2023-01-0576 Published 11 Apr 2023

Quantitative Resilience Assessment of GPS, IMU, and LiDAR Sensor Fusion for Vehicle Localization Using Resilience Engineering Theory

Johan Fanas Rojas, Parth Kadav, Nicolas Brown, and Rick Meyer Western Michigan University

Thomas Bradley Colorado State University

Zachary Asher Western Michigan University

Citation: Fanas Rojas, J., Kadav, P., Brown, N., Meyer, R. et al., "Quantitative Resilience Assessment of GPS, IMU, and LiDAR Sensor Fusion for Vehicle Localization Using Resilience Engineering Theory," SAE Technical Paper 2023-01-0576, 2023, doi:10.4271/2023-01-0576.

Received: 25 Oct 2022 Revised: 10 Jan 2023 Accepted: 03 Feb 2023

Abstract

Practical applications of recently developed sensor fusion algorithms perform poorly in the real world due to a lack of proper evaluation during development. Existing evaluation metrics do not properly address a wide variety of testing scenarios. This issue can be addressed using proactive performance measurements such as the tools of resilience engineering theory rather than reactive performance measurements such as root mean square error. Resilience engineering is an established discipline for evaluating proactive performance on complex socio-technical systems which has been underutilized for automated vehicle development and evaluation. In this study, we use resilience engineering metrics to assess the performance of a sensor fusion algorithm for vehicle localization. A Kalman Filter is used to fuse GPS, IMU and LiDAR data for vehicle localization in the CARLA simulator. This vehicle localization algorithm was then evaluated using resilience engineering metrics in the simulated multipath and overpass scenario. These scenarios were developed in the CARLA simulator by collecting real-world data in an overpass and multipath scenario using WMU's research vehicle. The absorptive, adaptative, restorative capacities, and the overall resilience of the system was assessed by using the resilience triangle. Simulation results indicate that the vehicle localization pipeline possesses a higher quantitative resilience when encountering overpass scenarios. Nevertheless, the system obtained a higher adaptive capacity when encountering multipath scenarios. These resilience engineering metrics show that the fusion systems recover faster when encountering disturbances due to signal interference in overpasses and that the system is in a disturbed state for a shorter duration in multipath scenarios. Overall these results demonstrate that resilience engineering metrics provide valuable insights regarding complicated systems such as automated vehicle localization. In future work, the insights from resilience engineering can be used to improve the design and thus performance of future localization algorithms.

Introduction

Partially automated and fully autonomous vehicles (AVs) are systems capable of navigating through different driving environments and making decisions with limited or no human input [1]. These systems exist to improve safety but the number of accidents have caused distrust from the society [2,3]. California's Department of Motor Vehicles reported the disengagement and accident reports of AVs from 2014-2018 and it was determined that many of the disengagements were due to other road users (pedestrians, cyclists and vehicles) and that those disengagements were perceived as high-risk events for the driver and passengers [4,5]. According to Wang et al. the number of disengagements varies amongst AV manufacturers and ranges from 2x10^{-4} to 3 disengagements per mile [6]. In general, the cause of failure in the system varies greatly and can occur from behavior prediction, object detection, the decision making of the vehicle, hardware failures or software failures. These accident reports state that in many cases the human had to take over control of the vehicle because the AV was exposed to unexpected objects or scenarios and wasn't capable of reacting to these conditions. The system was unable to track the object or interpret the information received from the sensor(s) [6]. Therefore, can be considered as systems with low operational resilience since they are unable to adapt to different scenarios. These studies and reports indicate the need for better methodologies for decision making, perception algorithms and fault tolerant solutions.

The architecture of an AV is composed of three subsystems (i) the perception subsystem, (ii) the planning subsystem and (iii) the control subsystem [7]. An AV is responsible for collecting data from sensors, perceiving the environment, as well as planning and vehicle control. The perception subsystem is one of the most crucial since a misinterpretation or misdetection of the environment can cause the vehicle to execute a dangerous maneuver. The perception subsystem of an autonomous system collects data through sensors such as radars, LiDAR, cameras, GNSS and IMU to acquire knowledge from the environment. For years, artificial intelligence and methodologies such as state estimation techniques have been used in the perception subsystem to perform object detection, sensor fusion, localization, free space detection and other derivatives [8]. Despite the incredible potential that artificial intelligence has in this field and the effectiveness of state estimation methods and thus sensor fusion, the real world data is subjected to noise and often results in poor detection or false-positive detections [9, 10].

Artificial intelligence based sensor fusion has been the subject of active research for many years and there are several techniques to evaluate its performance [11]. However, most of these metrics are developed in a specific application domain and often lack implementation in a more general practical domain. This means that there is a gap between the performance measures used in testing environments and the ones used in real world environments [12]. Sensor fusion metrics can be divided into fusion data with ground truth and without ground truth. Metrics that use ground truth are for example, root mean square error (RMSE), accuracy, association, detection performance, etc [12]. While metrics without ground truth can be fusion break rate, delay, etc. Some performance measures may be influenced by the user's subjectivity and the testing domain, which may vary depending on the application domain [11, 13]. The development of a metric subjected to all possible scenarios is very complex. The metrics for fusion systems depend on the user's needs and the current metrics are effective but do not operate as expected in the practical domain as evidenced by the AV disengagement reports discussed previously.

Resilience engineering (RE) is an emerging topic of systems engineering which is related to safety management. RE can improve AV's operation and development due to its ability to overcome limitations of existing safety management strategies. RE designs towards proactive safety rather than the current standard of reactive safety [14]. In general, RE provides techniques for improving the operational resilience of complicated systems. From a systemic view, resilience is the intrinsic ability of a system to adjust its functioning prior to, during, or following changes and disturbances to the system while sustaining required operations under both expected and unexpected conditions. Aspects that compose RE include soft redundancy, functional diversity, and control at instability. The RE community widely recognizes AVs as an exciting application, but the only studies currently existing are too limited in scope; they model AV effects on traffic rather than the AV system's performance [15, 16]. This is because the engineering required to develop all driving aspects for an AV is highly interdisciplinary, specialized, and a new skill set within the automotive community.

Despite the benefits that RE possesses to improve the operational resilience of a system compared to traditional methods, it has not been applied to improve the operational resilience of the perception subsystem of an autonomous system. Other papers such as [17] state that they used resilience assessment metrics; however, many of these metrics are not consistent with the four cornerstones of RE (responding, monitoring, anticipating, and learning) [18]. In addition, these metrics indicate the system's ability to complete a given task but do not provide information about the system's overall performance. This is problematic because the RE literature states that a system may fail internally and could still achieve the desired goal [19]. Therefore, these metrics are reactive instead of proactive, which goes against the key concepts of RE. In a previous study, our research group used RE metrics to assess the resilience of an AV's control subsystem [20] when performing path tracking. The application of RE for the operational improvement and assessment of an AV's perception subsystem has not been performed and it is imperative to achieve eventual resilient performance of the entire AV system. Thus there is a research gap between RE and assessment methods for perceptions algorithms.

