2020-01-1206 Published 14 Apr 2020

Driving Safety Performance Assessment Metrics for ADS-Equipped Vehicles

Jeffrey Wishart and Steven Como Exponent Inc., Arizona State University

Maria Elli Intel

Brendan Russo Northern Arizona University

Jack Weast Intel

Niraj Altekar University of Arizona

Emmanuel James Northern Arizona University

Yan Chen Arizona State University

Citation: Wishart, J., Como, S., Elli, M., Russo, B. et al., "Driving Safety Performance Assessment Metrics for ADS-Equipped Vehicles," *SAE Int. J. Advances & Curr. Prac. in Mobility* 2(5):2881-2899, 2020, doi:10.4271/2020-01-1206.

This article is from WCX™ 2020 SAE World Congress Experience.

Abstract

The driving safety performance of automated driving system (ADS)-equipped vehicles (AVs) must be quantified using metrics in order to be able to assess the driving safety performance and compare it to that of human-driven vehicles. In this research, driving safety performance metrics and methods for the measurement and analysis of said metrics are defined and/or developed.

A comprehensive literature review of metrics that have been proposed for measuring the driving safety performance of both human-driven vehicles and AVs was conducted. A list of proposed metrics, including novel contributions to the literature, that collectively, quantitatively describe the driving safety performance of an AV was then compiled, including proximal surrogate indicators, driving behaviors, and rules-of-the-road violations. These metrics, which include metrics from on- and off-board data sources, allow the driving safety performance of an AV to be measured in a variety of situations, including crashes, potential conflicts, and near misses. These measurements enable the evaluation of temporal flows and the quantification of key aspects of driving safety performance. The identification and exploration of metrics focusing explicitly on AVs as well as proposing a comprehensive set of metrics is a unique contribution to the literature. The objective is to develop a concise set of metrics that allow driving safety performance assessments to be effectively made and that align with the needs of both the ADS development and transportation engineering communities and accommodate differences in cultural/regional norms.

Concurrent project work includes equipping an intersection with a sensor suite of cameras, LIDAR, and RADAR to collect data requiring off-board sources and employing test AVs to collect data requiring on-board sources. Additional concurrent work includes development of artificial intelligence and computer vision-based algorithms to automatically calculate the metrics using the collected data. Future work includes using the collected data and algorithms to finalize the list of metrics and then develop a methodology that uses the metrics to provide an overall driving safety performance assessment score for an AV.

Introduction

One of the biggest questions facing the safe deployment and commercialization of ADS (ADS)-equipped vehicles (AVs) today is "What level of driving safety performance is required compared to that of a human-driven vehicle?"

In order to answer this question, a methodology to quantitatively measure and contrast the driving safety performance of AVs and human-driven vehicles is required. While efforts have been made to develop driving safety performance metrics for human-driven vehicles (e.g., [1, 2]), to the authors' knowledge, an overall methodology for a driving safety performance assessment does not exist. Moreover, since several of the metrics were developed with human-driven vehicles in mind, some of these metrics are less applicable to AVs. A need exists to develop metrics as a first step towards development of the driving safety performance assessment methodology.

© 2020 Institute of Automated Mobility. Published by SAE International. This Open Access article is published under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits distribution, and reproduction in any medium, provided that the original author(s) and the source are credited.

The Institute of Automated Mobility (IAM), created by Executive Order in Arizona to advance the AV industry through research into policy and technology, is undertaking a multi-phase project to develop the aforementioned driving safety performance methodology as part of an overall safety assurance framework. The first phase of the project, the subject of this paper, involves development of a proposed list of metrics that will be used in this methodology. The metrics on the list, some of which are specific to AVs while others were developed by the traffic engineering community for human-driven vehicles, are preliminary because the subsequent investigation may enable a determination that metrics should be eliminated or even added. This phase also includes concurrent development of a proof-of-concept method for performing measurements of the metrics that will allow for the driving safety performance assessment of AVs.

The concurrent proof-of-concept development work includes developing algorithms using artificial intelligence and computer vision techniques. Data collected at a signalized intersection in Anthem, Arizona and from test AVs are input to the developed algorithms and measurements of the proposed metrics are made. This work enables the assessment of the proposed list of metrics and will be expanded in future phases of the project to finalize the metrics list and develop the driving safety performance methodology that uses the metrics to assess AV driving behavior. This study is a pilot program with plans for expansion to diverse intersections and roadways throughout Arizona to capture useful data in categorizing traffic conflicts and help validate the safety of AVs. This work is a crucial step in boosting public perception of AVs and ensuring their safe implementation throughout development and commercialization.

Literature Review

The first step of the current work was a comprehensive literature review of metrics that have been proposed (and used) to quantify driving performance. There were 50 references identified that included one or more metrics and were evaluated during this work. A list of these references is included in Appendix A and a description of the metrics evaluated (some of which were not included in the preliminary list) is provided in Appendix B.

Traditionally, traffic safety has been measured using historical police-reported crash data in terms of crash frequency and/or severity. However, while use of these crash data is a critical tool in conducting traffic safety analyses, there are several limitations and potential drawbacks to this reliance for safety assessments as noted in the Highway Safety Manual (HSM) [1]. These include data quality and accuracy issues (e.g., coding errors, imprecise location information, subjectivity, etc.), crash reporting thresholds (e.g., minor crashes not meeting an agency's reporting threshold will not be recorded), and differences in crash reporting criteria between jurisdictions. Additionally, since crashes tend to be relatively rare and random, crash data often need to be aggregated over several years to obtain meaningful conclusions. To address these drawbacks, numerous studies focused on human driving behaviors have been conducted to identify "surrogate safety measures". These studies consist of observations primarily obtained from instrumented intersections allowing for the analysis of human driving behavior without relying on crash data. These measures also provide a promising avenue to assess ADS safety. Some of these surrogate safety measures are briefly discussed in the HSM [1], though, most transportation agencies still rely primarily on historical crash data for traffic safety analyses.

In 2003, Gettman and Head [2] conducted a study with traffic simulation models, employing surrogate safety measures including Post-Encroachment Time (PET), Time-to-Collision (TTC), and deceleration rates of reacting vehicles. The focus of this work centered around the design of intersections and using these surrogate safety measures in simulation models to determine the safety benefits of modifying intersections to reduce traffic conflicts.

Essa and Sayed have conducted similar studies in recent years, using both signalized and unsignalized intersections to evaluate the characteristics of vehicles traveling through the intersections [3]. Although previous work was conducted for purposes of human driven vehicles, applying surrogate safety measures to the analysis of these intersections provides a wealth of data that can help create methods for testing AVs by using baseline vehicle behaviors.

Mobileye researchers have defined a comprehensive safety model known as the Responsibility-Sensitive Safety (RSS) model [4]. These researchers have taken a "black box" approach to developing the mathematical formulae needed by AVs to formalize what it means to be a safe driver based on minimum longitudinal and lateral distances, and always performing proper responses when in dangerous situations. When followed, an AV should not be the cause of accidents. This "black box" approach, and the formal definitions of safe distances and required proper responses, could be useful in helping to qualify the various safety metrics chosen in this study to provide an overall safety assessment which could be compared between not only AVs of different manufacturers, but also human driven vehicles.

In 2017, Mahmud, et al. compiled a work of 38 major proximal indicators that have been used in previous traffic conflict simulation work [5]. This research is one of the most comprehensive sources found discussing mathematical summaries, formulations, and limitations of a large variety of proximal indicators used in previous studies. This work was established to contribute different approaches of analyzing traffic safety without relying on traditional methods of analyzing crash data. The scope of this paper aims to condense these and other metrics into a useful data set that can be applied to the observation of an intersection in Anthem, Arizona and expanded to additional intersections, to help quantify the safety of AVs.

The complete list of safety metrics identified in the literature review is included in Appendix B. These metrics were evaluated, compared, and condensed into a list of 14 to be considered in this study. It should be noted that some of these metrics (e.g., PET, TTC, MTTC, etc.) can be used to assess the safety of both AV and non-AV behavior, allowing for potential comparisons between the safety performance of human drivers and ADS. Another contribution by this work to the state of the art is to make the first attempt at a list of metrics that can be used in a driving safety performance assessment methodology. No single metric will be sufficient, and the future research on the methodology will refine the current list, with possible additions of new metrics and/or elimination of existing metrics. Additionally, it is possible that the metrics will be region dependent as a result of cultural or legislative variances characterizing the driving nature in a given country

or even state. The objective will be to find the best combination of metrics that are required for the driving safety performance assessment, where all things being equal, fewer metrics (and thus fewer data that must be captured) is a better option.

Driving Safety Performance Metrics

Upon completion of the literature review, the key driving safety performance metrics determined to be most significant in facilitating the evaluation of AV driving safety performance as well as established metrics currently used in the traffic safety community were identified and analyzed. The metrics developed by the traffic engineering community for human-driven vehicles are discussed separately; these metrics will be useful for comparison of AV driving safety performance with previous studies on human-driven vehicle driving safety performance. The following characteristics are listed for each metric to provide detailed information as well as justification for inclusion in this initial list:

1. Definition: Description of metric as proposed for use in this study.
2. Data Source Taxonomy: The metrics can be classified according to a taxonomy based on the data source(s): (1) The data for the metric are from *either* off- or on-board sources (the latter of which do not necessarily include ADS data), (2) the data for the metric *are from both* off-board sources and the on-board ADS, and (3) the data for the metric requires on-board ADS data. It should be noted that the off-board and on-board data sources are independent.
3. Subjectivity Taxonomy: The metrics can be classified according to whether there is subjectivity in: (1) the metric assumptions, (2) the metric threshold or (3) no subjectivity in either (i.e., objective).
4. Observable variables: Required variables (including units) needed to calculate metric.
5. Formulation: Mathematical model utilized to measure metric.
6. Subjective Assumptions/Thresholds: Subjective assumptions and thresholds are proposed (according to the Subjectivity Taxonomy above) for the purposes of this project; however, these assumptions and thresholds will be evaluated and potentially revised as a result of data collection.
7. Origin: The source of each metric, either from the literature review, adapted from the literature review, or novel.
8. Justification: Reason for inclusion of each metric in this study.

Minimum Safe Distance Violation (MSDV)

Definition: Minimum Safe Distance Violation (MSDV) is defined as an instance in which the actions of the ego vehicle result in encroaching upon its safe boundaries with another (safety-relevant) entity within the scenario environment, as defined by current velocities and acceleration capabilities of both entities.

