REQUIREMENTS AND DATA SOURCES NEEDED FOR VALIDATION OF COMPONENT PROPERTIES AND PERFORMANCE IN SIMULATION BASED BENEFIT ASSESSMENT OF DRIVER ASSISTANCE TECHNOLOGIES

Mikael Ljung Aust[1]
Tim Gordon[2]
Daniel Blower[2]
Hemant Sardar[2]
Irene Isaksson-Hellman[1]
Jan Ivarsson[1]
Lotta Jakobsson[1]

[1]Volvo Cars, Sweden
[2]UMTRI, USA
Paper Number 09-0438

ABSTRACT

In one of the Advanced Crash Avoidance Technology (ACAT) projects, a computational simulation approach has been used to assess the potential benefit of three advanced Driver Assistance Technologies in a lane departure scenario. The main advantage of a computational simulation approach to driver assistance technologies evaluation is that a wide range of conditions can be explored at a comparatively low cost. Also, though multiple data sources related to traffic safety are available, few approaches make systematic and integrated use of them. Using them to validate simulation components provides a way of integrating data from various sources into a reusable format.

When using simulation, the properties of each simulated component need validation. The objective of this paper is to describe data requirements for component validation, as well as how data which meet the requirements has been identified and extracted. The basic approach of the project is to look at each simulated component and determine which of its properties influence scenario outcome. Data sources which provide input on those properties are identified, and data from them is extracted and prepared for use in the simulation. To achieve a high level of detail and accuracy for all components, data from multiple sources are used including crash databases, field operational tests, testing on test-tracks and driving simulator experiments.

The research conducted in this project shows that sufficient data can be obtained to validate the properties of the simulation components. There are limitations in available data for some sources which raises questions of representativity, but these can in principle be overcome by extended data collection. The research also shows that while extensive effort may have to go into validation the first time a simulation is developed, similar subsequent projects will require much less validation effort since the simulation components can be reused.

INTRODUCTION

This paper describes part of the research performed in one of the projects funded by NHTSA under the Advanced Crash Avoidance Technologies (ACAT) program [1], performed by a team of researchers from Volvo Cars, Ford and UMTRI (referred to as the VFU-team). The underlying purpose of the ACAT program has been to address gaps in current knowledge about the performance and likely effectiveness of new and emerging active safety technologies in reducing crash numbers.

The VFU-team has focused its work on three advanced driver assistance technologies, developed by Volvo Cars, which address crashes initiated through lane departures. These crashes include road departure crashes, head-on collisions, sideswipes, and other crash modes. The technologies are *Driver Alert Control* (DAC), *Lane Departure Warning* (LDW), and *Emergency Lane Assist* (ELA). *Driver Alert Control* is designed to estimate the impairment level of a driver and inform the driver of his/her impaired state, where impairment is assessed through quality of lane keeping over time. The driver is informed of his/her state so as to support a decision to continue to drive. *Lane Departure Warning* is aimed at warning the drivers if they are inadvertently drifting out of their lanes. Under such a scenario, LDW supports the driver by generating a warning. LDW will not take any automatic action to prevent a possible lane departure. Responsibility for the safe operation of the vehicle remains with the driver. *Emergency Lane Assist* relies on the detection of the vehicle position with respect to the road lane markings as well as detection of vehicles (both oncoming and those being overtaken) in the adjacent lanes. If a lane drift or lane change maneuver is commenced and this implies a risk for collision with an oncoming or overtaking vehicle, ELA applies a

torque to the steering wheel in order to prevent collision and return the vehicle to its original lane.

All of these driver assistance technologies aim at detecting degraded driving and to provide suitable information, warning, or intervention. Together they form a logical chain of warnings and interventions. DAC is expected to influence the exposure of drivers to episodes of drowsy driving and hence operates in the earliest phase. LDW and ELA are relevant to the evolution of vehicle kinematics during a conflict, and operate in the early and late conflict stages.
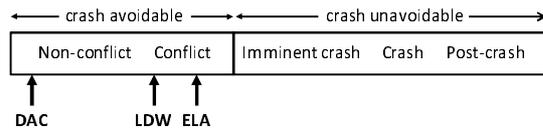

Figure 1. NHTSA crash phase timing.

THE VFU-TEAM APPROACH – COMPUTATIONAL SIMULATION

As described by NHTSA, an ACAT project should meet two goals. The first is to develop a formalized Safety Impact Methodology (SIM) tool to evaluate the ability of advanced crash avoidance technologies in full vehicle systems to solve specific motor vehicle safety problems. The second objective of the program is to demonstrate how the results of objective tests can be used by the SIM to establish the safety impact of a real driver assistance technology.

The VFU-team has chosen to address the first goal by developing a SIM tool which at the core uses a detailed mechanism-based (continuous-time simulation) approach to represent the potential influence of driver assistance technologies. The basic SIM procedure starts by exploring real-world crash mechanisms using both statistical and in-depth analysis of recorded crash events in order to understand contributory factors and event sequences, including the role of tiredness, distraction and judgment in actual crashes involving lane/road departure. This information is then used to develop a comprehensive set of Driving Scenarios (DS) which precede the crashes. The DS are then further parameterized in all aspects necessary to represent them in software via a computational model. This means that all components needed to evaluate the influence of a driver assistance technology in the DS (vehicle, driver, road environment and technology) are represented through computational sub-models interacting in a virtual environment rather than physical objects interacting in the real world. Following the definition and parameterization of the full DS set, multiple cases are sampled and run in a Monte-Carlo simulation. The computational

model time-steps from the starting point of each DS until the DS has run for a pre-defined time interval (for example 10 or 20 seconds).