This paper addresses this research gap by applying RE to evaluate the performance of a sensor fusion algorithm for AV localization. Vehicle localization was performed by fusing GPS, IMU and LiDAR data using a Kalman Filter (KF) in the CARLA simulator. The performance of the KF was assessed in an overpass, and a multipath scenario. These scenarios were developed by collecting real-world data using a fully instrumented 2019 Kia Niro [21] to analyze the effect on the GPS receiver. The RE metrics of absorptive, adaptive, and restorative capacities as well as the overall resilience are presented and applied to the fusion pipeline by applying the resilience triangle. The metrics presented in this study provide the user with insights about the overall performance of the system rather than an outcome assessment like traditional performance measures. Overall this is a novel contribution to both RE and AV research since RE evaluation of AV sensor fusion has not been previously addressed to our knowledge.

Methodology

The quantitative resilience assessment will be performed by fusing GPS, IMU and LiDAR data for an overpass and multipath scenario and evaluating its performance using RE metrics in CARLA. In this section we will present a brief overview of localization and sensor fusion. Next, the overall methodology of RE will be presented and the associated metrics chosen will be described. Lastly, the development of real-world representative scenarios in the CARLA simulator are presented.

Localization

Localization is the methodology within the perception subsystem to precisely estimate the position of the ego vehicle in the world [9, 22]. However, GPS sensors suffer from several drawbacks such as multipath and signal interference effects. Multipath errors occur when a vehicle traverses an area with towering buildings or other tall infrastructure with reflective surfaces, and the signal path between the satellite and the receiver is not direct. Instead, the signal bounces off reflective surfaces before reaching the receiver. This increases the accuracy of the receiver since the duration of the signal's propagation time increases. The signal interference effect occurs when the vehicle traverses areas that cause interference between the satellite and the receiver. These circumstances may involve tunnels, overpasses, trees, etc. The receiver determines the distance by measuring the signal's propagation time. Therefore, if the receiver has some interference, it cannot effectively compute the distance.

In autonomous driving (AD), localization is important because the estimation of the vehicle's position relies on GPS data. An increase in the positioning error of GPS sensors decreases the performance of the system and can lead to dangerous outcomes. Studies in the literature state that to achieve safe maneuvering, an accuracy of 10 cm is required [23], while other studies state that with less than 30 cm of accuracy, safe AV driving can be achieved [24]. This improvement in accuracy can be achieved using sensor fusion but the system is still greatly affected whenever the GPS signal is lost. To demonstrate the performance of a localization algorithm affected by poor GPS signal and multipath effects, we use RE metrics to measure the system's operational resilience. Next we will discuss sensor fusion and the concept of the KF used in this study.

Sensor Fusion

Sensor fusion is used to improve measurements of two or more sensors, since individually there can be significant noise. There are several methods that have been used to fuse data from two or more sensors. These methods are: (1) estimation methods based on Gaussian filters (e.g, KF or particle filters (PF)), (2) probabilistic inference methods (Bayes theorem), and (3) artificial intelligence methods based on machine learning algorithms [22]. These methods have been used over the years to accurately localize the ego vehicle for AD. In this research, we utilize a KF to perform vehicle localization. The concept of KF will be introduced and discussed next.

Kalman Filter

KF is an algorithm used in various fields to estimate the unknown state of the system. Several applications such as navigation, localization and object tracking have been developed with the implementation of a KF [25].

The traditional KF is used with linear systems and discrete processes. The KF possesses two phases, the prediction and the update phase. In the prediction phase, the KF uses the estimated state of the previous timestep to predict the current state. This predicted state is called the *a priori* state since it has no information about the current state. The equations for the prediction phase are the following:

$$\hat{x}_{k|k-1} = F_k x_{k-1|k-1} + B_k u_k \tag{1}$$

$$P_{k|k-1} = F_k P_{k-1|k-1} F_k^T + Q_k \tag{2}$$

Where $\hat{x}_k$ is the predicted state estimate, x_k is the best state estimate, B_k is the control input matrix, F_k is the matrix that possesses the dynamics of the system, P_k is the predicted covariance matrix, Q_k is the process noise covariance matrix and u_k is the input.

In the update phase, a new measurement is received from a sensor and used to compute the error between the current measurement and the *a priori* state (this is also called residual measurement). This residual measurement is then multiplied by the optimal Kalman gain to correct the previous state. This corrected state is also called the *posterior* state. The equations for the update phase are the following:

$$S_k = H_k P_{k|k-1} H_k^T + R_k \tag{3}$$

$$K_k = P_{k|k-1} H_k^T S_k^{-1} \tag{4}$$

$$x_{k|k} = \hat{x}_{k|k-1} + K_k \left(z_k - H_k \hat{x}_{k|k-1} \right) \tag{5}$$

$$P_{k|k} = \left(I - K_k H_k \right) P_{k|k-1} \tag{6}$$

Where S_k is the measurement residual covariance matrix, H_k is the measurement model matrix, R_k is the measurement noise covariance matrix, z_k is the measurement vector and K_k is the Kalman gain.

The KF is used for linear systems; nevertheless, most real-world applications are nonlinear. Therefore, further developments of the KF are the EKF and Unscented Kalman Filter (UKF), which both operate with nonlinear systems.

In this study a KF was used to fuse GPS, IMU and LiDAR odometry in order to produce a better estimate of the vehicle's position. Although EKF and UKF are more robust, a KF is sufficient for providing insight into the impact of using RE metrics with sensor fusion algorithms. Also, due to the linear nature of the KF, the results are easier to interpret compared to other methods for this first application of RE. The state vector of the KF used in this study is $x = [p_x, p_y, p_z, V_x, V_y, V_z, \phi, \theta, \psi]$. Where p_x, p_y and p_z are the position of the vehicle in the x, y and z axis. V_x, V_y and V_z are the velocity on the x, y and z axis. Lastly, ϕ, θ, ψ are the roll, pitch and yaw of the vehicle. The states of the KF are in Cartesian coordinates in the global coordinate system. The motion model used to perform the fusion is the following:

$$p_{k+1} = p_k + v_k \cdot \Delta t + \frac{1}{2} a_k \cdot \Delta t^2 \tag{7}$$

$$v_{k+1} = v_k + a_k \cdot \Delta t \tag{8}$$

$$\phi_{k+1} = \phi_k + \omega_k^x \cdot \Delta t \tag{9}$$

$$\theta_{k+1} = \theta_k + \omega_k^y \cdot \Delta t \tag{10}$$

$$\psi_{k+1} = \psi_k + \omega_k^z \cdot \Delta t \tag{11}$$

where a and ω are the linear acceleration and angular velocity. The linear acceleration and angular velocity are measured from the IMU sensor and are inputs to the motion model. The motion model represents the state space of the system and the control inputs represent the action space used to propagate the system's dynamics. To obtain LiDAR odometry, LiDAR and IMU data were fused using FAST-LIO (Fast LiDAR-Inertial Odometry) [26, 27]. Much research has been done on vehicle localization based on GPS/IMU/LiDAR fusion; therefore, specifics on this will not be covered in this study [28-30].

Resilience Engineering Overview

RE is a paradigm for safety management that focuses on complex socio-technical systems [31]. In the past, safe systems have been engineered by evaluating events that have already happened [14]. RE improves safety systems by considering events that have yet to happen thus engineering safety that is proactive. Additionally it is important to define what is meant by safety. Safety I focuses on preventing adverse conditions that might disrupt the system, but sometimes these conditions come with necessary adjustments, the system needs to enhance its performance. That is when Safety II arises. RE is considered a methodology under Safety II, which focuses on strengthening the system's ability to succeed during varying conditions [32].

