OE Optical Engineering

RESEARCH PAPER

Toward standardized testing of automotive lidars: year two results

Theus H. Aspiras,[a] Vijayan K. Asari,[a] Cullen Bradley,[b] Anna Gnacek,[b] Charlie Kershner,[c] Daniel A. LeMaster,[d] Paul F. McManamon,[b] Andrew D. Reinhardt,[e] Eddie Ruff,[b,*] Christopher R. Valenta,[f] and Andre Van Rynbach[e]

[a]University of Dayton Vision Lab, Dayton, Ohio, United States
[b]Exciting Technology LLC, Dayton, Ohio, United States
[c]National Geospatial-Intelligence Agency, Springfield, Virginia, United States
[d]U.S. Department of Transportation, Highly Automated Systems Safety Center of Excellence, Washington, District of Columbia, United States
[e]Sensors Directorate, Air Force Research Laboratory, Wright-Patterson Air Force Base, Ohio, United States
[f]Georgia Tech Research Institute, Atlanta, Georgia, United States

ABSTRACT. SPIE is working with SAE International to develop lidar measurement standards for active safety systems. This multi-year effort aims to develop standard tests to measure the performance of low-cost lidar sensors developed for autonomous vehicles or advanced driver assistance systems, commonly referred to as automotive lidars. SPIE is sponsoring three years of testing to support this goal. We discuss the second-year test results. In year two, we tested nine models of automotive grade lidars, using child-size targets at short ranges and larger targets at longer ranges. We also tested the effect of high reflectivity signs near the targets, laser safety, and atmospheric effects. We observed large point densities and noise dependencies for different types of automotive lidars based on their scanning patterns and fields of view. In addition to measuring point density at a given range, we have begun to evaluate the point density in the presence of measurement impediments, such as atmospheric absorption or scattering and highly reflective corner cubes. We saw dynamic range effects in which bright objects, such as road signs with corner cubes embedded in the paint, make it difficult to detect low-reflectivity targets that are close to the high-reflectivity target. Furthermore, preliminary testing showed that atmospheric extinction in a water-glycol fog chamber is comparable to natural fog conditions at ranges that are meaningful for automotive lidar, but additional characterization is required before determining general applicability. This testing also showed that laser propagation through water-glycol fog results in appreciable backscatter, which is often ignored in automotive lidar modeling. In year two, we have begun to measure the effect of impediments to measuring the 3D point cloud density; these measurements will be expanded in year three to include interference with other lidars.

Keywords: Lidar benchmarking standards; automotive lidar standards; lidar; ladar; autonomous vehicles; autonomous perception

Paper 20240160G received Feb. 15, 2024; revised Jun. 28, 2024; accepted Jul. 4, 2024; published Aug. 10, 2024.

*Address all correspondence to Eddie Ruff, eddie@excitingtechnology.com

1 Introduction

The purpose of this work is to develop standard tests for automotive lidar performance based on experience gained through vendor-independent field testing. Although we primarily consider automotive lidars, we will also include lidars for other short-range autonomous platforms that are ground- or air-based, such as robotic delivery, special purpose autonomous military vehicles, driver assist, and vehicle safety. Lidars in this class are also applied in transportation infrastructure applications, such as monitoring intersections and railroad grade crossings, to protecting vulnerable road users and prevent crashes. Developing standard test procedures will make it easier to compare one lidar's performance against another. Standard procedures are necessary as there will be a range of lidars with different capabilities built for different applications. Consequently, testing standards must be developed and be sufficiently flexible to cover various types of lidars used on different platforms for varying applications.

Although lidar vendors and vehicle manufacturers have performed their own internal testing and evaluation, these tests are not published or publicly discussed. The intent of this work is to develop similar tests in an open, multi-vendor environment in which all suppliers are invited to participate. Confidentiality is maintained with each lidar's specific performance only revealed to the supplier. The methods and lessons learned from this series of tests will inform the SAE International "Active Safety System Lidar Performance Task Force" in their definition and promulgation of standards. (For more information about the active safety system lidar performance task force, please refer to Ref. 1.)

The progress report from year one includes a literature review on related benchmarking efforts.[2] Additional standardization efforts include DIN SAE SPEC 91471 "Assessment methodology for automotive Lidar sensors" and IEEE P2936 "Standard for Test Methods of Automotive Lidar Performance." The former defines 13 key performance indicators in categories of detection distance, resolution, false positive detection, spatial error, and target separability. The latter, if it comes to fruition, will define scenario test methods for accuracy, precision, and resolution in range and angle detection statistics. Future plans for IEEE P2936 are unknown—the most recent update was made in July 2021.[3] Also in 2021, the China Automobile Industry Association published T/CAAMTB 58-2021 "vehicle-mounted lidar detection method," which includes standards for the measurement of detection distance, blind spots, angle accuracy and resolution, interference, and other topics. These works are foundational to a community goal to support an active safety standard.

In year one, we tested the performance of eight, unique automotive lidars. We used 10% Lambertian reflective child size targets that were 80 cm high by 15 cm wide. These targets were chosen to represent a child just old enough to walk as a corner case to test the limiting performance of automotive lidars. We then placed road signs next to the child-size targets. In many cases, the ability of the automotive lidar to distinctly separate the simulated child from the sign was compromised by the highly reflective road sign. Road signs contain small corner cubes in the paint and therefore reflect laser light back over a smaller angle, creating reflectivity much greater than 100% when compared with a Lambertian target. Having a very bright object near a much dimmer target creates a challenging dynamic range issue for automotive lidars.[2] This dynamic range challenge is further exasperated by the limited number of discrete returns produced by typical automotive lidars: typically brightest and last, first and brightest, etc. When more than two or three targets are detected by a single pulse from these lidars, some targets are intentionally ignored due to memory/throughput limitations.

In year two, we added larger targets, two lanes of testing, laser safety testing, and fog testing as a start in evaluating the effects of impediments measuring the required 3D point cloud. (We use the term "lane" throughout this paper; however, it is important to note that the testing ranges are unimproved grass lots and not actual paved road lanes. Year 3 testing may include paved surfaces as discussed in Sec. 6.1.) At short ranges for city driving situations, the child size target is an important corner case, but at longer ranges, it is difficult for many lidars to obtain sufficient point density or points per square meter on a small target, such as a truck tire fragment or debris that has fallen from a truck. Those could be some of the corner cases for fast highways, such as the autobahn or US freeways. The year two event also included laser safety testing of the lidars. Individually, automotive lidars will be required to meet laser safety standards, such as from the American National Standards Institute (ANSI Z136), the International Electrotechnical

Commission (IEC 60825), the Food and Drug Administration (FDA), or the US Code of Federal Regulation (CFR 1040). However, a critical aspect of laser safety for automotive lidars will be the occurrence of multiple lanes of automotive lidar lasers simultaneously lasing. Finally, the second-year testing also included experiments to test the concept of a smaller, field-capable fog chamber to evaluate lidar performance in adverse weather conditions. Further developed, this testing could complement or supplement usage of larger, more dedicated facilities. Our ideal inclement weather testing situation would be for a tier one supplier, or auto company, to be able to do moderate quality automotive lidar weather testing using similar capabilities, while relying upon the more established facilities for dedicated requirements, such as calibration of the field test.

2 Test Setup

2.1 Location

Similar to the first year of testing, year two was also conducted near the Bridging the Innovation Development Gap (BRIDG) center in Kissimmee, Florida (28.291525°N, −81.371776°W). The tests were conducted on April 29th and 30th, 2023.

Data were acquired on both testing days. The original intention was to use Saturday as a rehearsal to ensure that everything was nominal. Sunday was planned to be the actual test day, but persistent high winds displaced targets during data acquisition. Data were processed from both days, and we note that there are several uncontrollable conditions that were not monitored during the open-field testing. Being subject to Florida weather, in this case, high humidity, high winds, and scattered thunderstorms do not allow for an absolute comparison but instead a relative comparison between lidars.

Table 1 summarizes the weather conditions during the two test days.

The test area was comprised of the unimproved lot behind the BRIDG facility and Skywater Technologies buildings. An overhead quadcopter photo of the long-range testing lane along with several, but not all, benchmarking volunteers are shown in Fig. 1.

Table 1 On-site weather conditions for April 29, 2023, and April 30, 2023.

Day	Temperature (°F)	Humidity (%)	Wind speed (mph)	Visibility (miles)	Comment
1	66 to 86	65 to 98	4.9 to 12.5	10	Clear and sunny during the day, thunderstorm immediately after teardown
2	72 to 79	56 to 95	12.8 to 24.2	10	Clear and sunny, but persistent high winds

Fig. 1 Quadcopter image of lane 1 layout and benchmarking volunteers.

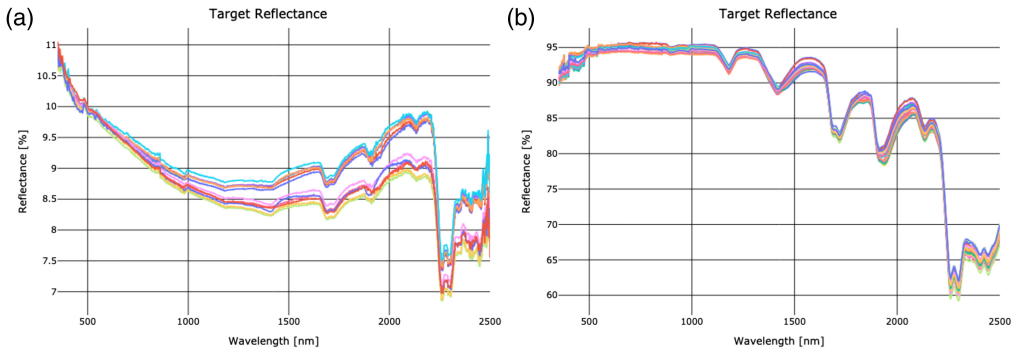

Fig. 2 Traces for 1 m^2 targets. (a) 12 low-reflectance (10%) and (b) 22 high-reflectance (94%).

2.2 Targets

Labsphere donated calibrated 10% and 94% Lambertian coated targets. Each target coating was given a unique ID and had its reflectivity characterized by Labsphere between 350 and 2500 nm. These targets were positioned <50 m from the lidar under test origin and are identical to the calibrated targets used previously in year one of our testing. The lidar origin was the fixed tripod location used to mount each lidar during data collection. Labsphere developed larger 1 m × 1 m calibrated targets for year two that were also mosaiced to create 2 m × 1 m and 3 m × 1 m targets that were then positioned >50 m down range. These larger targets enabled the various lidars to achieve more points on target compared with year one. Figure 2 shows several calibration traces for the high and low reflectance targets.

2.3 Testing Plan and Layout

Multiple testing lanes were constructed such that more than one device could be tested at any given time. Table 2 describes the target type and max range for each testing lane. For example, lidar A could be tested on lane 2, then immediately tested on lane 1, and followed by a fog test in the box truck. Lane 1 is the long-range testing lane with only calibrated Lambertian reflectors from Labsphere. Lane 2 is the short-range testing lane with similar calibrated targets, although some were oriented along the ground to mimic road debris and obstacles. Lane 3 is identical to lane 1 and occupies the same physical space but with the addition of large road signs with embedded corner cubes to act as confusers. Figure 3 shows a cartoon of the three lanes with the legend symbol key for the targets.