The DS precede the crashes but they are not pre-crash scenarios in the sense that a crash inevitably follows from DS development. Rather, crashes may or may not result from any given DS as it develops over time (this applies both to real driving and simulations). The DS are thus "coarse-grain" in the sense that they cover a broad range of situations which include the ones that lead to crashes but also a number of situations where no crash occurs. In other words, rather than looking at single case accident reconstruction, the aim is to generate an ensemble of crash/no-crash situations, and then study whether the crash avoidance technology under evaluation changes the overall proportions of crash/non-crash outcome for this ensemble. Running the simulation for all DS's therefore results in two distributions of virtual conflicts and crashes, one with the technology and one without. These distributions are then mapped to real-world crash types and frequencies using some form of crash metric, as illustrated in Figure 2:
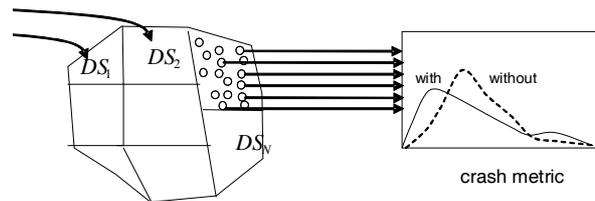
Figure 2. Driving Scenario resolution for Monte Carlo simulation with and without the specific safety technology being evaluated.

DATA REQUIREMENTS AND SOURCES IN A COMPUTATIONAL SIMULATION APPROACH

It can be argued that all approaches to evaluation of a driver assistance technology have the same representation issues which must be addressed. Basically, the characteristics of the four components necessary for evaluation (driver, vehicle, technology and evaluation environment) should be either identical or at least sufficiently similar to their counterparts in real world crash characteristics, otherwise evaluation results could be called into question.

Each of the components necessary for evaluation must be correctly represented in two aspects. One is the structural aspect. Taking the vehicle as an example, for the evaluation to be valid, one must first identify the characteristics of vehicles typically involved in the targeted crash type, and then find a

way to represent those characteristics in the actual evaluation. For physical measurements, this usually means bringing vehicles with those characteristics to the test track or out in the field. For computational simulation, it means implementing virtual vehicles with those characteristics in the simulation, with the general ability to interact with the other sub models.

The second aspect of representation is more functional or dynamic in character. Each component needs a definition of (1) its initial state at the beginning of the evaluation and (2) how it should respond to change. Again using the vehicle as an example, one must provide clear values for the vehicle's initial state in all test configurations (its initial speed, initial lane position, etc). One must also provide clear definitions of how it should react to, for example, a steering input from the driver (suspension settings, tire properties, etc).

In these representation issues, a computational simulation approach faces a somewhat different challenge compared to evaluations based on physical measurements. Due to the possibility of running tens of thousands simulations with different component configurations, representativeness is less of a problem in a simulation approach. While in physical measurements one usually must select just a few configurations to represent the crash problem due to limitations in resources and time, in a simulation approach one can run most (or even all) possible configurations. On the other hand, since driver, vehicle, environment and technology characteristics all are represented as sub-models in a virtual environment rather than through their physical counterparts, extensive work has to go into making these models and the environment act as they would have in real life in all relevant aspects.

The main challenge for a simulation approach therefore is one of validity rather than representativeness, and its outcome will depend on how well each sub-model represents its real life counterpart in the simulation of relevant aspects. This has consequences for how results from objective testing can be used in the SIM tool. Basically, to ensure that each sub-model represents its real life counterpart, all three sub-model aspects (functional structure, initial state and response to change) must be validated against real world data in some way. A substantial part of the VFU team's work has therefore been devoted to retrieving and processing the structure and performance data needed for such sub-model development and validation.

To achieve a high level of detail and accuracy for all components, data from multiple sources must be used. Though crash data from sources such as GES is a natural starting point for such work, and forms an essential part of defining the crash circumstances which the technologies under evaluation are meant to address, crash data in itself contains limited or no detail on a number of the pre-crash conditions or parameters which must be defined in order to perform reliable simulations.

To overcome some of the limitations of crash data, the methodology developed in this project has been to let crash data supply "one leg of the tripod", while the second and third leg is in naturalistic driving data and objective testing. Objective testing here refers both to testing of vehicle and technology performance ("technical testing") as well as to testing of human-technology interactions ("human factors testing").

This means that in relation to the second objective of the ACAT program (demonstrate how results of objective tests can be used by to establish the safety impact of a real driver assistance technology), the role of objective testing is driven towards calibration and validation of computational sub-models and their interaction in the simulation. Data sources used in this project include:

- Design information and algorithms associated with the driver assistance technologies

- Basic scientific knowledge about vehicle dynamics and driving dynamics

- Statistically valid crash databases and detailed investigations of crash causation (GES, CDS, etc.)

- Databases of naturalistic driving (obtained from previous Field Operational Tests)

- Databases of roadway characteristics

- Objective tests in the form of detailed technical tests of the vehicle and the driver assistance technologies, typically on a test track

- Objective tests designed to capture typical ranges of human performance where the driver is in the loop, typically on a test track or in a driving simulator.

