ADAS Reliability against Weather Conditions: Quantification of Performance Robustness

Taufiq Rahman, Andrew Liu, Daniel Cheema
National Research Council Canada
Canada

Victor Chirila
PMG Technologies Inc.
Canada

Dominique Charlebois
Transport Canada
Canada

Paper Number 23-0306

ABSTRACT

Advanced Driving Assistance System (ADAS) technologies provide an additional safety layer besides human drivers. Continual evaluation of the safety of the dynamic driving task enables ADAS to initiate a corrective (e.g., automated braking) and/or a preventative (e.g., audio-visual alerts) action if and when an unsafe roadway event is detected. To provide situational awareness, these safety systems principally rely on the vehicle mounted sensors whose performance can be greatly affected by weather events such as strong sunlight, atmospheric precipitation (rain, snowfall, fog), etc. Correspondingly, this study was conducted to characterize the performance of ADAS features in different weather conditions. Automated emergency braking (AEB) was selected as a representative ADAS feature. Two vehicles under test (VUT) were equipped with perception sensors such as LiDAR, RGB camera, infrared camera, radar, inertial measurement unit, GNSS, etc. Relevance and prominent use of these sensors in pre-production and developmental driving automation systems are widely reported in the literature. In addition, the data available through the OBD-II port of the VUTs was also recorded with temporal correspondence with the external sensors. Although weather related tests involving automotive systems have been traditionally performed in weather chambers, adoption of these test protocols for ADAS testing can be challenging. Because testing of ADAS must be performed dynamically, a runway of several hundred meters is necessary, and typical weather chambers cannot accommodate this requirement. Alternatively, this study utilized naturally occurring weather events to record AEB performance. For the purpose of this study, AEB tests performed under optimal weather conditions (sunny and bright) constituted the baseline performance. The same tests were performed in a number of different weather and roadway conditions; e.g., day/night, snow covered asphalt, persistent snowfall, overcast, rainfall etc. A number of metrics resulting from the test data analysis were used to quantify AEB performance in adverse weather conditions. These include distance of the test target when AEB system detected an imminent collision in different weather conditions, distance of the test target when AEB initiated an automated braking action in different road surface conditions (dry/wet asphalt vs snow covered asphalt), and whether AEB was successful in stopping a collision from happening in the test scenarios. These metrics helped to identify the failure modes of AEB in adverse weather conditions. It should be noted that quantification of ADAS performance robustness against adverse weather conditions is closely related to quantification of operational design domain (ODD), which is an emerging topic in driving automation systems literature. Nonetheless, observations and inferences made from this study will be used to design more comprehensive and elaborate test protocols for ADAS that are expected to improve in system capacity and ODD in near future.

INTRODUCTION

Although not formally recognized in SAE standard J3016 [1], the term ADAS (advanced driver assistance systems) commonly refers to SAE L0-L2 features/systems that assist human drivers in performing some aspect of the dynamic driving task (DDT). Implementation of these systems may take many forms. Examples include executing momentary interventional actions (e.g., automatic emergency braking - AEB), exerting sustained control over vehicle operation in limited scope (e.g., adaptive cruise control - ACC, lane support system - LSS), or simply alerting the driver when a potentially unsafe event is detected (e.g., forward collision warning - FCW). Improving safety and reducing the cognitive

load of the DDT for human drivers are the two main value propositions of ADAS. Sensors and perception algorithms work together to enable ADAS to monitor the dynamic driving environment in real-time by detecting, recognizing and classifying relevant objects and events so that appropriate responses can be prepared and executed. Correspondingly, ADAS can be broadly decomposed into two functionally distinct, yet closely related building blocks: detection & control. Due to sub-optimal visibility and challenging road traction characteristics, weather events (e.g., snowfall, rain, fog, etc.) may have adverse effects on both the detection and the control functions of ADAS. This paper studied the combined performance of both functions against varying weather conditions. To this end, two ADAS-equipped vehicles (2021 Volkswagon Jetta and 2021 Toyota RAV4) were instrumented with external perception sensors so that the vehicles' responses in different test scenarios can be recorded and analyzed. It should be noted that the deployed instrumentation did not significantly interfere with the ADAS functions because interference contributed by the active elements (LiDAR & radar) was considered minimal. Studying weather effects on ADAS features can be a challenging endeavor. Although weather testing for automotive has been traditionally performed in weather chambers, performance evaluation of ADAS features, especially in dynamic test scenarios, would generally require several hundred meters of testing roadway with simulated weather events. Developing such a test facility can be resource intensive. As an alternative, this study employed opportunistic data acquisition in different naturally occurring weather events. Tests were performed in sunny , rainfall, and snowfall conditions with the AEB system as a representative ADAS feature.

A brief literature review provided below indicates that performance quantification of ADAS in adverse weather condition is an emerging topic in the literature. Heuristic knowledge of how weather events affect sensor performance, and how degraded sensor data affect perception algorithms is still developing. As a result, this study was designed as an exploratory endeavor in response to the scarcity of prior instances of studies on ADAS performance degradation due to weather events. Correspondingly, a data acquisition system that can collect multi-modal sensor data with sufficiently accurate spatio-temporal correspondence was constructed, and data provided by the system was analyzed with a view to discover AEB failure modes related to weather events.