2.4 Instruments

2.4.1 *Reference lidar*

Similar to the previous year of testing, year two utilized a survey-grade Riegl VZ-400i terrestrial Lidar scanner. A summary of the Riegl instrument can be found in a previous paper.[2] Table 3 contains the specifications for this instrument.

One important contrast between the Riegl lidar and the test lidars is the scan rate. Each Riegl scan was collected using 0.02 deg horizontal and vertical angular sampling at a scan rate of 1.2 mHz or ∼14 min per 1 full scan rotation. Four complete reference datasets were collected, one for each lane configuration and the fog chamber. Figure 4 shows the complete, ground-truth reference point cloud for lane 1.

Table 2 Testing lane descriptions.

Lane	Targets	Max range
Lane 1	Calibrated Labsphere targets, 94% and 10% reflectance	208 m
Lane 2	Calibrated Labsphere targets, 94% and 10% reflectance	50 m
Lane 3	Calibrated Labsphere targets, 94% and 10%. And road sign confusers	208 m

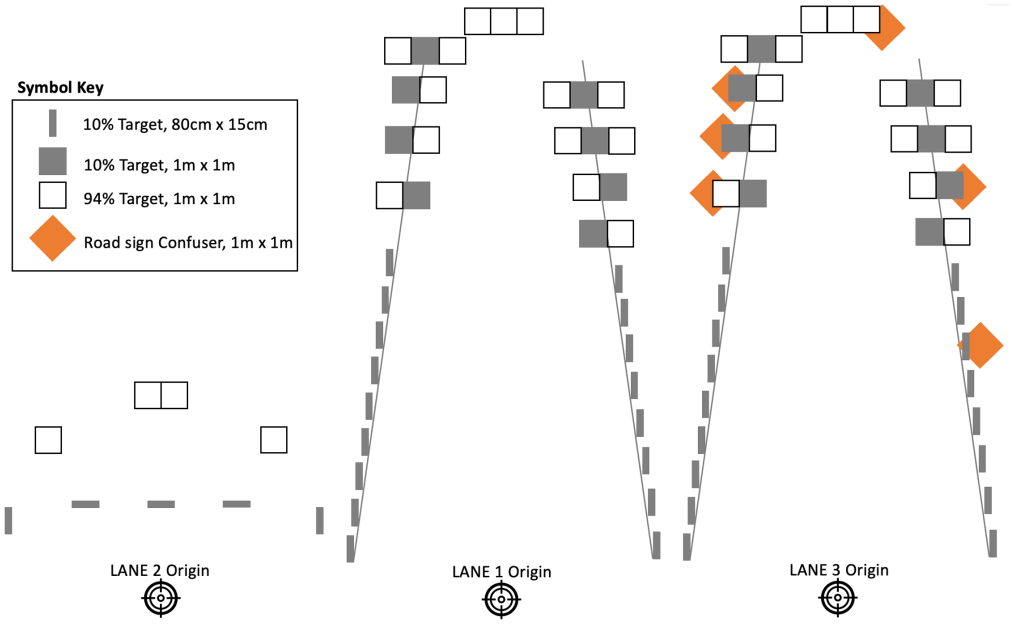

Fig. 3 Cartoon of all three lanes.

Table 3 Riegl VZ-400i specifications.

Parameter	Specification
Max measurement range ($\rho \geq 20\%$)	120 m
Max measurement range ($\rho \geq 90\%$)	250 m
Accuracy	5 mm at 100 m
Precision	3 mm at 100 m
Beam divergence	0.35 mrad
Max targets per pulse	4
Laser wavelength	1550 nm

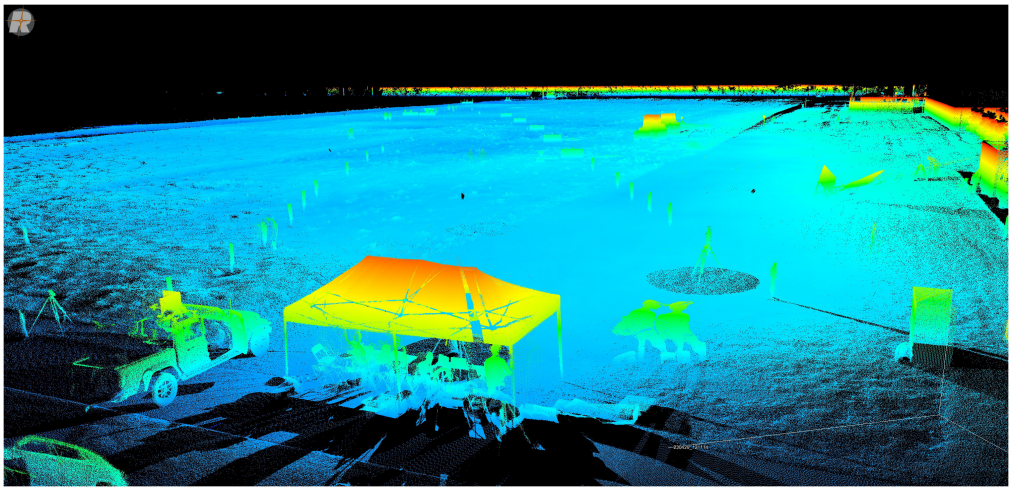

Fig. 4 Riegl ground-truth lidar scan for lane 1.

Table 4 Lidar test matrix. Due to weather and limited time, not all lidars were tested in each category.

Lidar	Lane 1	Lane 2	Lane 3	Fog	Laser safety	Comment
A	NP	✓	NP	✓	✓	NP: day 1 collection
B	✓	✓	✓	✓	✓	—
C	✓	✓	✓	✓	✓	—
D	✓	✓	—	✓	✓	Time constraints
E	NP	✓	✓	—	✓	NP: day 1 collection
F	✓	✓	—	✓	✓	Time constraints
G	—	—	✓	—	—	Time constraints
H	—	✓	—	—	—	Technical issues
I	✓	✓	—	✓	✓	Time constraints
Riegl	✓	✓	✓	✓	—	Survey Lidar, ground truth

2.4.2 *Test lidars*

Year two testing collected data for nine lidars in addition to the Riegl ground-truth lidar. Note that, although there was some overlap with the year one lidars under test, not all lidars tested in year one were also tested in year two. However, different lidars were tested in year two compared with year one. To maintain anonymity, the lidars in Table 4 are labeled as "A" through "I" in no particular order. The make and model of each lidar are not disclosed. The Riegl ground-truth is labeled as such in each plot. The "NP" label indicates that the collected data were not processed.

Year two included more lidar diversity in terms of wavelength, scanning pattern, and operating range. The general wavelength of operation was in the range of 800 to 1600 nm. Most lidar range precision was reported to be near 5 cm, and the general operating range spanned from as close as 10 m to as far as 300 m.

2.5 Data Processing

Commercial automotive lidars typically interface via local TCP networking using ethernet cables. The host machine, a laptop in this case, was configured to monitor and record the network traffic for the specified lidar IP address and subnet mask. Viewing and recording the lidar point cloud data were accomplished in three different ways. The first approach was to use a Linux operating system running Ubuntu 20.04 equipped with ROS Noetic. Once the ROS drivers and workspaces were configured, the corresponding point cloud data were then viewed in RViz and recorded to a rosbag file via the terminal application. This approach worked well for lidars that supported ROS. For other proprietary technologies that are not open source, the lidar manufacturer provided custom software with GUIs. A third approach for non-Linux machines was to use Wireshark to record the network traffic to a Pcap file. All three approaches were used in year two testing as the new data processing software accepts a wide range of file formats.

Year two introduced new data processing software, namely the automated technology for LADAR analysis system (ATLAS) labeling suite.[4] All point clouds were imported into this software, and with the aid of several semi-automated tools, the point clouds were aligned, the fiducials were selected, the targets were registered and labeled, and any metrics of interest were then exported.

2.6 Data Metrics

The two metrics observed for year two were the number of points on target per square meter and the range precision in centimeters.

Additional ATLAS outputs, such as accuracy and false positives, were omitted from the year two results. For year three, a comprehensive list of KPIs will be addressed with an open-source processing code.

The number of points per square meter was computed in the ATLAS processing suite using bounding boxes or custom boundaries around each target surface for a single acquisition. For conjoined targets of different reflectivity, a dividing line was manually inserted based on the known dimensions of the target. This approach allows us to separate side-by-side low-reflectivity targets from high-reflectivity targets.

The Riegl ground-truth point cloud was compressed or down-sampled before being imported into the ATLAS processing suite. Specifically, the ATLAS processing suite received a 10 cm octree dataset, meaning the point cloud was processed out to only keep 1 point per 10 cm in the horizontal and vertical dimensions, but the range direction was preserved as collected. Because each target was scanned multiple times as the Riegl moved down the lane, the point density varies. In general, the processed dataset is ∼90% reduced compared with the original dataset in terms of points per square meter. The down-sample factor was ∼10×, and the Riegl points per square meter data was multiplied by this factor to approximate the original, uncompressed point cloud. Without this factor, the survey-grade lidar is on the same order of magnitude in terms of points per square meter when compared with the automotive lidars. The plots in Sec. 3 account for the downsampled data and clearly show the Riegl ground truth lidar outperforming the automotive lidars in points per square meter.

The range precision required additional processing steps due to the high wind and slow scan time of the Riegl lidar as several targets moved during ground truth acquisitions. To mitigate this error, a best-fit target plane was constructed from the average of the Riegl scan for each target. The average distance to the best-fit target plane was then used to calculate the range precision for the tested lidars. This averaging of the ground truth was necessary due to the target displacement from high winds during testing. We show the range precision of the Riegl compared to the best-fit target plane for all range precision figures. The Riegl is specified for 3 mm range precision at a distance of 100 m, and our measurements agree with this spec when we utilize the best-fit plane technique.

2.7 Fog Testing

In year two, one goal was to determine whether artificial water-glycol fog is a viable substitute for naturally occurring fog in automotive lidar benchmarking. The advantages of the water-glycol approach include low cost, portability, and flexibility in terms of time, place, and repeatability. The year two experiments are an initial proof-of-concept demonstration, and future experiments will determine if these advantages can be realized in the face of practical difficulties.

Fog affects lidar performance primarily through backscatter and extinction.[5,6] In some cases, backscatter is ignored, and extinction is presented as the only salient consideration.[7,8] For each atmospheric constituent, the one-way transmission loss τ along the path, R, between the lidar and target is expressed by the Beer-Lambert law:

$$\tau = e^{-\int_0^R \alpha(r)\mathrm{d}r}, \tag{1}$$

where $\alpha(r)$ is the range dependent extinction coefficient. The round-trip transmission loss is τ^2. Figure 5 presents MODTRAN simulations of τ at lidar wavelengths 905 and 1550 nm for the approximate conditions of the April 2023 benchmarking trial (i.e., mid-latitude summer with rural aerosols) compared with the transmission loss for two categories of naturally occurring fog over the equivalent path. The result shows that transmission loss by fog type is effectively the same at both wavelengths. In addition note that, for clean air, the 1550 nm wavelength experiences less attenuation than the 905 nm. However, this advantage degrades as extinction increases.