In the following, the properties of each tripod leg will be described. The description will focus on how each tripod leg has been used to contribute to the development and validation of the computational sub-models.

THE FIRST LEG OF THE TRIPOD - DEFINING SUB-MODELS USING CRASH DATA

Data used directly to develop the sub-models used in the simulation include the National Automotive Sampling System General Estimates System (NASS GES) and NASS Crashworthiness Data System (CDS), crash data from the State of

Michigan, and roadway geometric information from the Highway Performance Monitoring System (HPMS). While HPMS is not strictly a crash data source, it is discussed here because its use is entwined with crash data.

Using Statistical Crash Data

The NASS GES is a nationally-representative sample of police-reported crashes, compiled by the National Center for Statistics and Analysis in the NHTSA. GES is a probability sample of motor vehicle crashes that occurred in the United States. The GES file covers crashes of all severities and all vehicle types. Police accident reports (PARs) are sampled from approximately 400 police jurisdictions within 60 primary sampling units and sent to a contractor for coding. The GES data includes a description of the crash environment, each vehicle and driver involved in a crash, and each person involved in a crash. GES data are coded entirely from police reports, without any supplemental investigation. Consequently, the data in GES is limited to what is available on a PAR. GES typically includes records for about 100,000 motor vehicles involved in 60,000 crashes.

In relation to sub-model development and validation, the crash data in the GES file is primarily useful to at a high level characterize the DS relevant to the technologies. This formed a very important part of the work in the project, because even if the crash data itself does not contain all details needed to run simulations, it provides a delimitation of the crash problem and thus the framework within which further parameters and details are necessary to work out.

The GES data include a set of variables that captures the sequence from just prior to the initiation of the "crash envelope" to the collision or other harm-inducing event. The crash envelope is defined as extending from the point in which the driver recognizes an impending danger or the vehicle was in an imminent path of collision with another vehicle, animal, or non-motorist to the point at which the driver either has successfully avoided the collision or the collision has occurred. Data elements record the vehicle maneuver immediately prior to the critical envelope (in the pre-crash maneuver variable), the event or condition that made the situation critical (critical event), the corrective action taken by the driver, and the stability of the vehicle after the maneuver. There is also an accident type variable that captures the relative position and movement of the vehicles leading to the first harmful event [2]. All of these variables appear in the GES and CDS data sets [3].

The approach to capturing crash events in GES (and CDS as well) is well-suited to a project focusing on evaluation of driver assistance technologies. Many other crash data systems focus on the first harmful event, or provide a sequence of events in the crash, which record the series of harmful events. But in crash avoidance research, information about the vehicle state prior to the initiation of the crash sequence and any harmful event is more interesting. The driver assistance technologies evaluated in this project all monitor vehicle position within the lane in normal driving, prior to any crash or conflict. Vehicle movement prior to the critical event (P_CRASH1) and critical event for this vehicle's first impact (P_CRASH2), in the GES file are therefore of primary interest in identifying the relevant crash types [4].

The identification of target crash types was accomplished primarily by the two variables, Vehicle movement prior to critical event (P_CRASH1) and critical event for this vehicle's first impact (P_CRASH2). However, a number of other variables were included to refine the identification of crashes that might be influenced by DAC, LDW or ELA. These variables record the number of vehicles in the crash, whether the vehicle was involved in the first harmful event in the crash, the travel speed of the vehicle, and whether the driver was under the influence of alcohol or drugs.

Based on the general crash characteristics identified as relevant to the DAC, LDW or ELA technologies, four dynamically-distinct crash types were identified as relevant to the technologies.

- Single-vehicle road departure

- Prior lane-keeping, lane departure

- Changing lanes, lane departure

- Other lane or road departure, prior lane-keeping or changing lanes

Furthermore, a number of vehicle, environmental and driver factors were examined in relation to the target crash types. The purpose was to identify factors associated with the crash types which could be used to specify the structural and functional aspects of the sub-models. For example, in relation to the environment, types of roadway, roadway alignment, weather, road surface conditions and light conditions were studied. In relation to the driver model, Driver fatigue was investigated.

It was found that road type and road curvature are stable and can be assumed to be reliably reported. Weather, road surface conditions, and light conditions are less stable but can still be considered sufficiently reliable in the crash data. Driver fatigue however is both very difficult to identify and transient. However, there is no feasible alternative source of information other than the crash data. While many cases of fatigue may be missed, it is assumed that the cases that are identified are true cases of fatigue.

In the end, a total of 36 crash scenarios were identified, accounting for 96.6 percent of target crashes. Of these, the top 25 crash scenarios (accounting for 90.5% of the target crash scenarios) were targeted for simulation using the Safety Impact Methodology (see Table 1).

Table 1.

