be shown to correctly estimate the collision frequency.

To address this, the measure that was more promising is investigated more in Paper 2 . To validate the correctness of the estimation using EVT, it is compared to an estimate from crash statistics. For the comparison to be valid, the data used for the EVT estimate is from a larger field test made up of 250 000 km driven by humans. The results from this confirmed the initial conclusions from Paper 1 that this measure gives credible results. It was also found that the EVT model could be fitted in two different ways resulting in some differences in the inferences drawn. By fitting the model to a few of the most extreme events, the drivers' performance showed to be significantly better than the average human. The conclusion is that this is what can be expected from data based on trained test drivers.

The thesis author was responsible for the problem formulation, implementation, analysis and writing the paper.

Paper 3

D. Åsljung, J. Nilsson and J. Fredriksson, Using Extreme Value Theory for Vehicle Level Safety Validation and Implications for Autonomous Vehicles, Accepted for publication in *IEEE Transactions on Intelligent Vehicles.*

The analysis of different types of threat measures made in Paper 1 was done on a limited amount of data, which makes the results preliminary. In Paper 2 it was shown that depending on what threshold is used for the EVT model, the inferences drawn could differ. As described in Chapter 3, this process is often performed manually by visual inspection. In order to to be able to efficiently use EVT for validation of safety requirements, this has to be done automatically.

In Paper 3, the same larger field test as in Paper 2 is used to verify the result received from Paper 1. The result from this larger field test is very similar to what was found in Paper 1, which further strengthens the conclusions that a measure that reflects the closeness to a point where a collision is unavoidable is the better choice.

The Paper also includes an evaluation of three different methods of automatically choose a threshold for the EVT model. All methods choose a probable threshold for both measures, suggesting that the whole process can be automatically performed.

The thesis author was responsible for the problem formulation, implementation, analysis and writing the paper.

Chapter 5

Summary and Future Work

The attached papers present a method that can be used to validate the safety of a vehicle's driving. Data captured during real traffic driving is used to evaluate the closeness to a collision, which is extrapolated into a collision frequency using EVT. Different types of measures for the closeness to a collision, as well as methods to correctly fit the EVT model to the data, has been evaluated. Based on these results, the usage of EVT for safety validation looks promising.

The papers included in this thesis only considers rear-end collisions. In order to use EVT for safety validation, there is a need for a set of measures that considers all types of situations where a collision can occur. The closeness to a collision also needs to be comparable between two situations of equal threat.

The data that is used as input to the methods is gathered using sensors that interpret the surroundings. These interpretations will always have some errors compared to the real environment. It needs to be investigated how these errors affect the estimations and the inferences drawn from the results.

The vehicles that have been used for data collection in the papers have been driven by humans. A reason for this is to be able to have a reference to compare the results from the methods against. As a next step, data from vehicles being in some form of automation should be investigated. It needs to be validated that the applicability of the method does not change when automated vehicles are to be evaluated instead.

Bibliography

[1] S. Singh, "Critical reasons for crashes investigated in the National Motor Vehicle Crash Causation Survey. (Traffic Safety Facts Crash•Stats. Report No. DOT HS 812 115)," National Highway Traffic Safety Administration, Washington, DC, Tech. Rep. February, 2015.

[2] nuTonomy Inc., "nuTonomy to Test its Self-Driving Cars on Specific Public Roads in Boston, MA," 11 2016. [Online]. Available: http://nutonomy.com/images/nutonomy/nuTonomy_Media_Release_Boston_Testing_Launch.pdf

[3] Uber Technologies Inc., "Pittsburgh, your Self-Driving Uber is arriving now," 9 2016. [Online]. Available: https://newsroom.uber.com/pittsburgh-self-driving-uber/

[4] Tesla Inc., "All Tesla Cars Being Produced Now Have Full Self-Driving Hardware | Tesla," 10 2016. [Online]. Available: https://www.tesla.com/blog/all-tesla-cars-being-produced-now-have-full-self-driving-hardware

[5] Volvo Car Group, "Volvo Cars and Autoliv team up with NVIDIA to develop advanced systems for self-driving cars - Volvo Car Group Global Media Newsroom," 2017. [Online]. Available: https://www.media.volvocars.com/global/en-gb/media/pressreleases/209929

[6] Waymo LLC, "Apply to be part of Waymo's early rider program," 4 2017. [Online]. Available: https://medium.com/waymo/apply-to-be-part-of-waymos-early-rider-program-5fd996c7a86f

[7] P. Koopman and M. Wagner, "Autonomous Vehicle Safety: An Interdisciplinary Challenge," *IEEE Intelligent Transportation Systems Magazine*, vol. 9, no. 1, pp. 90–96, 2017.

