Calibration and evaluation of responsibility-sensitive safety (RSS) in automated vehicle performance during cut-in scenarios

Shuang Liu\textsuperscript{a}, Xuesong Wang\textsuperscript{a, b, *}, Omar Hassanin\textsuperscript{a}, Xiaoyan Xu\textsuperscript{a}, Minming Yang\textsuperscript{a}, David Hurwitz\textsuperscript{c}, Xiangbin Wu\textsuperscript{d}

\textsuperscript{a} School of Transportation Engineering, Tongji University, Shanghai, 201804, China
\textsuperscript{b} The Key Laboratory of Road and Traffic Engineering, Ministry of Education, Shanghai, 201804, China
\textsuperscript{c} School of Civil and Construction Engineering, Oregon State University, Corvallis, OR 97331, United States
\textsuperscript{d} Intelligent Driving Lab, Intel Labs China, Beijing 100190, China

\textbf{ARTICLE INFO}

\textbf{Keywords:}
Responsibility-Sensitive Safety (RSS)
Adaptive cruise control (ACC)
Automated driving
Cut-in scenario
Surrogate safety measures (SSM)

\textbf{ABSTRACT}

The ability of automated vehicles (AV) to avoid accidents in complex traffic environments is the focus of considerable public attention. Intel has proposed a mathematical model called Responsibility-Sensitive Safety (RSS) to ensure AVs maintain a safe distance from surrounding vehicles, but testing has, to date, been limited. This study calibrates and evaluates the RSS model based on cut-in scenarios in which minimal time-to-collision (TTC) is less than 3 s. Two hundred cut-in events were extracted from Shanghai Naturalistic Driving Study data, and the corresponding scenario information for each event was imported into a simulation platform. In each scenario, the human driver was replaced by an AV driven by the model predictive control-based adaptive cruise control (ACC) system embedded with the RSS model. The safety performance of three conditions, the human driver, RSS-embedded ACC model, and ACC-only model, were evaluated and compared. Compared to the performance of human drivers and ACC-only algorithm respectively, the RSS model increased the average TTC per event by 2.86 s and 0.94 s, shortened time-exposed TTC by 1.34 s and 0.65 s, and reduced time-integrated TTC by 0.91 s\textsuperscript{2} and 0.72 s\textsuperscript{2}. These changes indicate that the RSS-embedded ACC model can improve safety performance in emergent cut-in scenarios. The RSS model can therefore be applied as a security guarantee, that is, to ensure the AV’s timely awareness and response to dangerous cut-in situations, thus mitigating potential conflict.

1. Introduction

As the automated vehicle (AV) moves closer toward general use, its safety has become the focus of significant public attention. According to the 2018 annual autopilot report released by California’s Department of Motor Vehicles, even the vehicle produced by Waymo, the most successful company in the field of automated driving, still requires manual intervention every 8.95 thousand kilometers (CA, 2018); that is, an accident may occur if the vehicle continues to be automatically driven without timely human intervention. To increase AV dependability and avoid traffic accidents, it is essential to standardize the AV’s safety assurance.

Adaptive cruise control (ACC), one of the more widely used advanced driver assistance systems (ADAS), is the most commonly
employed AV control algorithm. As it has been demonstrated to stably and efficiently control a vehicle’s longitudinal relative distance from the lead vehicle ahead (Gerhard et al., 2005; Magdici and Althoff, 2017), ACC has gradually become a common feature in vehicle configuration. However, since the goal of ACC is to maintain a safe distance in normal driving conditions, if an event requiring a quick response occurs, for example, if an adjacent vehicle suddenly changes lanes to cut in front of the subject vehicle, the ACC system will be deactivated and the driver will be required to take over control of the vehicle (Alipour-Fanid et al., 2017; Milanes et al., 2014).

A cut-in event, in which a vehicle changing lanes enters a space relatively close to vehicles in the targeted lane, while not within the definition of normal driving conditions, is a commonplace driving scenario. Studies have demonstrated that cut-in behavior is likely to cause traffic congestion, conflicts, and even crashes. In 2015, illegal lane changing and illegal overtaking accounted for 4.9% of the total number of crashes in China (Traffic Management Bureau of the Public Security Ministry, 2016). In the United States, the 240,000 to 610,000 traffic crashes each year due to improper lane changes account for 6% of all crashes (Hou et al., 2015). Because cut-in behavior deactivates ADAS, it is clear that cut-ins adversely affect the safety performance of AVs (Dou et al., 2016; Zhou and Peng, 2005). To mitigate crash risk, the ACC system needs to be improved to better respond to the cut-in scenario.

The Responsibility-Sensitive Safety (RSS) model, developed by Intel Mobileye, formalizes human drivers’ subjective and interpretable principles of safety in a rigorous mathematical model (Shashua et al., 2019). The model defines the safe distance for all driving scenarios, provides the AV with an appropriate response to evade dangerous situations, and assigns responsibility for traffic crashes accordingly, that is, to those vehicles that do not follow the defined appropriate response. Thus, RSS can enhance the AV’s safety by being applied as a constraint to the AV’s ACC algorithm. RSS has a range of customizable parameters that can be set according to the human driver’s personal driving style, thereby making the AV more comfortable and predictable, or “human-like,” and consequently safer (Matts et al., 2019). In this study, we integrated RSS into the ACC system, that is, we modified the ACC algorithm by integrating into it the RSS model (Takahama and Akasaka, 2018), for the purpose of evaluating RSS’s ability to improve the AV’s safety performance in suddenly emerging cut-in scenarios.

Naturalistic driving data provides an opportunity to observe cut-in events during real-world driving. The Shanghai Naturalistic Driving Study (SH-NDS), the first naturalistic driving study (NDS) in China, was conducted from 2012 to 2015. By recording drivers’ behavior with high-accuracy sensors and cameras during their daily driving, kinematic data such as acceleration, speed, relative speed, and distance from the surrounding vehicles were continuously collected (Eskandarian, 2012). By December 2015, 161,055 km of driving data from 60 drivers had been collected, a quantity that contained abundant kinematic and video data for cut-in scenario identification and analysis.