The safety boundaries (aka safety envelope) are defined by clear lateral and longitudinal distances defined by the RSS model [4] that the vehicle should maintain towards surrounding road users in order to prevent the ego-vehicle to be the cause of a road accident.

Data Source Taxonomy: 1; Data required to calculate MSDV may be obtained from either off-board sources consisting of a to-be-determined (TBD) sensor suite of cameras, RADAR, and/or LIDAR or on-board sources such as ADS localization and perception data for AVs as well as GNSS, IMU, or dashboard cameras for all vehicles.

Subjectivity Taxonomy: 1; Several of the parameters involved in the RSS minimum safe distance calculation are defined following reasonable, but subjective, assumptions on the behavior of other entities (e.g., the maximum longitudinal braking, $a^{long}_{\max,accel}$).

Additionally, these parameters may vary with environmental conditions, among other factors. The discussion in [4] suggests that these parameters need to be agreed upon by regulation.

Observable variables: To calculate the MSDV, the longitudinal distance between two vehicles, d^{long} [m], the lateral distance between two vehicles, d^{lat} [m], the longitudinal velocity, v^{long} [m·s^{-1}], and the lateral velocity, v^{lat} [m·s^{-1}], of the vehicles must be obtained.

Besides acquiring the current distances and velocities (lateral and longitudinal) of the two vehicles under consideration, the response time ρ (the time [s] it takes the AV or person to react to a situation, including reaction and actuation time), the minimum and maximum decelerations the entities could apply, $a^{long,lat}_{min,decel}$ and $a^{long,lat}_{max,decel}$ [m·s^{-2}] respectively, and the maximum acceleration an entity could apply during the response time period, $a^{long,lat}_{max,accel}$ [m·s^{-2}], must be assumed and defined in order to calculate the minimum safe distances.

Formulation: The formulation for the longitudinal distance that defines where the MSDV occurs for entities moving in the same direction (adopted from Section 3.1 of [4]) is:

$$d^{long,same}_{min} = \left[v_1^{long}\rho_1 + \frac{1}{2}a^{long}_{1,max,accel}\rho_1^2 + \frac{\left(v_1^{long} + \rho_1 a^{long}_{1,max,accel}\right)^2}{2a^{long}_{1,min,decel}} - \frac{\left(v_2^{long}\right)^2}{2a^{long}_{2,max,decel}} \right]_+ \quad (1)$$

where the ego vehicle (subscript 1) is following behind another entity (subscript 2) and both are moving in the same direction, and $[x]_+ := \max\{x, 0\}$.

The formulation for the longitudinal distance that defines where the MSDV occurs for entities moving in opposite directions towards each other (adopted from Section 3.3 of [4]) is:

$$d_{min}^{long,opp} = \frac{2v_1^{long} + \rho_1 a_{1,max,accel}^{long}}{2}\rho_1 + \frac{\left(v_1^{long} + \rho_1 a_{1,max,accel}^{long}\right)^2}{2a_{1,min,decel,correct}^{long}} + \frac{2\left|v_2^{long}\right| + \rho_2 a_{2max,accel}^{long}}{2}\rho_2 + \frac{\left(\left|v_2^{long}\right| + \rho_2 a_{2max,accel}^{long}\right)^2}{2a_{2min,decel}^{long}} \quad (2)$$

The formulation for the lateral distance that defines where MSDV occurs, for adjacent entities (adopted from Section 3.4 of [4]) is:

$$d_{min}^{lat} = \mu + \left[\frac{2v_1^{lat} + \rho_1 a_{1,max,accel}^{lat}}{2}\rho_1 + \frac{\left(v_1^{lat} + \rho_1 a_{1,max,accel}^{lat}\right)^2}{2a_{1,min,decel}^{lat}} - \left(\frac{2v_2^{lat} - \rho_2 a_{2,max,accel}^{lat}}{2}\rho_2 - \frac{\left(v_2^{lat} - \rho_2 a_{2,max,accel}^{lat}\right)^2}{2a_{2,min,decel}^{lat}} \right) \right]_+ \quad (3)$$

where the ego vehicle (subscript 1) is to the left of the other entity (subscript 2), d_{min}^{lat} is the distance between the right side of the ego vehicle and the left side of the other entity, and μ is the lateral fluctuation margin [m] , and $[x]_+ := \max\{x, 0\}$.

Based on the definitions above, the MSDV metric is defined as follows:

$$MSDV' = \begin{cases} 1 & if\ d^{lat} < d_{min}^{lat} \wedge\ d^{long} < d_{min}^{long^*} \\ 0 & else \end{cases} \quad (4)$$

$$MSDV = \begin{cases} 1 & if\ MSDV' = 1\ \wedge Originated\ by\ ego\ vehicle \\ 0 & else \end{cases} \quad (5)$$

where $d_{min}^{long^*}$ is either $d_{min}^{long,same}$ or $d_{min}^{long,opp}$, depending on whether the entities are driving in the same or opposite directions, and is the minimum longitudinal distance that will reduce the likelihood of an accident occurring, based on maximum longitudinal acceleration/deceleration for each entity and a given response time; d_{min}^{lat} is the minimum lateral distance that will reduce the likelihood of an accident occurring, based on maximum lateral acceleration and a given response time. An MSDV violation occurs only if the action causing the violation originates with the ego vehicle. Reference [4] has details on how the longitudinal and lateral minimum safe distances are derived in cases where the ego vehicle is approaching intersections or interacting with pedestrians.

Subjective Assumptions/Thresholds: The assumed reasonable acceleration capabilities, response times, fluctuation margin, and responses of other entities are assumptions that determine the minimum distances of the MSDV.

Origin: This is a metric adopted from the RSS model published in [4].

Justification: Quantifying risky distances is fundamental for the assessment of safe driving of any vehicle type. Violating these safety boundaries can lead to accidents that otherwise could have been avoided. The MSDV is an indication that safe driving distances are, or are not, being maintained by the ego vehicle.

Proper Response Action (PRA)

Definition: Proper Response Action (PRA) is defined as an instance of an action (longitudinal and/or lateral acceleration) taken by the ego vehicle to restore itself to its calculated safety boundaries after a safe distance violation has occurred. The PRA must occur at a pre-determined time and rate in order to be deemed a sufficient response.

Data Source Taxonomy: 1; Data required to calculate PRA may be obtained from either off-board sources consisting of a TBD sensor suite of cameras, RADAR, and/or LIDAR or on-board sources such as ADS localization and perception data for AVs as well as GNSS, IMU, or dashboard cameras for all vehicles.

Subjectivity Taxonomy: 1, 2; The Proper Response Action either occurs or it does not, resulting in an objective formulation. However, the PRA is based, in part, on the MSDV, and the latter has subjective assumptions, and the thresholds above which the restoration occurs at a sufficiently high rate (for both imminent collision and non-imminent collision situations) are a subjective determination.

Observable variables: Similarly to MSDV, to calculate the PRA, MSDV' must be calculated, which requires the longitudinal and lateral distances between the two entities involved (ego vehicle and a road user), d^{long} [m] and d^{lat} [m], respectively, the longitudinal velocity, v^{long} [m·s^{-1}], and the lateral velocity, v^{lat} [m·s^{-1}], of the two entities. Additionally, the lateral and longitudinal accelerations, a^{lat} and a^{long}, respectively, of the ego vehicle [m·s^{-2}] are required.

Formulation:

$$PRA = \begin{cases} 1 & if\ MSDV' = 1 \wedge \left(\begin{array}{l} a^{lat} / \in \left[a_{min,accel}^{lat}, a_{max,accel}^{lat} \right] \vee \\ a^{long} / \in \left[a_{min,accel}^{long}, a_{max,accel}^{long} \right] \end{array} \right) \\ 0 & else \end{cases} \quad (6)$$

where a^{lat} is the lateral acceleration and a^{long} is the longitudinal acceleration of the ego vehicle, $a_{min,accel}^{lat}$ and $a_{min,accel}^{long}$ are the assumed parameters for minimum longitudinal and lateral deceleration that indicate that the ego vehicle is attempting to restore its safety envelope at a sufficient rate. The maximum lateral and longitudinal accelerations, $a_{max,accel}^{lat}$ and $a_{max,accel}^{long}$, respectively are operational design domain (ODD)- and vehicle dynamics-dependent parameters. The PRA includes situations where a collision is imminent and an evasive maneuver is required; in this case acceleration(s) larger

than $a^{lat}_{min,accel}$ and/or $a^{long}_{min,accel}$ is/are likely required. It should be noted that the ego vehicle can accelerate (either longitudinally, laterally, or both) up to its maximum values during the response time ρ but must begin the PRA at the end of the response time.

The PRA metric could be either 1 (the ego vehicle is applying a proper response after the minimum safe distances have been violated) or 0 (the ego vehicle is not applying a proper response).

Subjective Assumptions/Thresholds: The subjective assumptions of the PRA are the same as for the MSDV: The assumed reasonable acceleration capabilities, response times, fluctuation margin, and responses of other entities are assumptions that determine the minimum distances of the MSDV. The subjective thresholds of the PRA are the minimum lateral and longitudinal accelerations, $a^{lat}_{min,accel}$ and $a^{long}_{min,accel}$, respectively.

Origin: This is a metric adopted from the "Proper Response" concept presented in the RSS model published in [4]. For more details about the execution of the proper response action in various situations, refer to [4].

Justification: Knowing only when a safety boundary has been violated is insufficient to evaluate the safe driving behavior of the ego vehicle. Knowledge of whether the ADS is acting properly after such a violation is also required. According to the minimum safe distance calculations, the vehicle should apply proper inputs that will restore the safe state of the ego vehicle.

Minimum Safe Distance Factor (MSDF)

Definition: Minimum Safe Distance Factor (MSDF) is defined as a multiple of the minimum lateral and longitudinal safe distances maintained by an ego vehicle from surrounding entities. It is the ratio of measured distances (lateral and longitudinal) to the calculated safe distances (lateral and longitudinal).

Data Source Taxonomy: 1; Data required to calculate MSDF may be obtained from either off-board sources consisting of a TBD sensor suite of cameras, RADAR, and/or LIDAR or on-board sources such as ADS localization and perception data for AVs as well as GNSS, IMU, or dashboard cameras for all vehicles.