[8] SAE International Surface Vehicle Recommended Practice, *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*, sep. 2016 ed. SAE Standard J3016.

[9] H. Winner, S. Hakuli, F. Lotz, and C. Singer, *Handbook of Driver Assistance Systems.* Springer, 2014.

[10] International Organizational for Standardization, *ISO 26262-8: Road vehicles — Functional safety.* International Organization for Standardization. Geneva, Switzerland., 2017.

[11] C. Bergenhem, R. Johansson, A. Söderberg, J. Nilsson, J. Tryggvesson, M. Törngren, and S. Ursing, "How to Reach Complete Safety Requirement Refinement for Autonomous Vehicles," in *CARS - Critical Automotive applications: Robustness & Safety*, 2015. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01190734

[12] J. Woodcock, P. G. Larsen, J. Bicarregui, and J. Fitzgerald, "Formal methods," *ACM Computing Surveys*, vol. 41, no. 4, pp. 1–36, 10 2009.

[13] J. Nilsson, J. Fredriksson, and A. Odblom, "Verification of Collision Avoidance Systems using Reachability Analysis," in *IFAC 19th World Congress*, Cape Town, South Africa, 2014.

[14] M. Althoff and J. M. Dolan, "Online Verification of Automated Road Vehicles Using Reachability Analysis," *IEEE Transactions on Robotics*, vol. 30, no. 4, pp. 903–918, 8 2014.

[15] N. Kalra and S. M. Paddock, "Driving to Safety: How Many Miles of Driving Would it Take to Demonstrate Autonomous Vehicle Reliability?" RAND Corporation, Tech. Rep., 2016.

[16] M. Distner, M. Bengtsson, T. Broberg, and L. Jakobsson, "City Safety-A System Addressing Rear-End Collisions At Low Speeds," in *21st Enhanced Safety Vehicles Conference*, Stuttgart, Germany, 2009.

[17] E. Coelingh, H. Lind, W. Birk, and D. Wetterberg, "Collision Warning with Auto Brake," in *FISITA World Congress*, Yokohama, Japan, 2006.

[18] J. Nilsson, A. C. E. Odblom, and J. Fredriksson, "Worst-Case Analysis of Automotive Collision Avoidance Systems," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 4, pp. 1899–1911, 4 2016. [Online]. Available: http://ieeexplore.ieee.org/document/7078887/

[19] J. Hillenbrand and K. Kroschel, "A Study on the Performance of Uncooperative Collision Mitigation Systems at Intersection-like Traffic Situations," in *IEEE Conference on Cybernetics and Intelligent Systems.* IEEE, 6 2006, pp. 1–6.

[20] D. Gruyer, S. Choi, C. Boussard, and B. D'Andrea-Novel, "From virtual to reality, how to prototype, test and evaluate new ADAS: Application to automatic car parking," *IEEE Intelligent Vehicles Symposium, Proceedings*, no. Iv, pp. 261–267, 2014.

[21] J. E. Stellet, M. R. Zofka, J. Schumacher, T. Schamm, F. Niewels, and J. M. Zöllner, "Testing of advanced driver assistance towards automated driving : A survey and taxonomy on existing approaches and open questions," *IEEE 18th International Conference on Intelligent Transportation Systems*, pp. 1455–1462, 2015.

[22] O. Gietelink, J. Ploeg, B. De Schutter, and M. Verhaegen, "Development of advanced driver assistance systems with vehicle hardware-in-the-loop simulations," *Vehicle System Dynamics*, vol. 44, no. 7, pp. 569–590, 2006.

[23] S. Coles, J. Bawa, L. Trenner, and P. Dorazio, *An introduction to statistical modeling of extreme values.* Springer-Verlag, London, UK, 2001, vol. 208.

[24] J. Jansson, "Collision Avoidance Theory with Application to Automotive Collision Mitigation," Ph.D. dissertation, Linköping University, 2005.

[25] A. Eidehall and L. Petersson, "Statistical Threat Assessment for General Road Scenes Using Monte Carlo Sampling," *IEEE Transactions on Intelligent Transportation Systems*, vol. 9, no. 1, pp. 137–147, 2008.

[26] D. Greene, J. Liu, J. Reich, Y. Hirokawa, A. Shinagawa, H. Ito, and T. Mikami, "An efficient computational architecture for a collision early-warning system for vehicles, pedestrians, and bicyclists," *IEEE Transactions on Intelligent Transportation Systems*, 2011.

[27] M. Althoff, O. Stursberg, and M. Buss, "Model-Based Probabilistic Collision Detection in Autonomous Driving," *IEEE Transactions on Intelligent Transportation Systems*, 6 2009.