From the SH-NDS data, this study retrieved emergent cut-in events, recovering the trajectory and speed of the subject vehicle (vehicle responding to the cut-in) to simulate virtual scenarios in which the subject vehicle was replaced by an AV controlled by the ACC algorithm embedded with RSS. The safety performance of the RSS-embedded ACC model was compared to that of the ACC-only model by estimating the time-to-collision (TTC) surrogate safety indicators of time-exposed TTC and time-integrated TTC. For evaluating the safety impact of RSS on emergent cut-in scenarios, this study found the optimal set of parameters that can make the RSS model achieve the best performance, thereby enhancing the safety of both the AV and the human-driven vehicles.

2. Data preparation

2.1. Shanghai Naturalistic driving study

The lane changing data used in this study were collected from the SH-NDS, which was jointly conducted by Tongji University, General Motors (GM), and the Virginia Tech Transportation Institute (VTITI). Five GM light vehicles were equipped for the study with Strategic Highway Research Program 2 (SHRP2) NextGen Data Acquisition Systems (DAS).

The DAS includes an interface box for collecting vehicle controller area network (CAN) data, an accelerometer, radar system, light meter, temperature/humidity sensor, GPS sensor, and four synchronized cameras. The accelerometer captures the longitudinal and lateral acceleration of the experimental vehicle, and the radar system measures the range and range rate to the vehicle in front, or lead vehicle (LV), and to vehicles in the adjacent lanes. The GPS sensor identifies the position of the experimental subject vehicle, and the four camera views can be used to verify the various sensor data (Wang et al., 2019).

By December 2015, 161,055 km of driving data from 60 drivers had been collected. The 60 participants in the Shanghai Naturalistic Driving Study (SH-NDS) were randomly sampled from the population of licensed Shanghai drivers. The distributions of their gender, age, and driving experience accord with those of the general Chinese driving population. The drivers had an average age of 38.43 and an average driving experience of 8.34 years. Just over 20% were women, which is comparable to the 23.48% proportion of women in the 2014 Chinese driving population (Traffic Management Bureau of the Ministry of Public Security). All participants were non-professional drivers who owned vehicles, had at least 2 years of driving experience, and had the need to drive daily (Zhu et al., 2018). They were relatively safe drivers: over half had received no traffic violations in the last 2 years, and of those violations, over half were for illegal parking. Just under half had been involved in crashes in the last 2 years, most of which were property damage only.

2.2. Cut-in scenario and extraction

The focus of this study is the emergent cut-in event in which a lane changing vehicle (LCV) cuts suddenly in front of the subject NDS vehicle, making the DAS-equipped NDS vehicle, for our purposes, the following vehicle (FV), and the LCV the lead vehicle (LV). The FV’s radar records basic information such as the position, velocity and acceleration of the LV and vehicles in adjacent lanes, permitting extraction of the LCV’s data.
Fig. 1 shows a typical NDS cut-in scenario and its AV simulation: the main difference is that the FV changes from the NDS vehicle to the simulated AV. Fig. 2 further illustrates the potential cut-in with images taken from an NDS FV’s forward camera. In SH-NDS data, T0 is the LV position directly in front of the NDS. As shown in Fig. 2, the red LV directly in front of the FV’s camera view is in the T0 position in the target lane. If the FV’s radar records a change in the position of a vehicle from the adjacent lane toward the T0 position, the vehicle is determined to be an LCV intending to execute a lane change into the targeted FV lane. In Fig. 2, the blue car is the LCV taking over the T0 position. If the lane change meets the X-Range critical condition defined below, it is determined to be a cut-in maneuver, and the LCV can be labeled a cut-in vehicle (CV) as in Fig. 2 (Wang et al., 2019).

Cut-in events were extracted for this study from the volume of naturalistic driving data, as illustrated in Fig. 3. The cut-in events were automatically extracted from the NDS data by applying a filter, an iterative process where initial criteria and thresholds followed the procedure adopted by Wang et al. (2019) and Wang and Xu (2019). Brief descriptions of the event extraction thresholds are as follows:

- FV and LCV speeds greater than 1 m/s, ensuring that both cars are always in motion.
- FV’s maximum lateral acceleration is less than 0.07 g, and the lane offset is less than 1.7 m, both of which ensure that the FV will not move laterally.
- The Y-range (lateral, or transverse, distance from LCV to FV) is less than 2.2 m, indicating that the LCV has begun to move to the FV’s (target) lane; and the Y-range is less than a maximum of 1.2 m to ensure that the LCV has become stable in the target lane. Together, these standards ensure that the LCV has changed lanes.
- The maximum X-Range (longitudinal distance from LCV to FV) is defined as 75 m. This threshold, determined by 200 observations in this study, prevents inclusion of lane changes that are so far ahead of the FV that the FV is unlikely to be affected by them.
- The minimum TTC less than 3 s limits extracted lane changes to sudden cut-in events.

The last two criterion define a cut-in. That is, they distinguish it from a relatively safe lane change. Through automatic extraction, we obtained an initial total of 5,339 cut-in events, defined as a lane change into a space shorter than 75 m from the FV in the target lane. The final criterion of TTC under 3 s was applied to limit the sample to the most sudden and therefore dangerous cut-ins, resulting in a final total of 200 cut-in events being selected for study. All of the 200 were manually verified by the forward roadway camera video, ensuring that all LCVs had indeed moved from adjacent lanes into the NDS FV target lanes. The manual verification process also confirmed the type of road and LCV (e.g., car or truck), as well as weather and lighting conditions, and turn signal usage at the time of the cut-in event (Yang et al., 2019).