The wavelength average extinction coefficients in this figure are 8.58 and 20.07 km^{-1} for naturally occurring radiative and advective fog, respectively. Rasshoffer et al.[6] used the method of Bohren and Huffman[9] with a single-scattering assumption to describe the extinction coefficient for Mie scattering particles, such as, fog as

$$\alpha = \frac{\pi}{8}\int_0^\infty D^2 Q(D,n)N(D)\mathrm{d}D, \tag{2}$$

where D is the droplet diameter, Q is the Mie extinction efficiency, n is the complex refractive index of the droplets, and N is the statistical distribution of droplet diameters. Code

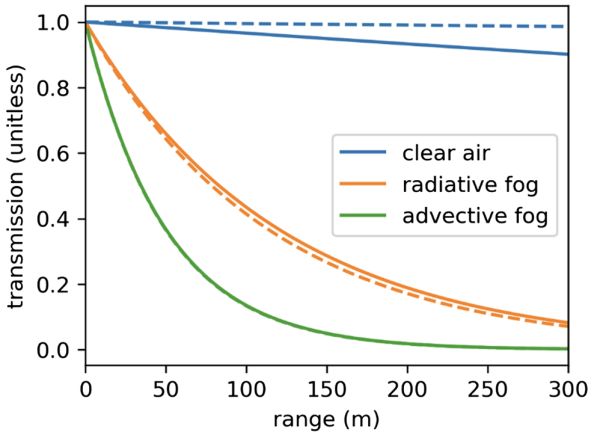

Fig. 5 One-way transmission loss τ for 905 nm (solid line) and 1550 nm (dash line) lasers in three different environments.

implementing the Bohren and Huffman calculation of Q is available in Ref. 10. Differences in the distributions of N are responsible for the differences in transmission between the radiative and advective fogs in Fig. 5. The peak droplet diameter for radiative fog is roughly between 6 and 10 μm with a narrower distribution, whereas the advective fog peaks between 16 and 20 μm with a distribution that is several times wider than the advective case.[11]

The optical properties of naturally occurring fog can be compared with the properties of artificially generated fog via Eq. (2). Artificial water or saltwater fog is used at various existing weather simulation facilities.[8,12] In our experiment, we employ water-glycol theatrical fog. Hagen recently reported on extinction, refractive index, and particle sizes for water-glycol fog in the visible and infrared regime.[13] He showed that extinction for natural and artificial fogs are essentially the same for any given droplet size but that the factors contributing to N are different. Specifically, the peak droplet diameter of water-glycol fog is 3 μm, which is smaller than the natural fogs described previously. Furthermore, the width of the artificial fog droplet diameter distribution is narrower, though this width is not quantified. Other researchers report a similar outcome for N when generating artificial saltwater fog.[12] Figure 6 shows that a narrowly distributed artificial fog with a peak droplet diameter around 3 μm may result in significantly different extinction coefficients at 905 and 1550 nm lidar wavelengths. This figure also illustrates why extinction at these two lidar wavelengths is effectively the same for natural fog: the broad N for advective fog (shown in this example) will suppress the effect of the oscillations in extinction efficiency. The distribution of artificial fog droplet diameters was not measured in the 2023 experiments, but this measurement is a priority for 2024.

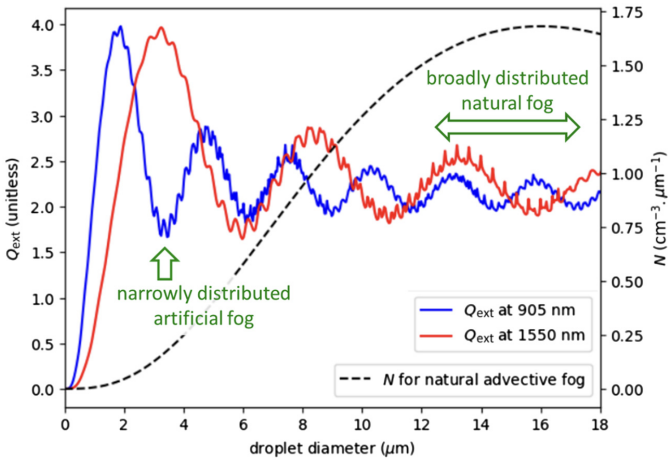

Fig. 6 Differences in extinction between artificial and natural fog are driven by the wavelength dependence of Mie efficiency, Q, and droplet diameter distribution, N.

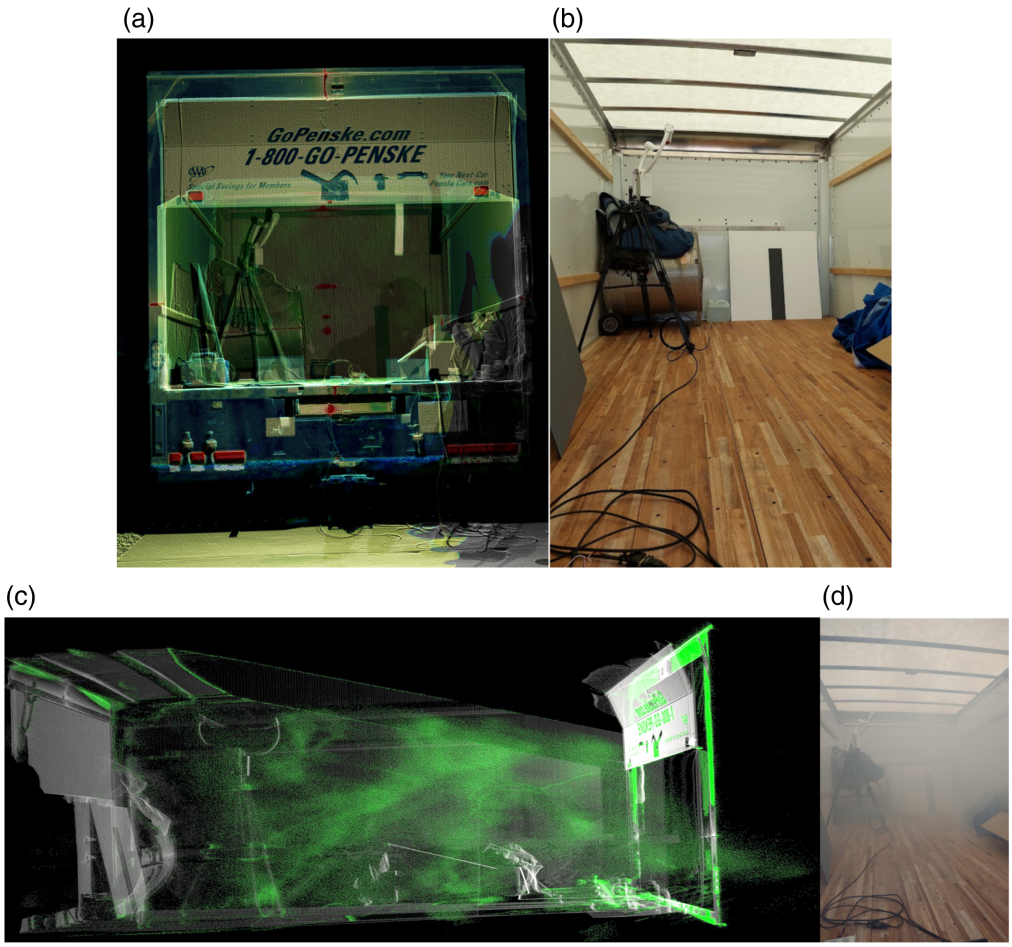

Fig. 7 Cargo area of the truck used for fog testing. (a) and (b) Without fog and (c) and (d) with fog. Ground truth lidar point cloud visualized in (a) and (c), with photographs of the test setup in (b) and (d). Note the vertical black target in front of the square white target in the back right corner of the truck.

For the 2023 trials, a cargo truck enclosure, 16 ft (4.9 m) × 8 ft (2.4 m) and 7 ft (2.1 m) in height, was used as a prototype fog chamber. The truck hatch was partially closed during testing by ∼2 ft to reduce the vapor loss rate during testing. Figure 7 shows ground-truth point clouds and photographs of the truck with and without fog, and Fig. 9 shows a layout of the experiment. This test setup was sufficient to provide point cloud data time-stamped to extinction data captured in 1-min intervals as fog density decreased after an initial 5-s pulse. The fog generator was an FQ-100 Performance Fog Generator. For each test, the manual operation button was held for 5 s, which was sufficient to fill the chamber. A forward scatter meter (Campbell Scientific model CS120A) collected extinction data during the duration of each test.

The fog generator had a maximum output rated to completely obscure and fill a volume of 20,000 ft^3 (566.34 m^3) within 1 min. The fog generator output was set to 2% of maximum, or 400 ft^3/min (11.3 m^3/min). This rate was still sufficient to fill the volumes of the tent and the truck cargo area within a 5-s burst.

The fog was funneled to the back of the truck via crude ducting material, in which a 1 m × 1 m 94% reflectivity target was located with a smaller 10% reflectivity target centrally located, effectively acting as a stripe on the target.

Six lidars were tested: lidars A, B, C, D, F, and I, in addition to the Riegl ground truth data. The reference lidar was used to examine the test area both during a fog event and without fogging. As depicted in Fig. 8, testing was conducted by placing a device under test horizontally centered in the cargo area and forward by ∼0.6 m from the back of the truck. Fog was dispersed for 5 s at a 2% maximum output. The resulting vapor cloud sufficiently filled the test area as it

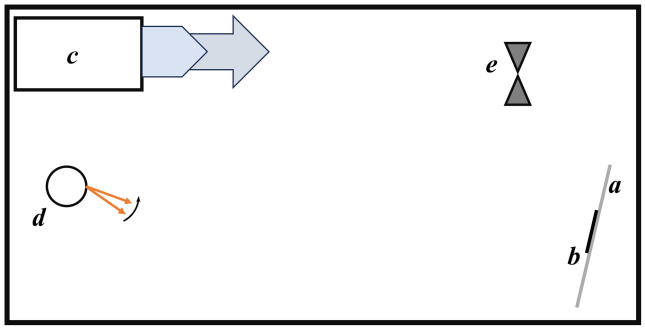

Fig. 8 Fog testing layout in the cargo area of the truck. A 1 m × 1 m 94% reflectivity target (a) was set behind a smaller, narrow 10% reflectivity target (b). A forward scatter meter (e) was placed close to the target without obstructing it from the field of view. The lidar, (d), was placed near the front of the test area, with an artificial fog generator, (c), placed toward the side near the front of the test area.