RELATED WORK & RATIONALE FOR PRESENTED STUDY

How adverse weather conditions affect performance of automotive perception sensors has generated strong interest from the research community in recent years. For example, Lambert *et. al.* in [2, 3] recorded static and dynamic performance of 10 different automotive LiDAR models in different controlled weather conditions (rain, fog and strong light) in a 200 meters long weather chamber to create the LIBRE dataset. In another weather chamber study, Judd *et. al.* experimentally evaluated imaging performance of multi-spectral sensors in foggy conditions in a 55 meters long weather chamber [4]. Infrared camera operating in different wavebands (LWIR: 7-14 μm, MWIR: 3-5 μm, SWIR: 0.95-1.7 μm), a regular RGB camera, and a Velodyne VLP-16 LiDAR were the sensors used in this study. Since the environmental enclosure constructed to protect the sensors from the simulated fog particles hindered the operation of the LiDAR, its imaging performance was excluded from the analysis. This study showed that, even in thick foggy condition that is completely opaque to the human eye, the LWIR sensor is the most capable sensor in terms of detecting pedestrians and cyclists (i.e., vulnerable road users - VRU). Unlike the aforementioned studies that employed weather chambers for simulating adverse weather conditions, the study in [5] utilized naturally occurring weather events (rain, heavy snow and fog) to study the performance of two LiDAR sensors (Velodyne VLP-32C and Ouster OS1-32). Their experimental data showed that while rain and snow had little effect, foggy conditions severely affected sensing performance. It is not surprising that radar sensors are not well represented in these studies. The underlying sensing modality of radar sensors render them the most resilient in adverse weather conditions [6]. The experimental data presented in [7] indicate that the minimum detection range for radar sensors in heavy fog is 260 meters i.e., (evaluated with signal to noise - SNR ratio threshold of 20 dB). A detection range of 260 meters can be characterized as sufficient for most contemporary ADAS applications. Correspondingly, radar and GNSS have been identified as the two sensing modalities that are relatively more robust against weather induced performance degradation in a qualitative evaluation presented in [6]. See Table 1 for a summary.

The aforementioned studies primarily focus on raw sensor performance. Consideration of other functional sub-systems of ADAS such as perception algorithms that enable detection, recognition and classification of relevant objects and events, and the control functions performing corrective driving maneuvers were found to be scarce. Nonetheless, aggregate performances of an *a priori* map-based place recognition system utilizing data from two sensing modalities under adverse weather conditions were evaluated in [8]. The same processing pipeline was used for both sensors for a

Table 1. Influence of weather on different sensing modalities as qualitatively evaluated in [6].

Modality	Light/Heavy rain		Dense smoke/Mist, Fog, Haze/Smog			Snow	Strong Light
	<4mm/hr	>25mm/hr	vis. <0.1km	vis.<0.5km	vis. >2km		
LiDAR	2	3	5	4	1	5	2
Radar	0	1	2	0	0	2	-
Camera	3	4	5	4	3	2 (dynamic) 3 (static)	5
Stereo Camera	almost same as monocular camera						
Gated NIR Camera[1]	2	3	2	1	0	2	4
FIR Camera[2]	2	3	3	1	0	2	4
GNSS	0	2	0	0	0	0	0

[1]NIR λ: 0.8-0.95μm [2]FIR λ: 15-1000μm

[3]Qualitative performance rubric:

0 - negligible: weather effects can almost be ignored.

1 - minor: weather effects barely cause detection error.

2 - slight: weather effects cause small errors in edge cases.

3 - moderate: weather effects cause perception error up to 30% of the time.

4 - serious: weather effects cause perception error 30-50% of the time.

5 - severe: weather effects can cause false positives or negatives.

fair comparison. This study found the LiDAR-based system to be more robust than the camera-based system. Since observations of this study is limited to the dataset used, more elaborate experiments must be performed before one can draw general conclusions.

Regardless of whether an ADAS performance study focuses on overall system performance or individual component performance, the body of literature dealing with pre-production/developmental ADAS is relatively richer than that involving production ADAS found in consumer road vehicles. Some examples of the latter include [9] and [10]. The elaborate experiment-based study [9] involving 36 ADAS equipped vehicles (model years including 2012-2018) from 15 different automakers exposed gaps in terms of performance variability. Even in some cases complete failure of the intended function was recorded (e.g., AEB systems designed to avoid collisions with pedestrians failing to do so in test scenarios). This study represents a total of ~100,000kms of on-road and test-track driving, and overall system performance of ADAS features was the main focus. It should be mentioned that commercial ADAS are proprietary in nature, and evaluating overall system performance is considered the only way to study their effectiveness because knowledge of their internal architecture is not readily available. The multi-year study [9] underscored the necessity of approaching ADAS as a stochastic system whose performance must be evaluated accordingly; i.e., by analyzing statistically significant datasets created from real-world tests and/or simulation models that represent real-world objects, events and their interactions with sufficient accuracy. The study presented in this paper can be considered a continuation of the study in [9] that attempted to expand the testing scope by integrating a data acquisition system composed laboratory grade automotive perception sensors so that: (a) the perception potential of a test scenario in adverse weather conditions using state-of-the-art sensing technologies can be independently assessed, and (b) greater insight into the failure modes of commercial ADAS can be obtained.