Once the LCV has taken over the T0 position, it becomes the lead vehicle (LV). The next step toward generating the simulated scenario was calculation of the LV’s longitudinal trajectory by removing the noise coming from the sensors. The trajectory can be calculated by using this equation:

\[
LV_{Position(i+1)} = LV_{Position(i)} + V_i \times t + \frac{1}{2} a_i \times t^2
\]  

In Equation (1), \(LV_{Position(i)}\) is the cumulative position of the LV in the longitudinal direction, \(V\) is the LV’s velocity, and \(a\) is its...
acceleration, all at timestamp $i$ with $t$ as time interval (0.1 s) and $i$ as the simulation step.

Noise in the data generated backward movements in the trajectory, which, as shown in Fig. 4, caused the LV to drift and become unstable during the simulation. By using Equation (1) to remove the noise, the backward movements are eliminated and the line is smoothed. The calculation also ensured that no timestamp had zero speed value, thereby permitting the simulation to work properly.

Finally, because the LV’s longitudinal trajectory data changed substantially in a short time, the simulated LV was longitudinally unstable. As shown in Fig. 5, locally estimated scatterplot smoothing (LOESS) was used with span = 0.1 to overfit and smooth the relative longitudinal distance.

3. Methodology

3.1. Safety distance in RSS model

In the RSS model, safety distance consists of two measures: longitudinal and lateral. During cut-in events, the rear vehicle needs to pay close attention to the front vehicle, the relative position of which changes both longitudinally and laterally (Shashua et al., 2019;
Let $\upsilon_r$, $\upsilon_f$ be the longitudinal velocities of the rear vehicle $c_r$ and the forward vehicle $c_f$, respectively. The minimum safe longitudinal distance between $c_r$ and $c_f$ is:

$$
\begin{align*}
    d_{\text{min}} &= \upsilon_r \rho + \frac{1}{2} a_{\text{max, accel}} \rho^2 + \frac{(\upsilon_f + \rho a_{\text{max, accel}})^2}{2a_{\text{min, brake}}} - \frac{v_f^2}{2a_{\text{max, brake}}}.
\end{align*}
$$

In Equation (2), $a_{\text{max, brake}}$ is the acceleration of $c_f$ and $a_{\text{max, accel}}$ is the acceleration of $c_r$ during the response time $\rho$. After $\rho$, $c_r$ will brake by at least $a_{\text{min, brake}}$ until it reaches a full stop so it will not collide with $c_f$, $|x| = \max(x, 0)$.

Let $\upsilon_1$ and $\upsilon_2$ be the lateral velocities of the vehicles $c_r$ and $c_f$, respectively, during the time interval $[0, \rho]$, as $c_r$ and $c_f$ perform $a_{\text{max, brake}}$ as they move toward each other. $\upsilon_{1, \rho}$ denotes $\upsilon_1 + a_{\text{lat, accel}} \rho$, and $\upsilon_{2, \rho}$ denotes $\upsilon_2 - a_{\text{lat, accel}} \rho$. The two vehicles will then apply lateral braking of $a_{\text{min, brake}}$ until they reach zero lateral velocity, after which the final lateral distance between them will be at least $\mu$. 

**Fig. 4.** Backward movements in LV trajectory and solution.

**Fig. 5.** Relative longitudinal distance smoothed with LOESS.
The minimal safe lateral distance between $c_f$ and $c_j$ is:

$$d_{min}^l = \mu + \left\lfloor \frac{v_1 + v_\rho}{2} + \frac{v_{1,\rho}^2}{2d_{min,\text{brake}}} - \left( \frac{v_2 + v_\rho}{2} - \frac{v_{2,\rho}^2}{2d_{min,\text{brake}}} \right) \right\rfloor$$  \hspace{1cm} (3)

In the current study, the longitudinal and lateral distances between the LCV and the NDS FV were detected by the NDS sensors. These distances were reproduced in the simulation as the distances between the simulated LCV and autonomous vehicle (AV) which replaced the NDS FV and which was modeled with the RSS-embedded ACC algorithm. At each simulation step, the RSS-defined safe longitudinal and lateral distances were calculated using the relative speed and distances detected between the LCV (front vehicle, in RSS terms) and the AV (rear vehicle). If the current distance was less than the RSS safety distance, the RSS algorithm would be triggered and the AV would decelerate at the RSS minimum deceleration. Otherwise, the AV would remain under the control of the ACC algorithm.

### 3.2. Surrogate safety measurements

As crashes are relatively infrequent, a sufficient sample size of crash data is often unavailable. Potential conflicts and near crashes are much more frequent and are positively related to crashes, so alternative approaches (Hydén, 1987; Oh and Kim, 2010) such as surrogate safety measurements (SSM) are commonly used to evaluate safety performance. SSM, also known as safety indicators, can be used to quantify the occurrence and severity of potential crash risk and vehicle conflict (Zheng et al., 2014). These indicators are derived from vehicle interaction analysis, including time-dependent vehicle movements, risk avoidance behavior, and safety margin of conflicts (Shi et al., 2018).

Time-to-collision (TTC) is the most widely used SSM (Zheng et al., 2014). Defined by the time it would take two vehicles to collide if they continued on their current course (Minderhoud and Bovy, 2001), a lower TTC value represents a more severe conflict. A TTC threshold is therefore commonly set to distinguish between safe and unsafe driving interactions (Nilsson et al., 1993; Hirst, 1997; Sultan et al., 2002; Li et al., 2014, 2016, 2017). Potential two-vehicle conflicts are thus generally identified from raw kinematic data by using a predefined TTC threshold, normally of 3 s. If the minimum TTC in a given dynamic conflict process is less than 1 s, the conflict is considered severe and would be detected (Shi et al., 2018).

#### 3.2.1. Time-exposed TTC (TET) and time-integrated TTC (TIT)

TTC thresholds ranging from 1 to 3 s had no significant impact on final results. In this study, the TTC threshold was therefore set to 3 s.