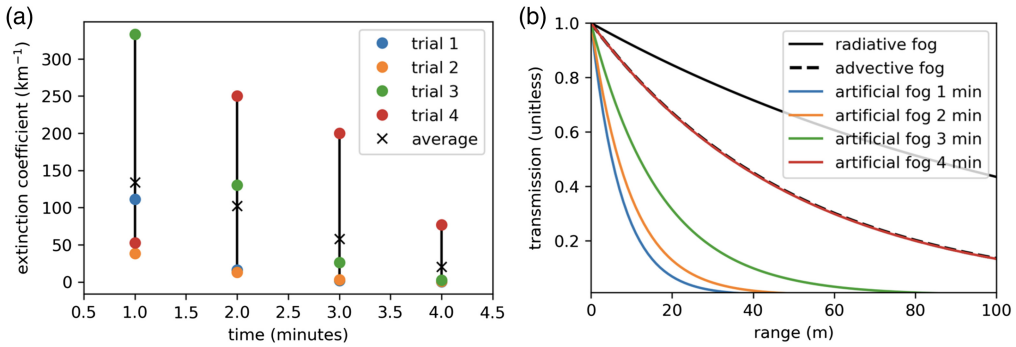

Fig. 9 (b) The artificial fog transmissions are modeled using Eq. (1) and (a) the average extinction values. (a) Artificial fog extinction coefficients averaged over 1-min intervals. (b) A comparison of effective one-way transmission in natural and artificial fogs.

evolved over several minutes. Data from the forward scatter meter show high variability in the extinction coefficient at 850 nm (i.e., the operating wavelength of the forward scatter meter) during the release of the fog into the volume as shown in Fig. 9(a). The forward scatter meter samples in a small region and is not representative of the entire volume, especially during the initial mixing of the vapor into the cavity. The artificial fog transmission over these four trials is extrapolated using the results from Fig. 9(a) with Eq. (1) (presuming a homogeneous mixture) and compared to natural fog transmission at 905 nm in Fig. 9(b). This comparison requires the approximation that Q at 850 and 905 nm are the same. These results show that the extinction conditions achieved in the fog chamber are comparable to natural fog conditions at ranges that are meaningful in terms of automotive lidar, at least for several minutes of testing.

2.8 Laser Safety

Individual lidar units are labeled as class 1 lasers on their specification sheets with either ANSI, IEC, FDA, or CFR standards. The laser classification is based on specific scan patterns and laser pulse characteristics known by the lidar vendor and typically considered proprietary. When many vehicles on the road are operating lidars simultaneously, laser safety becomes more of a concern as laser pulses of various origins and wavelengths overlap in time and space. The purpose of the laser safety study is to assign a safety factor and risk category for exposure to multiple lidars based on the American National Standard for Safe Use of Lasers (ANSI-Z136). ANSI specifies quantitative limits for a single laser source in terms of maximum permissible exposure (MPE), which can be interpreted for multiple laser sources. Laser safety results are not an official endorsement or laser classification from SPIE.

The calculated figures of merit are Safety Factor and Multi-lidar Risk Level. Table 5 summarizes the results for each lidar. We define these terms as:

Table 5 Laser safety results.

Lidar	Safety factor	Multi-lidar risk level	Comment
A	2.25	Low	—
B	>10	Low	—
C	1.75	Moderate	—
D	<1	Moderate to high	Limited by measurement device
E	<1	Moderate to high	Limited by measurement device
F	<1	Moderate to high	Limited by measurement device
G	—	—	Omitted testing
H	—	—	Omitted testing
I	<1	Moderate to high	Limited by measurement device
Riegl	—	—	Omitted testing

Safety factor: The safety factor is the ratio of calculated energy density to the ANSI calculated MPE. The ANSI MPE is based on 10 s of exposure and is 10 cm from the aperture. The power measurements are collected from 10 cm to align with the standard. The value is rounded to the nearest half-integer to account for experimental variations. The higher the value is, the lesser the potential hazard for multiple pulse exposure is.

Multi-lidar Risk Level: The multi-lidar risk levels are low, moderate, or high and defined by the benchmarking team. The categories are assigned based on the calculated safety factor, as well as any noticeable concerns with scan pattern, safety mitigation, effective pulse rate, etc., when considering multiple lidar exposure to the eye. This risk category considers the system beyond the MPE.

Data collection includes average power measurements and effective pulse repetition rate measurements from a 10 cm distance. The average power measurements are performed with an InGaAs photodetector integrating sphere (Thorlabs S145C) that spans an 800 to 1700 nm wavelength range. The effective pulse rate is measured with an InGaAs 2GHz fast photodetector (EOT-3000). The effective pulse repetition rate refers to the frequency of a pulse at a specific point in the scan scene. The measurement instruments are centered at the same position relative to the lidar. Being the first year of laser safety evaluation, the experimental setup only includes the lidar, measurement device, and data acquisition device, as shown in Fig. 10.

The MPE for a pulsed laser source depends on the wavelength λ, pulse width, exposure time T_{exp}, and effective pulse repetition rate PRF_{eff}. MPE is in units of Joules per square centimeter. All lidars had nanosecond class pulse widths, as measured with the fast detector and oscilloscope. The MPE equation derived from the ANSI-Z136 for a nanosecond class IR laser with a PRF of less than 55 kHz is given as

$$\text{MPE} = 5 \times 10^{2(\lambda - 0.700)} (\text{PRF}_{\text{eff}} T_{\text{exp}})^{-1/4} \times 10^{-7}. \tag{3}$$

The accuracy of the measurements was limited by the scan and pulse patterns, solar background, detector sensitivity, and active area of the two devices. Results are shown in the table below for seven of the nine lidars evaluated. Two lidars were omitted due to time constraints.

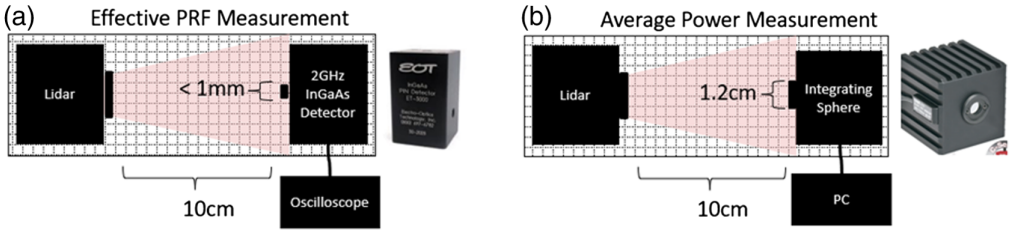

Fig. 10 (a) Effective PRF measurement layout. (b) Average power measurement layout.

Table 6 Safety factor variance for a single lidar using different operating modes.

Lidar	Operating mode	Safety factor	Power (μW)			Energy density (J/cm^2)	MPE (J/cm^2)
			Average	Min	Max		
A	A	1.37	328	152	940	6.44464E-08	8.82404E-08
A	B	1.89	327	153	940	4.18294E-08	7.92629E-08
A	C	1.95	745	428	1072	3.04958E-08	5.96152E-08
A	D	2.10	1017	819	1177	2.49779E-08	5.24681E-08
A	E	2.10	1016	817	1180	2.49533E-08	5.24681E-08
A	F	2.02	720	414	1024	2.94724E-08	5.96152E-08
A	G	2.03	718	410	1043	2.93906E-08	5.96152E-08
A	H	2.96	722	413	1050	1.77326E-08	5.24681E-08
A	I	3.22	674	176	1174	1.62291E-08	5.22090E-08

Calculated safety factors for lidars D, E, F, and I exhibited a mismatch of aperture size between the two measurement devices. This resulted in a lower safety factor. The multi-lidar risk level was subsequently adjusted from a single value to a value range of "moderate to high" because of the uncertainty in the safety factor calculations due to aperture mismatch among other acquisition factors.

To maintain the anonymity of the lidar vendors and their sensor models, we omit any revealing sensor specifications. For example, the wavelength, PRF, and pulse train characteristics are intentionally omitted from the laser safety tables. The final output of our calculations is a single parameter denoted as the safety factor. However, as shown in Table 6, even for a single lidar operating under different modes, the calculated safety factor proportionally changes with respect to the energy density and MPE. In this case, lidar A is tested through several different modes of operation. The modes vary the sensor FOV and scan rate among other parameters.

The laser safety measurements did not consider the unique scan pattern of each lidar; therefore, the measurement process was not ideal, and assumptions could have misled the risk category assignments. An ideal evaluation would require full knowledge of scan patterns, pulse characteristics, and beam characteristics. However, sharing this information may not be preferred by the participating companies. An improved process for future measurements includes an indoor environment to limit the solar background, multiple collection locations in the scan area, and identical aperture sizes on measurement devices. In addition, a thermal power sensor is more commonly used for measuring the average power of short pulses at high repetition frequencies than an integrating sphere, which can produce inconsistent outputs under these pulse conditions.

3 Results

This section shows the results for lane 1, lane 2, lane 3, and fog testing. Tabulated data and additional comments are given in Appendix 6.2.

3.1 Lane 1: Long Range, No Road Sign Confusers

Figure 11 shows the points per square meter for each lidar and lane 1 target. The logarithmic y-axis shows the points on target per square meter. The dual x-axis shows both the target name and range in meters.

The figures have been separated based on the target reflectivity. Absent results can be attributed to one or more of the following: target displacement due to high wind, target outside lidar FOV, and target outside lidar range.

We assume each lidar point cloud to be from a single 10-s persistent scan with ± 1 s scan time deviations. The point density plots from lane 1 show that, as the range increases, the points

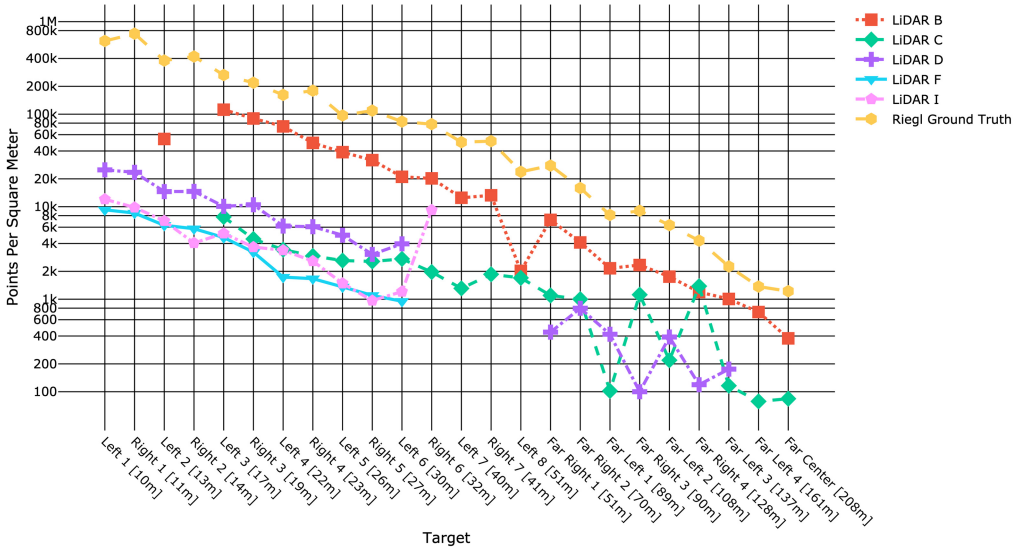

Fig. 11 Lane 1: range versus points per square meter.

per square meter decreases. Outliers in the negative slope of this trendline are due to different scan pattern techniques. For example, the pink trace lidar I had nearly as many points per square meter for its farthest target (right 6) when compared with its nearest target (left 1). The scan pattern of this particular lidar was likely more beneficial on right 6, thereby breaking the negative trendline. Without revealing specifics about lidar I, a more beneficial scan pattern could mean several things. For example, a line-scanning lidar might have additional lines on a further target compared to a closer target that might have only had a single line. Other lidars have embedded software routines, such as automatic gain control or power thresholding based on range, that can skew the trendlines of these results. Similar outliers can be seen on the B, C, and D traces.

Figure 12 shows the lane 1 range precision in centimeters. We know the specification of the Riegl ground truth lidar to have a 3 mm range precision at 100 m. Observing the yellow trace near targets left 7 and right 7, which are ∼100 m in range, we can confirm that the Riegl lidar is 0.3 cm or 3 mm in range precision.