Top 25 Crash Scenarios for Targeted Crash Types, From GES 2002-2006

Road type	Roadway alignment	Weather & road surface	Light condition	Driver fatigued	Percent	Rank
2 or more lanes, divided	Straight	Not adverse, dry	Daylight	No	19.7	1
2 or more lanes, undivided	Straight	Not adverse, dry	Daylight	No	9.9	2
2 or more lanes, divided	Straight	Not adverse, dry	Not daylight	No	9.1	3
2-lane, 2-way undivided	Straight	Not adverse, dry	Daylight	No	8.6	4
2-lane, 2-way undivided	Curve	Not adverse, dry	Daylight	No	5.8	5
2-lane, 2-way undivided	Straight	Not adverse, dry	Not daylight	No	5.2	6
2-lane, 2-way undivided	Curve	Not adverse, dry	Not daylight	No	4.1	7
2 or more lanes, undivided	Straight	Not adverse, dry	Not daylight	No	3.1	8
2 or more lanes, divided	Straight	Adverse, not dry	Daylight	No	2.5	9
2 or more lanes, divided	Curve	Not adverse, dry	Daylight	No	2.4	10
2-lane, 2-way undivided	Straight	Not adverse, dry	Daylight	Yes	1.9	11
2 or more lanes, divided	Straight	Adverse, not dry	Not daylight	No	1.9	12
2 or more lanes, divided	Straight	Not adverse, dry	Not daylight	Yes	1.8	13
2-lane, 2-way undivided	Curve	Adverse, not dry	Daylight	No	1.8	14
2 or more lanes, divided	Straight	Not adverse, dry	Daylight	Yes	1.7	15
2 or more lanes, divided	Curve	Not adverse, dry	Not daylight	No	1.6	16
2-lane, 2-way undivided	Straight	Adverse, not dry	Daylight	No	1.4	17
2 or more lanes, divided	Straight	Not adverse, not dry	Daylight	No	1.2	18
2 or more lanes, undivided	Straight	Adverse, Not dry	Daylight	No	1.2	19
2-lane, 2-way undivided	Straight	Not adverse, dry	Not daylight	Yes	1.2	20
2-lane, 2-way undivided	Straight	Adverse, not dry	Not daylight	No	1	21
2-lane, 2-way undivided	Curve	Adverse, not dry	Not daylight	No	1	22
2 or more lanes, undivided	Curve	Not adverse, dry	Daylight	No	1	23
2-lane, 2-way undivided	Straight	Not adverse, not dry	Daylight	No	0.8	24
2-lane, 2-way undivided	Curve	Not adverse, not dry	Daylight	No	0.7	25

Using In-depth Crash Data

The **NASS CDS** file was also used in the high level characterisation of the DS to identify relevant crash types. CDS is a data system complementary to the GES file. The CDS investigations go beyond the PARs to include on-site investigation and documentation of the scene, as well as measurements of all crash damage on the vehicles, extensive documentation of crash injuries using hospital and other records, and estimates of the change in velocity (delta *v*) for each vehicle in the

crash, where possible. Case materials for individual CDS crashes are available on the internet [5].

Since CDS uses the same set of variables and definitions as GES, the CDS cases can be used as a sample of the types of events that would be selected in GES. CDS cases that met the selection algorithms developed for GES were reviewed to see if these crashes had the characteristics relevant for the technologies. Because the original case materials for the GES cases are not available for review, the ability to review more in-depth cases in CDS also provided valuable insight into how crashes are classified. The review confirmed that the algorithms developed for GES identified the appropriate crashes.

Using Highway Performance Monitoring Data

Detailed roadway geometric data are needed to make a correct sub-model regarding the characteristics of the roadway where lane/road departure crashes occur, such as lane widths, lane markings, radius of curvature, shoulder width, and shoulder type. This information is not available in the GES data, which only distinguishes curved from straight roads, the number of travel lanes, and a few other details. The more detailed information is however available in roadway inventory files, such as the Highway Performance Monitoring System (HPMS) data on the road system. Highway Performance Monitoring System (HPMS) Data represents the national highway system and includes data on the extent, condition, performance, use and operating characteristics of the Nation's highways [6].

However, because GES crashes are not geolocated (i.e., located using a standard geographic reference system such as longitude and latitude), it is not possible to link the GES crashes directly to a roadway inventory file. Accordingly, it is necessary to obtain descriptions of the roadway geometry for the target crash types from some other source. The Michigan Crash files was able to provide a link to such data, because all crashes in Michigan are geolocated (located by latitude and longitude) and those locations can be linked to roadway inventory data, specifically Highway Performance Monitoring System (HPMS) Data.

The Michigan crash data captures information on crashes entered by police officers on the Michigan crash report, and includes data on all reportable crashes involving a motor vehicle. Reportable crashes involve a motor vehicle in transport on a roadway resulting in a fatality, injury, or property damage of $1,000 or more. This standard is reasonably comparable with reporting standards in most other states and thus with crashes in the GES file. Many of the variables in the Michigan crash file that describe the crash environment, such as

weather, road condition, number of lanes, speed limit, and travel speed, are compatible with similar variables in GES, even though it was necessary to develop comparable selection criteria for other variables.

The Michigan crash file was analyzed to identify a set of crashes which matched as close as possible to the target crash types identified in the GES file. In addition, selected standard characteristics of the Michigan road system were compared with national distributions. The purpose of these comparisons was to demonstrate that the Michigan roadway system is reasonably comparable with the national road system. A method was then developed to match the crash scenarios on the sampled segments in Michigan to national estimates from GES. This made it possible to identify detailed crash scenarios relevant to the DAC/LDW/ELA technologies in the Michigan data which are comparable to those in the GES crash data. Since all police-reported crashes in Michigan are geolocated, the identified crashes and their crash sites could be linked to the HPMS roadway files and the detailed roadway information they contain, thus providing a very detailed geometric description of the roadway at those crash locations. This description could then be used to calibrate and validate the roadway sub-model in the simulation.