#### 3.2.2. Surrogate safety measurements (SSM)

SSM to assess the safety performance of ACC systems. After conducting sensitivity analysis, all four studies found that setting different TTC thresholds ranging from 1 to 3 s had no significant impact on final results. In this study, the TTC threshold was therefore set to 3 s in order to select a reasonably broad representation of typical cut-in events. TTC is computed by the following equation:

$$TTC_i(t) = \begin{cases} \frac{x_{i-1}(t) - x_i(t) - L_{i-1}}{\nu_i(t) - \nu_{i-1}(t)} & \nu_i(t) > \nu_{i-1}(t) \\ \infty & \nu_i(t) \leq \nu_{i-1}(t) \end{cases}$$  \hspace{1cm} (4)

where $x_i(t)$ is the position of the FV ($i$) at timestamp $t$, $L_{i-1}(t)$ is the length of the LV ($i$-1), and $\nu_i(t)$ is the velocity of the FV ($i$) at timestamp $t$.

As minimum TTC is an instantaneous variable, time-exposed TTC (TET) and time-integrated TTC (TIT) are derived from TTC to evaluate the duration and integration of TTCs below a certain threshold (Horst, 1991). TET and TIT are calculated by

$$TET_i = \sum_{\tau_c=0}^{N} \delta_i(t) \cdot \tau_c$$  \hspace{1cm} (5)

$$\delta_i(t) = \begin{cases} 1 & \forall 0 < TTC_i(t) \leq TTC^c \\ 0 & \text{otherwise} \end{cases}$$  \hspace{1cm} (6)

$$TIT_i = \sum_{\tau_c=0}^{N} \left| TTC_i(t) - TTC^c \right| \cdot \tau_c$$  \hspace{1cm} (7)

$$\forall 0 \leq TTC_i(t) \leq TTC^c$$  \hspace{1cm} (8)

where $\delta_i(t)$ is a variable that switches between 1 (signaling a risky condition) and 0 (non-risky), $\tau_c$ is the observation time interval (e.g. 0.1 s), and $TTC^c$ is the TTC threshold value (e.g. 3 s) (Nie et al., 2017).

### 3.3. Simulation platform

The MATLAB Automated Driving Toolbox™ provides autonomous vehicle and automated driving algorithms. This study uses two of its modules for testing the RSS-embedded ACC model:

1. The Vehicle and Environment module is responsible for simulating the dynamics of the AV, or ego vehicle (EV), driver steering, and reading the scenario from the NDS. The EV detects its surrounding environment with a 150-m range camera and a 174-m range 250-degree radar detector, and then sends the information to the ACC module described below.
2. The Adaptive Cruise Control (ACC) with Sensor Fusion module is responsible for taking together the output information from the EV (AV) detectors, and the information the EV obtains from the NDS scenario. The module then analyses both sets of data to obtain the most important object (MIO), which is the point in the LV nearest to the EV. The relative distance and relative speed between the EV and the MIO are calculated, and the information is sent to module predictive control (MPC). MPC makes the decision to decelerate or accelerate, and the decision is then sent to the Vehicle and Environment module above.

To evaluate RSS, this study compares two versions of the ACC model: with and without RSS. The RSS-embedded ACC model uses the RSS parameters of Equation (2) in Section 3.1. The ACC-only model uses its own equation for calculating the safe distance ($d_{safe}$) between the LV and AV. The ACC equation, however, has been found to lead to overlarge distances between the LV and EV, and its deceleration rate is insufficient to prevent crashes or reduce the severity of a conflict. The equation for $d_{safe}$ is:

$$d_{safe} = d_{default} + T_{gap} \times V_x$$

where the default spacing $d_{default}$ and time gap $T_{gap}$ are ACC design parameters; $V_x$ is the AV’s longitudinal velocity.

To compare the models, two steps were taken:

1. Running the ACC-only model using the values in Table 1 for the 200 events.
2. Running the RSS-embedded ACC model after the RSS parameters of Equation (2) are calibrated on the simulation results from the ACC-only model.

The safety performance of the RSS-embedded ACC model is highly affected by parameters in both the ACC and RSS algorithms. ACC reaction time, especially in commercial vehicles, has been found to be slightly shorter than that of human drivers, with values from 0.5 to 1.0 s demonstrated in previous studies (Zhu et al., 2020; Chai et al., 2019). Additionally, maximum and minimum deceleration rates in the RSS and ACC algorithms conform with each other. Specifically, this conformity ensures that the minimum deceleration rate cannot be too low, as the model may still be functional enough to help the AV avoid potential danger. According to Equation (2), the maximum deceleration also cannot be too high; otherwise the RSS safety distance will be too large, forcing the AV to maintain an overlarge distance from the cut-in vehicle. Maintaining this distance would result in an overly conservative model and extravagant use of road resources.

Employing the simulation parameters listed in Table 1, the Simulink module of MATLAB 2019a was used to build the simulations. The 200 events selected from SH-NDS data were converted into Simulink scenarios: for each virtual scenario, the trajectory and speed of the cut-in vehicle, or LCV, and the initial speed and position of the subject vehicle, or FV, were maintained from the original NDS event. An AV controlled by ACC-only or RSS-embedded ACC replaced the original subject vehicle, from which it inherited initial speed and position.

3.4. Calibration of RSS model

The aim of calibrating the RSS-embedded ACC model is to use parameter values for Equation (2) that maximize safety while ensuring the model is not too conservative. As noted in Step 2 above, the calibration was conducted on the data generated by the ACC-only simulation. Each event had its own distinct duration, which was defined by the number of rows in the data generated by the simulation. The simulation interval, or step, was 0.1 s. The objective function, described further below, is defined as:

$$\text{Minimize} \sum_{i=1}^{E} \left[ \sum_{t=1}^{T_{2i}} \left( \frac{TTC - TTC_i(t)}{(TTC - TTC_i(t)).(\#ofsteps)} \right) \right]$$

$$+ \sum_{t=1}^{T_{2i}} \left[ \sum_{r=1}^{T_{2i}} \text{absolute}(d_{relative}(t) - d_{relative}(t))/(d_{relative}(t).(T_{2i} - T_{rn})) - (T_{2i} - T_{rn})/(T_{2i} - T_{rn}) \right]$$

where

$$\forall 0 \leq TTC_i(t) \leq TTC^*$$

$$\forall 1 \leq E, e \leq 200$$

$E$ is the total number of events that did not result in accidents, $TTC^*$ is critical time-to-collision (3 s), $t^*$ is the first time step when TTC was lower than $TTC^*$, $T_1$ is the last time step when $TTC < TTC^*$, and $(\#ofsteps)$ is the number of steps taken when TTC was lower than $TTC^*$ with respect to $\forall 0 \leq TTC_i(t) \leq TTC^*$.