Subjectivity Taxonomy: 1; The MSDF is based on the MSDV, and the latter has subjective assumptions.

Observable Variables: The observable variables required to calculate MSDF are identical to those needed for MSDV as this is simply a factor of the safety boundaries being maintained at any instant.

Formulation:

$$MSDF^{lat} = \frac{d^{lat}}{d^{lat}_{min}} \tag{7}$$

$$MSDF^{long} = \frac{d^{long}}{d^{long}_{min}} \tag{8}$$

MSDF, being a ratio of maintained distances to the calculated safe distances, has a minimum value of zero, that implies a collision. A value of 1 means that the maintained distances are exactly equal to the calculated minimum safe distances of RSS, whereas values more than one imply the additional distance maintained by the ego vehicle. Thus, the data range for MSDF is 0 to infinity, which is a different format than the other metrics. Conforming this data range will require real-world data, which will help elucidate which values of MSDF are safer than others. This research will result in a curve for safety as a function of distance of the ego vehicle from the safety-relevant entities in its surroundings.

Origin: The MSDF is a novel metric derived from the MSDV.

Subjective Assumptions: The subjective assumptions of MSDF are the same as for the MSDV: The assumed reasonable acceleration capabilities, response times, fluctuation margin, and responses of other entities are assumptions that determine the minimum distances of the MSDV.

Justification: Values of MSDF higher than 1 hint towards the degree of defensive nature of the driving style. It should be noted that higher values of this metric do not necessarily result in higher safety and further research will be used to elucidate whether this metric is useful in deriving an overall driving safety performance assessment.

Minimum Safe Distance Calculation Error (MSDCE)

Definition: Minimum Safe Distance Calculation Error (MSDCE) is defined by the accuracy of the ADS calculation of its safety boundaries with respect to other safety-relevant entities in comparison to ground truth. The safety-relevant entities are those that can potentially result in a MSDV′ = 1.

Data Source Taxonomy: 2; The calculation of MSDCE requires separate safety envelope calculations by off-board sources (for ground truth determination), and the ADS. The off-board data may be obtained from sources consisting of a TBD sensor suite of cameras, RADAR, and/or LIDAR. The ADS data will be from the localization and perception systems.

Subjectivity Taxonomy: 1; MSDCE is based on the MSDV, and the latter has subjective assumptions.

Observable variables: The same variables as MSDV for the calculation of d^{long}_{min}, $d^{long,opp}_{min}[m]$ and $d^{lat}_{min}[m]$ and the same variables as MSDV for the calculation of the ground truth values $d^{long}_{gt,min}$, $d^{long,opp}_{gt,min}[m]$ and $d^{lat}_{gt,min}[m]$, are required to calculate the MSDCE.

Formulation: The formulation for MSDCE is the error between the ground truth and ADS calculations:

$$MSDCE_{long} = \frac{\left| d^{long*}_{gt,min} - d^{long}_{min} \right|}{d^{long}_{gt,min}} \tag{9}$$

$$MSDCE_{lat} = \frac{\left| d^{lat}_{gt,min} - d^{lat}_{min} \right|}{d^{lat}_{gt,min}} \tag{10}$$

$$MSDCE = \sqrt{MSDCE_{long} + MSDCE_{lat}} \tag{11}$$

Subjective Assumptions: The subjective assumptions of MSDCE are the same as for MSDV: The assumed reasonable acceleration capabilities, response times, fluctuation margin, and responses of other entities are assumptions that determine the minimum distances of the MSDV.

Origin: MSDCE is a novel metric derived from the MSDV.

Justification: The MSDCE will allow for a quantification of the error introduced by an incorrect calculation of the perceived safety envelope by the ADS. This will provide evidence for the ability (or lack thereof) of the ADS to determine the minimum safe distances according to RSS. It should be noted that just because a MSDCE has occurred does not necessarily mean that a safety incident will occur (e.g., crash); rather, it speaks to the potential inability for the AV to accurately detect a MSDV, and therefore potentially cause a PRA violation as well. Depending on the behavior of the other entities in the environment, this error may or may not result in a safety incident. Further study in to how to quantify a failure that results in a safety incident in the real world is required. Further, it should be noted that this metric will account for false positives, false negatives, improper classifications, and imprecise tracking, thus providing an overall assessment of the ADS perception and classification capabilities.

Collision Incident (CI)

Definition: Collision Incident (CI) is defined as an instance of the ego vehicle being at fault in a collision, as determined by an examination of the data from the on-board and/or off-board sensors, potentially including the vehicle event data recorder (EDR), along with the police report (the severity will be determined using the KABCO scale [6]).

Data Source Taxonomy: 1; If a collision is reported to police and/or insurance, a report would exist documenting the collision. The extent of details within these reports vary depending on personnel, accident severity, and jurisdiction. If a collision is not reported (which is often-times the case for very minor collisions) the event may be obtained from either off-board sources consisting of a TBD sensor suite of cameras, RADAR, and/or LIDAR or on-board sources such as ADS localization and perception data for AVs as well as GNSS, IMU, dashboard cameras, or EDR data for all vehicles.

Subjectivity Taxonomy: 2; The thresholds separating severity levels on the KABCO index for this metric are slightly subjective because there may be minor variations depending on the individual and municipality evaluating the collision.

Observable Variables: CI is a binary variable. If a collision is reported to police or insurance, all pertinent information related to the incident, vehicles, persons, etc. involved in the incident would be available. The lateral and longitudinal distances to the other entity can also be measured.

Formulation: The formulation for CI is a measure of frequency based on binary variables (1 if a collision occurs, and 0 if not). It can also be related to the longitudinal and lateral distances to another entity by:

$$CI = \begin{cases} 1 & \text{if } d^{lat} = 0 \wedge d^{long} = 0 \wedge \textit{Ego vehicle at fault} \\ 0 & \text{else} \end{cases} \quad (12)$$

Subjective Assumptions/Thresholds: Collision reporting requirements vary in different jurisdictions with different thresholds based on monetary damage estimates (any collision resulting in injury or fatality should be reported by law). If a collision is reported to police, the severity of the collision can be measured by the most severely injured person involved in the collision using the KABCO scale defined as follows [6]:

- K (Fatal Injury) - Any injury that results in a death within thirty days of the incident.
- A (Incapacitating Injury) - Any injury, other than a fatal injury, which prevents the injured person from walking, driving, or normally continuing the activities the person was capable of performing before the injury occurred. Includes: severe lacerations, broken or distorted limbs, skull or chest injuries, abdominal injuries, and unconsciousness when taken from the crash scene.
- B (Non-Incapacitating Injury) - Any injury which is evident to observers at the scene of the crash in which the injury occurred. Examples: contusions (bruises), lacerations, bloody nose, lump on head, or abrasions.
- C (Possible Injury) - Complaint of pain without visible injury. Includes momentary unconsciousness, claim of injuries not evident, limping, complaint of pain, nausea, or hysteria.
- O (No Injury) - No complaint or treatment was required by the involved persons.

The data range for this metric is binary (either 0 or 1), along with the KABCO index score. This is a different format than the other metrics. Conforming this data range will require real-world data, which will help elucidate how the CI value impacts the overall driving safety performance assessment.

Origin: CI is an existing metric described in prior literature.

Justification: It is critical to capture instances of collisions (whether they are reported or not) as a safety metric. A collision represents the worst possible outcome in terms of interactions between vehicles, and may result in property damage, injuries, and/or loss of life.

Rules-of-the-Road Violation (RRV)

Definition: Rules-of-the-Road Violation (RRV) is defined as an instance of the ego vehicle violating a traffic regulation that would result in an infraction or citation. Possible road violations to consider include but are not limited to: disregarding a red light, disregarding a stop or yield sign, illegal crossover of designated lane markings, exceeding speed limits, and improper or non-use of turn signals. It should be noted that there may be situations where it is necessary to violate a rule of the road to achieve a safe outcome. For example, if a vehicle is instructed to drive over the centerline by a police officer directing traffic, the vehicle should follow such instructions even though doing so would initiate a rule of the road violation.

Data Source Taxonomy: 1; Data required to identify instances of RRV may be obtained from either off-board sources consisting of a TBD sensor suite of cameras, RADAR, and/or LIDAR or on-board sources such as ADS localization and perception data for AVs as well as GNSS, IMU, or dashboard cameras for all vehicles. Roadway characteristics such as speed limit or traffic control device type and location could be obtained from agencies or other sources (e.g., field collected or obtained from existing maps/databases).

Subjectivity Taxonomy: 2; The threshold for this metric is subjective because rules of the road may vary depending on

region, some rules may be guidelines while others are strict laws, and sometimes it is required to break a rule of the road (e.g., police officer directing traffic around an obstacle due to traffic incident).

Observable Variables: Observable variables required to capture RRV depend on violation type. For instance, a speed limit violation would occur if vehicle speed exceeded the designated roadway speed limit; whereas, a traffic control device violation would result if a vehicle disregarded a traffic signal, sign, or pavement marking. Each violation type will have its own observable variables. Some violations are easier to measure than others (e.g., speed limit violations and disregarding traffic control devices). However, other violations would be difficult to capture with automated thresholds.

Formulation: The formulation for RRV is a measure of frequency based on binary variables (1 if a violation occurs, and 0 if not).

$$RRV = \begin{cases} 1 & if\ Rule\ Violated \\ 0 & else \end{cases} \tag{13}$$

Subjective Assumptions/Thresholds: The thresholds for what constitutes an RRV will be locale dependent.

Origin: RRV is mentioned in previous literature as important considerations in the development of AVs; however, the application of this metric as a safety metric is novel.

Justification: The rules of the road in use today are crucial to safe driving for both AVs and non-AVs. Violating these rules not only poses a danger to other road users but reduces trust and causes problematic interactions between road users.

Achieved Behavioral Competency (ABC)

Definition: Achieved Behavioral Competency (ABC) is a confirmation that the ego vehicle has correctly executed a specific behavioral competency. This metric is designed to be used in collaboration with the ADS developer in a testing scenario in which the behavioral competency to be demonstrated is agreed upon a priori. An example of the behavioral competency metric is the list from the Waymo Vehicle Safety Self-Assessment document submitted to NHTSA in 2018 [7], which includes 28 behavioral competencies from NHTSA (originating from work done by the California PATH) and 19 additional behavioral competencies added by Waymo. It is yet to be decided which behavioral competencies will be required to validate safe operation.