THE SECOND LEG OF THE TRIPOD - DEFINING SUB-MODELS USING NATURALISTIC DATA

It is clear that all the parameters needed for a reasonably comprehensive representation of the DS needed for the SIM simulation are not available directly from GES crash data. For example, information about vehicle kinematics such as specific speeds, yaw rates, lane positions, etc, just prior to a road departure is not captured in crash databases. However, this information is needed in order to develop proper sub-models which can simulate the potential benefits of the driver assistance technologies as envisioned in the Safety Impact Methodology (SIM). Hence these must be derived from other sources, specifically from naturalistic driving data.

Naturalistic data from UMTRI's RDCW (Road Departure Crash Warning) Field Operational Test (FOT) was found to be sufficiently comprehensive in terms of the data for the purposes of populating the relevant elements of the scenarios. The RDCW FOT collected data from 78 drivers distributed evenly by gender and within three age groups. The total distance traveled was 83,000 miles, covering almost 2,500 hours and over 11,000 separate trips spanning a 10-month window that included summer, fall, and winter weather. The drivers used 11 specially instrumented passenger sedans equipped with the RDCW safety technologies

being evaluated and UMTRI's data acquisition system. The RDCW safety technologies targeted crashes involving vehicles that drift off the road edge or into occupied adjacent lanes, as well as those involving vehicles traveling too quickly into turns for the driver to maintain control. A detailed report on the FOT results, including technology effectiveness, driver responses, and other findings was submitted to the US DOT at the conclusion of the FOT [7].

It is clear that the distributions for parameters such as vehicle speeds and lane positions need only encompass the universe of the relevant crashes of interest (i.e., those crashes relevant to the driver assistance technology that is being simulated), and not the entire universe of driving behaviour. Therefore, the data mining from naturalistic databases was done using the high level DS characterisation obtained from crash data as the input variables (see Table 1 above).

A key element of data mining from naturalistic data is to ensure that the parameter distributions exclude driving for which the dominant crash type is non-technology-relevant. For example, in high traffic situations, the probability of a rear-end crash occurring because of multi-vehicle interactions (two or more) is higher than crashes due to single vehicle road departures, so high traffic situations should be excluded from the mining process.

The RDCW database includes a traffic density parameter that can be used to filter out high traffic situations. Additional resolution of traffic densities can be determined by looking at various other vehicle parameters such as the combination of repetitive braking and acceleration, inappropriate following distances and vehicle speeds not commensurate with roadway types (e.g. too slow for freeway driving).

Figure 3 shows the distributions of some of the key parameters for Crash Scenario 1 as identified from GES in Table 1 above. This scenario consists mainly of driving on limited access roads that are two or more lanes with divided medians (e.g. freeways & interstates), no adverse weather conditions, straight and dry roads, with the crashes occurring during daylight hours when the driver was not fatigued.
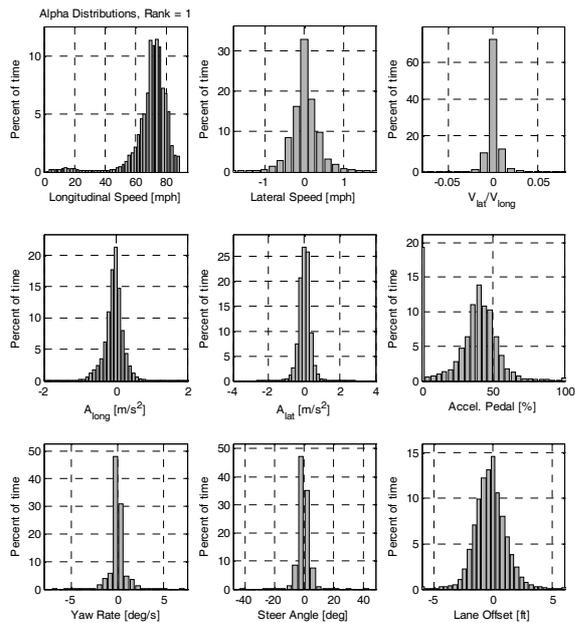
Figure 3. Distributions of vehicle parameters for Crash Scenario 1 in Table 1 above.

THE THIRD LEG OF THE TRIPOD - DEFINING THE SUB-MODELS USING OBJECTIVE TESTING

While GES crash data and naturalistic driving data are very valuable data sources, some characteristics of the sub-models can only be obtained through appropriate objective testing.

This in particular concerns the functional or dynamic aspects of the driver assistance technologies and the driver, and is driven by sub-model implementation issues. For the driver assistance technologies, there are two main ways of implementing them in the simulation. One can either use the actual technology software, or create a sub-model which replicates the technology's behaviour without being identical to the technology itself. Regardless of choice in this regard, because the virtual world in which the technology will run will not represent the full complexity of the real world, the performance of the technology in that virtual world needs to be calibrated against real world technology performance.

If this is not done, there is a risk that the technologies will over-perform during evaluation, since they would operate under the somewhat idealised conditions which exist in the computational simulation, and the driver model would not capture the behaviour of real drivers. For example, if the lane markings which a LDW technology depends on for lane tracking always are perfectly visible in the virtual world, the LDW sub-model will always have perfect lane tracking in the simulation. However, lane markings in the real world sometimes are faded or missing entirely, and LDW availability is therefore less than 100 percent in the real world, depending on lane marking quality. The LDW sub-model used in the simulation must therefore be calibrated to match the performance of a LDW running on real roads.