<table>
<thead>
<tr>
<th>Table 1: Simulation Algorithm Parameters for ACC-only model.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
</tr>
<tr>
<td>Time of deceleration</td>
</tr>
<tr>
<td>ACC deceleration rate</td>
</tr>
<tr>
<td>ACC reaction time</td>
</tr>
<tr>
<td>ACC default space ($d_{default}$)</td>
</tr>
<tr>
<td>ACC time gap ($T_{gap}$)</td>
</tr>
</tbody>
</table>
$d_{\text{min-rss}}(t)$ is the output from Equation (2), and $d_{\text{relative}}(t)$ is the distance between LV and EV. $e$ is the number of events without accidents and $d_{\text{min-rss}}(t) > d_{\text{relative}}(t)$ for at least one time step, which means RSS was ready to be activated. $t_{\text{rss-i}}$ is the step wherein RSS can be activated and $d_{\text{min-rss}}(t) > d_{\text{relative}}(t)$, and $T_2i$ is the total number of steps in the event (Time*10).

The objective function consists of three areas related to safety, conservativeness, and time of RSS activation. The example in Fig. 6 shows the parameters $(d_{\text{min-rss}}(t), d_{\text{relative}}, \text{TTC, TTC}^*)$ of the objective function generated by using the calibrated values:

Descriptions of the three areas and their processes are as follows:

1. Safety area:
   - $\sum_{TTC^* < TTC} (TTC^* - TTC_i(t))$ is the area between TTC* and TTC less than 3 s during the simulation of a specific event, depicted by the blue area in Fig. 6. The smaller the area, the safer the event.
   - The area is divided by TTC* to decrease the objective function unit and to describe it as a percentage.
   - The area is also divided by (# of steps) to obtain the percentage for each time step in order to assist comparison between the three areas of the objective function. Because the conservativeness area will be larger than the safety area, reducing the effect of the safety area in the calibration process, the percentage must be known to identify which area is most important.

2. Conservativeness area:
   - The distance between LV and EV should not be too large, that is, too conservative, which can happen when the difference between $d_{\text{min-rss}}$ and $d_{\text{relative}}$ is small. Because $d_{\text{min-rss}}$ is the safe distance at which the EV should remain so as to avoid a crash, the distance is almost enough if $d_{\text{min-rss}}$ is equal to $d_{\text{relative}}$.
   - The $\sum_{TTC^* < TTC} \text{absolute}(d_{\text{min-rss}}(t) - d_{\text{relative}}(t))$ area is calculated after $t_{\text{rss-i}}$ because it indicates the RSS will be activated. The “absolute” is necessary because if $d_{\text{min-rss}}(t) < d_{\text{relative}}(t)$, the area will have a negative value and will thus reduce the objective function incorrectly.
   - The area is divided by $d_{\text{relative}}$ to decrease the objective function unit and to describe it as a percentage.
   - The area is also divided by $(T_2i - t_{\text{rss-i}})$, which is the # of steps in the orange area, to make it easier to compare the three areas. The sum of this area’s events is $e$ rather than $E$ because $e$ includes only the events in which RSS must be activated; that is, non-dangerous events are excluded.

3. Time of RSS activation area:
   - The earlier the RSS model activates, the safer the vehicle will be. The TTC will not have entered the critical level, and the conservativeness area will be smaller because the RSS-embedded model will have attempted to minimize the area by making $d_{\text{min-rss}}(t)$ and $d_{\text{relative}}(t)$ almost equal.
   - The sign of this area is negative, which indicates that the larger the number of events that RSS activates, the smaller the objective function will be. The aim is to force the objective function to choose a set of parameters that activate RSS earlier for every event.
   - Early RSS activation means that $(T_2i - t_{\text{rss-i}})$ will be larger because $t_{\text{rss-i}}$ will be smaller; thus it is always better to activate RSS earlier.
   - This area is divided by $T_2i$ to decrease it and to describe it as a percentage.
   - This area is also divided by $e$ to obtain the average percentage per event and to make the area smaller than 1.0.

Fig. 6. Simulation results from ACC-only model for an event $i = 48$. 
This study makes some assumptions:

- The calibration calculates the \( d_{\text{min}} \) for the RSS, but does not activate the RSS model.
- All the events with accidents in the ACC-only model are removed from the calibration because after the crash occurs, the relative distance becomes negative and then zero because there is no longer any distance between the FV and LV.
- The steps with relative distance under 1 m are removed. In the objective function, the conservativeness area is divided by \( d_{\text{relative}} \), and will be very large if \( d_{\text{relative}} < 1 \); consequently, these steps would generate incorrect results for the objective function. Only 34 out of the 34,052 steps from all events have \( d_{\text{relative}} < 1 \).
- In Fig. 6, TTCs longer than 10 s are shown as 10 s for simpler visualization.

The genetic algorithm (GA) is a widely implemented calibration method. The GA is effective at finding a model’s optimal parameter combination as it possesses the advantage of being able to solve both constrained and unconstrained global optimization problems while avoiding local minima (Saifuzzaman et al., 2015; Zhu et al., 2018). GA proceeds as follows: 1) a population consisting of \( N \) individuals is initialized, and each individual refers to a random parameter that is set, for this study, in the RSS model; 2) the fitness of each individual is calculated by the predefined objective function; 3) crossovers between randomly selected individual pairs (parents) and mutations within randomly selected individuals are implemented to produce individuals of the next generation (children); and 4) steps 2 and 3 are repeated until the termination criteria are satisfied.

