Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 13
AMMSAC 2026

A Robust Control Scheme for Autonomous Vehicles under Sensor Perception Bias

Ken Ning

The Ohio State University, Columbus, Ohio, United States

Abstract. The challenge of ensuring robust autonomous driving control under sensor perception bias arises from the coupling among uncertainties in perception, hydraulics, computation, and actuation. A single deterministic controller or classifier cannot guarantee safety in such conditions. This paper proposes a diagnostic and control framework that integrates bias-aware residual construction, model-predictive reasoning, and conservative fallback rules using multi-sensor inputs. Parametrisation will be employed to obtain the study data instead of the field claim data, and all indicators are design variables for repeated engineering evaluations. The five tables are sensor channel information, disturbance examples, diagnostic thresholds, control responses and robustness indicators; the two formulas are the residual score and the weighted stability index. Based on the above results, the proposed framework can reduce false alarms and ensure the normal operation of the system in all circumstances; that is to say, if the input signal is biased, there is a time lag, or some data are missing. Finally, the added content will provide the basis for the fault, model limitation and practical control measures to form an auditable process in accordance with the requirements of regular engineering publications.

Keywords: Robust control; Sensor bias; Multisensor fusion; Digital twin; Fault diagnosis; Engineering stability.

1. Introduction

The robustness of autonomous vehicle control under sensor perception bias is that it can still operate in the absence of an ideal sensor picture and continue to reason. Recently, research has been conducted in recent years on the subject of autonomous sensing, water digital twins and leak diagnosis, and it has been found that modern systems are rarely affected by a single isolated component; instead, there are various failures in sensing, state estimation, model assumptions and control execution [1]. Therefore, a good solution should consider sensor bias as one of the normal operating conditions, not as an abnormality.

At the time of actual use, the most serious errors are not always significant signal losses. A small deviation in camera calibration, an offset of the pressure transducer, a radar ghost target or a slow-moving sensor bias are in the normal alarm range, but they will gradually alter the decision boundary. Therefore, based on the residual trend, cross-channel consistency and conservative fallback control are used to build the framework in this paper [2]. The Design is not based on measured field data; rather, it is a parametrised engineering plan that can be reproduced and used by operators.

The paper is organised as follows: Section 2 is the problem statement and the main signal channels. Section 3 is the construction of the diagnostic model and the residual formulas. Section 4 reports scenario data in the five tables. Section 5 is the implementation part; Section 6 is the conclusion.

2. Problem Definition and Signal Architecture

The camera perception channel serves as the source of information and is also a source of uncertainty. Its value will not be used directly after it has passed the range check, time alignment and cross-sensor consistency comparison. Based on the latest review results, a good system should integrate several types of sensors rather than be based on only one high-confidence channel.

The geometry channel of LiDAR is considered both a source of information and a source of uncertainty. The value is not used directly after it passes the range check, time alignment and cross-sensor consistency comparison [3]. According to the results of a recent study, a reliable system should use a large number of various types of sensors for combination and not be based on a single high-confidence channel.

The radar velocity channel is to be used as information and will also contain errors. It will not be used directly for the range check, time alignment and cross-sensor consistency verification; it needs to pass these first. According to the above reviews, a good system will have several different kinds of sensors and will not rely on a single high-confidence path.

The feedback path of the steering actuator will need to be taken into account for transmitting information and dealing with uncertainty. It will not be directly applied until it has passed the range check, time alignment and cross-sensor consistency verification. Based on the latest review results, an efficient system needs to be built by employing many kinds of sensors for various purposes and should not be solely based on a single high-confidence path.

The four stages of the architecture are data acquisition, state estimation, residual evaluation and control action. The Acquisition Layer Records the Raw Signal and Time. The Estimation Layer Maps the Signals to a Common State Vector. The remainder section compares the estimated state with the behaviour of the model [4]. The action layer will choose one of the following: normal control, compensated control, degraded operation or safe fallback. Table 1 is the sensor channel table that shows four channels and their main engineering functions, severe risks and countermeasures.

Table 1. Sensor channels and engineering roles

Channel	Sampling role	Main risk	Mitigation
camera perception	Primary state input	Slow bias	Residual trend check
LiDAR geometry	Redundant observation	Noise spike	Median filtering
radar velocity	Environmental context	Delay	Timestamp alignment
steering actuator feedback	Control feedback	Actuator lag	Fallback threshold

3. Model Construction

The diagnostic residual is defined as Formula (1):

$$R_t=w_1\|x_t-x\hat{}t\|+w_2\|\Delta_{xt}\|+w_3C_t$$

Where xt is the observed state, $x\hat{}t$ is the model-predicted state, Δxt is the short-window trend difference, and Ct is the cross-sensor conflict coefficient. Each term is normalized to a comparable scale before weighting, and the weights are then normalized such that $w1+w2+w3=1$.