Data Source Taxonomy: 3; The ABC data will be sourced from the ADS developer for the specific behavioral competency that is to be demonstrated, for example while during a validation test of the AV driving safety performance. Data required to determine whether the behavioral competency is achieved may be obtained from either off-board sources consisting of a TBD sensor suite of cameras, RADAR, and/or LIDAR or on-board sources such as ADS localization and perception data for AVs as well as GNSS, IMU, or dashboard cameras for all vehicles.

Subjectivity Taxonomy: 3; This is an objective metric because the determination of whether or not the ego vehicle executed the behavioral competency based on the predefined acceptance criteria is binary.

Observable Variables: Measurement of this safety metric requires behavioral competency itself to be observed and confirmed by the ADS planning system.

Formulation: The formulation for ABC is an assignment of binary variables (1 if the AV demonstrated the set behavioral competency, and 0 if it did not).

ABC is a binary measurement (0 or 1): either the specified behavioral competency is executed properly, or it is not.

Subjective Assumptions/Thresholds: Not applicable

Origin: This metric is novel in its use as a safety metric.

Justification: ABC has been included in the preliminary list of metrics because an assessment of driving safety performance should include achieving a behavioral competency with intent, and not by accident or luck.

ADS Active (ADSA)

Definition: ADS Active (ADSA) is a confirmation that the ADS is active while executing the behavioral competency. This metric is designed to be used in collaboration with the ADS developer in a testing scenario in which the ADS developer must prove that the ADS is active at all times during behavioral competency execution. This proof can be via vehicle actuation data or some other agreed-upon format.

Data Source Taxonomy: 3; The data for the ADSA metric will be sourced from the ADS developer.

Subjectivity Taxonomy: 3; This is an objective metric because the ADS is either active or inactive while executing the behavioral competency.

Observable Variables: The ADS actively executing the entire dynamic driving task (DDT) is observed [8].

Formulation: The formulation for ADSA is an assignment of binary variables (1 if the ADS is active, and 0 if it is disengaged at any point during the execution of the behavioral competency).

Subjective Assumptions/Thresholds: Not applicable

Origin: This metric is adapted from the disengagement report required of companies testing AVs in California [9].

Justification: This metric has been included in the preliminary list because in order for an assessment of driving safety performance to be made, knowledge of whether the ADS is active is needed. It should be noted that there is some controversy of this metric as implemented in California, where the 'miles per disengagement' is reported [10]. If the testing occurs only (or mainly) in simple situations, the metric could lead to a misleading assessment of the overall AV driving safety performance. The ADSA metric implemented in this work is different for several reasons. First, this metric is not used in isolation and will be complemented by the other metrics on the list. Secondly, the envisioned driving safety performance assessment methodology will include a quantification of each scenario in which the behavioral competency is being executed as part of the assessment so that testing in more challenging situations will be appropriately addressed.

Human Traffic Control Detection Error Rate (HTCDER)

Definition: Human Traffic Control Detection Error Rate (HTCDER) is a confirmation that the ego vehicle can detect the direction(s) of a human traffic control (HTC) actor (which may include a direction to violate one or more rules of the road).

Data Source Taxonomy: 2, Required data for HTCDER include information from the on-board ADS perception system as well as off-board sources (the latter for ground truth data) in order to establish whether the HTC actor and direction(s) were actually correctly detected and whether the ego vehicle successfully completed the directed maneuver. The off-board and on-board ADS data are compared in order to ensure that errors, such as false positives and false negatives, are accounted for in the on-board ADS data.

Subjectivity Taxonomy: 3; HTCDER is an objective metric because it is defined as the detection error to the HTC actor for any given scenario.

Observable variables: The required variables to measure the HTCDER are the actual number of human traffic control instructions (from off-board sources) and the number of instances where the human traffic control instructions are detected by the ego vehicle (on-board sources).

The formulation of HTCDER is as follows:

$$HTCDER = \frac{GTI - CDI}{GTI} \tag{14}$$

where:
GTI = Number *of Ground Truth Instructions*
CDI = Number *of Correctly Detected Instructions*

The data range is from 0 to 1, where a lower value indicates a higher number of HTC instructions were correctly detected.

Subjective Assumptions/Thresholds: Not applicable

Origin: A metric examining AV behavior when interacting with traffic control personnel was not found elsewhere in the literature, so the HTCDER metric is novel.

Justification: In special cases, there are times when law enforcement officers have to manually direct traffic to follow certain directions. This may also include instructions to violate one or more of the established rules of the road. The misdetection of such direction and non-compliance with detected directions can result in severe safety concerns. The HTCDER safety metric is a measure of how well the ADS is detecting instructions by HTCs, which is crucial to minimize safety issues in these situations.

Human Traffic Control Violation Rate (HTCVR)

Definition: Human Traffic Control Violation Rate (HTVR) is a confirmation that the ego vehicle, once it has detected direction(s) of an HTC actor (which may include a direction to violate one or more rules of the road), follows the instructions.

Data Source Taxonomy: 2, Required data for HTCVR include information from the on-board ADS as well as off-board sources (the latter for ground truth data) in order to establish whether the ego vehicle complied with the correctly detected traffic control direction(s). The off-board and on-board ADS data are compared in order to ensure that errors, such as false positives and false negatives, are accounted for in the on-board ADS data.

Subjectivity Taxonomy: 3; The HTCVR is an objective metric because it is defined as the ego vehicle compliance to the directions of HTC actor for any given scenario.

Observable variables: The required variables to measure the human traffic control adherence are the actual number of human traffic control instructions (from off-board sources), number of instances where the human traffic control instructions are correctly detected by the ego vehicle (on-board sources), and the number of instances where the human traffic control instructions are complied with (off-board and on-board sources).

The formulation of HTCVR is as follows:

$$HTCVR = \frac{CDI - CCI}{CDI} \tag{15}$$

where:
CDI = Number of Correctly Detected Instructions
CCI = Number of Correctly Complied Instructions

The data range is from 0 to 1, where a lower value indicates that a higher number of correctly detected instructions were complied with, for a safer situation when interacting with an HTC.

Subjective Assumptions/Thresholds: Not applicable

Origin: A metric examining AV behavior when interacting with traffic control personnel was not found elsewhere in the literature, so the HTCVR metric is novel.

Justification: In special cases, there are times when law enforcement officers have to manually direct traffic to follow certain directions. This may also include instructions to violate one or more of the established rules of the road. The misdetection of such direction and non-compliance with detected directions can result in severe safety concerns. The HTCDER and HTCVR metrics allow ADS developers to assess, monitor, and minimize the safety issues caused by violating the instructions by traffic control personnel.

Traffic Engineering-Developed Metrics

As mentioned, the traffic engineering community has developed driving safety performance metrics for human-driven vehicles. These metrics have limitations that have been discussed in the literature, (e.g., in [11]) such as the failure to account for reaction time of the driver/ADS. However, since many studies have been undertaken that include data with these metrics, the following metrics have been included in the preliminary list so that a comparison of the test AVs against the human-driven vehicle study data can be made. Such a comparison could prove useful, especially for future work that aims to answer the common question "What level of driving safety performance is required compared to that of a human-driven vehicle?"

Time-to-Collision (TTC)

Definition: TTC is defined as the time until a collision between two entities in the scenario environment that would occur if both continue with the present velocities. A violation occurs if the measured TTC is below a pre-determined threshold.

Data Source Taxonomy: 1; Data required to calculate TTC may be obtained from either off-board sources consisting of a TBD sensor suite of cameras, RADAR, and/or LIDAR or on-board sources such as ADS localization and perception data for AVs as well as GNSS, IMU, or dashboard cameras for all vehicles.

Subjectivity Taxonomy: 2; Although the formulation of TTC is objective, the threshold used to characterize a Time-to-Collision Violation (TTCV) (e.g., what constitutes an unacceptable conflict) is subjective.

Observable Variables: Measured variables needed to calculate TTC include relative position $X_L - X_F$ [m] and speed $v_F - v_L$ [m·s^{-1}] between the involved vehicles.

Formulation: The formulation is adapted from [5]:

$$TTC = \frac{X_L - X_F}{v_F - v_L} \tag{16}$$

where:
X_L: Leading vehicle position
X_F: Following vehicle position
v_L: Leading vehicle speed
v_f: Following vehicle speed

Note that the positions in Equation (11) are the relative positions of the two vehicles and is the shortest distance where the collision would occur.

A TTCV occurs if:

$$TTCV = \begin{cases} 1 & if\ TTC \le Threshold \\ 0 & else \end{cases} \tag{17}$$

Subjective Assumptions/Thresholds: Smaller TTC values indicates increasingly 'severe' conflicts, with TTC = 0 seconds indicating a collision occurred. There is no standard or agreed-upon threshold below which indicates a TTCV has occurred. Previous literature has considered threshold values ranging from 0.9 to 4.0 s [5]. Based on the results found in the literature review, the proposed threshold for this work is 3.0 s. Once sufficient data have been collected, this threshold will be finalized.

Origin: TTC is an existing metric discussed in prior studies, originally developed by Hayward in 1971[[12]]. More recent studies have been conducted with TTC as a surrogate safety measure with means of improving reliability and efficiency [5, 13].

Justification: TTC is a well-known and commonly used surrogate measure of driving safety metric in the traffic safety area. It is relatively easy to measure and interpret.

Modified Time-to-Collision (MTTC)

Definition: MTTC is defined as the time until a collision between two entities in the scenario environment would occur if both continue with the present velocities and accelerations. A violation occurs if the measured MTTC is below a pre-determined threshold.

Data Source Taxonomy: 1; Data required to calculate MTTC may be obtained from either off-board sources consisting of a TBD sensor suite of cameras, RADAR, and/or LIDAR or on-board sources such as ADS localization and perception data for AVs as well as GNSS, IMU, or dashboard cameras for all vehicles.

Subjectivity Taxonomy: 2; Similarly to TTC, the definition of MTTC is objective, however the threshold characterizing a MTTC violation is subjective.

Observable Variables: Measured variables needed to calculate MTTC include relative position $X_L - X_F$ [m], relative speed $v_F - v_L$ [m·s^{-1}], and relative acceleration $a_F - a_L$ [m·s^{-2}] between the involved vehicles.