The concept of RE has been adopted in safety management as a new tool for proactive safety. Resilience refers to managing unexpected changes and succeeding when the organization (or system) is under pressure [33]. The four cornerstones representing the resilience of a system are: responding, monitoring, anticipating, and learning [18]. These abilities are are defined as follows:

- **Ability to respond**: A system must be able to notice when a change has occurred and address it quickly.
- **Ability to monitor**: A system should be able to monitor its own performance and perceive changes outside the system that can potentially become an opportunity or a threat.
- **Ability to learn**: This ability provides the system with the means to learn from events that have caused accidents and others that could have potentially disrupted the system.
- **Ability to anticipate**: The purpose of this ability is to anticipate potential threats and opportunities for the system.

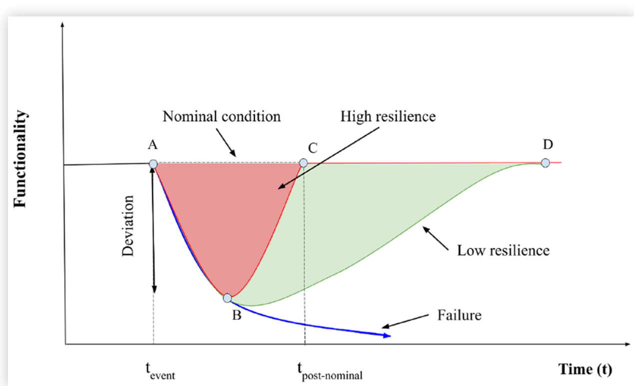
However, these four abilities can be summarized in the three capacities a resilient system must have [34]. These capacities are:

- **Absorptive capacity**: The absorptive capacity of a system refers to how well it can lessen the negative impacts of disruptions while maintaining better residual performance.
- **Restorative capacity**: Is the system's capacity to be quickly restored with minimal intervention and to reach higher levels of performance than in a disrupted condition.
- **Adaptive capacity**: Is the system's response to the stressor's effect, its ability to operate in a stable disturbed condition, and its preparedness to begin the recovery process following a failure event.

A resilient system must monitor the system's state and change its boundaries when drifting towards unsafe actions. Managing the decision-making process when having the goals and priorities well defined are key for a resilient system [35, 36]. RE improves a system's performance by correlating performance and safety instead of treating them as mutually exclusive. Even though RE has the potential to mitigate risks by improving the system's performance, safety is more complex than just a methodology to improve safety. RE has been applied as a safety management tool in several domains and successfully improved their performance. These domains include aviation, healthcare, petrochemical plants, manufacturing, railways, and construction [33]. With the use of RE, studies in the aviation industry have shown a reduction of approximately 40% of life cycle cost for aircraft control actuators [37]. Furthermore, RE has aided some healthcare systems to identify areas of weaknesses within their organization [38].

RE defines resilience in terms of how the system performs, not an attribute that a system possesses. Therefore, a system cannot be resilient but it can have the potential for resilient performance. Figure 1 shows the performance of a resilient system. In this Figure we see how a resilient system after failure (point A), is able to quickly return to its nominal condition and maintain its functionality or reliability. While a system with low resilience is more affected by the disturbance and takes longer to return to its nominal condition. The points A-B-C and A-B-D form a triangle called the resilience triangle and it is used to measure the resilience of the system to disturbances or events. This concept is used to present a series of metrics to evaluate the ability of the system to react to

FIGURE 1 General resilience triangle for a high and low resilience performance system and a system subjected to failure



disturbances. In this study, the resilience triangle is applied to evaluate a sensor fusion algorithm but the concept can be extended to other subsystems or applications. In the next section we will break down the analysis of the resilience triangle.

Evaluation Metrics from Resilience Engineering Principles

To evaluate the fusion algorithms and the performance of the system when exposed to adverse events (or failure), we will use the resilience triangle. The resilience triangle consists of measuring the performance of the system in the three capacities that a resilient system must have in the face of external disturbances [39]. To analyze the resilience triangle presented in the previous section, we present the phases of the resilience triangle in Figure 2 [34].

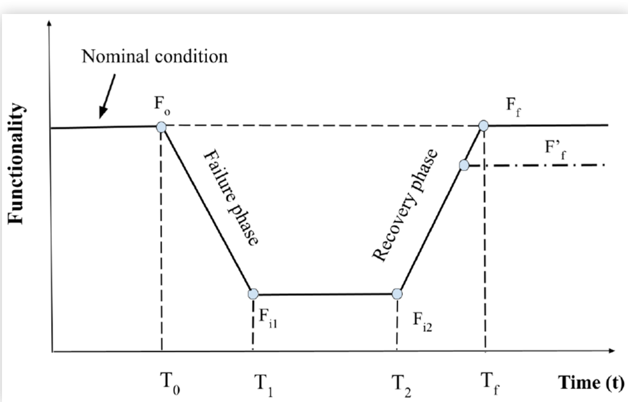
The resilience triangle has all three capacities described in the previous section. The absorptive capacity is the ability of the system to reduce the effects of failure and maintain high functionality. This is the fault phase represented between points F_o and F_{i1}. The absorptive capacity can be determined by the following equations:

$$Ab = \left| \frac{F_o}{F_{i1}} \cdot C_{ab} \right| \tag{12}$$

$$C_{ab} = 1 - \left(\frac{F_o}{F_0 - F_{i1}} \right) \tag{12a}$$

where F_o is the initial functionality before the failure and F_{i1} is the residual functionality of the system at time T_1. The capacity value C_{ab}, is used to compensate for the aging impact and associated functionality loss during the failure phase of the system. In other words, when the system fails, it does not always return to its original nominal state. Consequently, as the system experiences more and more failures, it degrades over time, a phenomenon known as the aging factor. Therefore, the capacity value C_{ab} compensates for this degradation effect.

FIGURE 2 Trapezoid describing the phases of the resilience triangle



The restorative capacity is the ability of the system to restore itself after failure quickly. In other words, it is the inverse of the difference between T_f and T_{i2}. If a system recovers quickly the time difference will approach zero; hence, achieve a higher score in the restorative capacity. Therefore, the restorative capacity can be computer by the following equation:

$$Res = \frac{1}{T_f - T_{i2}} \cdot C_T \cdot C_R \tag{13}$$

$$C_T = \frac{T_{i2} - T_o}{T_f - T_o} \tag{13a}$$

$$C_R = \frac{F_f}{F_f'} \tag{13b}$$

where T_f is the end time after recovery, T_2 is the time before the system starts to recover, T_o is the time the failure occurs, F'_f is the functionality post-recovery with a degradation due to aging, and F_{i2} is residual functionality before recovery. C_R is the ratio between the functionality obtained after recovery and the ideal post-recovery functionality. This is because a system does not always return to its previous state after a failure. We use C_R as a one in our analysis since we assume that the system will return to our nominal condition. The C_T factor is used for long recovery periods and prevents the duration from going to infinity.