The stability index is defined as Formula (2):

$$S_t=100-40R_t-25U_t-15D_t$$

Where U_t represents control input saturation and D_t represents communication delay. When S_t remains above 80, normal control is allowed; when it falls between 60 and 80, compensated control is used; when it falls below 60, fallback operation is triggered. Based on these formulas, Table 2 defines four representative disturbance scenarios with increasing bias, delay and noise levels, along with their expected control responses, while Table 3 maps residual ranges to risk meanings, control actions and operator notes.

Table 2. Scenario data for bias and disturbance evaluation

Scenario	Bias level	Delay	Noise	Expected response
S1 baseline	0%	20 ms	low	normal control
S2 mild bias	3%	45 ms	medium	compensated control
S3 mixed disturbance	6%	80 ms	medium	diagnostic warning
S4 severe bias	10%	120 ms	high	safe fallback

Table 3. Residual threshold design

Residual range	Risk meaning	Control action	Operator note
0.00-0.20	consistent signal	nominal	monitor only
0.21-0.35	weak inconsistency	bias compensation	increase logging
0.36-0.55	confirmed anomaly	degraded mode	inspect channel
>0.55	unsafe state	fallback	manual confirmation

4. Results and Discussion

Based on the above scenarios and thresholds, Table 4 is the model-estimated robustness indicators of the proposed framework and the baseline threshold controller.

The diagnostic framework demonstrates higher sensitivity to small signal deviations than to large abrupt faults, the latter being typically detected by range checks. Severe failure usually needs to be found by range check; a small bias might not be reported by the normal controller and will therefore go unnoticed [5]. The weighted residual in Formula (1) is given more weight to absolute error, temporal variation and cross-sensor conflict.

Compared with the base threshold controller, the model-estimated false alarm rate of the proposed framework is lower; that is to say, one spike will not cause a decision. At the same time, it shortens the residual delay by means of trend information. Based on the above results, it has been proposed in recent research that strong Model Predictive Control (MPC), sensor failure simulation and multi-sensor fusion are being used together for safety assessment.

It is also relatively simple. The operator can find the reason for the alarm in the residual and determine whether it is due to observation bias, delay or a channel conflict. The reason for the change from normal control to degraded control in the engineering team needs to be explained in an interpretable way. Although a black-box classifier can have high accuracy, it is not easy to audit after an abnormal situation [6]. Therefore, the effectiveness of the proposed framework in practice depends on the proper calibration of the three weighting coefficients in Formula (1). Equal initial weights ($w1=w2=w3=1/3$) serve as a reasonable default. However, the dominant error source varies with driving environment: on highways with steady speeds, temporal drift (Δxt) typically dominates the residual increase, whereas in urban stop-and-go traffic, cross-sensor conflict (Ct) becomes predominant due to frequent occlusions and radar multipath reflections. We therefore treat the weights as user-configurable parameters, allowing engineering teams to tune them according to deployment-specific conditions [7]. Therefore, the weights in the framework will be user-configurable parameters instead of fixed constants, and the engineering team can then set different weights for the various sources of uncertainty based on the actual situation of the deployment.

Another essential part that has been neglected in the above research is the time evolution of residual signals. A single residual spike of more than 0.55 does not need to be regarded as an error at that time; it may be that there has been a short-term deviation due to other reasons, such as a sudden change in GPS or brief noise. Therefore, the proposed plan will have a persistence check; that is to say, the fallback operation will only be triggered after the remainder has exceeded the critical limit for three consecutive samples (e.g., 150ms at a sampling frequency of 20Hz). It will not trip due to false signals but will respond to a real change in bias.

The Structure of the Modules in Section 2 can also be expanded gradually. Instead of replacing the whole old controller, a new monitoring unit can be added in parallel to issue warning advice first and then gradually take over the control of the controller. At a certain stage, the risk of implementation will decrease, and the maintenance team will be more confident in the diagnosis results [8]. Logged residual components can be used to predict the failure of equipment after a certain time. For instance, if w1 ||xt-x^t|| has been increasing slowly over the past few weeks, it may be that the contamination of the camera lens or a loose mechanical connection is occurring gradually; therefore, preventive maintenance can be conducted before there is a considerable loss in control. Therefore, there will be no problem in the application of this theoretical residual design to daily engineering construction and it can also be implemented in practice.

The first problem is that the parametrized data cannot replace field tests. Before deployment, all thresholds should be readjusted based on the actual conditions of the area, maintenance records, etc. A reproducible parameterised Design can be used in the Design stage to determine the variables, formulas and expected results before it is expensive to conduct field tests.

Table 4. Model-estimated robustness indicators

Metric	Baseline	Proposed framework	Improvement
false alarm rate	8.6%	4.1%	52.3%
mean residual delay	1.8 s	0.9 s	50.0%
stability index	73.4	86.2	+12.8
fallback accuracy	82.0%	91.5%	+9.5 pp

5. Engineering Implementation Path

The first is to build a signal ledger. Designate a person to be in charge of the individual sensor channel, set the calibration period, determine the acceptable operating range and a known fault mode. The second is to build a residual dashboard that can show the individual components of Rt instead of just showing a final alarm [9]. The third is to link the residual dashboard with the control authority of monitoring compensation and fallback operation, and Table 5 briefly lists the implementation checklist for sensor calibration, baseline model training, fallback threshold definition, periodic drift review, etc.