The same is true for the driver sub-model. In this case, the option of implementing the actual driver software is not available. Driver behaviour therefore has to be represented by a sub-model which imitates relevant aspects of driver behaviour in the relevant DS, such as inputs to brake pedal and steering wheel, without therefore claiming to represent the actual structure of human emotional, cognitive and/or motor processing. Calibrating the driver sub-model's performance in terms of action and reactions to the performance of real drivers is therefore very important for the validity of evaluation results.

Of course, the same principle applies also to the vehicle sub-model. However, in relation to the driver assistance technologies evaluated in this ACAT project, the performance of the vehicle sub-model is a smaller issue. Since none of the technologies are intended for scenarios associated with dynamic instability (i.e. the vehicle somehow starts to skid or lose traction), the influence of vehicle dynamics on technology performance will be limited. For example, the intervention provided by ELA (steering the vehicle back into the original lane if there is a risk of collision with a vehicle in an adjacent lane) involves only low lateral accelerations and speeds in order to avoid the risk of dynamic instability.

Regardless of whether the testing is performed in a driving simulator, on a test track or on public roads, the objective tests can be said to come in two forms. One is in the form of detailed technical tests of the vehicle and its driver assistance technologies (technical testing). The other is objective tests designed to capture typical ranges of human performance where the driver is in the loop (HMI testing). The technical tests are used to calibrate and validate performance of the simulation relative to the driver assistance technologies, while the HMI tests are used in a similar way to calibrate the simulation for the driver performance. Note that in the HMI tests, there is a great deal of variability in performance, so the driver sub-model developed in the project attempts to capture a range of driving behaviors rather than just specific values in single recorded events.

Physical tests were conducted on the track at Volvo in Sweden as well as on the field, and in Ford's VIRtual Test Track EXperiment (VIRTTEX), a hydraulically powered, 6-degrees-of-freedom moving base driving simulator [8-11]. The main goal of that testing was to retrieve relevant data for

calibrating the technology and driver sub-models performance in the simulation to real world technology and driver performance. The main focus of the track and the field testing was to validate the physical performance envelope of LDW and ELA. Also, some track testing with naïve subjects took place at Volvo, as part of the evaluation of DAC. In the driving simulator, the emphasis was on human factors tests with naïve subjects (e.g. distracted and drowsy driver tests with LDW) though again some controlled technical tests were included.

Technical Testing for the Technology Sub-Models

To calibrate the technology sub-models against real world performance, data from objective testing should cover true positive performance, false positive performance and availability. Testing the true positive performance of a driver assistance technology means establishing the extent to which it correctly detects and acts in the situation it was developed to address. For example, true positive performance for LDW can be measured as how often a LDW technology produces a lane departure warning when a lane departure in fact is occurring.

Testing the false positive performance of a driver assistance technology means establishing the extent to which it "cries wolf", i.e. the technology informs, warns or intervenes in situations which are not of the targeted type. Or put another way, false positives occur when the technology generates alerts that would not be seen as helpful by the driver. For example, false positive performance for LDW can be measured as how often LDW produces a lane departure warning even though no lane departure is about to occur.

Testing for availability means establishing for the extent to which a technology is able to function as intended under various road conditions. Technology availability can be defined as the percentage of time during a test drive that the technology is active and operable relative to the total drive time. Availability needs to be evaluated on a variety of roads under various environmental conditions. For example, testing availability for LDW could be to drive one or more vehicles equipped with LDW in the field under a range of weather conditions on different road types, while recording for which portions of the drives the LDW has sufficiently robust lane tracking to be able to detect a lane departure.

All these three areas must be covered, since they all affect the performance of the driver assistance technology, and thus the technology sub-model when integrated into the simulation. Availability can be used to determine for which DS one can expect the technology to be available. True positive performance indicates how the technology can be expected to perform within those DS it is available in. Finally, the number of false positives is a key indicator of the degree to which a driver may come to trust and rely on the technology, and therefore important for the tuning of how often, how fast and how much the driver sub-model should respond to an alert from the technology.

HMI-Testing for the Driver Sub-Model

In relation to development and validation of the driver sub model, data on several driver performance aspects is needed. These aspects can be split into two main categories: driver performance in the conflict driving phase, and driver performance in the non-conflict driving phase. These two categories have slightly different focus in the types of HMI testing needed to calibrate and validate the performance of the driver sub-model.

Driver Performance in the Non-Conflict Driving Phase - Driver Alert Control (DAC) is intended to elicit a response from the driver to take a rest break soon or let another driver take over, based on the technology-inferred "driver state" as determined by a broad set of sensor data. Basically, the DAC acts as a monitor for the driver in that it is analyzing vehicle state data and evaluating how well the car is being "controlled" by the driver. A warning signal is issued to the driver based on predicted future vehicle states.