Formulation: The formulation is adapted from [14]:

$$MTTC' = \frac{-\Delta\vec{V} \pm \sqrt{\vec{V}^2 + 2\Delta\vec{A}\vec{D}}}{\Delta\vec{A}}0 \tag{18}$$

where:
$\varnothing\vec{V}$: Relative velocity
$\varnothing\vec{A}$: Relative acceleration
$\vec{D}$: Initial relative space gap

An MTTC Violation (MTTCV) occurs if:

$$MTTCV = \begin{cases} 1 & if\ MTTC' \le Threshold \\ 0 & else \end{cases} \tag{19}$$

Subjective Assumptions/Thresholds: Smaller MTTC values indicates increasingly 'severe' conflicts, with MTTC = 0 seconds indicating a collision occurred. Similar to TTC, there is no currently existing standard or agreed-upon threshold which indicates an MTTCV has occurred. Previous literature has considered threshold values ranging from 1.0 to 4.0 seconds for MTTC [13]. The proposed threshold for this work is 3.0 s. Once sufficient data have been collected, this threshold will be finalized.

Origin: MTTC is an existing metric discussed in prior studies [13, 5, 15].

Justification: Like TTC, MTTC is a known metric in the traffic safety area, although not as common as TTC, and is relatively simple to calculate. While the concept of MTTC is similar to TTC, it is proposed as a more robust metric as it does not rely on the assumption that vehicles will maintain constant speeds [5].

Post-Encroachment Time (PET)

Definition: PET is defined as the time from the end of encroachment of the lead vehicle to the beginning of encroachment of the second vehicle to the potential point of collision (of a conflict). A violation occurs if the measured PET is below a pre-determined threshold.

Data Source Taxonomy: 1; Data required to calculate PET may be obtained from either off-board sources consisting of a TBD sensor suite of cameras, RADAR, and/or LIDAR or on-board sources such as ADS localization and perception

data for AVs as well as GNSS, IMU, or dashboard cameras for all vehicles.

Subjectivity Taxonomy: 2; Although the formulation of PET is objective, the threshold used to characterize a PET Violation (PETV) (e.g., what constitutes an unacceptable conflict) is subjective.

Observable Variables: The measured variables required to calculate PET are relative time $t_2 - t_1$ [s] of two vehicles passing through a conflict point and the conflict point itself.

Formulation: The formulation is adapted from [5]:

$$PET = t_2 - t_1 \tag{20}$$

where:

t_2: Arrival time of (any part of) Vehicle 2 at Conflict Point

t_1: Arrival time of (any part of) Vehicle 1 at Conflict Point

A PET violation (PETV) occurs if:

$$PETV = \begin{cases} 1 & if\ PET \le Threshold \\ 0 & else \end{cases} \tag{21}$$

Subjective Assumptions/Thresholds: Smaller PET values indicate increasingly 'severe' conflicts, with PET = 0 seconds indicating a collision occurred. Similar to TTC and MTTC, there is no currently existing standard or agreed-upon threshold which indicates an MTCC violation has occurred. Previous literature has considered threshold values ranging from 3.0 to 5.0 seconds for PET [16, 17, 18]. The proposed threshold for this work is 3.0 s. Once sufficient data have been collected, this threshold will be finalized.

Origin: PET is an existing metric discussed in prior literature [5, 18].

Justification: Similarly to TTC and MTTC, PET is a known metric in the traffic safety area, although not as common as TTC (but more common than MTTC), and is relatively simple to calculate and interpret.

Aggressive Driving (AD)

Definition: Aggressive Driving (AD) is defined as repeated instances of maneuvers (longitudinal and lateral accelerations) above specified thresholds completed by the ego vehicle that, cumulatively, are potentially less safe. The thresholds will vary by jurisdiction and/or culture.

Data Source Taxonomy: 1; Data required to calculate AD may be obtained from either off-board sources consisting of a TBD sensor suite of cameras, RADAR, and/or LIDAR or on-board sources such as ADS localization and perception data for AVs as well as GNSS, IMU, or dashboard cameras for all vehicles.

Subjectivity Taxonomy: 1, 2; AD either occurs or it does not, resulting in an objective formulation. However, AD is based, in part, on the MSDV, and the latter has subjective assumptions. The thresholds for this metric are highly subjective because they are dependent on the culture and the surrounding environment.

Observable Variables: The AD metric depends on the acceleration [g]. The measurement of the final metric is binary which will be calculated by determining if the lateral and/or longitudinal action meet the determined threshold.

Formulation: The formulation for AD is:

$$AD = \begin{cases} 1 & a^{lat} \ge \left|0.47\,g\right| \vee a^{long,accel} \ge \left|0.43g\right| \\ & \qquad \vee a^{long,decel} \ge \left|0.61g\right| \\ 0 & else \end{cases} \tag{16}$$

Subjective Assumptions/Thresholds: Proposed AD thresholds for lateral and longitudinal accelerations are based on the default values used by vehicle telematics provider Geotab to indicate aggressive driving maneuvers [19].

The overall AD Threshold is met if there are three (3) or more instances of accelerations (either all longitudinal, all lateral, or a combination of both) exceeding the acceleration thresholds for a duration of one second (or more) in any minute of driving.

Origin: This safety metric is not novel; however, what constitutes aggressive driving is very subjective and more research will be conducted to obtain a justified formulation and thresholds.

Justification: The AD metric indicates an inherent, potentially riskier natural driving behavior which should be considered when evaluating the driving safety performance of the ADS.

Concurrent Work

Data Collection

The data collection system is divided into two sub-modules including infrastructure (off-board) and vehicle (on-board). The sub-modules are responsible for collecting and managing the data coming from on-board and off-board data sources. The intersection sub-module collects data from the following possible sources available at each intersection:

1. Signal phase and timing (SPaT) data from the signal controller
2. Video data from cameras
3. Surrogate basic safety messages (BSMs) determined from video data for vehicles approaching or in the intersection
4. LIDAR sensor frames; LIDAR sensors have not yet been installed but are planned
5. RADAR sensor frames; RADAR sensors have not yet been installed but are planned
6. BSMs collected from surrounding connected vehicles

The sensor suite for the off-board sources is still in development. The objective is to develop the suite that is able to collect the required data but has the lowest cost both in terms of equipment and data size.

On September 5, 2019, an instrumentation suite was installed at the intersection of Daisy Mountain Drive and Gavilan Peak Parkway in Anthem, Arizona. This intersection was selected as a candidate for data collection due to accessibility, traffic patterns, and complexity to provide data of interest to this study. The instrumentation suite includes four (4) traffic cameras installed by Maricopa County Department of Transportation, each with a field of view encompassing a driving direction

approaching the intersection, for a monocular view of each direction. One of the traffic cameras at the Anthem intersection is shown in Figure 1.

The plan for this work includes involvement of several test AVs to provide the on-board source data. Efforts are underway to recruit a commercial ADS developer. Concurrently, two AVs from the university laboratories of two of the authors from Arizona State University (ASU) and University of Arizona (UArizona) will be used as test vehicles as well. The test AV from ASU is shown below in Figure 2.

It should be noted for the vehicle sub-module that raw sensor data except, perhaps, camera data, are often considered proprietary by ADS developers. Therefore, the required data for metrics under the Data Source taxonomy (2) (when considering ADS data) and (3) are not raw sensor data but rather are processed data that provide the inputs to the metrics as described in each of the metric subsections above. However, the test AVs will provide the raw sensor input for subsequent data processing.

The vehicle sub-module collects data from the following sources of the test AVs:

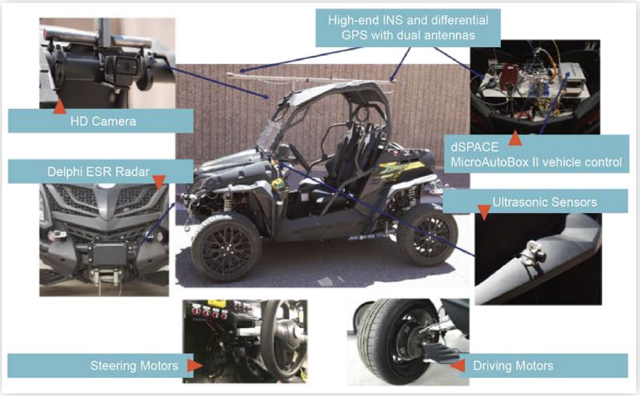
1. Data from the vehicle controller area network (CAN) bus
2. LIDAR sensor frames
3. RADAR sensor frames
4. Video camera frames
5. BSMs
6. ADS data
7. Vehicle driving commands

FIGURE 1 Instrumentation suite installed at intersection



© Institute of Automated Mobility.

FIGURE 2 ASU test AV



© Institute of Automated Mobility.

Synchronized Timestamping: Both sub-modules independently have access to a GPS receiver. The system clocks of data collection co-processors are driven by the time provided by corresponding GPS receivers. The system clock is then used to generate timestamps for each collected data point. This ensures that data collected from both submodules are time-synchronized.

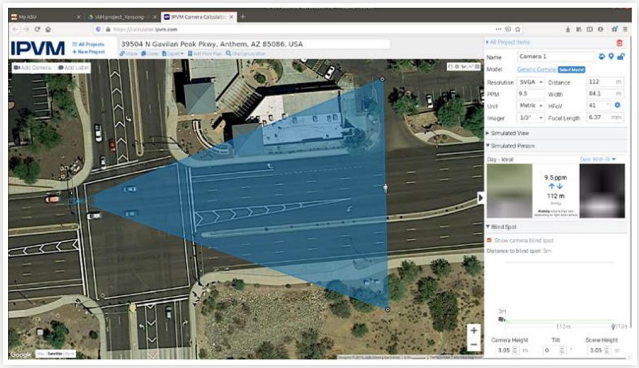
Data Storage: Data will be stored locally within each infrastructure location and in each vehicle. At pre-determined intervals, data will be streamed over an existing fiber optic network to a server located in the Maricopa County Department of Transportation where it can be pushed to a variety of consumers including research laboratories and into the regional active data warehouse (RADS). At pre-determined intervals, data are transferred over an existing fiber optic network to a dedicated server located in the Maricopa County Department of Transportation (MCDOT). From this server, the data are pushed to a collaborative data hosting platform, "CyVerse" [20], through which the data are shared with a variety of consumers, including collaborators of this project, Regional Active Data Warehouse (RADS) of the MCDOT, and other research laboratories.