DATA ACQUISITION SYSTEM

The two vehicles under test (VUT) were designated as VUT-1: 2021 RAV4 and VUT-2: 2021 Jetta. Each of the VUT was outfitted with a data acquisition system (DAQ). Except for the LiDAR sensor, the two data acquisition systems were identical. See Figure 1(a) for a pictorial of the two instrumented VUTs. The DAQ system was designed specifically

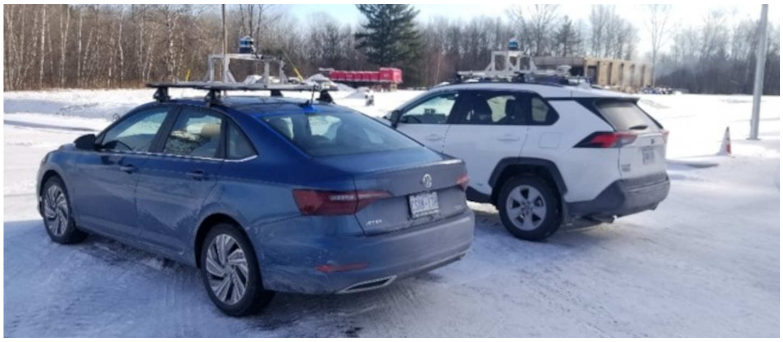
for supporting driving automation research in a cost-effective manner by taking advantage of internal resources (e.g., electro-mechanical & electronics design & fabrication, and in-house software development) and open-source software. This approach resulted in a DAQ system that is scalable, guarded against near-future obsolescence, and capable of any customization required to address project-specific needs.

Table 2. Sensor stack specifications.

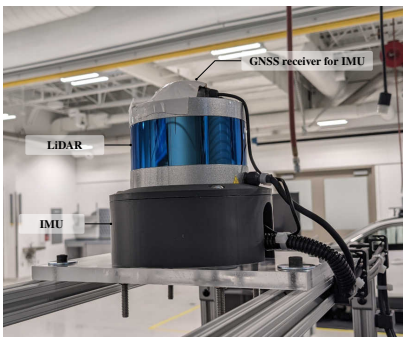
Sensors (# per VUT)	Specifications	
	VUT-1: 2021 RAV4	**VUT-2: 2021 Jetta**
LiDAR (1)	**Velodyne VLP-32C** 360° horizontal FOV, 40° vertical FOV 200m max range, 32 channels	**Velodyne VLP-16** 360° horizontal FOV, 30° vertical FOV 100m max range, 16 channels
Infrared camera (1)	**Intel RealSense D415** 850nm NIR, 1920 × 1080 active pixels 69.4° horizontal FOV, 42.5° vertical FOV	Same as VUT-1
RGB camera (1)	**Intel RealSense D415** 1920 × 1080 active pixels 69.4° horizontal FOV, 42.5° vertical FOV	Same as VUT-1
GNSS-RTK (1)	**SwiftNAV Piksi Multi** GPS, GLONASS, Galileio, BeiDou constellations, RTK relative accuracy: ~1cm horizontal, ~1.5cm vertical	Same as VUT-1
Radar array (3)	**Texas Instruments mmWave** 76-81GHz Single plane sensing w/±60° horizontal FOV	Same as VUT-1
IMU (1)	**Xsens 680G INS** Sensor fusion performance: Roll/pitch: 0.2° RMS Yaw/heading: 0.5° RMS Position: 1cm CEP, velocity: 0.05m/s RMS	Same as VUT-1
Instrument Cluster Camera (1)	720p consumer webcam	Same as VUT-1
CAN-bus monitor (1)	OBD-II CAN-bus monitor w/ SocketCan support	Same as VUT-1
In-vehicle power (1)	1000Wh rechargeable power station	Same as VUT-1

Hardware Stack

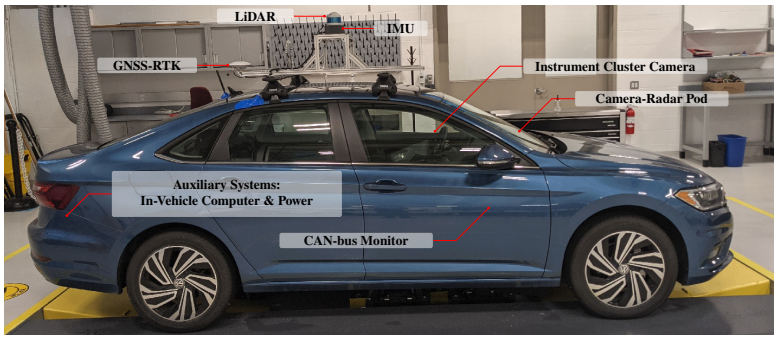
The hardware stack of the DAQ system can be categorized into three groups: (a) sensors, (b) computing platform to host the DAQ software stack, and (c) auxiliary components. The auxiliary components include a mobile power source and the cabling required for power and data transport. The computing platform and the mobile power source were installed inside the VUT trunk for convenience [see Figure 1(e)]. Specifications of the sensors in the DAQ system are provided in Table 2. LiDAR, IMU, GNSS-RTK sensors were installed on a rigid mechanical railing system on the roof of the VUT [see Figures 1(a) and 1(c)]. In addition, sensors installed inside the cabin include the camera-radar pod [see Figures 1(d) and 1(f)], a camera looking at the instrument cluster to record VUT ADAS responses, and a CAN-bus monitor plugged into the OBD-II port of the VUT. The IMU and the LiDAR sensors were assembled in a pod to ensure geometric alignment of sensors' z-axes to facilitate environmental characterization through the applicationof SLAM (simultaneous localization and mapping) techniques. Furthermore, the camera-radar pod [see Figure 1(d)] was built to facilitate multi-modal sensor fusion. The LiDAR-IMU and the camera-radar pods were centered on the vertical plane that bisects the lateral profile of the VUT.