It is important to choose the proper GA parameters for calibrating micro-simulation models. If the population is too small and sparsely spread, then the lack of genetic diversity may lead to quick convergence on local optima before the better optima may be visited. On the other hand, excessively large populations cause the GA to act like a random search algorithm (Ma and Abdulhai, 2001). Previous research has employed a variety of specific values (Table 2). Despite the caution from Ma and Abdulhai, the most recent studies, Saifuzzaman et al. (2015) and Zhu et al. (2018), used large values for the three GA parameters.

The genetic algorithm function in R was used in this study, and the GA parameter values used by Zhu et al. (2018) were adopted because both studies used the same data source, the Shanghai Naturalistic Driving Study (SH-NDS). The GA parameters for RSS calibration were thus specified as follows: population size 300, maximum number of generations 300, and number of stall generations 100. The maximum number of generations controlled the number of iterations. Since GA is a stochastic process, each optimization run produces a different solution. To find the solution closest to the global optimum, the optimization process was repeated 6 times, and the parameter combination with the minimum objective function was selected.

While running the GA on ACC-only scenarios, it was found that GA was reaching the global optima at a generation of 220–270 out of 300. This confirmed the study’s choice to use 300 as the maximum number of generations, as it gave the GA the ability to reach the global optima.

4. Results

4.1. RSS calibration results

Before running the GA, the upper and lower bounds of each parameter in the objective function needed to be estimated. Histograms were generated from the SH-NDS data to identify the real bounds of the RSS parameters during the cut-in events.

Using the standard maximum comfortable deceleration of 3 to 3.5 m/s\(^2\) (Xu et al., 2021; Bokare and Maurya, 2017; Maurya and Bokare, 2012) along with the acceleration data depicted in Fig. 7, the bounds for the four parameters were estimated. Table 3 shows the bounds and results for each parameter in each of the six GA calibration runs.

As can be seen in Table 3, Run 1 had the lowest minimum value for the objective function. Run 1 parameter values were therefore used to run the RSS-embedded ACC model and to calculate the minimum safety distance using the following equation:

\[
\begin{align*}
    d_{\text{min}} &= \left[ 0.496v_r + 1.542\rho^2 + \frac{(v_r + 3.084\rho)^2}{6.964} - \frac{v_i^2}{11.376} \right] \text{s}.
\end{align*}
\] (11)

The 95% quantile of lateral acceleration from the SH-NDS data was used as the maximum lateral acceleration in the RSS model, with a value set at 0.68 m/s\(^2\). In order to exclude crash events, the 75% quantile was used to set the RSS minimum lateral brake rate at 0.45 m/s\(^2\), and the minimal lateral distance was set at 0.2 m. The formula for calculating the transverse safety distance for the RSS model is, therefore, as shown in Equation (12).

<table>
<thead>
<tr>
<th>Research</th>
<th>Population Size</th>
<th>Maximum number of generations</th>
<th>Stall generations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhu et al. (2018)</td>
<td>300</td>
<td>300</td>
<td>100</td>
</tr>
<tr>
<td>Saifuzzaman et al. (2015)</td>
<td>200</td>
<td>600</td>
<td>100</td>
</tr>
<tr>
<td>Ma et al. (2007)</td>
<td>30</td>
<td>20</td>
<td>——</td>
</tr>
<tr>
<td>Kim et al. (2005)</td>
<td>30</td>
<td>150</td>
<td>100</td>
</tr>
<tr>
<td>Schultz and Rilett (2004)</td>
<td>30</td>
<td>40</td>
<td>——</td>
</tr>
</tbody>
</table>
\[ d_{\text{init}} = 0.2 + \frac{2\nu_1 + 0.337}{2} \times 0.496 + \left( \frac{\nu_1 + 0.337}{0.900} \right)^2 - \frac{\left( 2\nu_2 - 0.337 \right)}{2} \times 0.496 - \left( \frac{\nu_2 - 0.337}{0.900} \right)^2 \] 

4.2. Summary of simulation results

In the simulated 200 cut-in events, the AV reproduced the initial speed and position taken by the original NDS vehicle, but occasionally chose to accelerate and overtake the lane changing vehicle (LCV) instead of decelerating to open a safer gap. In these scenarios, the AV started to decelerate when the ACC algorithm detected an emergent cut-in, but if the level of deceleration was not enough to keep the AV behind the LCV, the AV quickly changed strategies to attempt to overtake the LCV. Consequently, either a crash occurred or the LCV abandoned its intent to change lanes. In either case, the intended cut-in failed, so we considered these scenarios invalid for our purposes. Because valid, or completed, cut-ins result from adequate AV deceleration, they tend to have lower initial AV speed and shorter relative longitudinal distance.

### Table 3
Summary of parameter estimates produced by the 6 GA optimization runs.

<table>
<thead>
<tr>
<th>Parameters (Units)</th>
<th>Bounds</th>
<th>Parameter estimates for each optimization run</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \rho ) (s)</td>
<td>[0.1, 0.5]</td>
<td>0.496 0.498 0.497 0.498 0.495 0.494</td>
</tr>
<tr>
<td>( a_{\text{max, accel}} ) (m/s(^2))</td>
<td>[1.5, 3.1]</td>
<td>3.084 3.047 3.037 3.046 3.068 3.053</td>
</tr>
<tr>
<td>( a_{\text{min, brake}} ) (m/s(^2))</td>
<td>[-4, -2]</td>
<td>3.482 3.592 3.376 3.591 3.600 3.668</td>
</tr>
<tr>
<td>( a_{\text{max, brake}} ) (m/s(^2))</td>
<td>[-7, -5]</td>
<td>5.688 6.031 5.531 6.041 6.110 6.479</td>
</tr>
</tbody>
</table>

### Table 4
AV’s initial speed and distance for valid and invalid events.

<table>
<thead>
<tr>
<th></th>
<th>Valid events</th>
<th>Invalid events</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ACC-only</td>
<td>RSS-embedded</td>
</tr>
<tr>
<td>Initial speed (m/s)</td>
<td>Range</td>
<td>[2.11, 30.87]</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>11.16</td>
</tr>
<tr>
<td></td>
<td>Std.</td>
<td>5.53</td>
</tr>
<tr>
<td>Initial longitudinal distance (m)</td>
<td>Range</td>
<td>[4.42, 52.85]</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>18.22</td>
</tr>
<tr>
<td></td>
<td>Std.</td>
<td>9.98</td>
</tr>
<tr>
<td>Number of events</td>
<td></td>
<td>185</td>
</tr>
</tbody>
</table>
Both valid and invalid events were analyzed for ACC-only and RSS-embedded ACC. Results are shown in Table 4.