The system's adaptive capacity is the ratio of operating duration from the initial functionality F_o to the new functionality F_f post-recovery. It can be computed using the following equation:

$$Ad = 1 - \frac{T_{i2} - T_{i1}}{T_f - T_o} \tag{14}$$

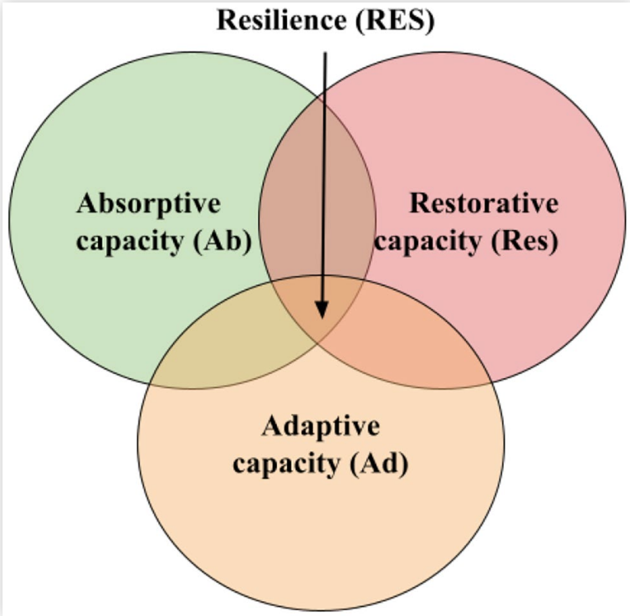
Finally the overall resilience of the system (RES) can be obtained by incorporating the system's capacities. The relationship between the capacities and resilience can be represented using the Venn Diagram shown in Figure 3. The overall resilience of the system is the overlap between the three capacities of the system. In other words, a system with resilient operational performance must possess all three of these capacities.

According to Yarveisy et al. [34] the relationship of the system's overall resilience can be represented by the following mathematical expression.

$$RES = Ab + Ad \cdot Res - Ab \cdot Ad \cdot Res \tag{15}$$

Yarveisy et al. [34] states that the system's absorptive capacity is the sole independent resilience trait that can be observed. Unlike restorative and adaptive capacities, the absorptive capacity is not reactive. It is the system's inherent ability to tolerate the negative effects of destructive events, and it only depends on the system's physical state [34]. Higher absorptive capacities mean that other capacities have less of an impact.

FIGURE 3 Venn diagram representing the logical relation between the three capacities of a resilient system



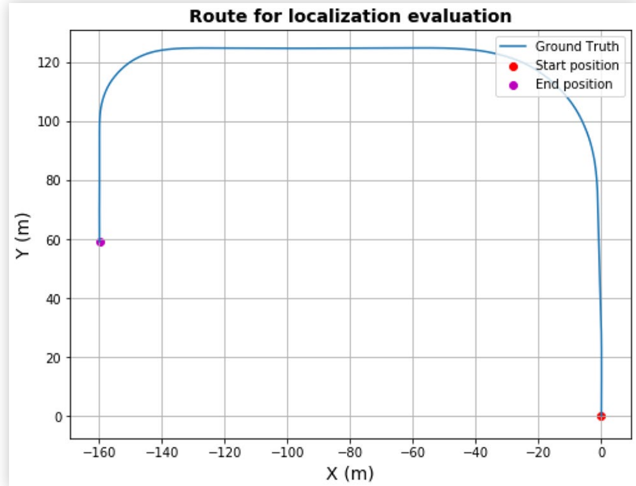
The goal of this study is to evaluate the performance of a KF in an overpass and multipath scenario. Therefore, the nominal condition used is a baseline scenario (which will be described in the following section) where the vehicle does not encounter any disturbances from the environment. Any deviation from this nominal condition will be considered as a performance loss and the resilience triangle will be applied to assess the resilience of the system. The performance loss of the overpass and multipath scenario can be seen as the error of the fused state with respect to the state of the baseline scenario (nominal state). In other words, deviations from this nominal state that exceed a user defined threshold (in this case 10 cm), is considered as a performance loss and the resilience triangle will be applied to showcase the performance of the system under unforeseen events. In the next section we will describe the simulation setup.

Simulation Setup

The CARLA simulator was used to test and evaluate the localization algorithm of an AV. CARLA is an open source simulator developed for training, validating, and testing algorithms for AV systems. It provides a simulation environment with a wide range of sensor specifications, environment conditions, vehicles, etc. It allows the user the ability to create customizable environments for testing and validating autonomous/ADAS driving behaviors. Scenarios like computer vision testing in adverse weather conditions or path planning while sharing the road with aggressive road users can be simulated [40-42]. Performance evaluation in simulation is an important stage before deploying AVs to real-world scenarios.

To perform vehicle localization in the CARLA simulator, a vehicle equipped with a GPS, IMU and LiDAR was used to perform sensor fusion. The GPS, IMU and LiDAR sensors were configured to operate at 10, 100 and 10 Hz, respectively.

FIGURE 4 Route for localization assessment

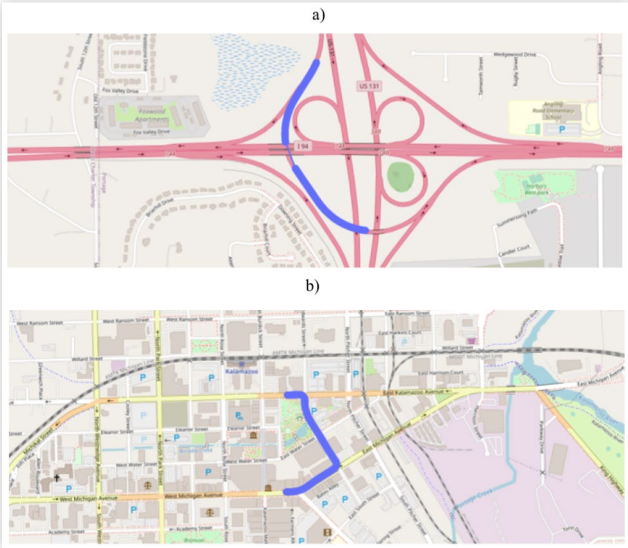


The vehicle was manually driven on the route shown in Figure 4 in the Town 10 of the CARLA simulator in order to collect and store sensor data. After the data was stored, a Gaussian noise was added to the IMU data with a standard deviation of 0.05 rad/s for the gyroscope and 0.2 m/s^2 for the accelerometer. To avoid transformations between the GPS, LiDAR and IMU frames, the sensors were spawned in the same position with respect to the vehicle. FAST-LIO [26,27] was used to perform SLAM and obtain LiDAR odometry for the fusion pipeline.

To evaluate the localization pipeline using the resilience metrics, three test scenarios were developed. Real-world data was collected using a fully instrumented 2019 Kia Niro in order to develop real-world representative scenarios in the CARLA simulator. Using the real world data as a guide, the scenarios were then replicated in the simulator. These scenarios are the following:

1. A baseline scenario where the vehicle uses GPS + Real Time Kinematics (RTK) and does not encounter deviations due to environmental disturbances. For this scenario we observed that the GPS + RTK configuration of our fully instrumented research vehicle in areas with no disturbances, operated under a 10 cm accuracy. Therefore, the GPS data is subjected to a Gaussian noise with a standard deviation of 10 cm throughout the whole route. This scenario will serve as the nominal condition for our resilience analysis.
2. A baseline scenario + signal interference, which will be called an overpass scenario. Our research vehicle was used to collect GPS data while traversing under an overpass in the I-94 highway in Portage, Michigan, in order to observe the effect on the GPS receiver. Figure 5a depicts the data collected as the vehicle traverses an overpass of approximately 50 meters.
3. A baseline scenario + multipath effect, which will be called multipath scenario. Similar to the previous scenario, WMU's research vehicle was used to collect data in downtown Kalamazoo, Michigan, to observe the effect of multipath errors on GPS receivers. Figure 5b depicts the path taken to analyze this effect.