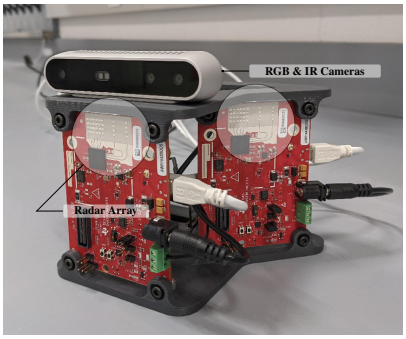
(a) Two instrumented VUTs.



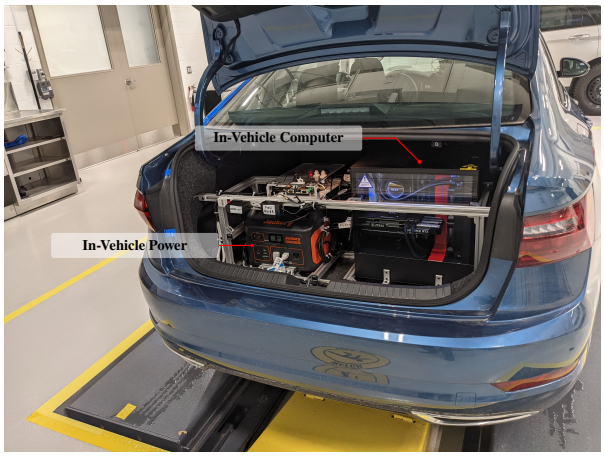
(b) LiDAR-IMU pod.



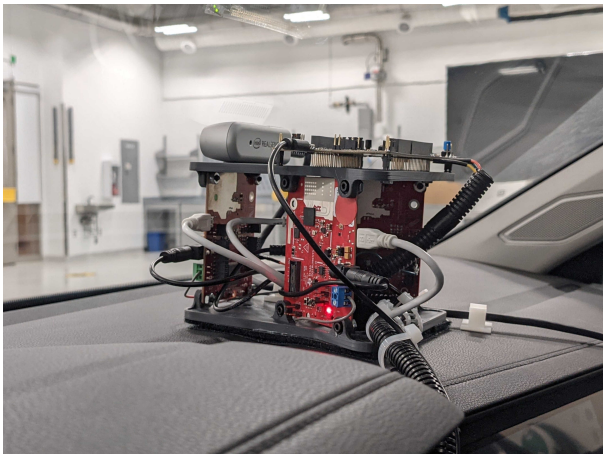
(c) Instrumentation of VUT.



(d) Camera-radar pod.



(e) Auxiliary systems of DAQ.



(f) Camera-radar pod mounted on the VUT dashboard.

Figure 1. Data acquisition system for ADAS performance evaluation.

Software Stack

The software stack of the DAQ is based on Robot Operating System (ROS) [11] hosted on a Linux (Ubuntu 18.04) computer. ROS is an open-source, meta-operating system originally developed for robotics reseach, and has been continually supported by a large, thriving community of developers. Industrial support for ROS has also been ubiquitous. OEMs of sensors and robotics platforms often provide open-source ROS drivers for their products that interface either directly with the target hardware or with a closed-source OEM-supplied driver package. ROS was chosen as the foundation of the DAQ software stack for a number of reasons. First of all, it provides a convenient way to interface with the sensor hardware utilizing OEM supplied drivers under a common framework. In addition, it is built for scalability that renders integration of new sensors into the DAQ a relatively easy task. ROS also ensures temporal

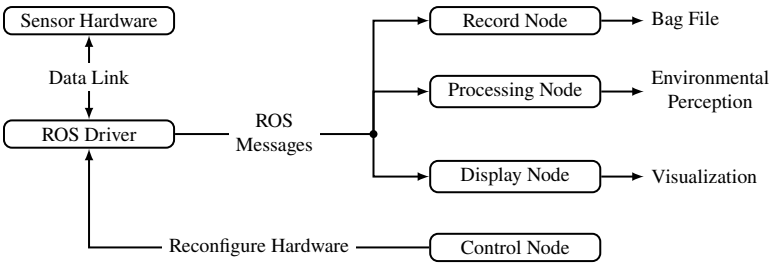


Figure 2. Typical ROS pipeline.

consistency among heterogeneous, multi-modal sensor data by enabling data acquisition under a common timestamp. In order to illustrate how ROS operates, a typical data pipeline in a ROS system is shown in Figure 2 where the sensor hardware is connected to the ROS driver *node* (i.e., a functional block in a ROS system) through a physical data link that transports binary data. The ROS driver node is programmed to parse the binary data and subsequently *publish* this parsed data under a *topic* as *messages*. Other nodes may *subscribe* to this topic and receive these messages concurrently to perform their individual functions. For example, the record node in Figure 2 subscribes to the topic advertised by the driver node, and can record the published messages in a *bag* file. Through this publish-subscribe paradigm of data exchange, two other nodes in Figure 2 also receive the same messages to deliver their intended functions (e.g., the processing node applies perception algorithms to characterize the environment, and the display node enables viewing of the data in real-time). In addition, the control node in Figure 2 can interact with the ROS driver node through a *separate* topic in order to reconfigure the sensor hardware on the fly.

As part of the software stack a GUI (graphical user front-end) was developed to enable operators to interact, monitor and record sensor data in bag files during tests. It should be mentioned that a bag file is a binary file that stores all or a subset of the messages available in a ROS system with appropriate timestamps. A bag file can later be played back so that recorded messages appear chronologically with accurate time deltas. The replay feature allows to review the test data on demand, and to perform analysis of the recorded data.

Table 3. AEB testing protocols.