As can be seen in Table 4, the higher number of valid events and lower number of invalid events for the RSS-embedded model implies that it was both more effective and conservative in dealing with dangerous cut-in scenarios. After comparing the mean and standard deviation of the AV’s initial longitudinal speed and distance for valid and invalid events, further observations can be made:

1) The average AV initial speed for ACC-only is 11.16 m/s for valid events, which is 3.53 m/s lower than the 14.69 m/s speed for the invalid events. For RSS-embedded, the average initial speed of the valid events is 11.40 m/s, approximately the same as for the invalid events at 11.45 m/s.

2) As noted above, the AV’s average initial longitudinal distance from the cut-in vehicle is shorter in the valid events for both models. For ACC-only, the average initial longitudinal distance of the valid events is 18.22 m, which is 17.47 m shorter than the 35.69 m distance for invalid events. For RSS-embedded, the distance for the valid events is 19.43 m, just 4.40 m shorter than the 23.83 distance of the invalid events.

The above findings indicate that the AV algorithm was more likely to perceive and react to the potential cut-in events when the AV began with a lower longitudinal speed and a shorter longitudinal distance from the cut-in vehicle.

4.3. Safety performance evaluation

The distributions of average longitudinal acceleration, speed, relative longitudinal distance, and relative lateral distance for each valid cut-in event are illustrated in Fig. 8 in the form of boxplots, where values for the RSS-embedded ACC and ACC-only models are compared with each other and with the NDS human driver model.

As shown in Fig. 8(a), the longitudinal acceleration of human drivers has the narrowest range, with a median of approximately $-0.02 \text{ m/s}^2$. The average acceleration of RSS-embedded ACC ($-0.80 \text{ m/s}^2$) is the lowest, with a median of $-0.69 \text{ m/s}^2$. The average acceleration of the ACC-only algorithm ($-0.59 \text{ m/s}^2$, with a median of $-0.41 \text{ m/s}^2$) differs little from RSS-embedded.

Fig. 8(b) indicates that the average longitudinal speed differs only slightly between human drivers and the AV algorithms, with the ACC-only being most similar. The distribution of human drivers’ speed is concentrated at 8.92 m/s and the average for ACC-only is 8.76 m/s, while that of RSS-embedded is 8.39 m/s.

![Boxplots of kinematic values](image)

Fig. 8. Average kinematics values for all valid events.
In Fig. 8(c), the average longitudinal relative distance and distribution range can be seen to increase left to right, which indicates that the AV algorithms improve safety by enlarging the cut-in gaps. RSS-embedded ACC is the most conservative.

Fig. 8(d) compares the lateral safety distance distributions of the three models, but it should be noted that the RSS lateral distance parameters were not calibrated with the GA algorithm because simulation results showed that none of the models triggered the RSS minimum final lateral distance, which, according to Equation (3), is set to $\mu = 0.2$ m. Because any collision of two vehicles requires extreme proximity both laterally and longitudinally, the lateral comparisons were, therefore, based on a range incorporating the RSS minimum relative longitudinal distance of 4.7 m, the average length of a vehicle. To compare the lateral safety distance distributions, Fig. 8(d) therefore uses a longitudinal relative distance within 4.7 m. It should also be noted that because longitudinal and lateral relative distances together signify the likelihood of a collision, an increase in one direction can permit a shorter distance in the other. As seen in Fig. 8(c), the RSS algorithm increases the longitudinal relative distance; as the longitudinal distance increases, the number of events that trigger RSS due to their short lateral relative distance is reduced. For human drivers, the relative lateral distance between vehicles is concentrated between 0.03 m and 1.19 m, with an average distance of 0.61 m; for ACC-only, the concentration is 2.34–3.38 m, and is 2.23–3.27 m for RSS-embedded. Both the ACC-only and RSS-embedded ACC algorithms can thus improve the lateral relative distance over that of the human driver.

Fig. 9 illustrates the distributions in the three models for minimum TTC (time-to-collision), TET (time-exposed TTC), and TIT (time-integrated TTC). As one of the criteria for selection of emergent cut-in events from NDS data was that the minimum TTC was lower than 3 s, the human driver minimum TTC is distributed between 0.5 and 3 s; it is concentrated between 1 and 3 s, with an average of 2.14 s. Minimum TTC distribution in the ACC-only model begins below 1 s, is concentrated above 5.15 s, and averages 4.06 s. The shorter minimum TTC at the low end of the ACC-only distribution indicates that it increased the risk in some cut-in events; that is, it would not be as effective in these events as human drivers. In the RSS-embedded ACC model, however, not only is the minimum TTC distribution concentrated in the area above 1 s, but also, the average is increased to 5 s. This increase implies that the safety risk is mitigated when RSS is integrated into ACC.

Since TET and TIT represent, respectively, the duration and integration of TTCs shorter than 3 s, the lower the TET and TIT values, the safer the cut-in event is considered. TET for human drivers is generally distributed between 0.5 and 6 s, averages 1.64 s, and for 92% of the events, is concentrated between 0.5 and 3 s. TET for ACC-only is concentrated between 0 and 6 s with an average of 0.95 s; that is, events with TTC less than 3 s had a duration of 0.95 s on average. In RSS-embedded ACC, TET is concentrated between 0 and just 2.5 s, has an average of 0.3 s, and for 78% of events it is concentrated at 0 s, indicating that the improved safety performance of RSS-embedded is significant. Although the distributions and averages in Fig. 9(b) indicate that both ACC-only and RSS-embedded can enhance safety performance as compared with human drivers, RSS-embedded performed more effectively in reducing TTC duration, with the TET distribution being barely a third of ACC-only.