FIGURE 5 GPS data collection with WMU's research vehicle for the overpass scenario (a) and the multipath scenario (b).



The collected data for both scenarios was analyzed and plotted. A colorbar was utilized to illustrate how the accuracy varies when the receiver encounters the aforementioned effects. As shown in Figure 6, there is a gap in the traveled path in the overpass scenario due to signal interference. Similarly, in the multipath scenario, the accuracy of the receiver decreases from 0.12 meters to 1.2 meters when passing through radio wave reflective surfaces on tall buildings.

FIGURE 6 Analysis of overpass and multipath scenario

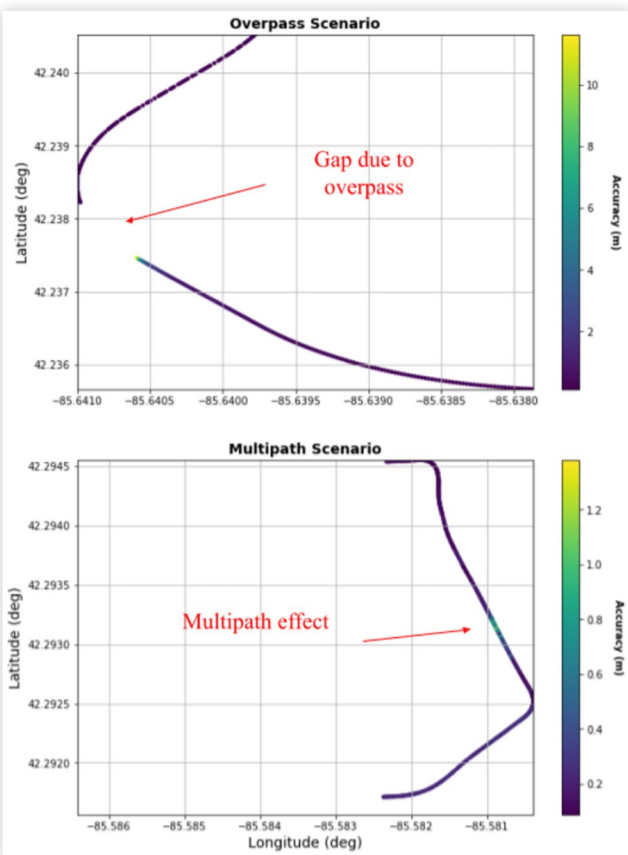
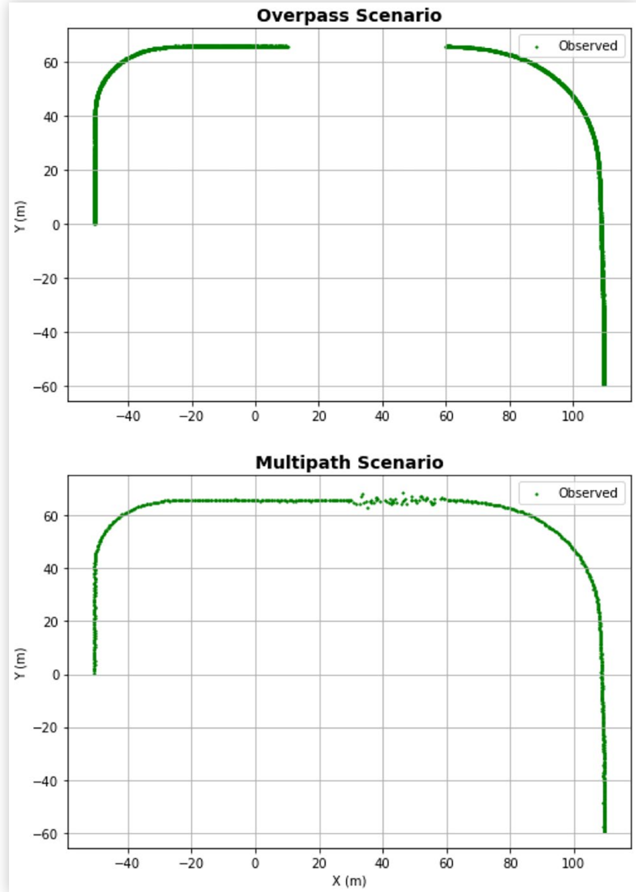


FIGURE 7 Overpass and multipath scenario in the CARLA simulator



To recreate these scenarios in the CARLA simulator, a 50 meter gap is created in the route shown in Figure 4. We assume in this gap, no GPS information is received and the fusion pipeline only relies on LiDAR and IMU. The created overpass scenario is shown in Figure 7. For the multipath scenario, we add a Gaussian noise with zero mean and standard deviation of 1.2 meters for 30 meters in the region shown in Figure 7. Now that the scenarios are generated, we can move forward with the sensor fusion implementation and the RE evaluation.

The RE metrics shown in the previous section were used to assess the performance of the system under the developed test scenarios.

Results

To compare conventional evaluation methods versus RE metrics, we used root mean square error (RMSE) to assess the estimated vehicle position against the ground truth. The evaluation of the KF using RMSE is shown in Figure 8.

In Figure 8 we observe that in the overpass scenario where there is no GPS signal and the fusion process relies on LiDAR odometry and IMU data, the system is capable of achieving