Metric Calculation Methodology

The metric calculation methodology is currently being developed. For the off-board MCDOT camera data, the following computer vision analysis will be applied to process the video data and then integrate these with other data to calculate the predefined driving safety performance metrics.

The main process to complete the computer vision analysis for camera data is summarized as follows. First, the pixels per meter (PPM) metric is calibrated for the cameras, using an online tool called IPVM, as shown in Figure 3.

FIGURE 3 An online tool of IPVM for PPM calibrations



Reprinted with permission. © IPVM.

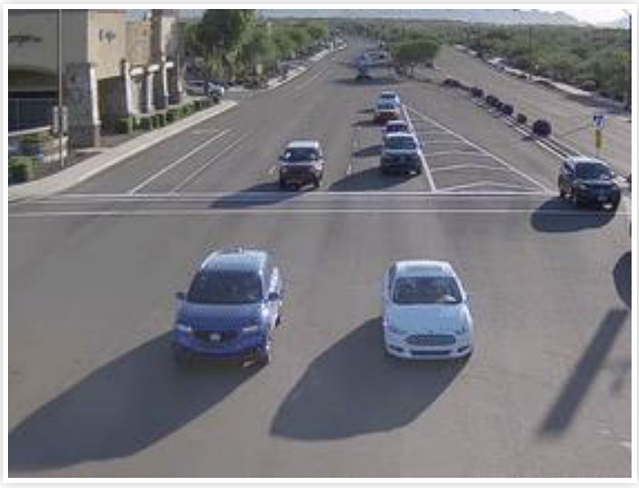
The camera location and the boundaries of camera images are identified in the map, and the blue triangle in Figure 3 is obtained. Then, the actual width of the road covered by the camera is obtained by an HD map system. Using this information, the PPM is calculated by the IPVM tool.

Second, the vehicle detection and tracking algorithm is developed by implementing a tracking and velocity estimation algorithm such as the Haar Cascade Vehicle Classifier (HCVC) [21]. The algorithm will be applied to detect and track the vehicle information from the camera data and to automate the process, which will be important when data are being collected from multiple intersections and vehicles.

Figure 4 shows a raw image data obtained from the continuous video data of a MCDOT camera. The vehicles close to the intersection are identified and labeled with bounding boxes by the algorithm mentioned previously, as shown in Figure 5. Subsequently, the number of pixels the vehicle traveled in one video frame to another is utilized to calculate the vehicle positions, speeds, and accelerations.

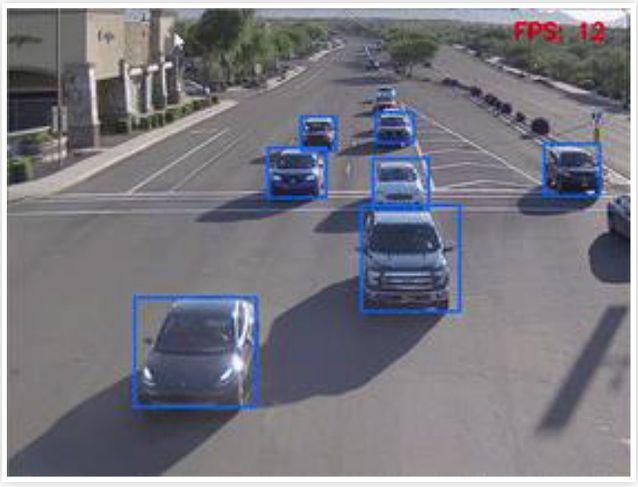
Figure 6 depicts the estimated speed of several vehicles in the frame. The vehicle position and motion estimates are then used to calculate the desired driving safety performance metrics. Several detection and tracking algorithms are currently being evaluated; the results shown in Figure 6 are used as an illustrative example. Similar methodologies will be developed for the other off-board sources as well as the on-board sources.

FIGURE 4 A raw image obtained from one MCDOT camera



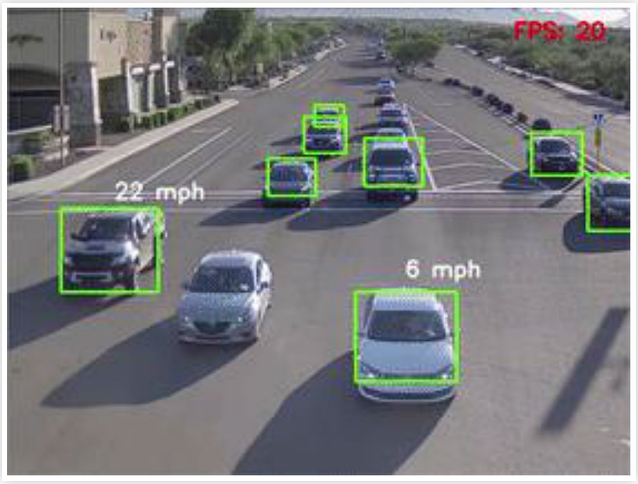
Reprinted with permission. © NoTraffic, Inc.

FIGURE 5 Vehicle detection based on the camera raw data



Reprinted with permission. © NoTraffic, Inc.

FIGURE 6 Speed calculation based on vehicle detection and tracking



Reprinted with permission. © NoTraffic, Inc.

Conclusions and Future Work

The work presented here represents, to the authors' knowledge, the first attempt at creating a comprehensive list of driving safety performance metrics for AVs. Each metric in the list was explored and details on how data will be collected was described. The justification for the inclusion of each metric in the list was also provided.

The data collection equipment and strategy being developed concurrently were also described. The data will be collected from both off-board sources from an intersection in Anthem, AZ as well as on-board sources from test AVs. Additional sensor equipment is to be installed at the

intersection, and additional test AVs are being sourced for the project. The initial methodology for metric calculations was also presented. The methodologies that will allow for automated calculation of each metric are currently being developed and will be presented in future publications.

As the off-board sensor suite becomes more capable and installed at additional intersections and more data from test AVs are obtained, the metrics will be evaluated using real-world data to determine if each should continue to be included in the list. Additional metrics will also be considered. The objective is to develop a concise set of metrics that allow driving safety performance assessments to be effectively made and that align with the needs of both the ADS development and transportation engineering communities and accommodate differences in cultural/regional norms. With a finalized list of driving safety performance metrics determined, the next phase of the project will be to develop the driving safety performance assessment methodology as a component of the ultimate goal of creating a safety assurance framework for AVs.

References

1. American Association of State Highway and Transportation Officials (AASHTO), *Highway Safety Manual* First Edition (Washington, DC, 2010).
2. Gettman, D. and Head, L., "Surrogate Safety Measures from Traffic Simulation Models," *Transportation Research Record* 1840:104-115, 2003.
3. Essa, M. and Sayed, T., "Simulated Traffic Conflicts: Do They Accurately Represent Field-Measured Conflicts?" *Transportation Research Record* 2514:48-57, 2015.
4. Shalev-Shwartz, S., Shammah, S., and Shashua, A., "On a Formal Model of Safe and Scalable Self-Driving Cars," arXiv.org:1708.06374, 2017.
5. Mahmud, S., Ferreira, L., Hoque, S., and Tavassoli, A., "Application of Proximal Surrogate Indicators for Safety Evaluation: A Review of Recent Developments and Research Needs," *International Association of Traffic and Safety Devices* 41:153-163, 2017.
6. Herbel, S., Laing, L., and McGovern, C., "Highway Safety Improvement Program (HSIP) Manual," FHWA-SA-09-029, Federal Highway Administration, 2010.
7. Waymo, "Waymo Safety Report: On the Road to Fully Self-Driving," 2018.
8. Society of Automotive Engineers, "Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles," J3016_201806, 2018.
9. State of California Department of Motor Vehicles, "Adopted Regulations for Testing of Autonomous Vehicles by Manufacturers," 2018.
10. Hawkins, A., "California's Self-Driving Car Reports Are Imperfect, but They're Better Than Nothing," *The Verge*, Feb. 13, 2019.
11. Nadimi, N., Behbahani, H., and Shahbazi, H., "Calibration and Validation of a New Time-Based Surrogate Safety Measure Using Fuzzy Inference System," *Journal of Traffic and Transportation Engineering* I(3):51-58, 2016.
12. Hayward, J., "Near Misses as a Measure of Safety at Urban Intersections," MS thesis, The Pennsylvania State University, 1971.
13. Ozbay, K., Yang, H., Bartin, B., and Mudigonda, S., "Derivation and Validation of New Simulation-Based Surrogate Safety Measure," *Transportation Research Record: Journal of the Transportation Research Board* 2083:105-113, 2008, doi:10.3141/2083-12.
14. Every, J., Barickman, F., Martin, J., Rao, S. et al., "A Novel Method to Evaluate the Safety of Highly Automated Vehicles," Transportation Research Center, 2019.
15. Essa, M. and Sayed, T., "Full Bayesian Conflict-Based Models for Real Time Safety Evaluation of Signalized Intersections," *Accident Analysis and Prevention* 129:367-381, 2019.
16. Drive Knowledge Tests, "Gap Selection Practice," www.driverknowledgetests.com/resources/gap-selection-practice, accessed Nov. 3, 2019.
17. Drive-Safely.net, "Keep a Safe Following Distance While Driving," www.drive-safely.net/safe-following-distance, accessed Nov. 1, 2019.
18. Morando, M., Tian, Q., Truong, L., and Vu, H., "Studying the Safety Impact of Autonomous Vehicles Using Simulation-Based Surrogate Safety Measures," *Journal of Advanced Transportation*, 2018.
19. Geotab, "Fleet Portal - Safety Parameters -- Default Values."
20. Merchant, N., Lyons, E., Goff, S., Vaughn, M., Ware, D., Micklos, D., and Antin, P., "The iPlant Collaborative: Cyberinfrastructure for Enabling Data to Discovery for the Life Sciences," *PLOS Biology*, 2016, doi:10.1371/journal.pbio.1002342.
21. Oliveira, M. and Santos, V., "Automatic Detection of Cars in Real Roads Using Haar-Like Features," in *8th Portuguese Conference on Automatic Control*, Vila Real, 2008.

Acknowledgements

The authors would like to thank the Institute of Automated Mobility for its generous funding of this work.