The range precision results from lane 1 show that, as the range increases, the range precision degrades. Due to the logarithmic y-axis scale, this positive slope trendline is most obvious with the Riegl ground truth as its range precision degrades by nearly 1 cm over the >200 m range.

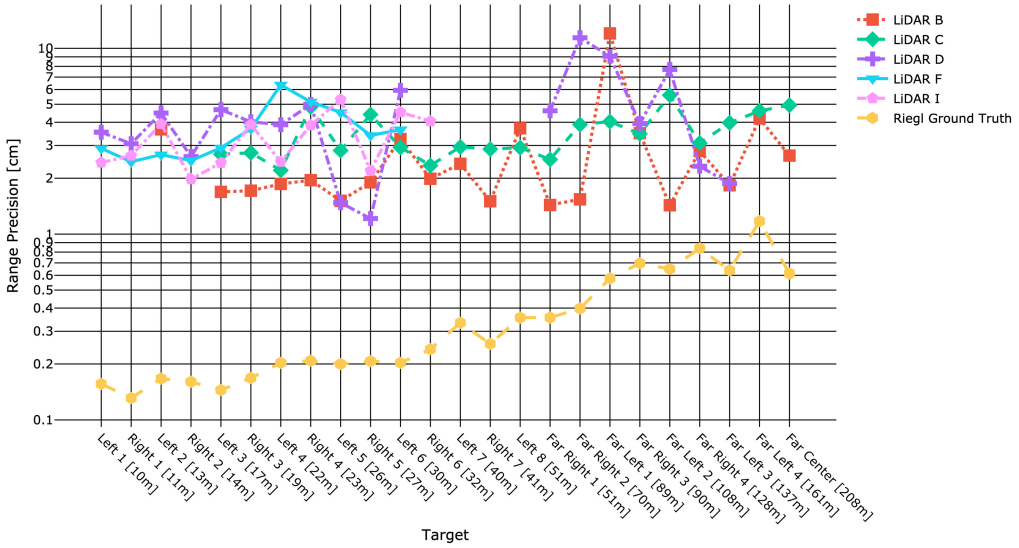

Fig. 12 Lane 1: range precision.

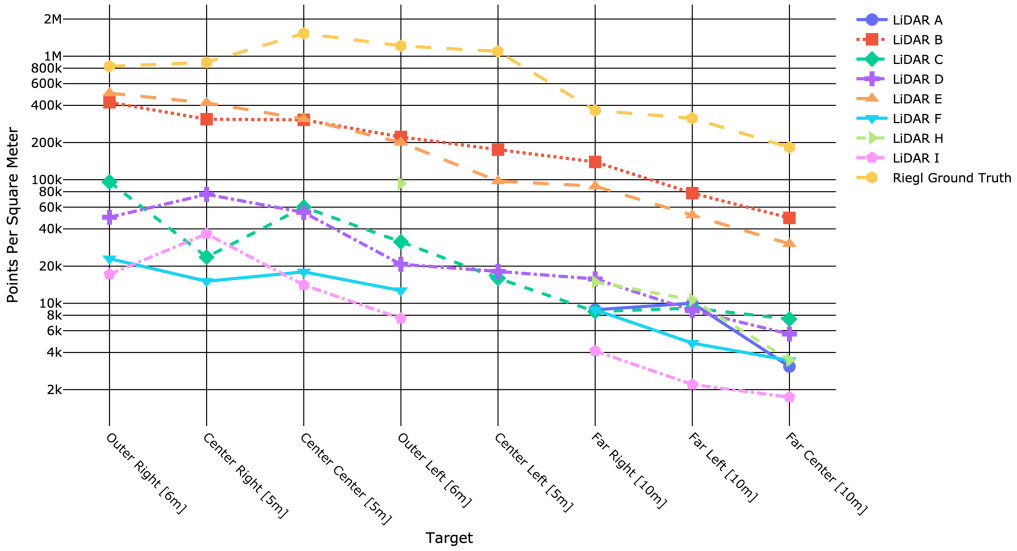

Fig. 13 Lane 2: range versus points per square meter.

The tested lidars follow similar trendlines increasing 1 cm or more. For example, a degradation from 2 to 3 cm for lidar B or from 3 to 4 cm for lidar F and lidar I.

3.2 Lane 2: Short Range, No Road Sign Confusers

The points per square meter for lane 2 show a similar trendline to lane 1. As shown in Fig. 13, when the range increases, the points per square meter decrease. The far center target corresponds to a range of ~50 m. When compared with lane 1, targets left 4 and right 4 were at similar ranges. The points on target are consistent between lanes 1 and 2 thus far. For example, lidar I has 2k points on both lane 2, far center, and lane 1, left 4.

Several lidars did not accumulate points on target for center left. This target was orientated with the 80 cm edge along the ground and only 15 cm in the air. Depending on the scan pattern of the lidar, certain low-profile targets could be omitted or out of the FOV. Targets center right and center center were also orientated as low-profile targets, but slightly offset in position from center left as shown in Fig. 3 previously.

As shown in Fig. 14, the range precision trendlines for lane 2 are not as clear as lane 1 due to the shorter range and low-profile targets. The Riegl ground truth was still below 1 cm on average,

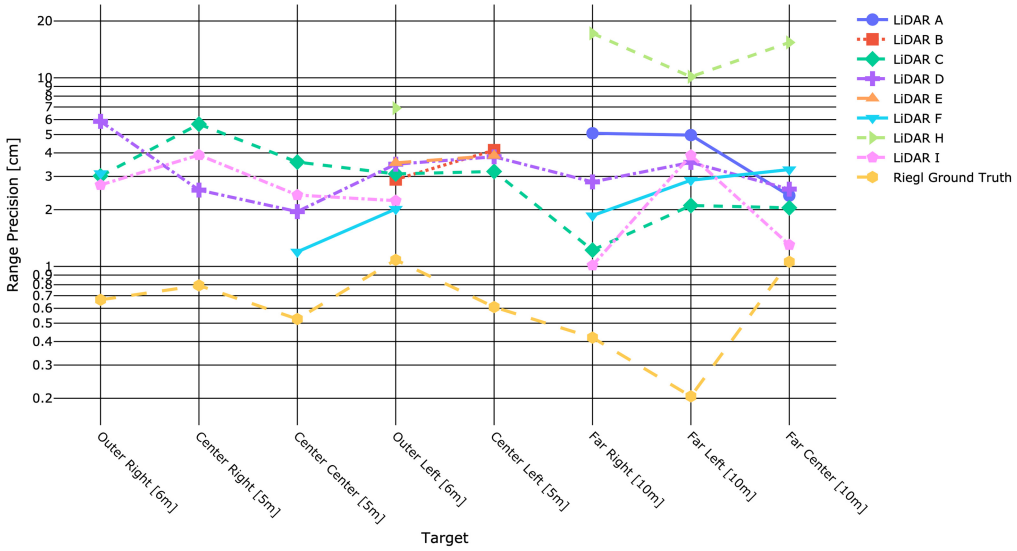

Fig. 14 Lane 2: range precision.

whereas the other lidars remain in the 2 cm to 3 cm range. Lidar H featured the highest range precision in the 10 to 20 cm range.

3.3 Lane 3: Long Range, with Road Sign Confusers

Figure 15 shows the points on target for the lidars that were tested on lane 3, namely B, C, E, and the Riegl. Lane 3 used the calibrated Labsphere targets and additional confusers or road signs positioned next to the calibrated targets.

In Fig. 15, we note the general trend of increased points on target for the retroreflective road sign. For example, the first target in the series, target right 6, accumulates ~2000 points for lidar B, whereas the road sign next to it receives 15 times more points on target. This same trend although not as extreme occurs at further ranges as well. The original right 6 target for lane 1 received a similar number of points on target as shown in Table 11. These results are expected as the embedded corner cubes in the road signs promote higher received photons. This test was intended to check the dynamic range and see if any sort of performance degradation occurs when highly directional reflectors are positioned next to low-reflectance Lambertian scatters.

In Fig. 16, we note again how the Riegl ground truth features the best range precision, despite adding confusers or road signs next to the targets.

When comparing two targets with and without a confuser next to them for three different lidars, one possible conclusion is that the points on target and the range precision decrease when a confuser is present. Tables 7 and 8 show this trend for the average, but it is not consistent for each lidar. This is a significant point of discussion as the addition of a road sign confuser is to simulate the corner case of a child in low reflectivity clothing near a highly reflective road sign. The embedded corner cubes in the road sign redirect a significant portion of the incident light back to the lidar, which could saturate the range bin returns and thereby degrade the dynamic range and the ability to detect the lower reflectivity human target. Although the dynamic range is not directly measured, Tables 7 and 8 show a degradation in performance with the addition of a road sign confuser. On average, the points on target decrease by 7.3% up to 10.5%, and the range precision decreases by 4.2% up to 19.1%.

With the usage of 10% and 94% reflectance targets, there is a considerable difference between both points on target and range precision. With 10% reflectance targets being of higher importance for recognizing targets of interest with respect to high reflectance backgrounds, this discrepancy is highly valuable to determine the capabilities of different lidars. Table 9 shows the discrepancies between two automotive lidars using a percentage comparison of points on target. For lidar B, it is shown that, for most target locations, there are less points available than the surrounding areas (lane 3, left 1 is considered noisy due to having an extra vertical line of points). For lidar C, there is an even larger discrepancy with target points, with some locations having

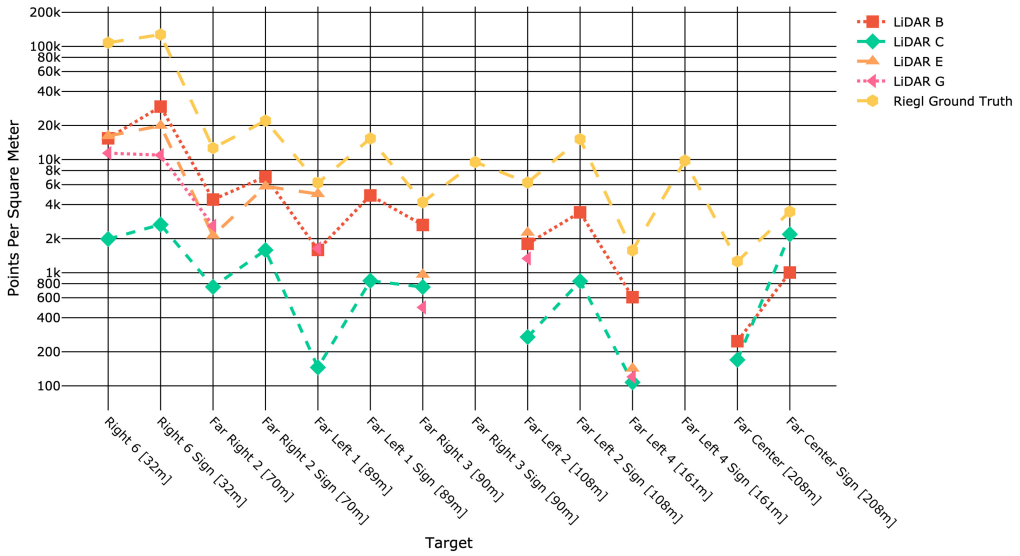

Fig. 15 Lane 3 range versus points per square meter.