FIGURE 8 RMSE of the state estimate versus ground truth

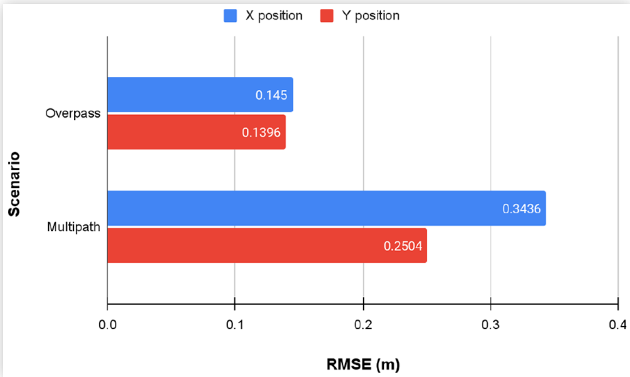


FIGURE 9 Savitzky-Golay smoothing filter applied to fused error measurements

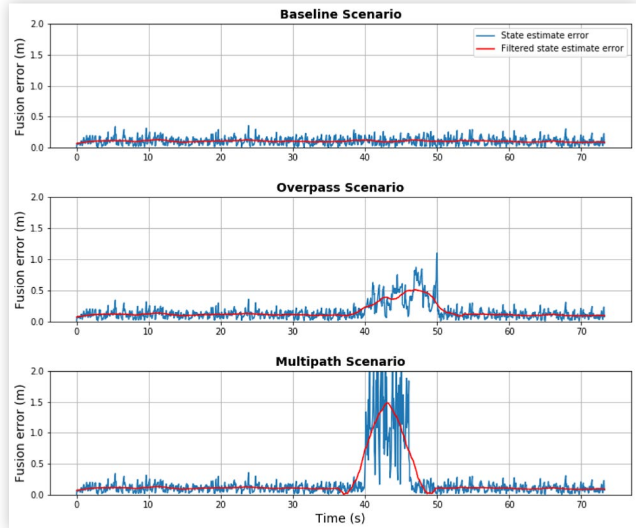
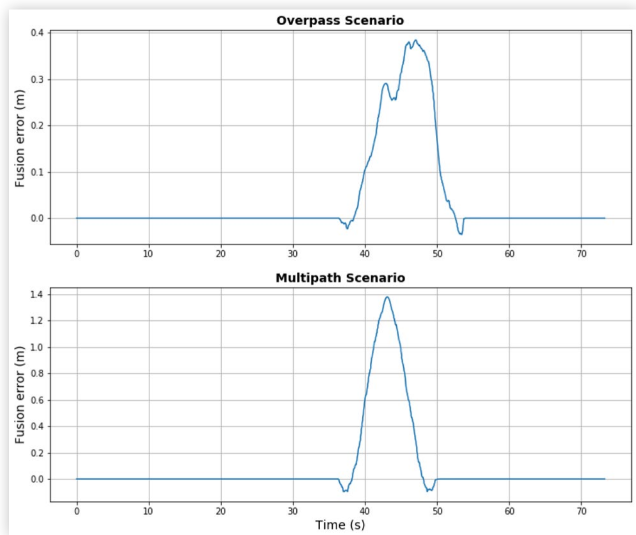


FIGURE 10 Isolated failure of the system for the overpass and multipath scenario



RMSE values of 0.145 and 0.1396 m in the x and y axis, respectively. This is due to LiDAR odometry being more precise than wheel odometry, even though it is also subjected to low drift over time [43]. On the other hand, we observe that the multipath scenario obtains a higher RMSE in both axes compared to the overpass scenario. This indicates that the performance of the multipath scenario degrades in the presence of increased accuracy due to reflective surface effects. Using RMSE as an evaluation metric demonstrates that the system performs better in overpass scenarios despite the absence of GPS measurements due to signal interference.

This metric is effective, but it provides no information other than the error relative to the ground truth. This is a problem because it is often difficult to obtain the ground truth in the real world since relied upon sensors are susceptible to environmental noise and conditions. Furthermore, it does not provide the user information regarding the performance of the system at the time of failure. The results are also not necessarily extensible to other similar test scenarios. Therefore, this metric may display false positives, biased information, or different system performance outcomes in alternative environments.

As a contrast, to perform the RE analysis and apply the resilience triangle, we plot the estimated state error versus the ground truth using the Euclidean distance equation. However, the behavior is nonlinear; therefore, we apply a Savitzky-Golay filter [44] to smooth the error between measurements. The error between both states and the smoothened output is shown in Figure 9.

To isolate the performance loss of the system, the baseline scenario is subtracted from the other scenarios. Figure 10 depicts the system's performance loss behavior when subjected to the corresponding scenario.

By identifying the four points of the trapezoid depicted in Figure 2 and fitting a line between each pair of points, we can obtain the three resilience phases of the system. The point of intersection between each pair of lines is then determined to form the trapezoid describing the resilience phases of the system. Figure 11 depicts the three lines that were used to form the trapezoid's three regions and the intersection points.

The trapezoidal vertices were used to determine the resilience triangle's parameters and compute the capacities and the overall resilience of the system. Table 1 displays the parameters obtained for each scenario from each estimated trapezoid shown in Figure 11. In Table 1, we see that the F_o and F_f values are 0.10 m. This is because the baseline scenario operated under 10 cm. Observing the behavior of the baseline scenario reveals that it fluctuates around 10 cm, but when subtracted from the other scenarios, it approaches zero. Therefore, our actual nominal condition is 10 cm. In other instances, the user may define this nominal condition based on the system's operation requirements.

Using the parameters shown in Table 1, the equations 12-15 are used to compute the system's capacities and overall resilience. Figure 12 depicts the results of our resilience evaluation.

Figure 12 shows that the system achieved a higher restorative capacity in the overpass scenario than in the multipath scenario. The restorative capacity indicates that when the system is exposed to overpass scenarios, it returns quicker to

FIGURE 11 Trapezoid regions of the resilience triangle for a) overpass scenario b) multipath scenario

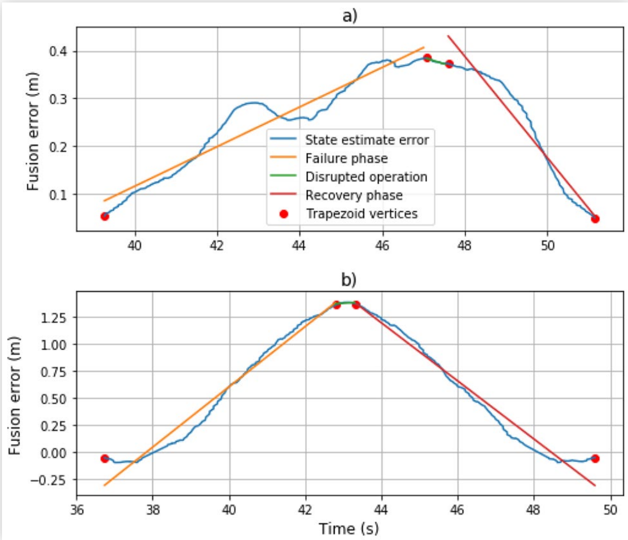
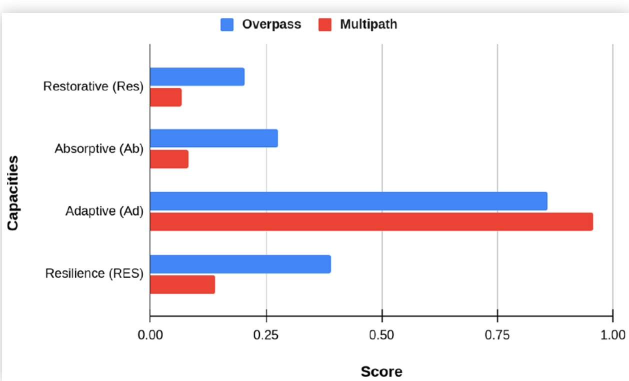


TABLE 1 Resilience parameters from each corresponding estimated trapezoid

Scenario/Params	Overpass scenario	Multipath scenario
F_o (m)	0.10	0.10
F_{i1} (m)	0.393	1.366
F_{i2} (m)	0.35	1.379
F_f (m)	0.10	0.10
T_o (s)	39.27	36.74
T_{i1} (s)	46.67	42.76
T_{i2} (s)	48.36	43.30
T_f (s)	51.15	49.60

FIGURE 12 Capacities and overall resilience of the system for the overpass and multipath scenario



its normal condition than when it is exposed to multipath effects caused by tall buildings and reflective surfaces. In other words, when an AV exits an overpass, it does not take long to receive a GPS measurement with RTK corrections and correct the state estimate of the vehicle's position. In contrast, when an AV navigates through tall buildings, the GPS signal with RTK corrections takes longer to receive because the accuracy of these measurements require clear skies. However, overpass or GPS interference scenarios are typically under clear sky conditions.