Definitions/Abbreviations

ABC - Achieved Behavioral Competency
AD - Aggressive Driving
ADS - automated driving system
ADSA - ADS Active
AI - artificial intelligence
ASU - Arizona State University
AV - ADS-equipped vehicle
BSM - basic safety message
DDT - dynamic driving task
EDR - event data recorder
ESC - Electronic Stability Control
GPS - Global Positioning System
HSM - Highway Safety Manual
HTCDER - Human Traffic Control Detection Error Rate
HTCVR - Human Traffic Control Violation Rate
IAM - Institute of Automated Mobility
IMU - inertial measurement unit

LIDAR - light detection and ranging
MSDCE - Minimum Safe Distance Calculation Error
MSDF - Minimum Safe Distance Factor
MSDV - Minimum Safe Distance Violation
MTTC - Modified Time-To-Collision
NHTSA - National Highway Traffic Safety Administration
PET - Post-Encroachment Time
PPM - pixels per meter
PRA - Proper Response Action
RADAR - radio detection and ranging
RADS - regional active data warehouse
RSS - Responsibility-Sensitive Safety
SPaT - signal phase and timing
TTC - Time-To-Collision

Appendix A: List of References from Literature Review

Publication ID	Year	Authors	Title	Publication
1	2003	Gettman, Douglas Head, Larry	Surrogate Safety Measures From Traffic Simulation Models	Transportation Research Record, 1840(1), pp. 104-115
2	2019	Every, Joshua L Barickman, Frank Martin, John Rao, Sughosh Schnelle, Scott Weng, Bowen	A Novel Method to Evaluate the Safety of Highly Automated Vehicles	25th International Technical Conference on the Enhanced Safety of Vehicles (ESV): Innovations in Vehicle Safety: Opportunities and Challenges
3	2017	Shalev-Shwartz, Shai Shammah, Shaked Shashua, Amnon	On a Formal Model of Safe and Scalable Self-driving Cars	arXiv:1708.06374v6 [cs.RO]
4	1993	Sayed, Tarek Brown, Gerald Drancis, Navin	Simulation of Traffic Conflicts at Unsignalized Intersections with TSC-Sim	Accident Analysis and Prevention,Vol.26, No.5, pp 593-607
5	2012	Autey, Jarvis Sayed, Tarek Zaki Mohammed H.	Safety evaluation of right-turn smart channels using automated traffic conflict analysis	Accident Analysis and Prevention 45 (2012) 120- 130
6	2011	Davis, Gary Hourdos, John Xiong, Hui Chatterjee, Indrajit	Outline for a causal model of traffic conflicts and crashes	Accident Analysis and Prevention 43 (2011) 1907- 1919
7	2010	El-Basyouny, Karim Sayed, Tarek	A method to account for outliers in the development of safety performance functions	Accident Analysis and Prevention 42 (2010) 1266-1272
8	2015	Essa, Mohammed Sayed, Tarek	Simulated Traffic Conflicts Do They Accurately Represent Field-Measured Conflicts?	Transportation Research Record: Journal of the Transportation Research Board, No. 2514, Transportation Research Board, Washington, D.C., 2015, pp. 48-57. DOI: 10.3141/2514-06
9	2018	Essa, Mohammed Sayed, Tarek	Traffic conflict models to evaluate the safety of signalized intersections at the cycle level	Transportation Research Part C 89 (2018) 289-302
10	2011	Ismail, Karim Sayed, Tarek Saunier, Nicholas	Methodologies for Aggregating Indicators of Traffic Conflict	Transportation Research Record: Journal of the Transportation Research Board, No. 2237, Transportation Research Board of the National Academies, Washington, D.C., 2011, pp. 10-19. DOI: 10.3141/2237-02
11	2018	Fraade-Blanar, Laura Blumenthal, Marjory Anderson, James Kalra, Nidhi	Measuring Automated Vehicle Safety: Forging a Framework	RAND Corporation
12	2018	Shastry, Akshay	Functional Safety Assessment in Autonomous Vehicles	Virginia Polytechnic Institute
13	2017	Kockelman, Kara Boyles, Stephen Stone, Peter Fagnant, Dan et. al.	An Assessment of Autonomous Vehicles: Traffic Impacts and Infrastructure Needs - Final Report	The University of Texas at Austin, Center for Transportation Research; Technical Report 0-6847-1. TXDOT Project Number 0-6847

© Institute of Automated Mobility.

(Continued)

Publication ID	Year	Authors	Title	Publication
14	2018	Guanetti, Jacopo Kim, Yeojun Borrelli, Francesco	Control of Connected and Automated Vehicles: State of the Art and Future Challenges	Annual Reviews of Control, Volume 45, pp. 18-40
15	2008	Ozbay, Kaan Yang, Hong Bartin, Bekir Mudigonda, Sandeep	Derivation and Validation of New Simulation-Based Surrogate Safety Measure	Transportation Research Record: Journal of the Transportation Research Board, No. 2083, Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 1105-113. DOI: 10.3141/2083-12
16	2013	Bagdadi, Omar Varhelyi, Andras	Development of a method for detecting jerks in safety critical events	Accident Analysis and Prevention 50 (2013) 83-91
17	2017	Every, Joshua L Barickman, Frank Martin, John Rao, Sughosh Schnelle, Scott Weng, Bowen	A Method for Evaluating Automated Vehicle Safety	SAE Government Industry Meeting; January 25-27, 2017. NHTSA
18	2017	Roesener, Christian Sauerbier, Jan Zlocki, Adrian Wang, Lei Fahrenkrog, Felix Varhelyi, Andras Dufils, Joris de Gelder, Erwin Breunig, Sandra Mejuto, Pablo Lanati, Jessica Tango, Fabio	A Comprehensive Evaluation Approach for Highly Automated Driving	European Commission Seventh Framework Programme (FP/2007-2013); 17-0259.
19	2019	APTIV Audi Apollo BMW Continental Daimler FCA Here Infineon Intel VW	Safety First for Automated Driving	https://newsroom.intel.com/wp-content/uploads/sites/11/2019/07/Intel-Safety-First-for-Automated-Driving.pdf
20	2012	Tarek Sayed, Karim Ismail, Mohamed H. Zaki, and Jarvis Autey	Feasibility of Computer Vision-Based Safety Evaluations Case Study of a Signalized Right-Turn Safety Treatment	Transportation Research Record: Journal of the Transportation Research Board, No. 2280, Transportation Research Board of the National Academies, Washington, D.C., 2012, pp. 18-27. DOI: 10.3141/2280-03
21	2017	Mahmud Sohail, Ferreira Luis, Hoque Shamshul Md., Tavassoli Ahmed	Application of proximal surrogate indicators for safety evaluation: A review of recent developments and research needs	2017 International Association of Traffic and Safety Sciences.
22	2001	Michiel M. Minderhoud, Piet H.L. Bovy	Extended time-to-collision measures for road traffic safety Assessment	Accident Analysis and Prevention 33 (2001) 89 - 97
23	2011	Giuseppe Guido, Frank Saccomanno, Alessandro Vitale, Vittorio Astarita, Demetrio Festa	Comparing Safety Performance Measures Obtained from Video Capture Data	Journal of Transportation Engineering, Vol. 137, No. 7
24	2015	Kevin Dopart	Automated and Connected Vehicles	Summary of the 9th University Transportation Centers Spotlight Conference; November 4-5, 2015, Washington, D.C.
25	2018	G. Fainekos H. B. Amor A. Shrivastava A. Dokhanchi E. Tuncali	RSS Model: Formalization with Temporal Logics for testing, verification, and control	Intel-ASU Automated Mobility Group Meeting; 12-13-2018.
26	2019	NHTSA	Removing Regulatory Barriers for Vehicles with Automated Driving Systems	Department of Transportation, National Highway Traffic Safety Administration; 49 CFR Part 571, Docket No. NHTSA-2019-0036.
27	2015	Calvin Pin Tarek Sayed Mohamed H. Zaki	Assessing Safety Improvements to Pedestrian Crossings Using Automated Conflict Analysis	Transportation Research Record 2514; Pages 58-67.

© Institute of Automated Mobility.

(Continued)

Publication ID	Year	Authors	Title	Publication
28	2016	Emanuele Sacchi Tarek Sayed	Conflict-Based Safety Performance Functions for Predicting Traffic Collisions by Type	Transportation Research Record 2583; Pages 50-55.
29	2018	Amir Rasouli John K. Tsotsos	Autonomous Vehicles that Interact with Pedestrians: A Survey of Theory and Practice	arXiv:1805.11773 [cs.RO]
30	2003	Vogel Katja	A comparison of headway and time to collision as safety indicators	Accident Analysis and Prevention 35 (2003) 427-433
31	2018	Danyang Tian Guoyuan Wu Kanok Boriboonsomsin Matthew J. Barth	Performance Measurement Evaluation Framework and Co-Benefit/Tradeoff Analysis for Connected and Automated Vehicles (CAV) Applications: A Survey	IEE Intelligent Transportation Systems Magazine, DOI 10.1109/MITS.2018.2842020
32	2018	Mark M. Morando Qingyun Tian Long T. Truong Hai L. Vu	Studying the Safety Impact of Autonomous Vehicles using Simulation-Based Surrogate Safety Measures	Journal of Advanced Transportation (2018), DOI: 10.115/2018/6135183
33	2003	Vogel Katja	A comparison of headway and time to collision as safety indicators	Accident Analysis and Prevention 35 (2003) 427-433
34	2009	Andrew Tarko Gary Davis Nicolas Saunier Tarek Sayed Simon Washington	Surrogate Measures of Safety	ANB20(3) Subcommittee on Surrogate Measures of Safety ANB20 Committee on Safety Data Evaluation and Analysis
35	2009	El-Basyouny, Karim Sayed, Tarek	Collision prediction models using multivariate Poisson-lognormal regression	Accident Analysis and Prevention 41 (2009) 820-828
36	2010	El-Basyouny, Karim Sayed, Tarek	A method to account for outliers in the development of safety performance functions	Accident Analysis and Prevention 42 (2010) 1266-1272
37	2019	Arvin, Ramin Kamrani, Mohsen Khattak, Asad	How instantaneous driving behavior contributes to crashes at intersections: Extracting useful information from connected vehicle message data	Accident Analysis and Prevention 127 (2019) 118-133
38	2018	NHTSA (Thorn, E., Kimmel, S., and Chaka, M.)	A Framework for Automated Driving System Testable Cases and Scenarios	Report No. DOT HS 812 623. Washington, DC: National Highway Traffic Safety Administration.
39	2015	Mohamed Essa Tarek Sayed	Simulated Traffic Conflicts: Do They Accurately Represent Field-Measured Conflicts?	Transportation Research Record 2514; Pg. 48-57
40	2015	Scott Smith Jeffrey Bellone Stephen Bransfield Amy Ingles George Noel Erin Reed Mikio Yanagisawa	Benefits Estimation Framework for Automated Vehicle Operations	Report No. FHWA JPO 16 229. Cambridge, MA: U.S. Department of Transportation
41	2017	Li Meng Wang Xiao Han Shi Ke	Traffic Conflict Identification Technology of Vehicle Intersection Based on Vehicle Video Trajectory Extraction	The 7th International Symposium on Frontiers in Ambient and Mobile Systems; Pg 963-968.