Similar observations can be made for the TIT indicator. TIT is distributed between 0 and 3.5 s$^2$ for RSS-embedded ACC, which is clearly less than that of the human driver. The average value is 0.19 s$^2$, which is less than one-fifth that of ACC-only (0.90 s$^2$) and less than one-sixth that of the human driver (1.10 s$^2$).

Fig. 10 shows the TIT distribution for the human driver compared to the reduced distribution for ACC-only and RSS-embedded. The horizontal and vertical coordinates represent the initial speed and longitudinal distance, respectively, of each set of the corresponding events. The color at the various coordinates illustrates the extent to which the human driver TIT is improved by the two algorithms: the 7 on the color scale indicates greatest improvement, so the warmer the color, the greater the degree of improvement.

Fig. 10(a) shows that human driver TIT is improved by ACC-only within certain limited ranges, as indicated, for example, by the yellow areas at an initial distance of 30–40 m with an initial speed lower than 10 m/s, and an initial distance larger than 30 m with an initial speed higher than 20 m/s. The fewer number and lower intensity of blue areas in Fig. 10(b) illustrate the more overall improvement with RSS-embedded. Additionally, Fig. 10(b)’s generally larger and warmer yellow areas in the upper right quadrant, depicting events with larger initial distance and higher initial speed, show RSS’s more significant improvement over the human driver than is evident for ACC-only. The difference between the models is even more apparent in the upper left quadrants, that is, in events with large initial distance but low initial speed. Calculation shows that RSS-embedded decreased the average TIT per event by 0.91 s$^2$ and 0.72 s$^2$ as compared to the performance of human drivers and ACC-only algorithm, respectively.

5. Discussion and conclusion

Adaptive cruise control (ACC), commonly used for automated vehicle (AV) control, has demonstrated success in normal driving conditions. The Responsibility-Sensitive Safety (RSS) algorithm can help the AV to respond safely to more urgent events, but testing has been limited. This study has therefore evaluated RSS’s safety impact on a vehicle in the case of a suddenly emergent cut-in. The trajectory and speed of the cut-in vehicle and the initial speed of the subject vehicle were extracted from 200 cut-in events observed in the Shanghai Naturalistic Driving Study (SH-NDS) data. The event information was imported into a simulation platform in which the human driver was replaced by an AV controlled by the ACC algorithm with and without RSS. Three models were developed, human driver, ACC-only, and RSS-embedded ACC, and were tested through the MATLAB Simulink module.

The safety performance of the RSS-embedded ACC algorithm was found superior to both the performance of the SH-NDS human drivers and the ACC-only algorithm. The RSS-embedded ACC simulation generated a greater number of valid cut-ins, that is, events in which the AV decelerated to open a safer gap, rather than the more dangerous response of accelerating to overtake the prospective cut-in vehicle. Additionally, RSS-embedded ACC generated events with an average initial speed and relative distance of 11.40 m/s and 19.43 m, respectively, both larger values than human driver or ACC-only events.

Time-to-collision (TTC) values were improved by the RSS-embedded algorithm. Compared to the performance of human drivers
and the ACC algorithm on its own, the RSS model increased the average TTC per event by 2.86 s and 0.94 s, respectively. The AV decelerated rapidly after RSS was triggered, reducing the required longitudinal safety distance and decreasing the time-integrated TTC (TIT) by an average of 0.91 s for human NDS events and 0.72 s for ACC-only events, thereby mitigating the potential emergency. Time-exposed TTC (TET), decreased an average of 1.34 s and 0.65 s for RSS-embedded events, and its distribution was within 0–2.5 s, narrower than that of the human NDS or ACC-only events. For 78% of events in the RSS model, TET was concentrated at 0 s, a better ratio than that of the human NDS or ACC-only events. For 78% of events in the RSS model, TET was concentrated at 0 s, a better ratio than that of the human NDS or ACC-only events. As the longitudinal relative distance between the two vehicles increased in the RSS-embedded model, the risk of lateral collision during a cut-in scenario was simultaneously reduced.

In addition to this study’s positive results, RSS, as an algorithm independent safety guarantee model, can be applied to guarantee the safety performance of multiple autonomous driving algorithms. However, neither ACC-only nor RSS-embedded ACC can mitigate all hazards related to emergent cut-in events. Although ACC is currently marketed in the automobile industry as a driving algorithm for low levels (not fully autonomous) of automated vehicles, ACC is fundamentally a longitudinal vehicle control algorithm. The RSS parameters are fixed, so the AV’s lateral movement cannot be controlled in RSS-embedded ACC simulation scenarios. Since the RSS minimum lateral distance between the two vehicles was not triggered in the simulation, the lateral parameters of the RSS model were selected according to the data statistics, requiring recalculation for correction based on RSS lateral safety distance. Although simulation results show that the selected parameter sets were effective and not too conservative for the simulated scenarios, in future studies, more combinations of RSS parameters should be tested and compared for optimization.

A last limitation is that the TTC threshold was set to 3 s in this study in order to select a broad representation of cut-in events. In future research, a greater variety of emergent cut-in scenarios should be simulated.
Acknowledgement

This study was jointly sponsored by the Chinese National Science Foundation (51878498), the Science and Technology Commission of Shanghai Municipality (18DZ120020), and the European Commission H2020 project LEVITATE (Societal level impacts of connected and automated vehicles) under grant number 824361. The authors are grateful to Barbara Rau Kyle for her helpful edit.

References


Shashua, A., S. Shalev-Shwartz, and S.S., n.d. Implementing the RSS Model on NHTSA Pre-Crash Scenarios Implementing the RSS Model on.