#	Ref.	Objective	Description	Expected Outcome
A	[12, 13]	Evaluate performance of forward collision avoidance involving a stationary target.	VUT travels at speeds of 20kph to 60kph towards a stationary target vehicle (see Figure 3-A) or a stationary VRU target (see Figure 3-C&D) positioned at x% of the lateral length of the front bumper ($x = 10, 25$). EuroNCAP Pedestrian Target Child - EPTc & EuroNCAP Pedestrian Target Adult - EPTa are used as VRU targets. The VRU tests are not *stricto sensu* part of EuroNCAP test protocols.	AEB is expected to detect the potential forward collision event, and activate the brakes accordingly to avoid collision.
B	[12]	Evaluate performance of forward collision avoidance involving a moving target.	VUT travels at speeds of 30kph to 60kph towards a target vehicle moving at 20kph (see Figure 3-B).	AEB is expected to detect the potential forward collision event, and activate the brakes accordingly to avoid collision.

TEST DATA ACQUISITION & ANALYSIS

Test Execution & Data Recording

Test procedures used in this study were developed from the UNECE R152 [12] and the EuroNCAP VRU [13] test

Figure 3. Test protocols used in this study.

Table 4. Tests performed for VUT-1 & 2 in different weather conditions.

Test Conditions	Test Target					
	Vehicle		EPTa		EPTc	
	VUT-1	VUT-2	VUT-1	VUT-2	VUT-1	VUT-2
Clear	20	17	0	0	16	0
Precipitation	7	0	15	20	18	18

protocols. A total of 2 types of tests were designed to provide coverage for various AEB scenarios. See Table 3 for details of these two test protocols. A total of 131 AEB tests involving both VUTs were conducted at different speeds in different combinations of these ambient conditions: day/night, clear sky/overcast, and dry/precipitation (snowfall & rain). Out of these 131 tests, 76 were performed on the VUT-1 vehicle and 55 were performed on VUT-2. In addition, 78 (~60%) of these tests were opportunistically performed in natural precipitation conditions. It should be noted that the naturally occurring precipitation conditions could not be kept uniform for all the test samples. Regardless, data points from these tests can be considered very valuable because of their potential to provide deep insight into AEB performance degradation in adverse weather.

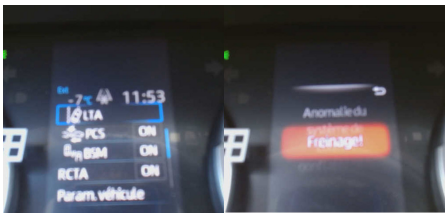
Data Analysis

The 131 AEB tests produced ~1 TB of binary data, and presenting an analysis of the entire dataset was considered to be beyond the scope of this paper. The data was recorded in the bag binary format, which has been developed for ROS systems. The bag format enabled "play-back" of the timeline of each test in real-time for analysis. When closer scrutiny of the data was necessary, play-back was performed at slower speeds. In the planning phase of the data analysis, several automated data processing methods involving image and pointcloud processing were prototyped and evaluated. Because of the diversity of the ambient test conditions, the multi-modal data showed a lot of variability, and the robustness of the prototyped data processing algorithms were considered insufficient for the data analysis activity. However, the ability to review the data visually at a slower frame rate was proved to be an effective method for evaluation. For example, the image frame where the VUT collided with a vehicle target in test sceneario "A" was isolated for the purpose of determining the exact time of collision (see Figure 4). It should be mentioned that the inertial force signature of collision with a soft target was too small for the IMU to register the event reliably.

Figure 4. Moment of collision with vehicle target in test scenario "A" (depression in the inflated target is the only indication of collision).

Since it was observed the AEBs of both VUTs failed in tests performed at speeds of 50-60kph, these tests were selected for analyzing AEB failure . Although radar sensors are more robust to weather effects, the units installed on the VUTs have limited range. Therefore, their performance was considered to be inadequate for the AEB experiments, and LiDAR was selected as the primary ranging sensor. An intermittent malfunction of the LiDAR unit installed on VUT-2 was discovered during data analysis. This rendered analysis of the VUT-2 AEB tests using LiDAR ranging untenable. Nonetheless, a total of 21 data points involving testing speeds of 60kph for the VUT-1 were analyzed and presented in this paper.

(a) YOLOv3 DNN model detecting and classifying target objects (left:input, right: output).



(b) Instrument cluster transitioning from normal state to AEB warning state.

Figure 5. Detection of events E1 and E2.

For the purpose of performance evaluation of AEB systems, each individual test timeline was broken into three separate events: (a) E1: when the test target is close enough for a typical perception algorithm to detect it, (b) E2: when the AEB system provides a warning to the instrument cluster of the VUT, and (c) E3: when an automated braking action (if any) is initiated by the AEB. The seminal deep neural network (DNN) model developed for object detection & classification in images [14] was selected as a typical perception algorithm for determining the occurrence of event E1. Specifically, event E1 was considered to have taken place when the DNN model (YOLOv3) first accurately detected & classified the test target in the images from the camera installed on the VUT dash. Continuity of the detection performance in subsequent frames was not considered for the sake of simplicity. The YOLOv3 DNN model was configured to detect and classify vehicle and people with an arbitrarily chosen prediction confidence score of 0.3 or higher [see Figure 5(a)]. In order to detect the occurrence of E2, a simple image processing algorithm was developed that determine the exact moment when the instrument cluster start showing an AEB warning message. To this end, images from the instrument cluster camera were segmented based on HSV (hue, saturation, value) representation of the distinctive color of the warning message shown on the instrument cluster [see Figure 5(b)]. Finally, event E3 was detected from the VUT's motion states (see Figure 6). For this purpose, the IMU reported velocity was fitted to a spline curve, and its first derivative provided acceleration values, and the second derivative provided jerk values. It should be mentioned that