In addition, the overpass scenario scored higher in the absorptive capacity compared to the multipath scenario. This indicates that the system reduces environmental effects in overpass scenarios more effectively. This is due to the fact that in the absence of GPS measurements, the system relies more on LiDAR and IMU measurements and uses the previous state estimate (which operated under the nominal conditions) to obtain predicted state estimates until a new GPS measurement is obtained. However, in multipath scenarios, the system receives noisier GPS measurement; hence, if the measurement and process covariance matrices are not changed under these conditions, it negatively affects the performance of the localization pipeline. These covariance matrices are frequently challenging to tune [45], and if your system lacks redundant sensors or an algorithm for adaptive adjustment of noise covariance matrix [46], it will deviate significantly from the nominal condition (even in the presence of reliable LiDAR and IMU measurements).

The system obtained a higher adaptive capacity score in the multipath scenario than in the overpass scenario. This means that the system spends less time in a disturbed state (from F_{i1} to F_{i2}). This is consistent with Figure 11, as the system is subject to this state of disruption for brief durations. In the multipath scenario, the fusion error shown in Figure 11 has only one concave point, whereas the overpass scenario has many. Therefore, as soon as the system departs from its nominal condition and enters a disturbed state, it attempts to return to its nominal state. This makes sense because, in multipath scenarios, the system still receives GPS measurements (although they are not as precise), so the KF attempts to reduce the state estimate error. Therefore, when an AV traverses areas with tall buildings, the GPS may or may not receive accurate measurements (because the sensor is affected by multipath effects). This causes the system to respond more quickly and require less time in a disturbed state. As the system does not receive GPS measurements in overpass scenarios, it is slower to react and in a disturbed state.

Finally, we can observe that the system achieved a higher overall resilience score for the overpass scenario is approximately 0.40, whereas for the multipath scenario is 0.16. The outcomes of our resilience evaluation indicate that AVs have greater operational resilience in overpass scenarios. Therefore, when an AV encounters unanticipated GPS interference conditions, the system achieves a higher operational performance and mitigates the environmental effects more effectively. In addition, simulation results indicate that LiDAR odometry should be utilized in overpass scenarios to reduce GPS interference due to its more accurate odometry predictions and low drift.

Conclusions

Despite the advances in AV perception tasks such as the free space detection, object detection and localization, accidents have occurred in the past where perception algorithms are not able to recognize certain objects and the vehicle makes the unsafe decisions. Perception algorithms use performance measures that are subjected to the testing or application domain, but they lack performance in the practical domain. This study fills that research gap by proposing RE metrics that assess a system's resilient performance. RE is a new methodology that consists of proactive safety methods to improve the operational resilience of a system. In this study, we evaluated a KF that fuses GPS, IMU and LiDAR data to perform vehicle localization in the CARLA simulator using RE metrics and RMSE. This fusion pipeline was assessed in an overpass and multipath scenario. The overpass and multipath scenarios were created by collecting data with WMU's research vehicle, analyzing their effect on the GPS receiver and replicating it in the CARLA simulator. The overpass and multipath scenarios were evaluated by computing the absorptive, adaptive, restorative capacities and the overall resilience of the system using the resilience triangle. Simulation results demonstrate that the overpass scenario outperformed the multipath scenario by achieving a lower RMSE along the x and y axes. Nevertheless, this metric provides no information about the system at the time of the failure beyond the error relative to the ground truth. On the other hand, the RE metrics provide us with information regarding the system's behavior when it is disturbed and indicate that the overpass scenario possessed a greater absorptive and restorative capacity. This means that when tunnels or overpasses are present, the system can mitigate the effects of a failure and recover more quickly. However, the system achieved greater adaptability in the multipath scenario, resulting in a shorter failure duration. The system obtained a higher overall resilience in the overpass scenario meaning that AVs have greater operational resilience in GPS interference circumstances.

RE metrics possess the tools to identify where and how complicated systems fail to meet performance requirements. By exploiting areas of weakness in a system, the user can apply countermeasures in order to account for more practical scenarios traditional metrics overlook. It is imperative that research and development of AV systems use RE metrics for evaluation to ensure resilience in real world performance. For future work, we will expand this study by performing the resilience assessment on real-world data. Specifically, the vehicle localization assessment pipeline will be applied to the shown areas in Kalamazoo, MI. In addition, a correlation between the two studies will be presented to provide insight on the RE metrics evaluating operational performance between testing and the practical domain.

References

1. Brodsky, J.S., "How an Uncertain Legal Landscape May Hit the Brakes On Self-Driving Cars," *Berkeley Technol. Law J.* 31, no. 2 (2016): 851-878.
2. Collingwood, L., "Privacy Implications and Liability Issues of Autonomous Vehicles," *Null* 26, no. 1 (2017): 32-45.
3. Taeihagh, A. and Lim, H.S.M., "Governing Autonomous Vehicles: Emerging Responses for Safety, Liability, Privacy, Cybersecurity, and Industry Risks," *Null* 39, no. 1 (2019): 103-128.
4. Lv, C., Cao, D., Zhao, Y., Auger, D.J. et al., "Analysis of Autopilot Disengagements Occurring During Autonomous Vehicle Testing," *IEEE/CAA Journal of Automatica Sinica* 5, no. 1 (2018): 58-68.
5. Dixit, V.V., Chand, S., and Nair, D.J., "Autonomous Vehicles: Disengagements, Accidents and Reaction Times," *PLoS One* 11, no. 12 (2016): e0168054.
6. Wang, J., Zhang, L., Huang, Y., and Zhao, J., "Safety of Autonomous Vehicles," *Journal of Advanced Transportation* 2020, doi:10.1155/2020/8867757.
7. Schwarting, W., Alonso-Mora, J., and Rus, D., "Planning and Decision-Making for Autonomous Vehicles," *Annu. Rev. Control Robot. Auton. Syst.* 1, no. 1 (2018): 187-210.
8. Park, Y., Dang, L.M., Lee, S., Han, D. et al., "Multiple Object Tracking in Deep Learning Approaches: A Survey," *Electronics* 10, no. 19 (2021): 2406.
9. Van Brummelen, J., O'Brien, M., Gruyer, D., and Najjaran, H., "Autonomous Vehicle Perception: The Technology of Today and Tomorrow," *Transp. Res. Part C: Emerg. Technol.* 89 (2018): 384-406.
10. Vialatte, J.-C. and Leduc-Primeau, F., "A Study of Deep Learning Robustness Against Computation Failures," *arXiv [cs.NE]* (2017).
11. Blasch, E.P., Pribilski, M., Daughtery, B., Roscoe, B. et al., "Fusion Metrics for Dynamic Situation Analysis," *Signal Processing, Sensor Fusion, and Target Recognition XIII*, SPIE (2004): 428-438.
12. García, J., Molina, J.M., and Trincado, J., "Real Evaluation for Designing Sensor Fusion in UAV Platforms," *Inf. Fusion* 63 (2020): 136-152.
13. Llinas, J. and Hall, D.L., "An Introduction to Multi-Sensor Data Fusion," *ISCAS '98. Proceedings of the 1998 IEEE International Symposium on Circuits and Systems (Cat. No.98CH36187)* 6 (1998): 537-540.
14. Hollnagel, E., Woods, D.D., and Leveson, N., *Resilience Engineering: Concepts and Precepts* (Ashgate Publishing, Ltd., 2006), ISBN:9780754681366
15. Madni, A.M., Sievers, M.W., Humann, J., Ordoukhanian, E., D'Ambrosio, J., and Sundaram, P., "Model-Based Approach for Engineering Resilient System-of-Systems: Application to Autonomous Vehicle Networks," *Disciplinary Convergence in Systems Engineering Research*, Springer International Publishing 365-380, 2018.
16. Marshall, C., Roberts, B., and Grenn, M., "Intelligent Control & Supervision for Autonomous System Resilience in Uncertain Worlds," in *2017 3rd International Conference on Control, Automation and Robotics (ICCAR)*, 438-443, 2017.
17. Jha, S., Banerjee, S.S., Cyriac, J., Kalbarczyk, Z.T., and Iyer, R.K., "AVFI: Fault Injection for Autonomous Vehicles," in *2018 48th Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshops (DSN-W)*, 55-56, 2018.
18. Hollnagel, E., "The Four Cornerstones of Resilience Engineering," *Resilience Engineering Perspectives, Volume 2*, CRC Press: 139–156, 2016.