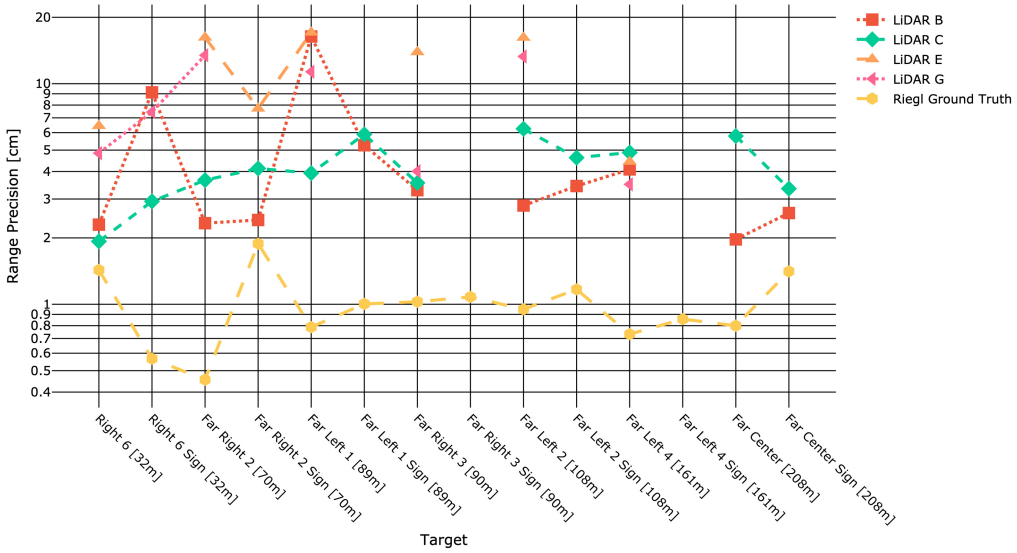

Fig. 16 Lane 3: range precision.

Table 7 Confuser comparison: points on target.

Lidar	Lane 1 right 6	Lane 3 right 6	% change	Lane 1 far center	Lane 3 far center	% change
B	2425	1848	−23.7%	1131	744	−34.2%
C	236	238	+1.0%	252	509	+50.4%
Riegl	935	1293	+27.6%	366	378	+3.1%
Average	1198	1126	−10.5%	583	543	−7.3%

Table 8 Confuser comparison: range precision.

Lidar	Lane 1 right 6	Lane 3 right 6	% change	Lane 1 far center	Lane 3 far center	% change
B	1.99	2.30	−13.0%	2.64	1.97	+34.0%
C	2.34	1.93	+21.2%	4.94	5.79	−14.6%
Riegl	0.24	1.43	−83.2%	0.61	0.80	−23.7%
Average	1.52	1.88	−19.1%	2.73	2.85	−4.2%

Table 9 Low reflectance target confusion comparison: points on target.

Lidar	Lane 3 right 1	Lane 3 right 2	Lane 3 right 3	Lane 3 right 4	Lane 3 left 1	Lane 3 left 2	Lane 3 left 3	Lane 3 left 4
B	58.2%	86.0%	85.0%	75.2%	212.5%	94.9%	99.8%	74.6%
C	87.0%	76.3%	32.4%	0.0%	9.8%	0.0%	0.0%	0.0%

an invisible low reflectance target when paired with the high reflectance target. Table 10 shows the percentage range precision on the 10% reflectance targets when compared with the 94% reflectance targets. There is a considerable increase in the range deviation target, which can lead to the possible mis-detection of targets of interest.

Table 10 Low reflectance target confusion comparison: range precision.

Lidar	Lane 3 right 1	Lane 3 right 2	Lane 3 right 3	Lane 3 right 4	Lane 3 left 1	Lane 3 left 2	Lane 3 left 3	Lane 3 left 4
B	79.1%	146.5%	135.1%	229.0%	128.9%	118.8%	213.9%	730.6%
C	224.4%	249.7%	165.3%	N/A	92.4%	N/A	N/A	N/A

3.4 Backscattering in Water-Glycol Fog

The purpose of this section is to highlight the prevalence of backscatter in the lidar returns captured in our water-glycol fog chamber. The total number of points in the point cloud is the sum of the backscattered points from the fog, n_{bs}, and points detected on hard targets in the scene, n_t, i.e., $n_{\text{tot}} = n_{bs} + n_t$. The ratio of backscattered points in the volume is therefore n_{bs}/n_{tot}. A point is considered to be backscatter (and included in n_{bs}) when it is outside all points registered in the scene prior to fog release and of any points within 25 cm of these points. The points n_t are all points that do not meet the criteria to be included in n_{bs}.

Figure 17 displays the backscatter ratio as it evolves over 500 scans for one of the lidars tested in this research. The backscatter ratio displays a marked drop initially and then increases to a maximum before the backscatter ratio decreases abruptly twice during the test. This may be suggestive of either adaptive processing in the specific lidar under test or possibly wind turbulence nonuniformly reducing the density of the water-glycol fog within the testing chamber. The cause of these differences is under investigation.

Figure 18 displays the dissipation of the fog from the time of release in three dimensions, as captured by test lidars B and C.

As shown in Fig. 18, the fog degrades the capability of lidars B and C to detect points on target. This behavior was seen for all lidars tested for the fog test, including lidars A, B, C, D, I, and F. Figures 18(a) and 18(c) display the results previous to the release of fog, and Figs. 18(b) and 18(d) display the point cloud at the time when scattering is at maximum for lidars B and C, respectively. This should correlate with when the fog is at maximum density and extinction is most significant. As displayed in Figs. 18(b) and 18(d), a secondary return in range closer to the lidar is encountered, which is characteristic of backscattering. Scattered returns are detected in a planar cluster: in the case of lidar B, close to a range of 3.0 m and for lidar C, closer to 2.0 m. The target and back wall that should be detected are both closer to 4.0 m from the lidar under test in both cases; thus a greater than 1 m error is introduced at short ranges. As shown in Fig. 18, scattered returns are prevalent shortly after the release of the fog and fill the volume. In this same time frame, the number of points reaching the back of the target area has diminished. The visualizations in this section are evidence that performance in artificial fog is driven by both increased backscattered returns and a reduction in target returns, whereas most models of lidar performance only consider the latter effect to be important.

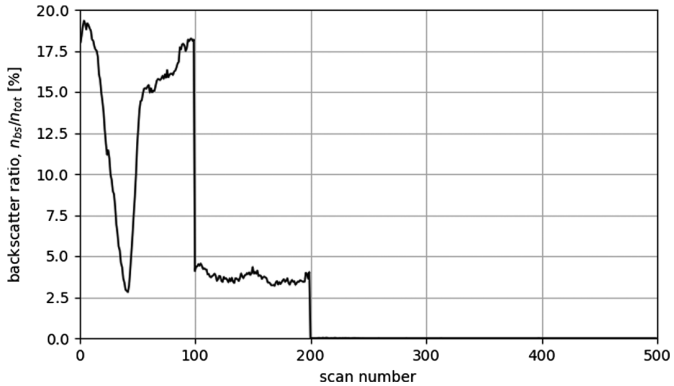

Fig. 17 Ratio of backscattered points in the volume, n_{bs}/n_{tot}, plotted for lidar B.

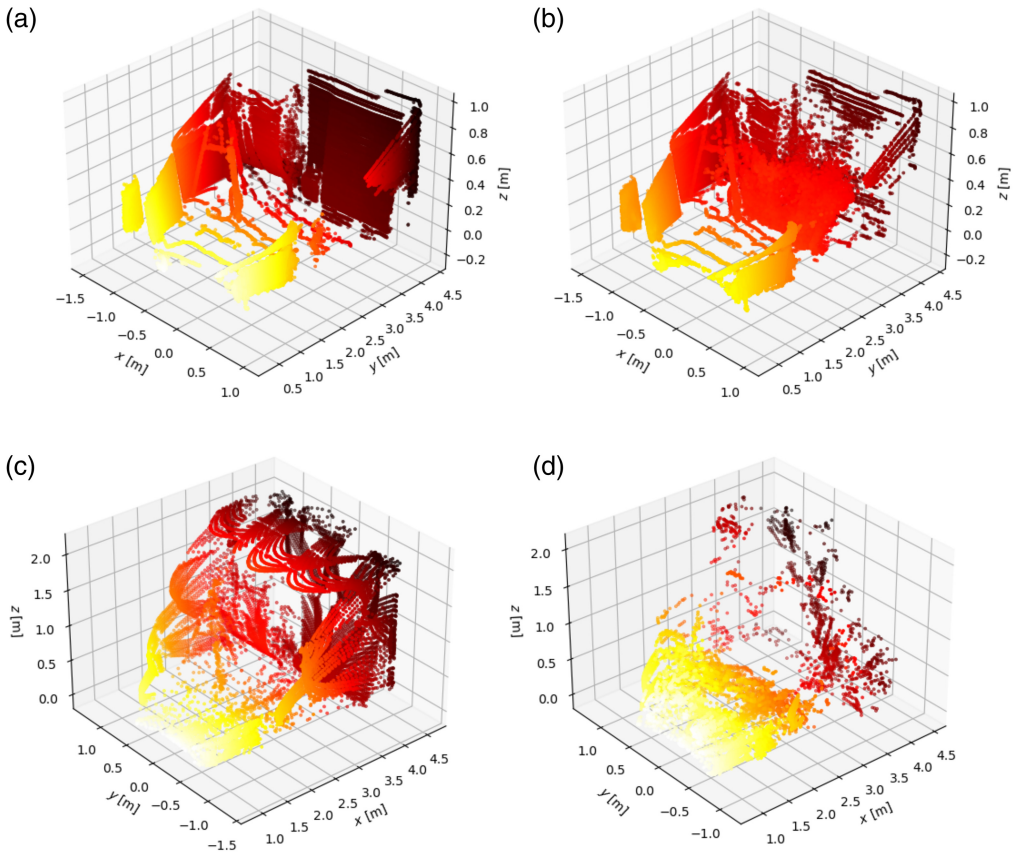

Fig. 18 Return for lidar B (a) without fog and (b) with fog. The return for lidar C (c) without fog and (d) with fog.

4 Conclusion

This paper presented the benchmarking results for nine automotive-grade lidars for three different testing lanes in addition to artificial fog tests and laser safety.

The lane tests showed that, as the range increased, the range precision degraded and the points per square meter decreased. On average, the addition of road signs placed alongside low-reflectivity targets resulted in 4.2% to 19.1% degraded range precision and decreased points per square meter compared with the same targets with no road signs. This result confirmed the assertion that road signs have the potential to act as confusers and degrade lidar performance when placed near low-reflectivity targets. The lane tests are analogous to a small child wearing low-reflectivity clothing next to a road sign with embedded corner cubes, and the results showed how the lidar's ability to detect the low-reflectivity target degrades.

The fog tests showed that extinction in the water-glycol fog chamber was comparable to natural fog conditions at ranges that are meaningful for automotive lidar. Water-glycol fog also exhibited appreciable backscatter, which violates a common assumption in lidar modeling. More work is required to determine if this outcome scales at longer ranges and for natural fog particle distributions. For 2024, we will attempt to improve the spatial and temporal uniformity of the fog, increase the temporal resolution of our measurements, and measure the fog particle diameter and density. This improved characterization is required before we can state definitive conclusions on the applicability of artificial fog in automotive lidar testing.