The key point here is that the DAC provides a warning to the driver in the non-conflict driving phase. The effects of DAC are therefore best represented in the simulation by estimates of the probability of driver compliance with the recommendation to take a rest break. Driver compliance deals with the effect of a warning, which may result in a variety of driver actions. At best, the driver will take some form of action to avoid a potential conflict driving phase in the future, for example, taking a break from driving or switch drivers if that is possible. At worst, a fatigued driver could ignore the DAC warning altogether. Depending on the level of driver compliance, the alert may reduce the frequency of drowsy driving scenarios, thus reducing crash risk.

While testing the true and false positive performance of the DAC technology still forms an integral part of the technical testing, the biggest challenge in objective testing for DAC is to find ways of establishing determinants for compliance and rates of compliance with DAC warnings. This challenge is not easily met. There are many factors which may influence compliance with DAC warnings, including the driver's perceived urgency in reaching a certain destination, the physical possibilities of actually taking a rest break or switching driver (finding a suitable place to stop at

reasonably close), whether the driver thinks the assessment made by DAC is correct, etc.

A full evaluation of all these factors were determined to be outside the scope of the current project both in terms of time and resources. However, some limited testing was carried out to determine drivers' responses to DAC warnings. A test track HMI clinic with drowsy drivers was performed. In this clinic all but one of the drivers received a DAC warning during the drive. When asked to give feedback on how they perceived this warning (questionnaire study) a large majority of the drivers felt the feedback from DAC was useful. They also reported that the DAC feedback influenced how they drove and that it made them more awake. Some of the drivers were surprised by the feedback from DAC. The drivers did not perceive the feedback as annoying or frightening.

So far, the testing results therefore seem to indicate that drivers who receive a DAC warning will be motivated to stop and do something about their drowsiness. If this is the case, then drivers will take action before the driving situation enters a conflict phase. This means that the DAC technology basically can be treated as a filter in the SIM-tool; by having the technology in the vehicle the number of drivers experiencing an unintended lane departure due to drowsiness will be reduced to a substantial degree.

Driver performance in the conflict driving phase - In relation to basic driver performance in a conflict driving phase, the driver sub-model needs to capture and be calibrated for two main aspects of driver behaviour. One is typical driver reaction times for the type of conflict evaluated. For this project, this means that objective testing must be carried out to determine how long it takes before a driver begins a steering correction when discovering or being warned of a lane departure. The other main aspect is the intensity and speed of the driver response, i.e. how fast a driver steers back into the lane when correcting for an unintended lane departure. This also has to be determined through objective testing in order to provide driver sub-model development and validation data.

For evaluation purposes, the driver sub-model also must be able to represent a range of driver behaviours rather than a single average behaviour. For example, in this project it was found that crash data commonly cites driver fatigue as a contributing factor underlying unintended lane departures. For a correct evaluation, the influence of that factor should be possible to represent in the driver sub-model along with the typical behaviour of alert, non-drowsy drivers. Put slightly differently, it must be possible to tune one or more parameters which influence the driver sub-model's

control over the vehicle in a manner which can be made consistent with both the driving performance displayed by drowsy drivers, as well as alert and non-drowsy drivers.

To exemplify, in this project, one way in which the influence of fatigue was captured in the driver sub-model was by including variable time delays in lane-keeping control process. For example, to represent visual distraction (the driver closing his eyes in a micro sleep), one can introduce a delay in the driver sub-model's processing of visual information. More specifically, if the driver sub-model in a non-distracted state responds to new information on lane boundaries as soon as it is given, then in a drowsy state, a time delay is introduced before new information on lane boundaries is processed, even though the main simulation process keeps on running. This means that the driver sub-model in its drowsy state will begin a steering correction calculation later than it would in its non-drowsy state.

To calibrate and validate the time delay settings in the driver sub-model, as well as the corresponding behaviour of alert drivers, testing of drivers' responses to imminent lane departure events with and without driver assistance technologies activated were needed. A number of driving simulator studies were carried out, focusing on both alert and drowsy driver's reactions to, and acceptance of, different HMI solutions in lane departure situations. Participants drove under a variety of simulated conditions including night and daytime driving on interstate roads, narrower city roads and country roads. In order to increase the number of situations that activated LDW, artificial "yaw deviations" were introduced, sometimes in combination with secondary tasks that increased the likelihood of driver distraction [12].

From the log files of these studies, data could be extracted to determine a typical range of alert and drowsy driver reaction times and response types (e.g., steering/braking input to the vehicle, maximum lateral exceedence, etc) to a lane departure event. Since the log files from VIRTTEX include both the no warning condition as well as drivers getting a warning, typical responses could be established both for drivers with and without the technology available.

A further aspect of driver sub-model calibration and validation for this project concerns driver response to an ELA intervention. The basic question is whether drivers will interfere with the intervention in a way which counteracts what the technology is designed to do. Interference may be more or less deliberate, for example, if drivers perceive the steering input from ELA to be some sort of vehicle malfunction rather than a driver assistance intervention, drivers may fight the

steering torque applied by ELA. If this is the case, then ELA will not be able to successfully intervene in all situations it could handle if drivers did not interfere.

Some limited testing for this type of driver interference was carried out on a test track. The results indicate that drivers' acceptance of the ELA intervention is good. There was a high approval rate among test drivers for the intervention provided, and no indications of driver interference with the intervention. More testing needs to be carried out to gain a fuller understanding of driver responses to this type of intervention, but for this project, it was decided to assume that driver interference is of limited import for the evaluation of ELA.