19. Hollnagel, E. and et al., "Prologue: The Scope of Resilience Engineering," Resilience Engineering in Practice: A Guidebook xxix-xxxix, 2011.
20. Fanas Rojas, J., Brown, N., Rupp, J., Bradley, T. et al., "Performance Evaluation of an Autonomous Vehicle Using Resilience Engineering," SAE Technical Paper 2022-01-0067 (2022), https://doi.org/10.4271/2022-01-0067.
21. Brown, N.E., Rojas, J.F., Goberville, N.A., Alzubi, H. et al., "Development of an Energy Efficient and Cost Effective Autonomous Vehicle Research Platform," *Sensors* 22, no. 16 (2022).
22. Rosique, F., Navarro, P.J., Fernández, C., and Padilla, A., "A Systematic Review of Perception System and Simulators for Autonomous Vehicles Research," *Sensors* 19(3), 2019, doi:10.3390/s19030648.
23. Reid, T.G.R., Houts, S.E., Cammarata, R., Mills, G. et al., "Localization Requirements for Autonomous Vehicles," *arXiv [cs.RO]* (2019).
24. Lu, Y., Ma, H., Smart, E., and Yu, H., "Real-Time Performance-Focused Localization Techniques for Autonomous Vehicle: A Review," *IEEE Trans. Intell. Transp. Syst.* (2021): 1-19.
25. Alatise, M.B. and Hancke, G.P., "A Review on Challenges of Autonomous Mobile Robot and Sensor Fusion Methods," *IEEE Access* 8 (2020): 39830-39846.
26. Xu, W. and Zhang, F., "FAST-LIO: A Fast, Robust LiDAR-Inertial Odometry Package by Tightly-Coupled Iterated Kalman Filter," *IEEE Robotics and Automation Letters* 6, no. 2 (2021): 3317-3324.
27. Xu, W., Cai, Y., He, D., Lin, J. et al., "FAST-LIO2: Fast Direct LiDAR-Inertial Odometry," *IEEE Trans. Rob.* 38, no. 4 (2022): 2053-2073.
28. Kanhere and Gao, "Integrity for GPS/LiDAR Fusion Utilizing a RAIM Framework," in *Proceedings of the 31st International Technical*, 2018.
29. Meng, X., Wang, H., and Liu, B., "A Robust Vehicle Localization Approach Based on GNSS/IMU/DMI/LiDAR Sensor Fusion for Autonomous Vehicles," *Sensors* 17(9), 2017, doi:10.3390/s17092140.
30. Shetty and Gao, "Covariance Estimation for gps-Lidar Sensor Fusion for Uavs," in *Proceedings of the 30th International Technical*, 2017.
31. Righi, A.W., Saurin, T.A., and Wachs, P., "A Systematic Literature Review of Resilience Engineering: Research Areas and a Research Agenda Proposal," *Reliab. Eng. Syst. Saf.* 141 (2015): 142-152.
32. Provan, D.J., Woods, D.D., Dekker, S.W.A., and Rae, A.J., "Safety II Professionals: How Resilience Engineering can Transform Safety Practice," *Reliab. Eng. Syst. Saf.* 195 (2020): 106740.
33. Woods, D.D., "Resilience Engineering: Redefining the Culture of Safety and Risk Management," Human Factors and Ergonomics Society Bulletin (2006).
34. Yarveisy, R., Gao, C., and Khan, F., "A Simple Yet Robust Resilience Assessment Metrics," *Reliab. Eng. Syst. Saf.* 197 (2020): 106810.
35. Hoffman, R.R. and Woods, D.D., "Beyond Simon's Slice: Five Fundamental Trade-Offs that Bound the Performance of Macrocognitive Work Systems," *IEEE Intell. Syst.* 26, no. 6 (2011): 67-71.
36. Hollnagel, E., *The ETTO Principle: Efficiency-thoroughness Trade-off : why Things that Go Right Sometimes Go Wrong* (Ashgate Publishing, Ltd., 2009), ISBN:9780754676775
37. Youn, B.D., Hu, C., and Wang, P., "Resilience-Driven System Design of Complex Engineered Systems," *J. Mech. Des.* 133, no. 10 (2011): 101011.
38. Chuang, S., Ou, J.-C., and Ma, H.-P., "Measurement of Resilience Potentials in Emergency Departments: Applications of a Tailored Resilience Assessment Grid," *Saf. Sci.* 121 (2020): 385-393.
39. Furuta, K., "Resilience Engineering," Reflections on the Fukushima Daiichi Nuclear Accident 435–454 (2014).
40. Dosovitskiy, A., Ros, G., Codevilla, F., Lopez, A. et al., "CARLA: An Open Urban Driving Simulator," *arXiv [cs.LG]* (2017).
41. Prescinotti Vivan, G., Goberville, N., Asher, Z., Brown, N. et al., "No Cost Autonomous Vehicle Advancements in CARLA through ROS," SAE Technical Paper 2021-01-0106 (2021), https://doi.org/10.4271/2021-01-0106.
42. Gómez-Huélamo, C., Del Egido, J., Bergasa, L.M., Barea, R., López-Guillén, E., Arango, F., Araluce, J., and López, J., "Train Here, Drive There: Simulating Real-World Use Cases with Fully-Autonomous Driving Architecture in CARLA Simulator," *Advances in Physical Agents II*, Springer International Publishing: 44-59, 2021.
43. Zhang, J. and Singh, S., "Low-Drift and Real-Time Lidar Odometry and Mapping," *Auton. Robots* 41, no. 2 (2017): 401-416.
44. Luo, J., Ying, K., and Bai, J., "Savitzky–Golay Smoothing and Differentiation Filter for Even Number Data," *Signal Processing* 85, no. 7 (2005): 1429-1434.
45. Matisko, P. and Havlena, V., "Noise Covariances Estimation for Kalman Filter Tuning," *IFAC Proceedings Volumes* 43, no. 10 (2010): 31-36.
46. Akhlaghi, S., Zhou, N., and Huang, Z., "Adaptive Adjustment of Noise Covariance in Kalman Filter for Dynamic State Estimation," in *2017 IEEE Power & Energy Society General Meeting*, ieeexplore.ieee.org: 1-5, 2017.

ISSN 0148-7191