The laser safety testing defined safety factors and risk categories for the tested lidars. Although manufacturers are aware of using wavelength to their advantage in addition to power scaling techniques and altering the FOV, these laser safety considerations are only for a single device. Each individual lidar was advertised as being eye safe; however, the assigned safety factors and risk categories allows us to place laser safety in the context of multiple lidar exposure. We intend to expand and scale our laser safety testing to simulate corner cases, such as multi-lane

highways, in which several eye-safe devices could potentially scan a target in a short enough amount of time that it becomes a hazard for the human eye.

These year two tests expanded upon previous year one testing efforts with the addition of larger 1 m^2 calibrated targets and a wider variety of lidars. In trying to quantify the performance differences and verify the vendor claims for each lidar, we intend to motivate this industry toward developing testing standards such that every lidar can be assessed and compared to a known standard.

5 Year 3 Expected Approach

A chief concern in 2024 will be knowledge of the extinction coefficient for water-glycol fog. This requirement on α dictates knowledge of n and some degree of control over N. The fog generator used in the 2023 experiment offers a potential means to adjust N by increasing the fog density through a timed-release mechanism. For 2024, this timed release can be driven to produce the desired N through feedback from a particle counter. Ideally, the particle counter data will be cross-referenced with transmissometer or forward scatter meter data, as was used in the 2023 experiment. A transmissometer would be preferred because these measurements accumulate scattering along the entire path, whereas the forward scatter meter effectively samples at a point. In both cases, the artificial fog will need to be temporally stable and adequately mixed in the volume of air seen by the lidar and the transmissometer over the measurement duration.

If these experiments support the assumption that extinction is the only important consideration when evaluating automotive lidar performance in fog, then it may be possible to eliminate the need to test in real/artificial fog entirely. Instead, the fog effect could be reproduced with a filter at the receive aperture, scaled according to the Beer-Lambert law for the desired range, fog type, and wavelength. We do not have enough data currently to determine if this is a viable approach, but it is the logical outcome if this common assumption in the literature is true.

6 Appendix

6.1 Weather Log

Figures 19 and 20 show the temperature, precipitation, wind speed and gust for the two testing days.

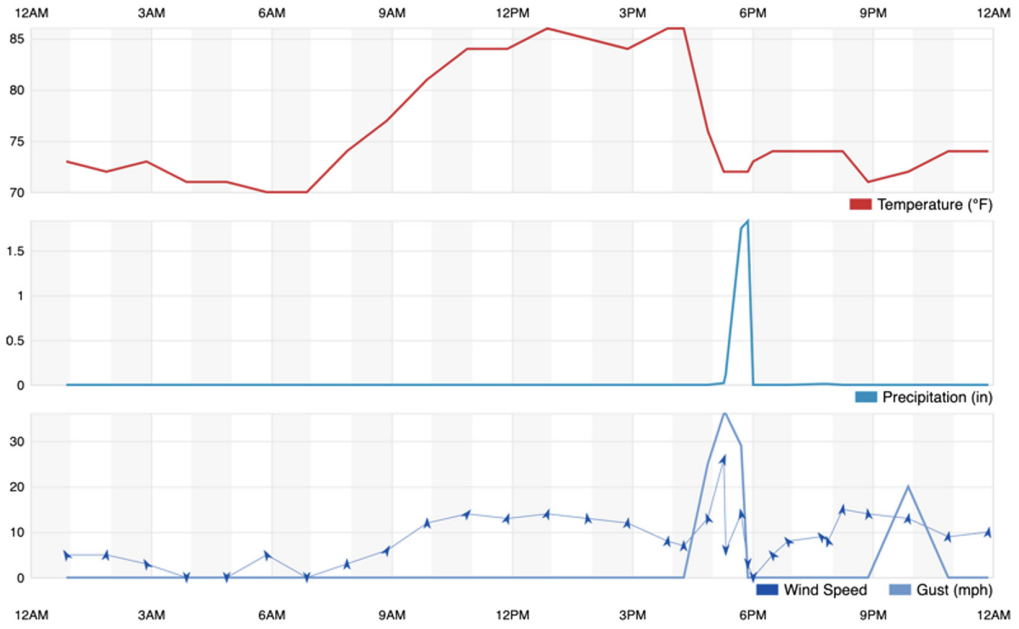

Fig. 19 April 29, 2023, Saturday weather conditions, nominal day with brief thunderstorm after teardown.[14]

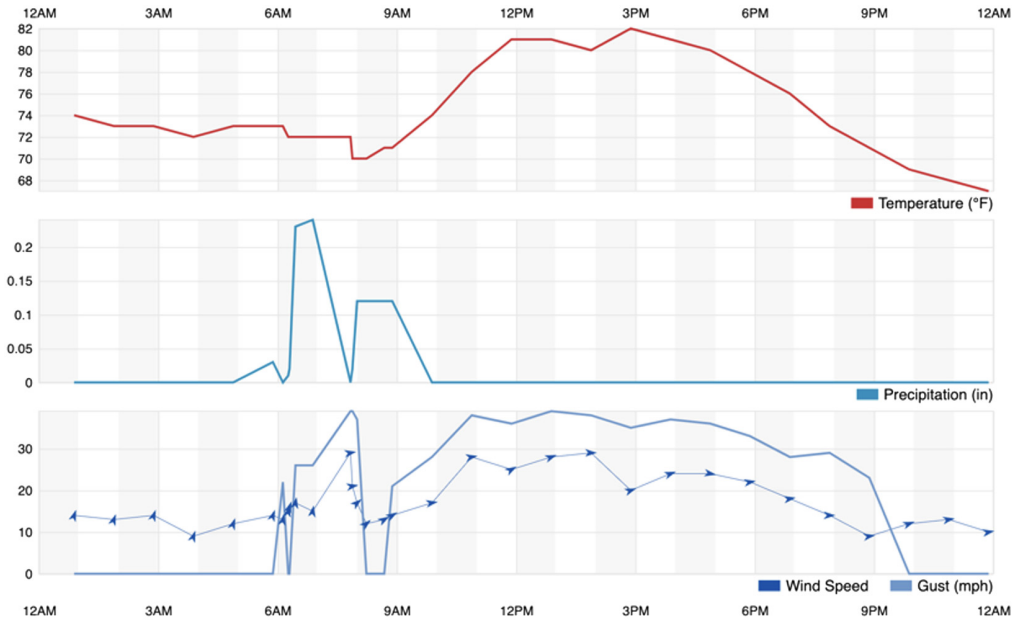

Fig. 20 April 30, 2023, Sunday weather conditions, clear day but with persistent high winds.[15]

6.2 Tabulated Results

Tables 11–16 are the points on target and range precision processed results for the lidar lane tests.

Table 11 Lane 1: points on Target.

Target	A	B	C	D	E	F	G	H	I	Riegl
Left 1	—	—	—	3000	—	1110	—	—	1444	7411
Left 2	—	6473	—	1747	—	755	—	—	843	4558
Left 3	—	13,412	930	1210	—	557	—	—	617	3176
Left 4	—	8853	412	741	—	208	—	—	406	1940
Left 5	—	4660	314	590	—	163	—	—	178	1162
Left 6	—	2521	327	476	—	114	—	—	145	998
Left 7	—	1497	157	—	—	—	—	—	—	598
Left 8	—	243	204	—	—	—	—	—	—	286
Right 1	—	—	—	2820	—	1024	—	—	1176	8889
Right 2	—	—	—	1755	—	694	—	—	486	5040
Right 3	—	10,771	537	1259	—	387	—	—	439	2623
Right 4	—	5865	352	729	—	200	—	—	310	2154
Right 5	—	3825	307	366	—	132	—	—	116	1312
Right 6	—	2425	236	—	—	—	—	—	1103	935
Right 7	—	1593	223	—	—	—	—	—	—	613
Far left 1	—	4314	204	836	—	—	—	—	—	1626
Far left 2	—	3506	438	780	—	—	—	—	—	1263
Far left 3	—	2014	232	351	—	—	—	—	—	452

Table 11 (*Continued*).

Target	A	B	C	D	E	F	G	H	I	Riegl
Far left 4	—	2177	236	—	—	—	—	—	—	411
Far right 1	—	14,413	2196	881	—	—	—	—	—	5564
Far right 2	—	8216	2000	1596	—	—	—	—	—	3182
Far right 3	—	7024	3357	300	—	—	—	—	—	2710
Far right 4	—	3581	4147	357	—	—	—	—	—	1288
Far center	—	1131	252	v	—	—	—	—	—	366
Average	—	5167	853	1099	—	485	—	—	605	2439

Table 12 Lane 1: range precision (cm).

Target	A	B	C	D	E	F	G	H	I	Riegl
Left 1	—	—	—	3.54	—	2.89	—	—	2.43	0.16
Left 2	—	3.66	—	4.48	—	2.68	—	—	3.92	0.17
Left 3	—	1.69	2.69	4.68	—	2.92	—	—	2.42	0.14
Left 4	—	1.86	2.21	3.88	—	6.36	—	—	2.45	0.20
Left 5	—	1.51	2.82	1.47	—	4.52	—	—	5.30	0.20
Left 6	—	3.25	2.93	5.94	—	3.65	—	—	4.53	0.20
Left 7	—	2.39	2.95	—	—	—	—	—	—	0.33
Left 8	—	3.71	2.92	—	—	—	—	—	—	0.36
Right 1	—	—	—	3.07	—	2.45	—	—	2.65	0.13
Right 2	—	—	—	2.61	—	2.49	—	—	1.98	0.16
Right 3	—	1.71	2.74	4.02	—	3.69	—	—	3.89	0.17
Right 4	—	1.95	4.84	4.99	—	5.15	—	—	3.86	0.21
Right 5	—	1.90	4.40	1.21	—	3.40	—	—	2.19	0.21
Right 6	—	1.99	2.34	—	—	—	—	—	4.05	0.24
Right 7	—	1.50	2.87	—	—	—	—	—	—	0.26
Far left 1	—	12.03	4.03	9.06	—	—	—	—	—	0.58
Far left 2	—	1.43	5.59	7.69	—	—	—	—	—	0.65
Far left 3	—	1.83	3.97	1.87	—	—	—	—	—	0.64
Far left 4	—	4.19	4.61	—	—	—	—	—	—	1.17
Far right 1	—	1.43	2.52	4.61	—	—	—	—	—	0.36
Far right 2	—	1.54	3.89	11.41	—	—	—	—	—	0.40
Far right 3	—	3.49	3.47	3.94	—	—	—	—	—	0.70
Far right 4	—	2.80	3.09	2.32	—	—	—	—	—	0.84
Far center	—	2.64	4.94	—	—	—	—	—	—	0.61
Average	—	2.79	3.49	4.49	—	3.65	—	—	3.30	0.38

Table 13 Lane 2: points on target.