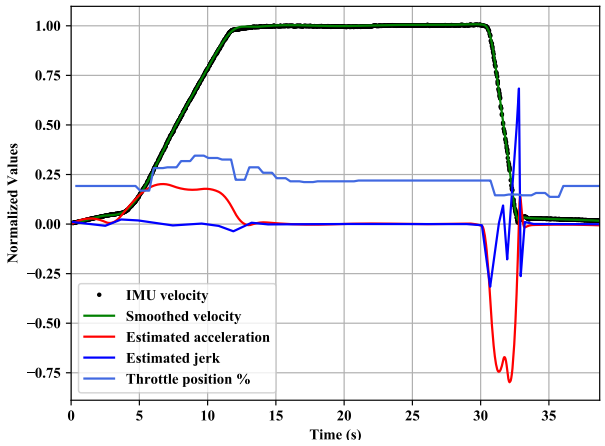


Figure 6. Detection of event E3 based on motion state of the VUT.

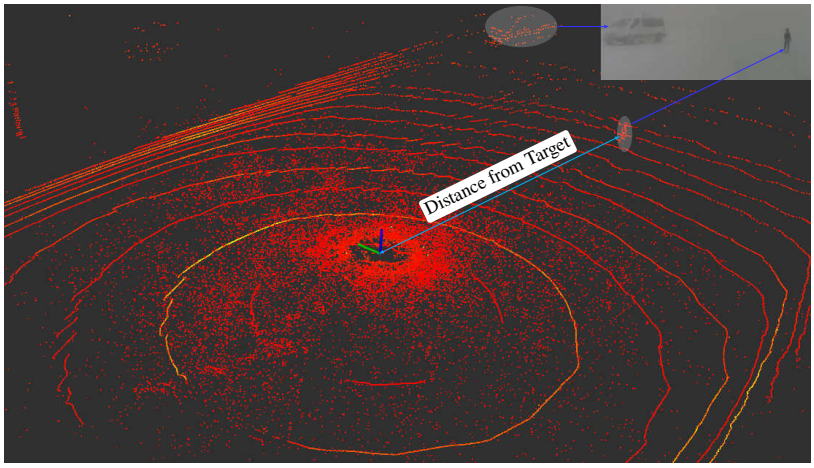


Figure 7. Determination of distance from target in the LiDAR data in snowfall conditions (false-positive noise generated by light pulses reflected from falling snow particles can be seen).

Figure 8. DNN-based object recognition and classification demonstrated irregular performance (left: input image from test conducted at night, middle: relevant objects detected and classified in the left image, right: detection failed even in sunny conditions).

except for the throttle position % parameter, other signals shown in Figure 6 were normalized over a range provided by the respective minimum and maximum values to facilitate easy visualization of the data. Furthermore, the OBD-II port of the VUT was queried with the appropriate CAN messages from the DAQ computer to obtain three vehicle states: engine RPM, vehicle speed, and throttle position %. Although the throttle position can be used as an indication of an automated braking event, it is generally not considered a reliable one because other sub-systems/features may cause it to change. In addition, the IMU reported speed was considered to be more accurate than the speed reported

through the OBD-II port. In order to determine the exact braking time, the steady-state speed prior to a braking event was considered. A braking event was considered to have happened when the instantaneous vehicle speed went below 2% of the steady-state speed. The 2% margin enabled the algorithm to remain insensitive to the measurement errors. The distance from the test target at these three events was determined from the LiDAR data (see Figure 7).

Test Results

In order to characterize the performance of OEM AEB systems, data collected from the external sensors was analyzed and compiled in Tables 5 and 6. The performance of the YOLOv3 DNN model was found to be irregular in all test scenarios. Correspondingly, detection of event E1 showed a lot of variability, which underscores the distinction between two relevant terms: *sensing* and *perception* (i.e., being able to sense is not analogous to being able to perceive). Nonetheless, application of the DNN model provided this study with an independent way of evaluating the "perception potential" of a test scene.

Table 5. Distance between VUT-1 and target in different "A" test scenarios.

#	Target	Speed	Test Conditions	Distance (m) from Target at Events			Collision?
				E1	E2	E3	
1	Vehicle	60kph	Day, overcast, high visibility & slightly wet road surface.	110.57	×	20.27	Yes
2	Vehicle	60kph	Day, overcast, high visibility & slightly wet road surface.	116.91	×	19.92	Yes
3	EPTc 10%	60kph	Day, heavy snowfall, limited visibility & snow covered road surface.	32.46	19.84	×	Yes
4	EPTc 10%	60kph	Day, slight snowfall, moderate visibility & snow covered road surface.	42.57	17.53	×	Yes
5	EPTc 10%	60kph	Day, slight snowfall, moderate visibility & snow covered road surface.	37.95	F	×	Yes
6	EPTc 25%	60kph	Day, moderate snowfall, limited visibility & snow covered road surface.	46.55	23.06	×	Yes
7	EPTc 10%	60kph	Day, moderate snowfall, limited visibility & snow covered road surface.	42.84	F	×	Yes
8	EPTc 10%	60kph	Day, heavy snowfall, limited visibility & snow covered road surface.	38.02	26.60	×	Yes
9	EPTc 25%	60kph	Day, heavy snowfall, limited visibility & snow covered road surface.	48.10	28.70	×	Yes
10	EPTc 25%	60kph	Day, heavy snowfall, limited visibility & snow covered road surface.	50.20	27.46	×	Yes
11	EPTc 10%	60kph	Day, heavy snowfall, limited visibility & snow covered road surface.	44.14	27.75	×	Yes
12	EPTc 25%	60kph	Night, no precipitation, limited visibility & snow covered road surface.	23.23	×	×	Yes
13	EPTc 10%	60kph	Night, no precipitation, limited visibility & snow covered road surface.	32.91	×	×	Yes
14	EPTa 25%	60kph	Night, moderate rainfall, poor visibility & snow covered road surface.	58.59	×	×	Yes