DISCUSSION

It has not been an explicit goal of this project to provide a universal analysis tool for estimating safety benefits for all vehicle safety technologies, or even all active safety technologies. The goals of the study are already ambitious, so the VFU-team's approach has been to focus on specific technologies and to develop an approach that seems appropriate for those. For a different technology a slightly different approach might be preferred. The current technologies operate in the non-conflict phase or early in the conflict phase and involve a relatively high degree of technology interaction with the driver. If for another technology (e.g. frontal crash mitigation by automatic braking) there is little interaction (technology performance is largely unaffected by driver actions) then fewer DS may need to be explored.

The research goals have also not explicitly included effects of vehicle type, driver age and skill. For the vehicle a single target vehicle type (mid-sized sedan class) was adopted and assumed to be "representative" in some sense. Also, though driver behaviour has been derived from a range of test subjects (including a wide age range and both male and female drivers) the population has not been resolved further in this study.

The computational simulation approach as described above is quite complex, and certainly contains more elements than simply "test and evaluate". This is because the crash environment – incorporating interaction between driver, vehicle, driver assistance technology and environment – is itself complex. The best approach to such complex problems seem to be by pooling and integrating available data sources rather than fixing on a single data source.

While extensive effort may have to go into sub-model validation the first time a computational simulation approach is developed, it has the advantage of being highly reusable; if a new technology addressing a particular crash type needs evaluation, only a few sub-models, or parameters of the sub-models, need to be updated, the rest can be largely reused.

Apart from reusability, this modularity will also allow for a great deal of future enhancement and refinement, as research in any of the sub-model areas can be applied to that sub-model without having to change the overall structure of the SIM. Future research and findings can thus easily be integrated into the basic version of the SIM.

CONCLUSIONS

When developing a computational simulation approach, the relevant properties of each simulated component need validation, and the major challenge is not so much achieving representativity as validity. The experiences from the current project show that in order to achieve a high level of detail and accuracy in the development and validation for all simulated components, data from multiple sources must be used. These sources include crash databases, naturalistic driving data from field operational tests, objective testing on test-tracks and driving simulator experiments.

The research also shows that sufficient data can be obtained to validate the properties of the simulation components. There are limitations in available data for some sources which may raise questions of representativity, such as for some of the objective testing of driver compliance with alerts and warnings issued by the driver assistance technologies. In principle though, these can be overcome by extended data collection.

Though multiple data sources related to traffic safety are available, few approaches make systematic and integrated use of them. Using them to validate simulation components provides a way of integrating data from various sources into a reusable format.

While extensive effort may have to go into sub-model validation the first time a simulation is developed, subsequent projects will require much less validation effort since the simulation components can be reused.

REFERENCES

[1]   Advanced Crash Avoidance Technologies (ACAT) program, funded by DOT/National Highway Traffic Safety Administration, DTNH22-06-H-00011

 [2] *National Automotive Sampling System (NASS) General Estimates System (GES) Analytical User's Manual, 1988-2007.* U.S. Department of Transportation, National Highway Traffic Safety Administration.

[3] For a detailed discussion of the approach to coding variables in the crash envelope, see *General Estimates System Coding and Editing Manual, 2006*. US DOT, NHTSA. See especially pp. 320-335.

[4] This approach is analogous to aspects of the "Universal Description," described in Burgett, et al., *A Methodology for Estimating Potential Safety Benefits for Pre-Production Driver Assistance Systems*. U.S. DOT, National Highway Traffic Safety Administration. Washington DC. DOT HS 810 945. May 2008. The approach taken here, however, begins with the pre-crash maneuver and is tailored to the specific operations and capabilities of the technologies being evaluated.

[5] http://www-nass.nhtsa.dot.gov/BIN/ NASSCASELIST.EXE/SETFILTER

[6] Federal Highway Administration Publication, Highway Performance Monitoring Field Manual, May 2005 Chapter I, page 1.

[7] LeBlanc, D., Sayer, J., Winkler, C., Bogard, S., Devonshire, J. Mefford, M., Hagan, M., Bareket, Z., Goodsell, R., and Gordon, T. (2006) "Road departure crash warning system (RDCW) field operational test final report." UMTRI reports no. 2006-9-1 and 2006-9-2.

[8] Cathey, L., Artz, B., Greenberg, J., Blommer, M., and Grant, P., "Characterization of the Visual Display system used in the Ford VIRTTEX Driving Simulator," Proc. of the Image 2003 Conference, Scottsdale, Arizona.

[9] Grant, P., Artz, B., Greenberg, J., and Cathey, L., "Motion Characteristics of the VIRTTEX Motion System," Conf. Proceedings, HCTSC 2001, Iowa City, Iowa, November 2001.

[10] Greenberg, J., Artz, B., and Cathey, L., "The Effect of Lateral Motion Cues During Simulated Driving," Proceedings of DSC North America, Dearborn, MI. October 2003.

[11] Blommer, M. and Greenberg, J., "Realistic 3D Sound Simulation in the VIRTTEX Driving Simulator," Proceedings of DSC North America, Dearborn, MI. October 2003.

[12] Blommer, M., Curry, R., Kozak, K., Greenberg, J., and Artz, B., "Implementation Of Controlled Lane Departures And Analysis Of Simulator Sickness For A Drowsy Driver Study," Proc. of 2006 Driving Simulation Conf., Paris, France, October 2006.