Target	A	B	C	D	E	F	G	H	I	Riegl
Outer left	—	26,648	3773	2483	24,010	1522	—	11,189	905	14,564
Outer right	—	50,811	11,527	5964	60,268	2757	—	—	2061	9942
Center left	—	21,047	1914	2176	11,712	—	—	—	—	13,116
Center center	—	36,646	7203	6528	37,288	2156	—	—	1689	18,355
Center right	—	37,106	2839	9139	50,298	1810	—	—	4382	10,704
Far left	10,028	77,896	9143	8757	51,621	4758	—	10,656	2201	31,441
Far center	6149	98,141	14,905	11,276	60,571	6887	—	6934	3479	36,583
Far right	8860	139,396	8570	15,741	88,708	8860	—	14,909	4117	36,406
Average	8345	60,961	7484	7758	48,059	4107	—	10,922	2690	21,388

Table 14 Lane 2: range precision (cm).

Target	A	B	C	D	E	F	G	H	I	Riegl
Outer left	—	2.89	3.09	3.48	3.54	2.02	—	6.91	2.23	1.08
Outer right	—	—	3.01	5.88	—	3.11	—	—	2.70	0.67
Center left	—	4.13	3.18	3.81	3.86	—	—	—	—	0.61
Center center	—	—	3.57	1.95	—	1.19	—	—	2.39	0.53
Center right	—	—	3.67	2.54	—	—	—	—	3.90	0.79
Far left	4.97	—	2.10	3.56	—	2.86	—	10.13	3.88	0.20
Far center	2.37	—	2.04	2.57	—	3.26	—	15.40	1.30	1.06
Far right	5.08	—	1.22	2.81	—	1.86	—	17.20	1.01	0.42
Average	4.14	3.51	2.73	3.33	3.70	2.39	—	12.41	2.49	0.67

Table 15 Lane 3: points on target.

Target	A	B	C	D	E	F	G	H	I	Riegl
Right 6	—	1848	238	—	1937	—	1367	—	—	1293
Right 6 sign	—	29388	2663	—	19922	—	10936	—	—	12,728
Far left 1	—	3181	291	—	9925	—	3279	—	—	1242
Far left 1 sign	—	4816	848	—	—	—	—	—	—	1535
Far left 2	—	3602	540	—	4522		2671	—	—	1252
Far left 2 sign	—	3414	840	—	—	—	—	—	—	1516
Far left 4	—	1822	322	—	427	—	360	—	—	472
Far left 4 sign	—	—	—	—	—	—	—	—	—	979
Far right 2	—	8859	1498	—	4256	—	5201	—	—	2527
Far right 2 sign	—	7059	1585	—	5794	—	—	—	—	2203
Far right 3	—	7919	2239	—	2886		1480	—	—	1258
Far right 3 sign	—	—	—	—	—	—	—	—	—	950
Far center	—	744	509	—	—	—	—	—	—	378
Far center sign	—	1004	2186	—	—	—	—	—	—	346
Average	—	6138	1146	—	4298	—	3613	—	—	2048

Table 16 Lane 3: range precision (cm).

Target	A	B	C	D	E	F	G	H	I	Riegl
Right 6	—	2.30	1.93	—	6.44	—	4.84	—	—	1.43
Right 6 sign	—	9.14	2.93	—	—	—	7.44	—	—	0.57
Far left 1	—	16.37	3.93	—	17.06	—	11.32	—	—	0.79
Far left 1 sign	—	5.26	5.89	—	—	—	—	—	—	1.00
Far left 2	—	2.80	6.24	—	16.16	—	13.28	—	—	0.95
Far left 2 sign	—	3.44	4.62	—	—	—	—	—	—	1.17
Far left 4	—	4.09	4.88	—	4.43	—	3.49	—	—	0.73
Far left 4 sign	—	—	—	—	—	—	—	—	—	0.86
Far right 2	—	2.33	3.65	—	16.20	—	13.45	—	—	0.46
Far right 2 sign	—	2.41	4.13	—	7.72	—	—	—	—	1.88
Far right 3	—	3.29	3.55	—	13.90	—	4.00	—	—	1.03
Far right 3 sign	—	—	—	—	—	—	—	—	—	1.08
Far center	—	1.97	5.79	—	—	—	—	—	—	0.08
Far center sign	—	2.59	3.34	—	—	—	—	—	—	1.41
Average	—	4.67	4.24	—	11.70	—	8.26	—	—	1.01

Code and Data Availability

The processed data that support the findings of this article are publicly available and listed in Sec. 6.2.

Additional data requests can be directed to eddie@excitingtechnology.com.

Acknowledgments

We would like to thank Labsphere and Chris Durell for once again providing both calibrated targets and personnel to assist in benchmarking activities. The University of Dayton Vision Lab for processing the point cloud data. Riegl for loaning Charlie Kershner of NGA a survey-grade lidar capable of being our ground truth lidar. Eyas Ayesh and Jacquelyn Van Zyl for helping with target setup and teardown. Oles Fylypiv for assisting with fog data acquisition. Tianyue Yu for assisting with point-cloud data acquisition. All of the organizations that provided personnel to volunteer on this benchmarking effort. SPIE for supporting these tests as a community service. Special thanks to Peter Hallett and Randy Arabie for their continued support. Those companies that provided test lidars. We believe this is an excellent service to the community.

References

1. M. Zar, "SAE active safety system LiDAR performance task force," SAE International, https://standardsworks.sae.org/standards-committees/active-safety-system-lidar-performance-task-force (accessed 17 July 2024).
2. Z. Jeffries et al., "Toward open benchmark tests for automotive lidars, year 1: static range error, accuracy, and precision," *Opt. Eng.* **62**(3), 031211 (2023).
3. C. Li and Y. Cheng, "7th WG meeting minutes (unapproved)" (2021). https://sagroups.ieee.org/2936/wp-content/uploads/sites/350/2022/02/709_Draft_Meeting-Minutes-P2936-7th-WG-meeting-20210802.pdf (accessed 21 July 2021).
4. T. H. Aspiras et al., "Towards a deep-learning aided point cloud labeling suite," *Proc. SPIE* **PC12101**, PC121010A (2022).
5. M. Kutila et al., "Automotive LiDAR performance verification in fog and rain," in *21st Int. Conf. Intell. Transport. Syst. (ITSC)*, IEEE (2018).

6. R. H. Rasshofer, M. Spies, and H. Spies, "Influences of weather phenomena on automotive laser radar systems," *Adv. Radio Sci.* **9**, 49–60 (2011).
7. Y. Li and J. Ibanez-Guzman, "LiDAR for autonomous driving: the principles, challenges, and trends for automotive LiDAR and perception systems," *IEEE Signal Process. Mag.* **37**(4), 50–61 (2020).
8. 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," *Proc. SPIE* **11002**, 110021F (2019).
9. C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles*, John Wiley & Sons (2008).
10. Scatlib index, http://atol.ucsd.edu/scatlib/index.htm (2008).
11. E. P. Shettle and R. W. Fenn, "*Models for the Aerosols of the Lower Atmosphere and the Effects of Humidity Variations on their Optical Properties*," Tech. Rep. 676, Optical Physics Division, Air Force Geophysics Laboratory (1979).
12. J. B. Wright et al., "Optical characterization of the Sandia fog facility," *Proc. SPIE* **10197**, 1019704 (2017).
13. N. Hagen, "The influence of natural and artificial fogs on visible and infrared imaging," *Proc. SPIE* **10765**, 1076504 (2018).
14. TWC Product and Technology LLC, "Weather conditions day 1," (2023). https://www.wunderground.com/history/daily/us/fl/orlando/KMCO/date/2023-4-29.
15. TWC Product and Technology LLC, "Weather conditions day 2," (2023). https://www.wunderground.com/history/daily/us/fl/orlando/KMCO/date/2023-4-30.

Theus H. Aspiras is a research engineer in the Electrical and Computer Engineering Department at the University of Dayton. His current research areas include brain signal analysis, object detection and tracking, machine learning, deep learning, and unsupervised/semi-supervised learning. He holds two patents with collaborators at AFRL and has published over 30 research papers in computer vision, pattern recognition, and machine learning.

Vijayan K. Asari is a professor in electrical and computer engineering and Ohio Research Scholars Endowed Chair in wide area surveillance at the University of Dayton. He directs the Center of Excellence for Computational Intelligence and Machine Vision. His research focuses on algorithms for object detection, recognition, and tracking. He has five U.S. patents and over 750 publications and has supervised numerous PhD dissertations. He is a senior member of IEEE and a fellow of SPIE.

Cullen Bradley is the research operation manager for Exciting Technology LLC and a researcher at the University of Dayton. He grows, designs, and fabricates EO crystal devices for non-mechanical beam steering. He received his MS degree in electro-optics from the University of Dayton in 2013 and his BS degree in physics with a math minor from St. John Fisher College in 2010. He has over 10 years of experience in electro-optics engineering.

Anna Gnacek is a research engineer at Exciting Technology, LLC. She received her BS and MS degrees in electrical engineering from the University of Dayton in 2019 and 2020, respectively. Her research focuses on LiDAR system signal processing. She has worked on various applications, including sense and avoid, space-situational awareness, and 3D mapping. She handles all stages of LiDAR design, from trade studies and system modeling to data analysis and performance evaluation.

Charlie Kershner: Biography is not available.

Daniel A. LeMaster is the senior scientist for sensors and perception at the U.S. Department of Transportation's Highly Automated Systems Safety Center. He previously held leadership roles at the Air Force Research Laboratory. Within SPIE, he has served as a senior editor for *Optical Engineering* and a publications committee member. His awards include recognitions from the Air Force, NATO, and SPIE. He is a fellow of SPIE and the Military Sensing Symposia (MSS).

Paul F. McManamon is a president of Exciting Technology LLC, chief science officer at Nuview, and technical director of the Lidar and Optical Communications Institute at the University of Dayton. He chaired the US National Academy of Sciences study on "Laser Radar" (2014) and was Uber's expert witness in the lawsuit against Google/Waymo. He has written two books on lidar and is a fellow of SPIE, IEEE, Optica, AFRL, DEPs, MSS, and AIAA.

Andrew D. Reinhardt is a research associate through the National Academy of Sciences at the Air Force Research Lab, sensors directorate. He received his PhD in electro-optics and photonics from the University of Dayton in 2021. He is the founder of QuatroSens, LLC. His research interests span passive and active sensing, optical system design, and advanced topics in optical sensing. He is a member of SPIE.

Eddie Ruff received his MS degree in electro-optics and photonics from the University of Dayton in 2018 and his BS degree in electrical and computer engineering in 2016. Currently, he works for Exciting Technology LLC. He has expertise in high-power supercontinuum pulsed lasers, EO detectors, rapid prototyping, tomographic reconstruction, and real-time signal processing. He is proficient in C++, Julia, Python, and Matlab and maintains the software and hardware interfaces for Exciting Technology's FSOC systems.

Christopher R. Valenta, P.E. is a principal research engineer at the Georgia Tech Research Institute and an adjunct professor at Georgia Tech. His research interests include remote sensing, electro-optics, lidar, and more. His work has been featured in Space News, Aviation Week, Wired, and other outlets. He is a senior member of IEEE and SPIE, a member of OSA and DEPS, and has won several awards, including the 2015 IEEE Microwave Magazine Best Paper Award.

Andre Van Rynbach is a research electronics engineer at the Air Force Research Laboratory in the Sensors Directorate at Wright-Patterson AFB, Ohio, USA. His research spans many areas of multi-spectral sensing technologies, including lidar system design, optical imaging, and 3D point cloud processing. He is a member of SPIE and the Society for Imaging Science and Technology.