× = Data not available F = System failed E1 = Target detected by YOLO
E2 = Warning shown on instrument cluster E3 = Automated braking initiated

In Table 5 results from test scenario"A" are presented. Images from the camera pointed at the instrument cluster were absent from some of the tests, and these cases were appropriately indicated in Table 5. For some tests, the AEB system did not provide any warning, and such an occurrence was considered a failure (tests 5 & 7). In some cases, the VUT was manually steered away from the target for safety reasons (tests 3-11), correspondingly event E3 did not have a chance to happen. In tests 12-14, images from the instrument cluster were absent, and the VUT was manually steered away from the target. In all of the tests in Table 5 with images of the instrument cluster available (except tests 5 & 7), the AEB provided a warning to the driver when the target was at a distance range of 17.53 meters to 27.75 meters, even when visibility was challenged by precipitation conditions. In tests 1 & 2, there was clear indication of collision with the target. In other tests collision did not take place because of the safety action involving steering away from the target. Nonetheless, it was determined that a collision would have happened if the safety action was not taken.

Results from test scenario "B" are presented in Table 6. All tests were performed in sunny conditions with excellent visibility. The poor performance of the DNN model, as shown by the variability in the E1 column of Table 6, can be attributed to the sunny ambient condition. In some cases the direction of the sunlight created artefacts in the image which potentially caused irregular detection performance. Nonetheless, the AEB performance indicated by events E2 and E3 was fairly consistent. Warnings were provided to the instrument cluster when the vehicle target was 27.91-31.70 meters away from the VUT. Automated braking was initiated when the vehicle target was 13.18-14.91 meters away. Collision only occurred when the road surface was covered in snow (test 7), and failure of the AEB control functions, instead of the perception functions, can be attributed as the principal cause for collision. Later than average initiation of braking along with challenging road conditions rendered it difficult for the VUT to come to a stop safely.

Table 6. Distance between VUT-1 and target in different "B" test scenarios.

#	Target	Speed	Test Conditions	Distance (m) from Target at Events			Collision?
				E1	E2	E3	
1	Vehicle	60kph	Sunny, excellent visibility & dry road surface.	122.46	31.02	13.68	No
2	Vehicle	60kph	Sunny, excellent visibility & dry road surface.	150.56	31.70	14.25	No
3	Vehicle	60kph	Sunny, excellent visibility & dry road surface.	157.86	29.16	13.98	No
4	Vehicle	60kph	Sunny, excellent visibility & dry road surface.	119.58	31.18	14.91	No
5	Vehicle	60kph	Sunny, excellent visibility & dry road surface.	53.88	31.4	13.9	No
6	Vehicle	60kph	Sunny, excellent visibility & dry road surface.	121.02	27.91	13.77	No
7	Vehicle	60kph	Sunny, excellent visibility & snow covered road surface.	33.91	×	13.18	Yes

× = Data not available E1 = Target detected by YOLO E2 = Warning shown on instrument cluster
E3 = Automated braking initiated

DISCUSSION, LIMITATIONS

Market penetration of ADAS features are expected to widen exponentially in near future. As system capabilities increase and drivers become more reliant on these systems, they will continue to be exposed to sub-optimal driving conditions including weather events that can degrade their performance. System availability in adverse weather conditions is one of the avenues where operational design domains (ODD) of ADAS can potentially see rapid advancements. Correspondingly, quantifying the effects of weather on ADAS performance has become increasingly important so that drivers can operate these systems safely within their safety limits. To this end, more complex and elaborate test procedures and protocols must be developed to evaluate these systems more extensively. Specifically, the safety and

the functionalities of these systems in different weather conditions must be experimentally evaluated to substantiate claims of robust performance. This study attempted to innovate ADAS testing by using external perception sensors so that the test environment could be characterized with more definition than what was possible with test procedures implemented before. This facilitated analysis of the failure modes of ADAS with deeper insight. The demonstrated ability to acquire rich spatio-temporal data of the test environment open up numerous possibilities for the next evolution of ADAS testing, which may range from formulating more complex tests involving multiple dynamic roadway elements to conducting tests without having to install extensive support infrastructure. As a first step toward this direction, the AEB system was focused as a representative ADAS feature.

Acquiring multi-modal sensor data in a temporally consistent manner, and subsequently making inferences from them are challenging tasks. This study achieved both by constructing a DAQ system that was robust enough to operate in adverse weather conditions, and also could withstand the rigors of test protocols involving collisions with soft/inflatable targets at high speeds. By leveraging open-source software and by exploiting internally developed intellectual properties involving software & hardware components, the DAQ system was constructed to be easily scalable and rapidly reconfigured or upgraded to address the demands of the application-specific needs.

The tests presented in this paper involved recording AEB performance in different naturally occurring weather conditions in a controlled test environment (i.e., a test-track). Admittedly, the test data collected for this paper represents a small fraction of vehicles from different manufacturers, models and model years driven on the road. In addition, the test conditions are also a small representation of the large permutations of roadway and weather conditions these vehicles will encounter in real-life. Therefore, observations from the presented test results cannot be reliably extrapolated to draw general conclusions. Nonetheless, as ODDs expand for these systems rapidly with the launch of new models and model years of vehicles, test procedures and protocols must evolve accordingly. This study is a step towards this goal.