© Institute of Automated Mobility.

(Continued)

Publication ID	Year	Authors	Title	Publication
42	2019	Mohammed Essa Tarek Sayed	Full Bayesian conflict-based models for real time safety evaluation of signalized intersections	Accident Analysis and Prevention 129 (2019) 367-381
43	2013	Huanyun Zhou Fei Huang	Development of traffic safety evaluation method based on simulated conflicts at signalized intersections	13th COTA International Conference of Transportation Professionals (CICTP 2013)
44	2013	Tarek Sayed Mohamed Zaki Jarvis Autey	A Novel Approach for Diagnosing Road Safety Issues Using Automated Computer Vision Techniques	16th Road Safety on Four Continents Conference
45	2019	Nicolas Saunier Tarek Sayed	Automated Analysis of Road Safety with Video Data	Transportation Research Record: Journal of the Transportation Research Board, No. 2019, Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 57-64.
46	2018	Deepti Muley Mohammed Ghanim Mohamed Kharbeche	Prediction of Traffic Conflicts at Signalized Intersections using SSAM	The 9th International Conference on Ambient Systems, Networks and Technologies
47	2018	Ding Zhao Xianan Huang Huei Peng Henry Lam David J. LeBlanc	Accelerated Evaluation of Automated Vehicles in Car-Following Maneuvers	IEEE Transactions on Intelligent Transportation Systems, Vol. 19, No. 3, 2018, pp. 733-744.
48	2015	Jaehyun So In-Kyu Lim Young-Jun Kweon	Exploring Traffic Conflict-Based Surrogate Approach for Safety Assessment of Highway Facilities	Transportation Research Record: Journal of the Transportation Research Board, No. 2513, Transportation Research Board of the National Academies, Washington, D.C., 2015, pp. 56-62.
49	2013	Kray, J.H. Horst, A.R.A. van der Oppe, S.	Manual conflict observation technique DOCTOR	Report 2013-1, Foundation Road safety for all, The Netherlands
50	2014	Nadimi, N. Behbahani, H. Shahbazi, H.	Calibration and Validation of a New Time-Based Surrogate Safety Measure using Fuzzy Inference System	Journal of Traffic and Transportation Engineering (English Edition), 2016; 3 (I) pp.51-58

© Institute of Automated Mobility.

Appendix B: Metrics from Literature Review

Metric	Definition
Minimum Safe Distance Violation	Minimum Safe Distance Violation (MSDV) is defined as an instance in which the actions of the ego vehicle result in encroaching upon its safe boundaries with another (safety-relevant) entity within the scenario environment, as defined by current velocities and acceleration capabilities of both entities.
Proper Response Action	Proper Response Action (PRA) is defined as an instance of an action (longitudinal and/or lateral acceleration), but not necessarily an evasive action, taken by the ego vehicle to restore itself to its calculated safety envelope after a safety envelope violation has occurred. The PRA must occur at a pre-determined rate in order to be deemed a sufficient response.
Minimum Safe Distance Factor	Minimum Safe Distance Factor (MSDF) is defined as a multiple of the minimum lateral and longitudinal safe distances of the safety envelope maintained by an ego vehicle from surrounding entities. It is the ratio of measured distances (lateral and longitudinal) to the calculated safe distances (lateral and longitudinal).
Minimum Safe Distance Calculation Error	Minimum Safe Distance Calculation Error (MSDCE) is defined by the accuracy of the ADS calculation of its safety boundaries with respect to other safety-relevant entities in comparison to ground truth. The safety-relevant entities are those that can potentially result in a MSDV' = 1.
Collision Incident	Collision Incident (CI) is defined as an instance of the ego vehicle being at fault in a collision (the severity will be determined using the KABCO scale [6]).
Rules-of-the-Road Violation	Rules-of-the-Road Violation (RRV) is defined as an instance of the ego vehicle violating a traffic regulation that would result in an infraction or citation. This metric focuses only on moving violations and would not include violations for parking or lack of proper registration paperwork, for example. Possible road violations to consider include but are not limited to: disregarding a red light, disregarding a stop or yield sign, illegal crossover of designated lane markings, exceeding speed limits, and improper or non-use of turn signals. It should be noted that there may be situations where it is necessary to violate a rule of the road to achieve a safe outcome. For example, if a vehicle is instructed to drive over the centerline by a police officer directing traffic, the vehicle should follow such instructions even though doing so would initiate a rule of the road violation.
Achieved Behavioral Competency	Achieved Behavioral Competency (ABC) is a confirmation that the ego vehicle has correctly executed a specific behavioral competency. This metric is designed to be used in collaboration with the ADS developer in a testing scenario in which the behavioral competency to be demonstrated is agreed upon a priori. An example of the behavioral competency metric is the list from the Waymo Vehicle Safety Self-Assessment document submitted to NHTSA in 2018 [7], which includes 28 behavioral competencies from NHTSA (originating from work done by the California PATH) and 19 additional behavioral competencies added by Waymo. It is yet to be decided which behavioral competencies will be required to validate safe operation.
ADS Active	ADS Active (ADSA) is a confirmation that the ADS is active while executing the behavioral competency. This metric is designed to be used in collaboration with the ADS developer in a testing scenario in which the ADS developer must prove that the ADS is active at all times during behavioral competency execution. This proof can be via vehicle actuation data or some other agreed-upon format.
Human Traffic Control Detection Error Rate	Human Traffic Control Detection Error Rate (HTCDER) is a confirmation that the ego vehicle can detect the direction(s) of a human traffic control (HTC) actor (which may include a direction to violate one or more rules of the road).
Human Traffic Control Violation Rate	Human Traffic Control Violation Rate (HTVR) is a confirmation that the ego vehicle, once it has detected direction(s) of an HTC actor (which may include a direction to violate one or more rules of the road), follows the instructions.
Time-to-Collision	TTC is defined as the time until a collision between two entities in the scenario environment that would occur if both continue with the present velocities. A violation occurs if the measured TTC is below a pre-determined threshold.
Modified Time-to-Collision	MTTC is defined as the time until a collision between two entities in the scenario environment would occur if both continue with the present velocities and accelerations. A violation occurs if the measured MTTC is below a pre-determined threshold.
Post-Encroachment Time	PET is defined as the time from the end of encroachment of the lead vehicle to the beginning of encroachment of the second vehicle to the potential point of collision (of a conflict). A violation occurs if the measured PET is below a pre-determined threshold.
Aggressive Driving	Aggressive Driving (AD) is defined as repeated instances of maneuvers (longitudinal and/or lateral accelerations) above specified thresholds completed by the ego vehicle that, while not resulting in safety envelope violations or evasive actions, cumulatively are still potentially less safe. The thresholds will vary by jurisdiction and/or culture.

© Institute of Automated Mobility.

(Continued)

Metric	Definition
Evasive Action*	Evasive Action (EA) is defined as an instance of a maneuver (longitudinal or lateral acceleration) above specified thresholds completed by the ego vehicle (including non-ADS-controlled active safety features) to avoid an imminent crash (based on velocities and accelerations) with another entity within the scenario environment (or vice versa).
Multiple Object Detection Error Rate*	Multiple Objects Detection Error Rate (MODER) is defined as the error rate of the ADS in tracking safety-relevant entities within the ego vehicle's field of view in the scenario environment in comparison to the ground truth tracking.
Multiple Object Tracking Precision*	Multiple Object Tracking Precision (MOTP) is defined as the ability of the ADS to determine the position of safety-relevant entities within the ego vehicle's field of view in the scenario environment in comparison to the ground truth positions.
Entities Classified Error Rate*	Entities Classified Error Rate (ECER) is defined as the error rate of the ADS in classifying safety-relevant entities within the ego vehicle's field of view in the scenario environment in comparison to the ground truth classifications.
Rear-End Collisions*	Number of accidents caused by bullet vehicle impacting rear of target vehicle.
Following Distance*	The distance between any two moving or non-moving consecutive vehicles.
Driver Cutoffs*	Situation in which one vehicle encroaches on a defined safety envelope of another vehicle from the front resulting in unsafe following distance.
Number of Evasive Actions*	Maneuvers taken by a vehicle to avoid a conflict which would be imminent based on the evaluated vehicle speed and trajectory.
Maximum Queue Length*	Maximum amount of time a vehicle waits at a red light before the light turns green.
Platoon Ratio*	The ratio of the number of vehicles arriving at the green phase of a traffic light compared to the green interval of the total cycle.
Number of Pedestrians/ Cyclists*	The number of identified pedestrians/cyclists identified by infrastructure sensors compared to the total number of pedestrians/cyclists identified by traffic cameras
Flow Density*	A measurement to evaluate traffic congestion based on the average vehicle travel speed compared to roadway speed limit.
Instantaneous Safety Metric (ISM)*	The probability of an unavoidable collision occurring.
Near Misses*	A situation which almost resulted in a crash measured by a vehicle encroaching within the safety envelope of another vehicle without resulting in actual contact.
Rapid Acceleration/ Deceleration*	A sudden increase or decrease in speed, most likely resulting from an evasive maneuver.
Delta Speed*	Maximum speed of a conflict event as defined by the independently determined maximum speed for each involved vehicle.
Gap Time*	Maximum difference in velocity between conflicting vehicles during a conflict event.

© Institute of Automated Mobility.

* **Denotes a metric not included in the preliminary list**

© 2020 Institute of Automated Mobility. Published by SAE International. This Open Access article is published under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits distribution, and reproduction in any medium, provided that the original author(s) and the source are credited.

Positions and opinions advanced in this work are those of the author(s) and not necessarily those of SAE International. Responsibility for the content of the work lies solely with the author(s).

e-ISSN 2641-9645