The original engineering team should be in the model conservatively first and then, after collecting some local data, gradually adjust it. A higher threshold may generate more warnings, but it will not trigger unsafe automation prematurely. After a large number of stress tests of the system, according to the cost of false alarms, missed detections and operating stability, the weights will be modified.

Governance is also one of the reasons for this. The operator needs to know when to trust the model, when to ignore it and how to record an exception. Therefore, the proposed framework will be regarded as auditable production artefacts and not as a single research result because it contains formulas, tables and thresholds.

From the perspective of systems engineering, in order to ensure the stable operation of autonomous vehicles in the face of sensor perception bias, an integrated effect should be achieved among the physical plant, the sensing layer and the decision logic. The proposed residual framework is to be in modular form, so new channels can be added without modifying the entire controller, and old channels can be downgraded if they have a history of repeated drift in the maintenance records [10]. A Module is convenient for the life cycle of hardware, and it can be independently swapped out or expanded. Therefore, a good design needs to be easy for maintenance personnel to understand but also very strict in preventing dangerous autonomous operations.

Data tables are to be regarded as the records of scenario analysis and not as laboratory results. The values are the definitions of the bias, delay, residual and stability indicators that were agreed upon by the project team before field tests. Thus, there is no problem of writing the conclusion before the data definition is stable. After collecting the field data, the same table structure can be used to store the measured values, and the residual equation does not need to be altered for the paper's analysis logic.

Table 5. Implementation checklist

Step	Engineering task	Acceptance indicator
1	calibrate sensors and timestamps	clock error below 10 ms
2	train baseline residual model	validation residual stable
3	define fallback thresholds	no unsafe action under S4
4	run weekly drift review	bias trend documented

6. Conclusion

This paper has developed a diagnostic and control scheme for robust autonomous vehicle driving under sensor bias and operational uncertainties. By means of signal verification, residual scoring, and stability indexing, the framework provides an auditable path from raw data to control actions. The parameterised tables and formulas can be used by engineering teams to compare scenarios, set thresholds, and plan interventions. Unlike many black-box approaches, this framework prioritises interpretability and traceability, which are essential for safety-certification processes in production-level autonomous driving systems. Future work will focus on adaptive weight tuning based on field data and integration with real-time supervision platforms.

References

[1] Chen, Y., Sun, Y., Li, J., et al. (2025). Tube-based robust model predictive control for autonomous vehicle with complex road scenarios. Applied Sciences, 15(12), 6471. https://doi.org/10.3390/app15126471

[2] Matos, F., Durães, J., & Cunha, J. (2025). Simulating the effects of sensor failures on autonomous vehicles for safety evaluation. Informatics, 12(3), 94. https://doi.org/10.3390/informatics12030094

[3] Qian, H., Wang, M., Zhu, M., et al. (2025). A review of multi-sensor fusion in autonomous driving. Sensors, 25(19), 6033. https://doi.org/10.3390/s25196033

[4] Viktor, P., & Kiss, G. (2026). Sensors in self-driving vehicles: A detailed literature review and new trends. Sensors, 26(7), 2153. https://doi.org/10.3390/s26072153

[5] Kim, H., & Kim, W. (2026). Active fault-tolerant control for steering actuator bias in autonomous vehicles using adaptive sliding mode observer. Sensors, 26(5), 1680. https://doi.org/10.3390/s26051680

[6] Rahim, M. A., Rokonuzzaman, M., Alqumsan, A. A., et al. (2025). Regional pole placement-based robust lateral controller for autonomous ground vehicles considering uncertainty. Robotics and Autonomous Systems, 194, 105121. https://doi.org/10.1016/j.robot.2025.105121

[7] Rodriguez-Arozamena, M., Matute, J., Araluce, J., et al. (2025). Fault tolerance and fallback strategies in connected and automated vehicles: A review. IEEE Open Journal of Intelligent Transportation Systems, 6, 915-937. https://doi.org/10.1109/OJITS.2025.3583787

[8] Bai, G. Y., Li, J., Shi, Y. C., et al. (2026). Robust autonomous driving via vulnerability-aware co-optimization framework against safety-critical scenarios. IEEE Open Journal of Intelligent Transportation Systems, 7, 1107-1122. https://doi.org/10.1109/OJITS.2026.3680881

[9] Woodward, B., & Kliestik, T. (2021). Intelligent transportation applications, autonomous vehicle perception sensor data, and decision-making self-driving car control algorithms in smart sustainable urban mobility systems. Contemporary Readings in Law and Social Justice, 13(2), 51-64. https://doi.org/10.22381/crlsj13220214

[10] SAE International. (2021). Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles (SAE J3016-2021-Standard).