FUTURE WORK

Since this study was a first step towards developing more innovative test procedures and protocols for ADAS, there are a number of ways this work can grow. The foundational ability to characterize the dynamic roadway environment and to establish temporal correspondence with the recorded vehicle states can be exploited to provide deeper insights into current test procedures. For example, identifying control functions of the AEB system as the more probable root cause of failure in test 7 of Table 6 was made possible with the help of data provided by the DAQ system.

The DAQ system used in this study can provide rich situational awareness of the evolving roadway environment without the aid of elaborate support infrastructure only available in test-tracks. This opens up the possibility of conducting more naturalistic evaluation of ADAS features wherein a VUT can be driven on public roads under a mileage accumulation exercise. This will enable the ADAS to be exposed to numerous ADAS-relevant events and objects that structured test protocols do not cover in different weather and roadway conditions. By recording & subsequently evaluating vehicle and driver behavior in these events, performance of ADAS can be characterized with unprecedented statistical significance.

As driving automation features expand their capabilities, quantified specification of ODD is expected to be an increasingly important topic in driving automation testing activity. The multi-modal data provided by the DAQ system can facilitate research & development activities in this topic. In one example, the data can be used to examine ADAS performance as a function of ODD. In another example, the data can accelerate the development of objective methodologies for specifying ODD.

CONCLUSION

In a typical ADAS system, the sensor hardware provide raw data to the perception components that analyze the sensor data in order to detect and track relevant objects and events. If the perception components identify an unsafe driving condition, the control elements initiate a remedial action that may involve a warning provided to the human driver and/or a corrective action in the form of automated driving maneuver. Weather events affect all functional aspects of an ADAS system: sensing, perception, and control. Therefore, it is important to develop robust data acquisition methods to support test protocols that identify ADAS failure modes in adverse weather conditions. This study was conducted as a first step towards this goal.

This study utilized external perception sensors to characterize the test environment with rich definition, and characterized the performance of a representative ADAS system in different weather conditions. The opportunistic nature of this study did not allow for more tests. Nonetheless, lessons learned from this study will be invaluable in developing and designing the next evolution of ADAS testing.

REFERENCES

[1] Society of Automotive Engineers, "Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles J3016_202104 (p. 41)," *SAE International*, 2021.

[2] J. Lambert, A. Carballo, A. M. Cano, P. Narksri, D. Wong, E. Takeuchi, and K. Takeda, "Performance analysis of 10 models of 3D LiDARs for automated driving," *IEEE Access*, vol. 8, pp. 131 699–131 722, 2020.

[3] A. Carballo, J. Lambert, A. Monrroy, D. Wong, P. Narksri, Y. Kitsukawa, E. Takeuchi, S. Kato, and K. Takeda, "Libre: The multiple 3D LiDAR dataset," in *2020 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2020, pp. 1094–1101.

[4] K. M. Judd, M. P. Thornton, and A. A. Richards, "Automotive sensing: Assessing the impact of fog on LWIR, MWIR, SWIR, visible, and LiDAR performance," in *Infrared Technology and Applications XLV*, vol. 11002. SPIE, 2019, pp. 322–334.

[5] J. Abdo, S. Hamblin, and G. Chen, "Effective range assessment of LiDAR imaging systems for autonomous vehicles under adverse weather conditions with stationary vehicles," *ASCE-ASME J Risk and Uncert in Engrg Sys Part B Mech Engrg*, vol. 8, no. 3, 2022.

[6] Y. Zhang, A. Carballo, H. Yang, and K. Takeda, "Autonomous driving in adverse weather conditions: A survey," *arXiv preprint arXiv:2112.08936*, 2021.

[7] V. Sharma and S. Sergeyev, "Range detection assessment of photonic radar under adverse weather perceptions," *Optics Communications*, vol. 472, p. 125891, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0030401820303631

[8] K. Żywanowski, A. Banaszczyk, and M. R. Nowicki, "Comparison of camera-based and 3d lidar-based place recognition across weather conditions," in *2020 16th International Conference on Control, Automation, Robotics and Vision (ICARCV)*. IEEE, 2020, pp. 886–891.

[9] E. Meloche, D. Charlebois, B. Anctil, G. Pierre, and A. Saleh, "Adas testing in canada: could partial automation make our roads safer?" in *26th International Technical Conference on the Enhanced Safety of Vehicles (ESV): Technology: Enabling a Safer TomorrowNational Highway Traffic Safety Administration*, 2019.

[10] M. Dollorenzo, V. Dodde, N. I. Giannoccaro, and D. Palermo, "Simulation and post-processing for advanced driver assistance system (ADAS)," *Machines*, vol. 10, no. 10, p. 867, 2022.

[11] *Robot Operating System - Melodic Morenia*, 2018, Available: https://www.ros.org.

[12] UN ECE, *Uniform provisions concerning the approval of motor vehicles with regard to the Advanced Emergency Braking System (AEBS) for M1 and N1 vehicles*, United Nations, Geneva, Switzerland, 2021. [Online]. Available: https://unece.org/sites/default/files/2021-02/R0152r1am1e.pdf

[13] European New Car Assessment Programme, *Test Protocol - AEB VRU Systems*, EURO NCAP, 2020. [Online]. Available: https://cdn.euroncap.com/media/58226/euro-ncap-aeb-vru-test-protocol-v303.pdf

[14] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.